Environmental Research Letters

LETTER

Human contribution to the record-breaking June and July 2019 heatwaves in Western Europe

Robert Vautard[1] , Maarten van Aalst[2], Olivier Boucher[1], Agathe Drouin[3], Karsten Haustein[4], Frank Kreienkamp[5], Geert Jan van Oldenborgh[6] , Friederike E L Otto[4] , Aurélien Ribes[3], Yoann Robin[3], Michel Schneider[3], Jean-Michel Soubeyroux[3], Peter Stott[7], Sonia I Seneviratne[8], Martha M Vogel[8] and Michael Wehner[9]

[1] Institut Pierre-Simon Laplace, Paris, France
[2] University of Twente and Red Cross Red Crescent Climate Centre, The Netherlands
[3] Météo-France, Toulouse, France
[4] Environmental Change Institute, University of Oxford, Oxford, United Kingdom
[5] Deutsche Wetterdienst, Offenbach, Germany
[6] KNMI, de Bilt, The Netherlands
[7] U.K. Met. Office, Exeter, United Kingdom
[8] ETH Zürich, Zürich, Switzerland
[9] Lawrence Berkeley National Laboratory, Berkeley, United States of America
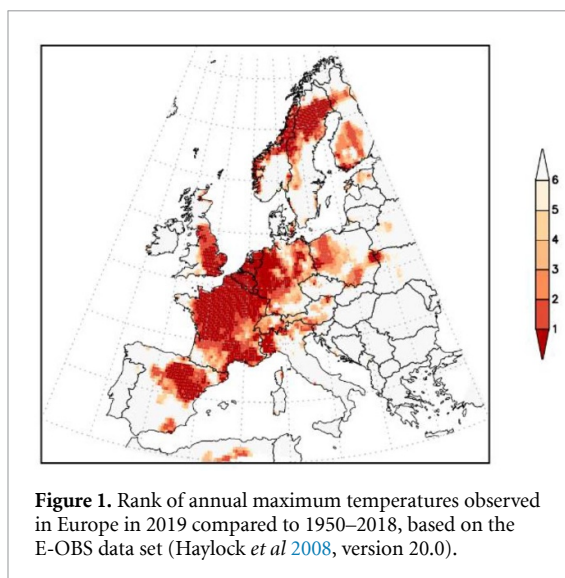
E-mail: robert.vautard@lsce.ipsl.fr

Abstract

Two extreme heatwaves hit Western Europe in the summer of 2019, with historical records broken by more than a degree in many locations, and significant societal impacts, including excess mortality of several thousand people. The extent to which human influence has played a role in the occurrence of these events has been of large interest to scientists, media and decision makers. However, the outstanding nature of these events poses challenges for physical and statistical modeling. Using an unprecedented number of climate model ensembles and statistical extreme value modeling, we demonstrate that these short and intense events would have had extremely small odds in the absence of human-induced climate change, and equivalently frequent events would have been 1.5 °C to 3 °C colder. For instance, in France and in The Netherlands, the July 3-day heatwave has a 50–150-year return period in the current climate and a return period of more than 1000 years without human forcing. The increase in the intensities is larger than the global warming by a factor 2 to 3. Finally, we note that the observed trends are much larger than those in current climate models.

1. Introduction

Two record-breaking heatwaves struck Western Europe in June and July 2019. These heatwaves were recognized as the deadliest disaster of 2019 in the world (CRED 2020). A first event took place in the last week of June 2019. The event broke several historical records at single locations, including France, Switzerland, Austria, Germany, the Czech Republic, Italy and Spain. In particular, the all-time temperature record for any single station in metropolitan France (old record 44.1 °C, Conqueyrac) was broken on June 28 by almost 2 °C with a new record of 46.0 °C, established near the city of Nîmes. In Switzerland,
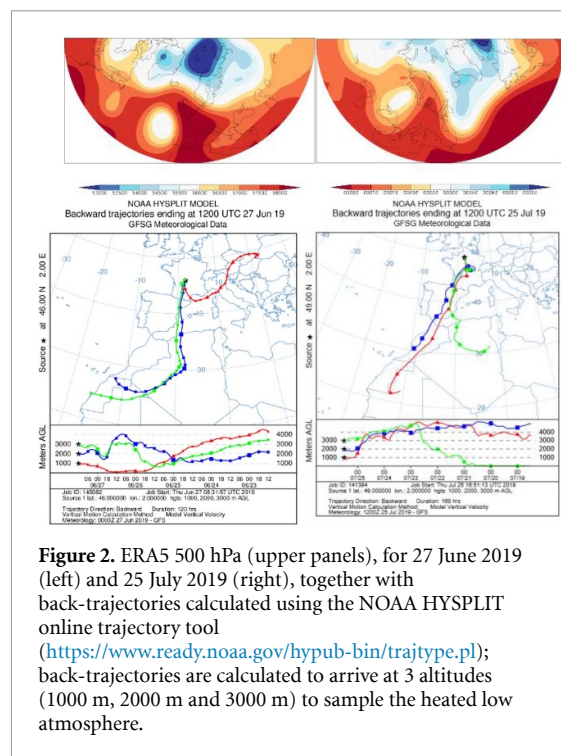
more than 40 stations experienced record daily maximum temperatures for June. In Austria, The Netherlands, Germany and even Europe the whole month of June 2019 was the warmest ever recorded (https://climate.copernicus.eu/record-breaking-temperatures-june). Such extreme heatwaves usually occur in mid-summer, when they have less impact on school days and professional activities than in June or September. In France, due to the heat in June 2019, the government decided to postpone one national school exam, inducing organizational challenges at large scale. In the hottest areas of Europe, a number of wildfires took place, and train tracks were damaged in Switzerland.

Figure 1. Rank of annual maximum temperatures observed in Europe in 2019 compared to 1950–2018, based on the E-OBS data set (Haylock *et al* 2008, version 20.0).

Figure 2. ERA5 500 hPa (upper panels), for 27 June 2019 (left) and 25 July 2019 (right), together with back-trajectories calculated using the NOAA HYSPLIT online trajectory tool (https://www.ready.noaa.gov/hypub-bin/trajtype.pl); back-trajectories are calculated to arrive at 3 altitudes (1000 m, 2000 m and 3000 m) to sample the heated low atmosphere.

A second short (3–4 d) record-breaking heatwave struck Western Europe and Scandinavia at the end of July 2019. Records were broken again, albeit in different areas. In France, the highest amplitudes of the heatwave were found in Northern and Central parts of the country, with records of either 1947 or 2003 broken by a large departure on July 25. For instance, the historical record of Paris (Station Paris-Montsouris) of 40.4 °C became 42.6 °C and a temperature of 43.6 °C was measured in the Paris suburbs. In Belgium and the Netherlands for the first time ever temperatures above 40 °C were observed. In Germany, the historical record of 40.3 °C was surpassed at 14 stations, with one station reaching 42.6 °C (Lingen). In the UK, a new highest ever maximum temperature of 38.7 °C was measured in Cambridge. Further west, where the heatwave was slightly less intense, the record from 1932 (35.1 °C) at the historic Oxford Radcliffe Meteorological Station (continuous measurements for more than 200 years; Burt and Burt 2019) was broken by more than one degree, with a new record maximum temperature of 36.5 °C. These high temperatures caused hundreds of extra deaths in Europe (see section 5).

Taking into account both episodes, the spatial extent of broken historical records is large and includes most areas of France, the Benelux, Switzerland, Germany, the Eastern U.K. and Northern Italy (figure 1). A few days after each of the events, reports of attribution to human influences were made (van Oldenborgh *et al* 2019, Vautard *et al* 2019a). In this article, we collect these results in a single study to draw common conclusions.

Both heatwaves occurred due to a ridge across western Europe, together with a low-pressure system developing offshore the Iberian peninsula, as shown in figures 2(a) and (c). These weather patterns induced intense advection of hot air from North Africa across Spain to France as shown by the NOAA HYSPLIT back-trajectories (figures 2(b) and (d)). Soil conditions across Europe were not anomalously dry prior to the June event, which rules out a large warming amplification by soil-atmosphere feedbacks. By contrast, the July heatwave was accompanied by severe drought conditions in areas such as France, parts of the Netherlands and Germany, which probably contributed to heat development given that dry soils have been shown to cause an additional temperature increase at regional scales due to land-atmosphere feedbacks (e.g. Seneviratne *et al* 2010). Interestingly, this may also point to a link between the two events, since the dry soils prior to the July event were largely the result of the June event. A similar mechanism played a role in the twin 2003 June and August heatwaves in Europe, with the June heatwave likely enhancing the intensity of the August heatwave because of its effect on soil drying (Seneviratne *et al* 2012).

We present here the results of an attribution analysis following the same methodology used in previous analyses (e.g. Otto *et al* 2018, Philip *et al* 2018, Kew *et al* 2019). We also refer to these studies and van Oldenborgh *et al* (2019) and Vautard *et al* (2019a) for a detailed explanation of methods and models.

2. Event definition, observations and trends

In both cases, we use an event definition that illustrates potential impacts on human health, by combining both daytime and nighttime heat and the persistence of the episode as multi-day events have been shown to have disproportionately larger health risks

in Europe (D'Ippoliti, 2010). For the June case we defined the event as the highest 3-day averaged daily mean temperature for the month of June each year (TG3x-Jun) and for the July case we used the all-year 3-day maximum (TG3x). The time span of the indicator almost corresponds or exceeds the length of the heatwave period. While the 3-day average maximum is slightly lower than the single day maximum, we are expecting it to be more sensitive to global warming (Tebaldi and Wehner 2018).

For the June case, the analysis is limited to France where it was most intense while for the July case it is extended to several European countries: France, Germany, The Netherlands and the U.K. These are countries in which a number of temperature records were broken and data were readily availability through study participants or public websites. The locations considered are single weather stations shown in supplementary table 1 (available online at stacks.iop.org/ERL/15/094077/mmedia). In both cases we also used the average over metropolitan France as obtained from the E-OBS data base (Haylock *et al* 2008). It is close to the value of the official French thermal index (also used), which averages temperature over 30 sites well distributed over the metropolitan area and is used to characterize heatwaves and cold spells at the scale of the country.

The rest of the analysis is based on a set of 6 individual weather stations, with the purpose to make the analysis more concrete for effects at local scale, which is the scale relevant for impacts, and also some of the selecte stations had records with a long history. We selected the stations based on the availability of data, the relevance to the heatwave, their series length (at least starting in 1951) and avoidance of urban heat island and irrigation cooling effects, which result in non-climatic trends. The locations considered (Toulouse for June, and Lille-Lesquin, de Bilt, Cambridge, Oxford, Weilerswist-Lommersum for July) all witnessed a historical record both in daily maximum and in 3-day mean temperature (apart from Oxford and Weilerswist-Lommersum where only daily maximum temperatures set a record). Further, the selected stations are either the nearest station with a long enough record to where the study authors reside, or representing a national record. Most of these daily temperature time-series have been quality controlled, and do not exhibit major homogeneity breaks at the monthly time-scale. However, no homogenization procedure is applied, as homogenization of daily time-series remains a challenging task (Mestre *et al* 2011). As a consequence, breaks related to changes in the measurement procedure can still affect these data, and in particular observed trends.

There is a clear trend in observed annual values of the event indicators in each case (see e.g. supplementary figure 1 for the July case stations with TG3x), and the 2019 values represent a large excursion away from the average indicator value which is already a

yearly maximum. The trend in observed series is then quantified using the properties of the fit of a generalized extreme value (GEV) analysis with a covariate (smoothed Global Mean Surface Temperature, GMST) representing an indicator of climate change (from anthropogenic and natural factors) on the position parameter, keeping the scale and shape parameters constant.

For extreme heat, the GEV has a negative shape parameter, which describes an upper bound to the distribution. This bound is however increased by global warming. If the temperature in 2019 is above the bound in 1900, the probability of the event occurring without the warming trend is zero and the probability ratio (PR) formally infinite, subject to the assumptions made and sampling uncertainties. Results for each station are shown in table 1.

In June, as observed in France at the country scale, the exceedance of observed TG3x-Jun has a current-climate return period of 30 year (15 year to 200 year) (table 1). This is roughly 180 times more than it would have been around 1901 (at least 12 times more). The increase in TG3x-Jun since 1901 is estimated to be 4.0 °C (3.0 °C to 5.2 °C). This implies a much higher warming trend in France in June hot extremes compared to that of the average European land summer temperature, which has warmed by about two degrees. For the station of Toulouse, similar results are found.

In July, the change in intensity for similarly likely heatwaves varies between 2 °C and 3.5 °C depending on the location. The return periods range from about 8 years in Oxford to 80 years in Lille. For the metropolitan France average, best estimates of the return periods are of the order of 130 years, even taking the trend into account. In France, Benelux and Germany the return periods for individual stations are relatively similar (60–80 years). In Germany for the selected station we find a return period of 12 years. This relatively low return period could be due to the fact that the station is located slightly on the eastern edge of the affected region and the core event was shorter than 3 d. In the U.K., return periods are shorter because the event was in fact shorter than 3 d and 3-day averages there mix hot temperatures with cooler ones. As seen in table 1, uncertainties on the return period are very large which leads to similarly large uncertainties for the PRs with many cases where an upper bound is infinite. In a few cases, the best fit also gives zero probability in 1900 thus only a lower bound can be given.

3. Models and their evaluation

The observations give a trend, but do not allow to attribute the trend to a cause in the traditional Pearl interpretation (Pearl 1988, Hannart *et al* 2016). For the attribution analysis we used a large set of 8 climate model ensembles including the multi-model

Table 1. Statistical quantities linked to the trend in the observed values of the indicator.

Location	Value 2019 (°C)	Return Period 2019 (Year)	Proba-bility Ratio	Change in intensity (°C)
June case	TG3x-Jun			
France Avg.	Météo-Fr: **27.5**	30*[15–200]*	>12	4.0*[3.0–5.2]*
Toulouse	30.1	54*[16–700]*	>10	4.3*[2.9–6.0]*
July case	TG3x			
France Avg.	E-OBS: 28.2	134 *[>30]*	>5	2.5 *[1.5–3.4]*
	Météo-Fr: 28.7			
Lille Lesquin	29.1	78 *[>20]*	>20	3.5 *[2.3–4.6]*
De Bilt	28.0	60*[20–1400]*	>60	2.9 *[2.0–3.7]*
Cambridge BG	26.0	28 *[11–200]*	250*[9—∞]*	2.3 [1.4–3.4]
Oxford	25.0	7.7 *[4.6–16]*	12*[5–290]*	2.1 *[1.3–2.9]*
Weilerswist- Lommersum	28.7	12 *[6–60]*	430 [18—∞]	3.4 *[2.2–4.9]*

ensembles EURO-CORDEX and CMIP5, single-model ensembles from the CMIP5 and CORDEX generation (EC-EARTH, RACMO), specific attribution ensembles (HadGEM3-A, weather@home) as well as two single-model ensembles from the CMIP6 generation (IPSL-CM6-LR and CNRM-CM6.1) that were available at the date of study. Supplementary table 2 summarizes the characteristics of the model ensembles and references. Note that one of the model ensembles, CNRM-CM6.1, was not used in the June case. Extraction of station points is done using a nearest neighbour method unless specified otherwise. As the grid spacing is smaller than the decorrelation scale for heatwaves the details do not make a difference.

To evaluate these models we test whether the statistics of extreme heat in these models are consistent with the observed statistics. The test consists of fitting the models to the same GEV distribution as in the observations and comparing the scale (σ) and shape (ξ) parameters of the fits to the model data with the parameters of the fits described in the section observational analysis. We do not consider the position parameter (μ) as biases in this parameter can easily be corrected without affecting the overall results. All results from this comparison are shown in supplementary figure 2.

In the June case, The EURO-CORDEX and CMIP5 multi-model ensembles and IPSL-CM6A-LR models overestimate the scale parameter by about 50%, weather@home by a factor two. EC-Earth and the dependent RACMO model would pass a test based on this parameter for both the county scale and the Toulouse site. The shape parameter is generally negative for heatwaves, but in France the parameter is less negative than in most regions (Vautard *et al* 2019b, see their figure 4). Although the uncertainty in shape parameter estimates can be large, models's shape parameters are collectively more negative than observations'. This induces a positive bias in thePR. Taking the two tests together we find that barely any ensemble passes the test that the fit parameters have to be compatible with the parameters describing the observations, in line with issues encountered

for area-averaged heatwaves in the eastern Mediterranean (Kew *et al* 2019). We found no evidence of atmospheric dynamics biases, and the cause of the overestimation of variability is still unknown (see also Leach *et al* 2020).

Hence, we are formally left for the present analysis with no real suitable ensemble to use for the attribution (though we did not check the suitability of each single CMIP5 or EURO-CORDEX models and cannot exclude that some might be suitable for both parameters). Given this, we decided not to give a synthesis result drawn from observations and models as in previous studies but still proceed with analyzing all models, noting that the results are only indicative at best when drawing conclusions.

In the July case (see supplementary table 2), the same conclusions hold regarding models skill as in our analysis of the June heatwave. Models have a too high variability and hence overestimate the scale parameter, sometimes by a large amount (factor 1.5 to 2.5). This is particularly marked for the France average. However, HadGEM3-A, EC-EARTH, IPSL-CM6-LR and CNRM-CM6.1 appear to have a reasonable departure from observations. For the other models the 95% confidence intervals on the scale parameter does not overlap with the confidence interval on the scale parameter from the observations. For individual stations studied here, shape parameters are simulated within observation uncertainties. The discrepancy for the scale parameter is also reduced except for weather@home where variability remains too high.

4. Attribution

The attribution was carried out using different methods for each model ensemble, and also between the June and July cases, due to the nature of the simulations and the availability of methodologies and production teams in real time. For transient simulations and the July case (EC-EARTH, RACMO, EURO-CORDEX, CMIP5, HadGEM3-A, IPSL-CM6-LR, CNRM-CM6.1), estimations are obtained from a GEV fit with the smoothed GMST covariate as an

indicator of climate change and human activities. The training period for the fit is taken as the largest possible period between 1900 and 2018 in order not to include the extreme event itself, as it would lead to a selection bias (see supplementary Material for the reference periods). For some model ensembles the fit was made over a shorter period as the data were not available back to 1900 (such as for RACMO, EURO-CORDEX and HadGEM3-A). For weather@home, due to the large ensemble size, a non-parametric comparison of the observed event in the simulation of the present day climate with the same event in a counterfactual climate performed. In the June case, the same methods were used except for EURO-CORDEX and IPSL-CM6 where a non-parametric comparison was also made. Despite these methodological differences, attribution is made comparable in all cases by comparing return periods or values for the exact same reference dates (1900 and 2019).

For ensembles where bias correction was applied prior to the analysis (the multi-model CMIP5, EURO-CORDEX), the estimation of PRs or intensity changes is made based on events exceeding the observed value of the index. For the non bias-corrected ensembles, the estimation is made based on events with similar return period as in the observations.

A synthesis is made based on observations and the model ensembles that passed the evaluation by weighting the results for the July case, and based on all ensembles for June, but without synthesis between observations and models due to model/observations inconsistency. For June, for each model ensemble, uncertainty only considers sampling uncertainty (representing natural variability). For the July case, the same is represented, but we also added the 'model uncertainty' obtained as the model spread (inter-model variability in the estimate) in addition to each model's sampling uncertainty (open bars in figure 3). This allows a more realistic uncertainty for each model result. In the model/observation synthesis (purple + open bars in figure 3), individual model results are combined with the observed estimate in two ways: a weighted average (by the inverse of the variances) denoted by the colored bar and an unweighted average denoted by the open bar. Model spread is added to the model synthesis without reduction due to the number of models. The unweighted average thus puts more weight on observations.

We present all results for the PRs between the 2019 and 1900 climates and for the change in intensity in figure 3 for the two cases.

4.1. June case
Despite model/observations discrepancies, in June, the observations and almost all models show a large increase in the probability of heatwaves like the one observed in June 2019 (as described by the 3-d mean temperature, both averaged over all of France and in

one specific city, Toulouse). For both the average over France and the Toulouse station we find that the probability has increased by *at least* a factor five (excluding the model with very strong bias in variability). However, observations indicate a much higher factor of a few hundreds. Similarly, the observed trend in temperature of the heat during an event with a similar frequency is around 4 °C, whereas the climate models show a much lower trend (about 2 degrees).

We note that while we are very confident about the positive trend and the fact that the probability has increased by *at least* a factor five. It is impossible to assign one specific number (a 'best guess' based on all models and observations) on the extent of the increase, given the large uncertainties in the observed trends (due to the relatively short time series from 1947–2019) and systematic differences between the representation of extreme heatwaves in the climate models and in the observations.

4.2. July case
For the France average, the heatwave was an event with a return period estimated to be 134 years. As for the June case, except for HadGEM-3A, which has a hot and dry bias, the changes in intensity are systematically underestimated, as they range from 1.1 °C (CNRM-CM6.1) to 1.6 °C (EC-EARTH). By combining information from models and observations, we conclude that the probability of such an event to occur for France has increased by a factor of at least 10 (see the synthesis in figure 3). This factor is very uncertain and could be two orders of magnitude higher. The change in intensity of an equally probable heatwave is between 1.5 degrees and 3 degrees. We found similar numerical results for Lille, with however an estimate of change in intensity higher in the observations, and models predict trend estimates that are consistently lower than observation trends, a fact that needs further investigation beyond the scope of this attribution study. We conclude for these cases that such an event would have had an extremely small probability to occur (less than about once every 1000 years) without climate change in France. Climate change had therefore a major influence to explain such temperatures, making them about 100 times more likely (at least a factor of ten).

For Germany, we analyzed Weilerswist-Lommersum. The changes in temperature are largely underestimated by the models compared to observations by all but the HadGEM3-A model. Based on observations and models, we find that the effect of climate change on heatwave intensity was to elevate temperatures by 1.5 degrees to 3.5 degrees. Because the event was less rare, the PRs are also less extreme. Again all models except HadGEM3-A multi-model ensemble underestimate the trend up to now. This leads to (much) lower PRs in these models than in the observations. The combination of models and

Figure 3. Changes in intensity (left panels) and probability ratios (right panels) obtained for all models and stations (one station per row). From top to bottom: June: France average; Toulouse; July: France average, Lille Lesquin, Weilerswist-Lommersum, De Bilt, Cambridge, Oxford; 95% confidence intervals are given with the red bar for models and blue bar for observations. The 'CMIP5 Météo France' calculation presented here for comparison is based on a different method described in Ribes *et al* (2020). See text for uncertainty estimates.

observations leads to an increase of a factor of about 10 (at least 3).

In de Bilt, the change in temperature of the hottest 3 days of the year is 2.9 °C ± 1.0 °C in the observations and around 1.5 °C in all models except HadGEM3-A (which has a dry and warm bias) and EURO-CORDEX (which has no aerosol changes except for one of the models). The large deviation of HadGEM3-A from the other models gives rise to a large model spread term (white boxes, which increases the uncertainty on the model estimate so that it agrees with the observed trend). Without the HadGEM3-A the models agree well with each other but not with the observations. The overall synthesis provides, as for France, an intensity change in the range of 1.5 degrees to 3 degrees. For the PR, we arbitrarily replaced the infinities by 10 000 year and 100 000 year for the upper bound on the PR of the fit to the observations. As expected the models show (much) lower PRs, due to the higher variability and lower trends. The models with the lowest trends, EC-Earth and RACMO, also give the lowest PR, around

10. Combining models and observations gives a best estimate of 300 with a lower bound of 25.

For U.K. stations, only four (Cambridge) and three (Oxford) model ensembles were kept in the analysis based on our selection criteria. As for the other locations, PRs cover a wide range. Combining observations and models lead us to a PR of ~20 in Cambridge (at least a factor of 3). For Oxford on the other hand, the heatwave was less extreme in TG3x and the PR numbers are lower. Interestingly, the change in intensity is better simulated than for other continental locations. Based on all information we find a rather similar range of temperature trends, from slightly less than 1.5 to ~2.5 degrees. The range is slightly higher for Cambridge than for Oxford.

In all cases bias-corrected ensembles do not appear to exhibit a different behaviour from non bias-corrected ensembles. This stems from the lack of obvious relation between biases and response to anthropogenic changes. Other methodological differences such as of the reference time periods selected and the method used (GEV fit vs.

nonparametric method) do not seem to affect results either, all model ensembles appearing to have similar behavior after standardization to the reference dates (1900 and 2019).

5. Vulnerability, exposure and adaptation

Heatwaves are amongst the deadliest natural disasters facing humanity today and their frequency and intensity is on the rise globally. Combined with other risk factors such as age, certain non-communicable diseases, socio-economic disadvantages, and the urban heat island effect, extreme heat impacts become even more acute with climate change (Kovats and Hajat 2008).

The most striking impacts of heatwaves, deaths, are not fully understood until weeks, months or even years after the initial event. However first estimates have shown that the two heatwaves led to a 50% extra death above normal during the alert periods in France (about 1500 extra deaths) (Santé Publique France 2019). Similar orders of magnitudes, but smaller numbers have been reported for July in The Netherlands (400 extra deaths), Belgium (400 extra deaths), U.K. (200 extra deaths).

Excess mortality is derived from statistical analysis comparing deaths during an extreme heat event to the typical projected number of deaths for the same time period based on historical record. (Mcgregor *et al* 2015) Those at highest risk of death during a heatwave are older people, people with respiratory illnesses, cardiovascular disease and other pre-existing conditions, homeless, socially isolated, urban residents and others (Mcgregor *et al* 2015). Deaths among these populations are not attributable to instances of extreme heat in real time but become apparent through a public health lens following the event.

Compared to the 2003 heatwave in Europe (estimates of 70 000 extra deaths), the numbers appear much reduced. While this would require a specific analysis for a good interpretation of these numbers, adaptation measures could have played a significant role (de Donato *et al* 2015). Following Europe's extreme heat event of 2003 many life saving measures have been put in place. The Netherlands established a 'National Heatwave Action Plan', France established the 'Plane Canicule', in Germany a heatwave warning system has been established and The United Kingdom established 'The Heatwave Plan for England'. Collectively these plans include many proven good practices such as: understanding local thresholds where excess heat becomes deadly, establishing early warning systems, heat protocols for public health and elderly care facilities, bolstering public communications about heat risks, ensuring people have access to cool spaces for a few hours a day, such as cooling centers, fountains and green spaces, and bolstering health systems to be prepared for a surge in demand (Ebi *et al* 2004, Fouillet *et al* 2008, Public Health England 2019)

However while these strong examples exist, on a whole, Europe is still highly vulnerable to heat extremes, with approximately 42% of its population over 65 vulnerable to heat risks (Watts *et al* 2018), and as evidenced by the significant excess mortality during the heat episode discussed in this paper. In addition to life saving measures during a heatwave, it is also crucial to catalyze longer-term efforts to adapt to raising heat risks in Europe (Bittner *et al* 2014). This includes increasing urban green spaces, increasing concentrations of reflective roofs, upgrading building codes to increase passive cooling strategies, and further bolstering health systems to be prepared for excess case loads (Singh *et al* 2019). Adaptation measures are developing in European cities, such as for example by the Paris City, with measures such as: ensuring everyone is within a 7-min walk from a green space with drinking water; incorporating durable water cooling systems into the urban landscape (fountains, reflecting pools, misting systems etc.); planting 20 000 trees; establishing 100 hectares of green roofs; integrating passive cooling measures into new and existing buildings and updating building codes (de Paris 2015).

6. Synthesis and discussion

The heatwaves that struck western Europe were rather short lived (3–4 d), yet very extreme as far as the highest temperatures are concerned: many all-time records were broken in most countries of Western Europe, including historical records exceeded by 1–2 degrees). The events were found to have, under current climate conditions, return periods in the range of 10 years to 150 years depending on locations studied. However, return periods can vary by large amounts from place to place.

Eight model ensembles, including two of the new CMIP6 models, were analyzed using the same event definition (3-d average of mean daily temperature) and methodology, together with observations, for attributing the changes in both intensity and probability of the event at six locations in France, Germany, the Netherlands and U.K.

At all locations analyzed, the combination of observations and model results indicate that temperature trends associated to this extreme event are in the approximate range of 1.5 degrees to 3 degrees, despite the fact that in June observations and model values have a large discrepancy. This indicates that without human-induced climate change heatwaves as exceptional as these one would have had temperatures about 1.5 to 3 degrees lower. Such temperature differences result in a substantial change in morbidity and mortality (Baccini *et al* 2008).

At all locations analyzed, the change in probability of the event is large. In France and the Netherlands, we find changes of at least a factor 10. Without climate change, this event would have been extremely improbable (return period larger than

about 1000 years). For the other locations, changes in probabilities were smaller but still very large, at least a factor of 2–3 for the U.K. station, and 3 for the German station. Differences found across countries are due to several factors, among which processes involved (e.g. soil moisture feedback), as well as the level of observed temperature obtained combined with the sensitivity due to the negative shape of the distribution.

One other finding is a significant difference in the trend in heat extremes between the observations and the models. While the models generally have too large a variability compared to observations, in June the observations have a heavier tail than the models (which have too negative a shape parameter compared to the observations). Specifically for the June 2019 case, the observations show a much stronger extreme temperature trend than the models, with a factor up to two in France. Further work is needed to understand this discrepancy, and determine whether the observed trends might be affected by measurement errors (e.g. homogenization), or if models just fail to capture a real emerging feature.

This analysis triggers several key research questions, which are: (i) What are the physical mechanisms involved in explaining the common model biases in the extremes (i.e. too high variability, too small trends in Europe where trends are large)? (ii) Would one obtain similar results using different statistical methods (only two methods have been applied here), and other conditionings? (iii) Are models improving from CMIP5 to CMIP6 (Wehner *et al* 2020)? (iv) Has climate change induced more atmospheric flows favorable to extreme heat, and, vice versa, for similar flows what are the changes in temperatures, and what design of numerical experiments could inform on this question? (v) How to attribute the co-occurrence of the two extreme heat waves?

While more research is needed to address these yet unsolved questions which cannot be developed in this letter, we speculate that for (i) models may have difficulties to correctly simulate land-atmosphere interactions, resulting in a deficit of skill for the simulation of heatwaves especially in regions where evapotranspiration regimes undergo transitions from energy-limited to soil-moisture limited regimes. This effect is known to be strong in southern France and other regions with Mediterranean climate and is getting stronger in central Europe with global warming (because of decreased evaporational cooling if soil moisture levels become limiting for plants' transpiration; e.g. Seneviratne *et al* (2010), Mueller and Seneviratne (2012). Analyses of CMIP5 ESMs have shown that a subset of the CMIP5 models have a clear tendency to overestimate soil moisture-temperature coupling, which leads to a bias of the overall ensemble (Sippel *et al* 2017, Vogel *et al* 2018). This bias is possibly reduced in the CMIP6 ensemble in Central Europe and Central North America (Seneviratne

and Hauser, submitted). Consistent with the CMIP5 bias, preliminary investigations into the deficits of weather@home have shown that cloud cover is often biased low in the model, which leads to unrealistically high hot extremes due to excessive soil moisture depletion during relatively short periods of simulated blocking. In contrast, low cold extremes in wet years continue to be simulated in weather@home. Another possible dynamical cause is that western Europe may occasionally be influenced by advection of hot and dry air from Spain and North Africa, leading to large excursions of temperature which models might not capture well.

Regarding (ii), Robin and Ribes (2020) analysed the same July event over France using a different statistical approach, and also provide a synthesis between models and observations. They report a return period for that event of 30 year (16–125 year), which is significantly less than in this study. However, the attribution diagnoses (risk ratio and change in intensity) are well consistent with our results, suggesting some robustness in these findings.

Above all, the results of this study show that, when considering extreme heatwaves, models exhibit a large spread of response to current climate change. This calls for a systematic investigation using multiple models to answer the attribution question, and a communication in qualitative terms can be considered. This also calls for a more process based evaluation of models and selection.

Acknowledgments

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID iDs

Robert Vautard ⬤ https://orcid.org/0000-0001-5544-9903
Geert Jan van Oldenborgh ⬤ https://orcid.org/0000-0002-6898-9535
Friederike E L Otto ⬤ https://orcid.org/0000-0001-8166-5917

References

Baccini M, Biggeri A, Accetta G, Kosatsky T, Katsouyanni K, Analitis A and Michelozzi P 2008 Heat effects on mortality in 15 European cities *Epidemiology* **19** 711–9

Bittner M, Matthies E, Dalbokova D and Menne B 2014 Are European countries prepared for the next big heat-wave? *Eur. J. Public Health* **24** 615–9

Burt S and Burt T 2019 *Oxford Weather and Climate since 1767* (Oxford University Press)

CRED 2020 Disaster year in review 2019 (available at: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjp47ml8OLpAhUaC2MBHX9oBFgQFjABegQIAxAB&url=https%3A%2F%2Fcred.be%2Fsites%2Fdefault%2Ffiles%2FCC58.pdf&usg=AOvVaw1NSZJQn_hWEu8jLlbpVR-F)

De' Donato F K *et al* 2015 Changes in the effect of heat on mortality in the last 20 years in nine European cities. results from the PHASE project *Int. J. Environ. Res. Public Health* **12** 15 567–15 583

D'Ippoliti D *et al* 2010 The impact of heat waves on mortality in 9 European cities: results from the EuroHEAT project *Environ. Health* **9** 37

Ebi K L, Teisberg T J, Kalkstein L S, Robinson L and Weiher R F 2004 Heat watch/warning systems save lives: estimated costs and benefits for Philadelphia 1995–98 *Bull. Am. Meteorol. Soc.* **85** 1067–74

Fouillet A, Rey G, Wagner V, Laaidi K, Empereur-Bissonnet P, Le Tertre A and Jougla E 2008 Has the impact of heat waves on mortality changed in France since the European heat wave of summer 2003? A study of the 2006 heat wave *Int. J. Epidemiol.* **37** 309–17

Hannart A, Pearl J, Otto F, Naveau P and Ghil M 2016 Causal counterfactual theory for the attribution of weather and climate-related events *Bull. Am. Meteorol. Soc.* **97** 99–110

Haylock M R, Hofstra N, Klein Tank A M G, Klok E J, Jones P D and New M 2008 A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006 *J. Geophys. Res.* **113** D20119

Kew S F, Philip S Y, Jan van Oldenborgh G, van der Schrier G, Otto F E and Vautard R 2019 The exceptional summer heat wave in southern Europe 2017 *Bull. Am. Meteorol. Soc.* **100** S49-S53

Kovats S and Hajat S 2008 Heat stress and public health: a critical review *Annu. Rev. Public Health* **29** 41–55

Leach N, Li S, Sparrow S, van Oldenborgh G J, Lott F C, Weisheimer A and Allen M R 2020 Anthropogenic influence on the 2018 summer warm spell in Europe: the impact of different spatio-temporal scales. BAMS (available at: http://ametsoc.net/eee/2018/8_Leach0201_w.pdf)

de Paris M 2015 Adaptation strategy: Paris climate & energy action plan (available at: https://api-site.paris.fr/images/76271)

Mcgregor G R, Bessemoulin R, Ebi K and Menne B eds 2015 *Heatwaves and Health: Guidance on Warning-system Development* vol 1142 (Geneva: World Meteorological Organization and World Health Organisation) http://bit.ly/2NbDx4S

Mestre O, Gruber C, Prieur C, Caussinus H and Jourdain S 2011 SPLIDHOM: a method for homogenization of daily temperature observations *J. Appl. Meteorol. Climatol.* **50** 2343–2358.qq

Mueller B and Seneviratne S I 2012 Hot days induced by precipitation deficits at the global scale *Proc. Natl Acad. Sci.* **109** 12398–403

Otto F E L, van der Wiel K, van Oldenborgh G J, Philip S, Kew S F, Uhe P and Cullen H 2018 Climate change increases the probability of heavy rains in Northern England/Southern Scotland like those of storm Desmond - a real-time event attribution revisited *Environ. Res. Lett.* **13** 024006

Pearl J 1988 *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* 2nd edn (Morgan Kaufmann)

Philip S, Kew S F, van Oldenborgh G J, Aalbers E, Vautard R, Otto F, Haustein K, Habets F and Singh R 2018 Validation of a rapid attribution of the may/june 2016 flood-inducing, precipitation in france to climate change *J. Hydrometeorol.* **19** 1881–98

Public Health England 2019 Heatwave plan for England. London, UK, Crown copyright (available at: http://bit.ly/31XMamQ)

Ribes A, Thao S and Cattiaux J 2020 Describing the relationship between a weather event and climate change: a new statistical approach *J. Clim.* **33** 6297–314

Robin Y and Ribes A 2020 Non-stationary GEV analysis for event attribution combining climate models and observations (ASCMO) submitted

Santé Publique France 2019(available at: https://www.santepubliquefrance.fr/determinants-de-sante/climat/fortes-chaleurs-canicule/documents/bulletin-national/systeme-d-alerte-canicule-et-sante.-bilan-de-mortalite-des-episodes-de-chaleur-de-juin-et-juillet-2019)

Seneviratne S I, Corti T, Davin E L, Hirschi M, Jaeger E B, Lehner I, Orlowsky B and Teuling A J 2010 Investigating soil moisture-climate interactions in a changing climate: a review *Earth Sci. Rev.* **99** 125–61

Seneviratne S I *et al* 2012 Swiss pre-alpine Rietholzbach research catchment and lysimeter: analysis of 32-year hydroclimatological time series and 2003 drought *Water Resour. Res.* **48** W06526

Seneviratne S I and Hauser M 2020 Regional climate sensitivity of climate extremes in CMIP6 and CMIP5 multi-model ensembles *Earth's Future* https://doi.org/10.1029/2019EF001474

Singh R, Arrighi J, Ijemba E, Strachan K, Spires M and Kadihasanoglu A 2019 Heatwave guide for cities Red Cross Red Crescent Climate Centre(available at: http://www.climatecentre.org/downloads/files/IFRCGeneva/RCCC%20Heatwave%20Guide%202019%20A4%20RR%20ONLINE%20copy.pdf)

Sippel S, Zscheischler J, Mahecha M D, Orth R, Reichstein M, Vogel M and Seneviratne S I 2017 Refining multi-model projections of temperature extremes by evaluation against land–atmosphere coupling diagnostics *Earth System Dynam.* **8** 387–403

Tebaldi C and Wehner M F 2018 Benefits of mitigation for future heat extremes under RCP4.5 compared to RCP8.5 *Clim. Change* **146** 349–61

van Oldenborgh G J *et al* 2019 Human contribution to record-breaking june 2019 heatwave in France. report from the world weather attribution team (available at: https://www.worldweatherattribution.org/wp-content/uploads/WWA-Science_France_heat_June_2019.pdf)

Vautard R *et al* 2019b Evaluation of the HadGEM3-A simulations in view of climate and weather event human influence attribution in Europe *Clim. Dyn.* **52** 1187–210

Vogel M M, Zscheischler J and Seneviratne S I 2018 Varying soil moisture–atmosphere feedbacks explain divergent temperature extremes and precipitation in central Europe *Earth Syst. Dynam.* **9** 1107-1125

Watts N, Amann M, Arnell N, Ayeb-Karlsson S, Belesova K, Berry H and Campbell-Lendrum D 2018 The 2018 report of the lancet countdown on health and climate change: shaping the health of nations for centuries to come *Lancet* **392** 2479–514

Wehner M, Gleckler P and Lee J 2020 Characterization of long period return values of extreme daily temperature and precipitation in the CMIP6 models: part 1, model evaluation *Weather Clim. Extremes* in revision