LETTER • OPEN ACCESS

Evaluating heat extremes in the UK Climate Projections (UKCP18)

View the article online for updates and enhancements.

ENVIRONMENTAL RESEARCH
LETTERS

LETTER

Evaluating heat extremes in the UK Climate Projections (UKCP18)

A T Kennedy-Asser[1,2] , O Andrews[1,2], D M Mitchell[1,2] and R F Warren[3]

[1] Bristol Research Initiative for the Dynamic Global Environment (BRIDGE), School of Geographical Sciences, University of Bristol, Bristol, United Kingdom
[2] Cabot Institute for the Environment, University of Bristol, Bristol, United Kingdom
[3] Tyndall Centre for Climate Change Research and School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, United Kingdom

E-mail: alan.kennedy@bristol.ac.uk

Keywords: UKCP18, model evaluation, heatwaves, heat stress, climate

Supplementary material for this article is available online

Abstract

In recent years, UK summer heatwaves have resulted in thousands of excess deaths, with both extreme temperatures and high humidity increasing health risks. Here, the UK Climate Projections 2018 (UKCP18) are compared to observational (HadUK-Grid) and reanalysis data (ERA5) to quantify model performance at capturing mean, extremes (95th to 99.5th percentiles) and variability in the climate state and heat stress metrics (simplified wet bulb global temperature, sWBGT; Humidex; apparent temperature). Simulations carried out for UKCP18 generally perform as well as or better than CMIP5 models in reproducing observed spatial patterns of UK climate relating to extreme heat, with RMSE values on average ~30% less than for the CMIP5 models. Increasing spatial resolution in UKCP18 simulations is shown to yield a minor improvement in model performance (RMSE values on average ~5% less) compared to observations, however there is considerable variability between ensemble members within resolution classes. For both UKCP18 and CMIP5 models, model error in capturing characteristics of extreme heat generally reduces when using heat stress metrics with a larger vapour pressure component, such as sWBGT. Finally, the 95th percentile of observed UK summer temperature is shown to have ~60% greater interannual variability than the summer mean over the recent past (1981–2000). This effect is underestimated in UKCP18 models (~33%) compared to HadUK-grid and ERA5. Compared to projected future changes in the global mean temperature, UK summer mean and 95th percentile temperatures are shown in increase at a faster rate than the global mean.

1. Introduction

Climate change is projected to increase extreme heat exposure risk around the world e.g. (Zhao *et al* 2015, Coffel *et al* 2017, Andrews *et al* 2018). UK heatwaves in recent decades have resulted in thousands of excess heat-related deaths (Johnson *et al* 2005, Green *et al* 2016, Public Health England 2019b), morbidity (Arbuthnott and Hajat 2017) and economic disruption across multiple sectors (Costa *et al* 2016). These risks could all increase in the future with climate change (Mitchell *et al* 2016, Vicedo-Cabrera *et al* 2018). As a result, since 2004, the UK has implemented a Heatwave Plan at the national level in England, with a warning system for at risk areas and advice on how to stay safe in hot weather (Public Health England 2019a).

Heat stress occurs under environmental conditions where humans (or other organisms) are unable to maintain stable internal body temperatures. Where capacity to regulate core body temperature is reduced, symptoms such as heat exhaustion and heat stroke can occur, along with other potential medical complications (Kovats and Hajat 2008). For humans, the principal physiological coping method for heat stress is sweating to reduce body temperature by evaporative cooling, meaning heat stress metrics often include both a temperature and humidity component, as, at higher humidity, sweating becomes a less efficient method of cooling the body (Kovats and Hajat 2008, Sherwood and Huber 2010).

There is no universal definition for what constitutes a heatwave that is appropriate for all situations and locations (Perkins and Alexander 2013),

resulting in a variety of definitions used in scientific literature, technical reports and policy plans. A generalised definition would be a 'prolonged period of n or more days with temperatures exceeding a given threshold, x'. Temperatures could be defined in terms of daily maximum, mean or minimum (Perkins and Alexander 2013), or using a metric of heat stress (Chen *et al* 2019). The temperature threshold could be defined in absolute terms that are regionally specific (Mccarthy *et al* 2019a) or, more generally, the threshold could be described in relative terms, such as a percentile of the long-term climate of a region (Russo *et al* 2017, Arnell *et al* 2019).

The latest UK Climate Projections (UKCP18) produced by the UK Meteorological Office (UKMO) (Lowe *et al* 2018) are available at spatial resolutions ranging from ∼60 km with global coverage down to convective permitting 2.2 km for the British Isles region (Murphy *et al* 2019, Kendon *et al* 2019). UKCP18 simulations provide an opportunity to understand climate risks associated with high-impact heatwave events over the UK for the recent past and future (Murphy *et al* 2019) and to explore the interaction between high temperature, high humidity and heat stress based on approaches developed in global studies (Fischer and Knutti 2013, Russo *et al* 2017, Di Napoli *et al* 2020). To date, there are only limited previous studies focussing on temperature changes using this data besides the official UKMO reports (Lowe *et al* 2018, Murphy *et al* 2019, Kendon *et al* 2019). There are a number of reports focussed on specific regions, for example for the Cairgorms and Bristol, both using the probabilistic forecasts (Rivington *et al* 2019, Arup *2020*), and some impacts studies focussed on urban heat islands (Lo *et al* 2020) and heat extremes at selected coastal locations (Edey *et al* 2020; in review).

Here, we evaluate the UKCP18 performance over the period 1981–2018 against observational and reanalysis datasets. Through the model evaluation, this paper addresses the following questions:

- How well do UKCP18 simulations capture the observed variability in heat stress related climate variables?
- What are the past spatiotemporal characteristics of heatwaves in the UK and does UKCP18 capture this variability?
- How are heat extremes projected to change for the UK relative to the global mean?

2. Methods

The UKCP18 model ensemble has four subsets which are assessed throughout this study (Lowe *et al* 2018, Murphy *et al* 2019, Centre for Environmental Data Analysis 2018), summarised in table 1. These include HadGEM3-GC3.05 simulations with global coverage (hereafter 'GCM subset'), along with a nested

regional climate model HadREM3-GA705 (hereafter 'RCM subset') covering the European domain, and within that a nested convective permitting model HadREM3-RA11M (hereafter 'CPM subset') covering the British Isles (Kendon *et al* 2019). A perturbed physics ensemble (PPE) was performed for the GCM and RCM subsets (Murphy *et al* 2019). In these PPEs 47 parameters were modified that relate to convection, boundary layer, gravity wave drag, clouds (radiative and microphysical properties), aerosols and the land surface (a full description of this method is available in Murphy *et al* 2019). All CPM simulations have consistent model physics (i.e. not a PPE), with the differences between simulations driven purely from the input conditions applied from the GCM and RCM simulation that each CPM simulation is nested within (Kendon *et al* 2019).

Although 15 simulations were available in the GCM subset, only 12 RCM and CPM simulations driven by these experiments were included in UKCP18, selected based upon their performance (Murphy *et al* 2019). Only these 12 subset members were assessed in this paper. Although we have shortened the ensemble subset names for brevity, the 12 subset members are named consistently with the convention of Murphy *et al* (2019). Additionally, eight CMIP5 simulations were analysed. These were included as part of the UKCP18 reports and were regridded by the UKMO to the common global grid of HadGEM3-GC3.05 (Lowe *et al* 2018, Murphy *et al* 2019).

The highest resolution CPM has the shortest data window available (1981–2000), and so this is taken as the common climatological baseline period. All UKCP18 simulations follow the CMIP5 historical climate forcing (natural plus anthropogenic) up to 2005 before changing to the RCP8.5 forcing, resulting in slightly greater warming than has been observed over the period 2006–2019 (Murphy *et al* 2019). Here, UK summer is taken as 1st June to 15th September as this is when England's Heatwave Plan is active (Public Health England 2019a).

HadUK-Grid gridded observations (Hollis *et al* 2019) and ERA5 reanalysis data (Hersbach *et al* 2020) were obtained for model validation. Where necessary for comparisons, HadUK-Grid and ERA5 were regridded to the three UKCP18 spatial grids using nearest neighbour interpolation. There was no special treatment for coastal points, which could introduce some cool coastal biases particularly for ERA5, however the impact on UK-wide averages is low. For comparisons with HadUK-Grid, monthly means of daily mean temperature (T_{mean}) and vapour pressure were calculated for UKCP18 and ERA5 (to match the temporal resolution of HadUK-Grid data for these variables). Details on these datasets as well as processing that was carried out for the comparison are summarsied in table 1. Some additional

Table 1. Overview of datasets used in this study.

Dataset	No. of simulations	Horizontal spatial resolution	Temporal coverage	Processing	Reference
UKCP18 GCM subset	12 PPE variants used	~60 km (global)	1981–2080 (daily)	Temporal averaging of T_{mean} and vapour pressure to monthly	Murphy *et al* (2019)
UKCP18 RCM subset	12 PPE variants used	~12 km (Europe)	1981–2080 (daily)	Temporal averaging of T_{mean} and vapour pressure to monthly	Murphy *et al* (2019)
UKCP18 CPM subset	12 driven from GCM and RCM variants	~2.2 km (British Isles)	1981–2000, 2021–2040, 2061–2080 (daily)	Temporal averaging of T_{mean} and vapour pressure to monthly	(Kendon *et al* 2019)
CMIP5	8 different models used	~60 km (global)	1981–2100 (daily)	Temporal averaging of T_{mean} and vapour pressure to monthly	Murphy *et al* (2019)
HadUK-Grid gridded observations	N/A	~60, 12, 1 km (UK land only)	1981–2018 (daily/-monthly)	Regridding from 1 km to 2.2 km	Hollis *et al* (2019)
ERA5 reana-lysis	N/A	0.25° (global)	1981–2018 (hourly)	Regridding to 60, 12 and 2.2 km; temporal averaging to daily (T_{max}) and monthly (T_{mean} and vapour pressure)	(Hersbach *et al* 2020)

datasets were used for analysis and discussion; further details on these and data processing are included in the supplementary information (available online at stacks.iop.org/ERL/16/014039/mmedia).

Three heat stress metrics are calculated here: simplified wet bulb globe temperature (sWBGT) (Australian Bureau of Meteorology 2010, Buzan *et al* 2015), apparent temperature (AT) (Steadman 1979, Zhao *et al* 2015) and Humidex (Masterton and Richardson 1979, Buzan *et al* 2015). sWBGT is an empirical algorithm for estimating the wet bulb globe temperature: a physically based heat stress metric derived for setting occupational safety thresholds (NIOSH 1986, ISO 2017). AT and Humidex were developed by meteorological agencies as more generic algorithms for thermal comfort (Buzan *et al* 2015, Sherwood 2018). It is acknowledged that these metrics will have limitations, not least their actual relevance for defining physiologically based levels of heat stress risk (Sherwood 2018). However, they broadly capture the range of variability from more compicated metrics (Sherwood 2018) and offer comparable results to many other heat stress studies (Zhao *et al* 2015, Buzan *et al* 2015, Mitchell *et al* 2016, Matthews *et al* 2017, Kjellstrom *et al* 2018). Heat stress is estimated for the warmest part of the day using vapour pressure and daily maximum temperature (T_{max}; see supplementary information).

As discussed in section 1, heatwaves can be generalised as *n* or more days exceeding a given temperature threshold, *x*. For each dataset, heatwave events were assessed in terms of their frequency and spatial extent. Four intensity thresholds (*x*) were used: 95th, 98th, 99th and 99.5th, along with four duration thresholds (*n*): 2, 4, 6 and 8 or more consecutive days, for T_{max}

and each of the three heat stress metrics. Percentiles are calculated for each grid cell over the reference period 1981–2018. Heatwave Exposure, $HE_{(x,n)}$, is defined for each metric as the average fractional area of the UK that experiences a heatwave exceeding *x* intensity and *n* duration threshold per summer. This metric allows the modelled heatwave variability to be compared to the observed. For reference, the observed frequencies of events of different example *HE* magnitudes and intensities are shown in supplementary figure 1.

3. Analysis

3.1. Dataset overview
The frequency distribution of average T_{max} over the Greater London region for 1981–2000 summers is shown for all datasets and UKCP18 subsets in figure 1(a). Other UK administrative regions are shown in supplementary figure 2. There is a notable offset of 1 °C–2 °C between ERA5 and HadUK-Grid data due to ERA5 having more muted diurnal variability, and hence lower T_{max} values compared to HadUK-Grid. HadUK-Grid also records lower daily minimum temperatures than ERA5 (shown in supplementary figure 3). This muted variability in ERA5 is likely due to the daily maximum and minimum values being calculated here from hourly data. For the period 1981–2000, the overall distribution of T_{max} values from GCM and RCM subsets fall between HadUK-Grid and ERA5. CPM simulations generally fall closer to the HadUK-Grid observations with a slight warm bias at the highest temperatures, while CMIP5 models are more similar to ERA5 values but
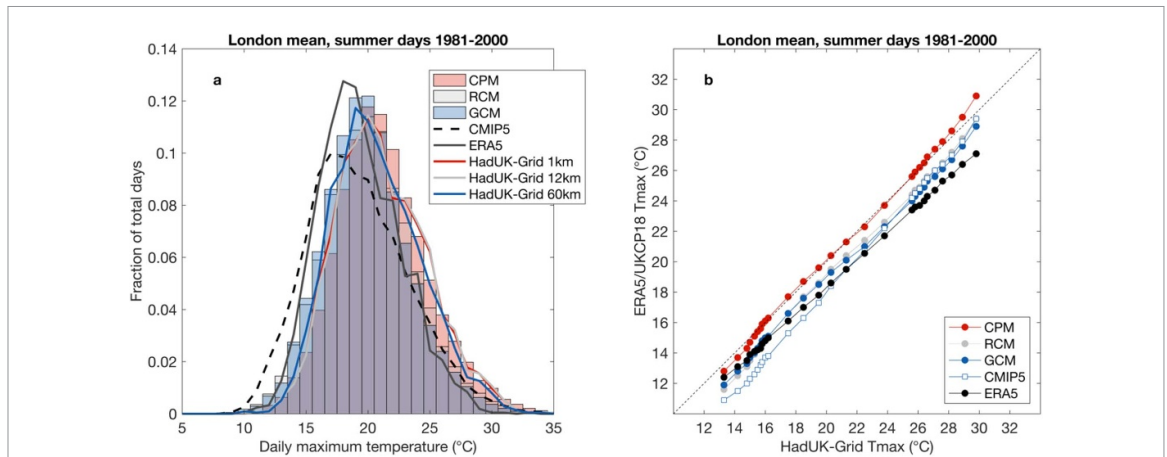
Figure 1. (a) Frequency distribution of T_{max} (°C) for all summer days 1981–2000 in the Greater London region. UKCP18 GCM, RCM and CPM subsets are shown by the coloured bars, the CMIP5 subset by the dashed lines and HadUK-Grid and ERA5 data by the solid lines; (b) P-P plot showing HadUK-Grid observations, ERA5 and UKCP18 subsets T_{max} (°C) for all summer days 1981–2000 in the Greater London region. Percentiles shown are the 1st, 2nd, 3rd, …, 9th, 10th, 20th, 30th,…, 80th, 90th, 91st, 92nd, …, 98th and 99th.
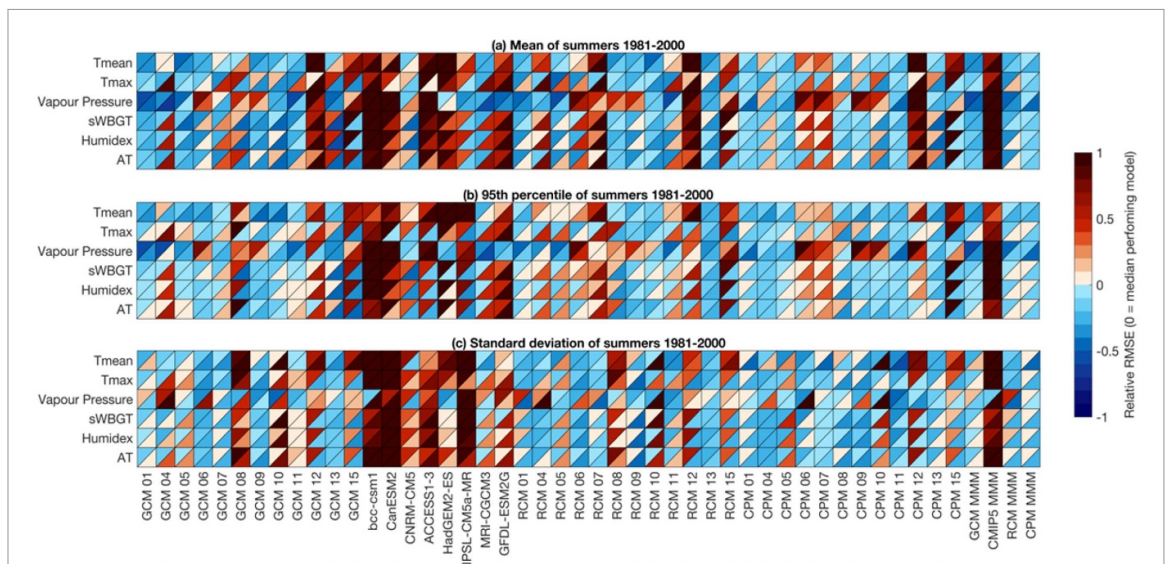

Figure 2. RMSE relative to the median performing simulation over UK land areas for UKCP18 model simulations compared to and normalised relative to ERA5 reanalysis data (upper left triangles) and HadUK-Grid observational data (lower right triangles) for summers during the period 1981–2000. RMSE for climate variables and derived heat stress metrics are shown in terms of the (a) mean, (b) 95th percentile and (c) temporal standard deviation. The multi-model means (MMM) for each subset of the UKCP18 ensemble are shown. Ensemble members are named consistently with Murphy *et al* (2019) with members 02, 03 and 14 not included in the analysis here as they were not available for the RCM and CPM.

with a greater spread and a notable overestimation of the fraction of days at low temperatures.

Resolution dependent differences in HadUK-grid T_{max} distributions are most notable between the 60 km (blue line) and 12 km (grey line) datasets. This is related to errors introduced during aggregation of the data (Hollis *et al* 2019). Supplementary figure 2 shows that in Wales, colder temperatures are captured in the CPM subset compared to the GCM and RCM subsets, likely due to the finer representation of upland areas. Overestimation at the higher end of the T_{max} range shown for Greater London by the CPM subset (figure 1(a)) is also most pronounced across southern and eastern regions of England compared to GCM and RCM subsets (supplementary figure 2).

This regional warm bias is possibly due to reduced cloud cover and soil moisture in the CPM (Kendon *et al* 2019).

Figure 1(b) shows a percentile-percentile (P-P) plot for each of the UKCP18 subsets and ERA5 in comparison to HadUK-Grid observations for the Greater London region. ERA5 does not capture as much variability in daily summer T_{max} compared to HadUK-Grid observations, particularly at higher percentiles. The CPM subset generally shows good agreement with HadUK-Grid, with only a slight cold bias at low temperatures and slight warm bias at high temperatures. The GCM and RCM subsets show a reasonably uniform cold bias across all percentiles. CMIP5 models show greater variability than

HadUK-Grid, with a significant cold bias at low temperatures, but only a minor cold bias at high temperatures. Although figure 1(b) shows there are general biases in the model and reanalysis data compared to the observations, there are no major inflections at the upper percentiles.

3.2. Evaluation of model performance

The relative root mean square error (RMSE) for selected UKCP18 climate variables and derived heat stress metrics compared to HadUK-Grid and ERA5 is shown in figure 2. RMSE is shown for the mean ($RMSE_{mean}$), 95th percentile ($RMSE_{95}$) and standard deviation ($RMSE_{SD}$) of all summer days for each variable or metric. The multi-model mean (MMM) is also shown for each UKCP18 subset. Summary statistics derived from this evaluation are presented in supplementary tables 1–4.

Averaged across all members for each resolution, RMSE scores are reasonably consistent between the GCM, RCM and CPM subsets (e.g. absolute T_{max} $RMSE_{mean}$ range 1.03 °C–1.35 °C; supplementary tables 1 and 2). Taken together across all climate variables and RMSE metrics model performance improves slightly with increasing resolution compared to HadUK-Grid and ERA5, with the CPM having absolute RMSE values ∼2%–5% less on average than the GCM and RCM subsets. The subset with the lowest $RMSE_{mean}$ and $RMSE_{95}$ depends on whether ERA5 or HadUK-Grid is used as a benchmark. For example, the absolute T_{max} $RMSE_{mean}$ compared to ERA5 is least for the RCM (1.031 °C); whereas compared to HadUK-Grid it is least for the CPM (1.089 °C). $RMSE_{SD}$ shows the most consistent improvements with increasing resolution compared to both reference datasets.

The range in relative and absolute RMSE values is generally much greater between individual ensemble members averaged across resolution classes (e.g. T_{max} $RMSE_{mean}$ ranges 0.68 °C–2.36 °C for different ensemble members; supplementary tables 3 and 4). Again, which members perform well depends to some extent on which dataset is used as a benchmark. Certain members perform relatively poorly at all resolutions (e.g. member 12 at all resolutions is poor for most variables in terms of the $RMSE_{mean}$ compared to both ERA5 and HadUK-Grid), while others perform well compared to only one dataset (e.g. member 15 performs well for all RMSE metrics and variables compared to HadUK-Grid, but poorly compared to ERA5). It is likely that the variability between subset members is due to errors introduced in the GCM carrying through into the higher resolution RCM and CPM that it forces (Kendon *et al* 2019). The similarity in errors between CPM simulations and the equivalent GCM and RCM simulations suggest this is the case. As noted in Kendon *et al* (2019), the CPM simulations do not have perturbed

physics, meaning the major differences between these simulations come from the driving model.

There is more variability between CMIP5 model simulations, with this subset typically having absolute RMSE values ∼50% larger than the other subsets (see supplementary table 1). However, there are exceptions; for example CNRM-CM5 performs comparably to many of the GCM simulations.

Zonal mean T_{max} biases in the 95th percentile for GCM, RCM and CPM simulations generally show an exaggerated latitudinal temperature gradient with northern regions too cold and southern regions too hot compared to ERA5 and HadUK-Grid (supplementary figures 4(a)–(c)). HadUK-Grid zonal mean T_{max} is ∼2 °C warmer in general than ERA5, so simulations are more consistent with HadUK-Grid further south, but produce a greater cold bias in the north of the UK. At all resolutions there are some simulations which are anomalously warm or cold across all latitudes. Spatial distributions of summer T_{max} biases highlight the same meridional structure but also reveal larger differences relative to the observations in coastal and upland areas (supplementary figures 4(d)–(f)).

3.3. Heatwave variability

In addition to capturing general climatic characteristics of the UK over the recent past, it is desirable that the UKCP18 simulations also capture realistic magnitude and frequency of heatwave events. The mean summer UK Heatwave Exposure, *HE*, was calculated for the period 1981–2018 for HadUK-Grid, ERA5, GCM, RCM and CMIP5 simulations in terms of T_{max} and sWBGT. As shown in figures 3(a) and (b) for HadUK-Grid, shorter, less intense UK heatwaves are more common than longer, more intense heatwaves (e.g. in terms of T_{max}, $HE_{(95,2)}$ is 0.40, while $HE_{(98,4)}$ is 0.05).

The error in *HE* between the GCM, RCM and CMIP5 simulations relative to HadUK-Grid is shown in figures 3(c)–(h). Model subsets, particularly CMIP5, generally overestimate heatwave events of all durations and intensities compared to HadUK-Grid (and ERA5; not shown) when heatwaves are defined in terms of T_{max}. Given the infrequent nature of longer, more intense events, in relative terms the model *HE* error is generally greater for more extreme events, as shown in supplementary figure 5. Short lived events (2–3 d up to 99th percentile) and less intense events (95th percentile up to 6–7 d) are best simulated by UKCP18.

When heatwaves are defined in terms of sWBGT (or other heat stress metrics; not shown), the general observed characteristics of *HE* remain the same (figure 3(b)), however, the errors between the model subsets relative to the observations change (figures 3(d), (f), (h)). Using sWBGT generally reduces the *HE* overestimation seen for T_{max}, particularly in CMIP5 simulations. For the GCM and RCM subsets, there
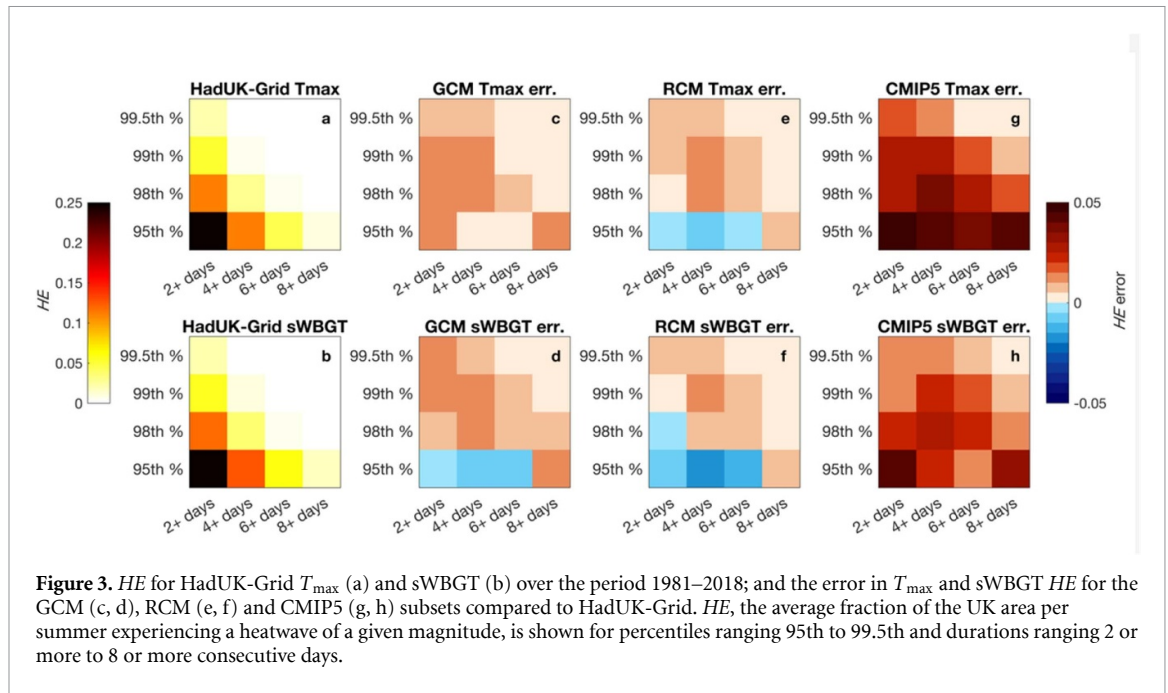
Figure 3. *HE* for HadUK-Grid T_{max} (a) and sWBGT (b) over the period 1981–2018; and the error in T_{max} and sWBGT *HE* for the GCM (c, d), RCM (e, f) and CMIP5 (g, h) subsets compared to HadUK-Grid. *HE*, the average fraction of the UK area per summer experiencing a heatwave of a given magnitude, is shown for percentiles ranging 95th to 99.5th and durations ranging 2 or more to 8 or more consecutive days.
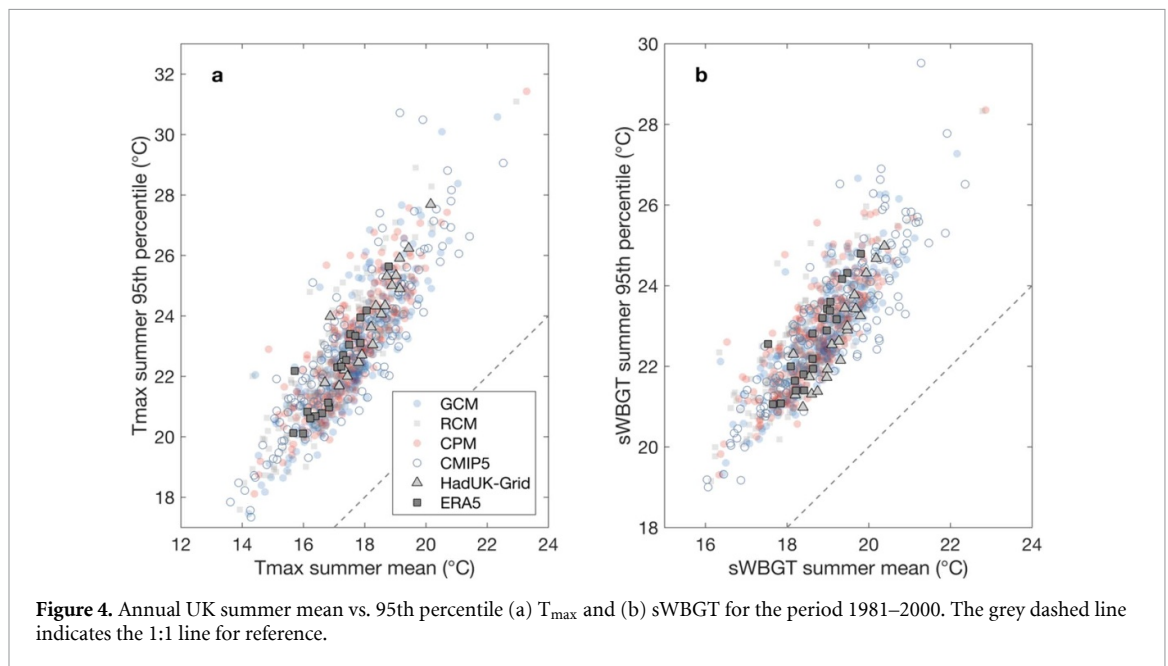

Figure 4. Annual UK summer mean vs. 95th percentile (a) T_{max} and (b) sWBGT for the period 1981–2000. The grey dashed line indicates the 1:1 line for reference.
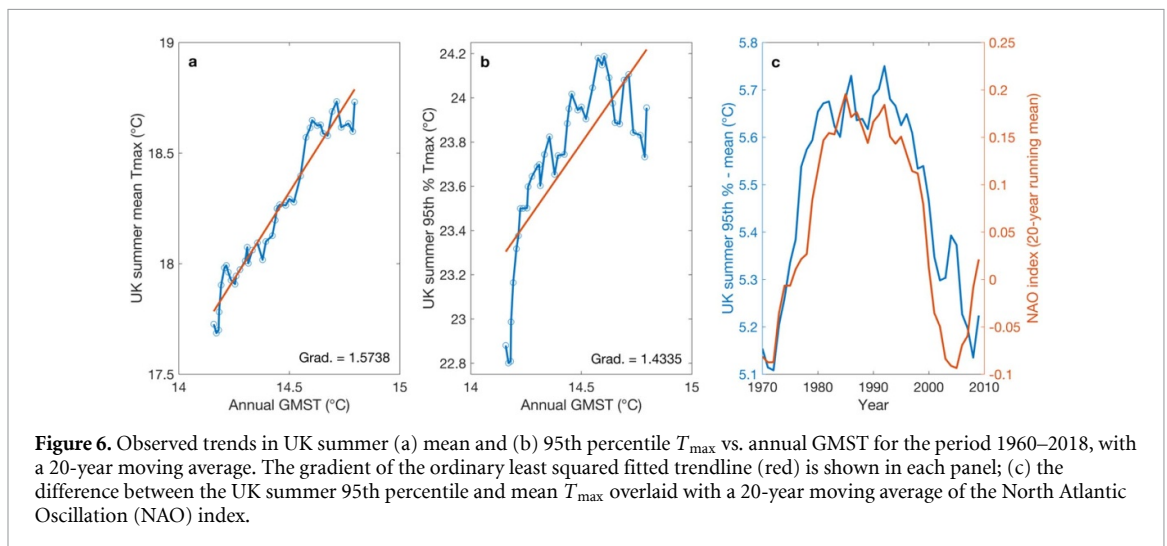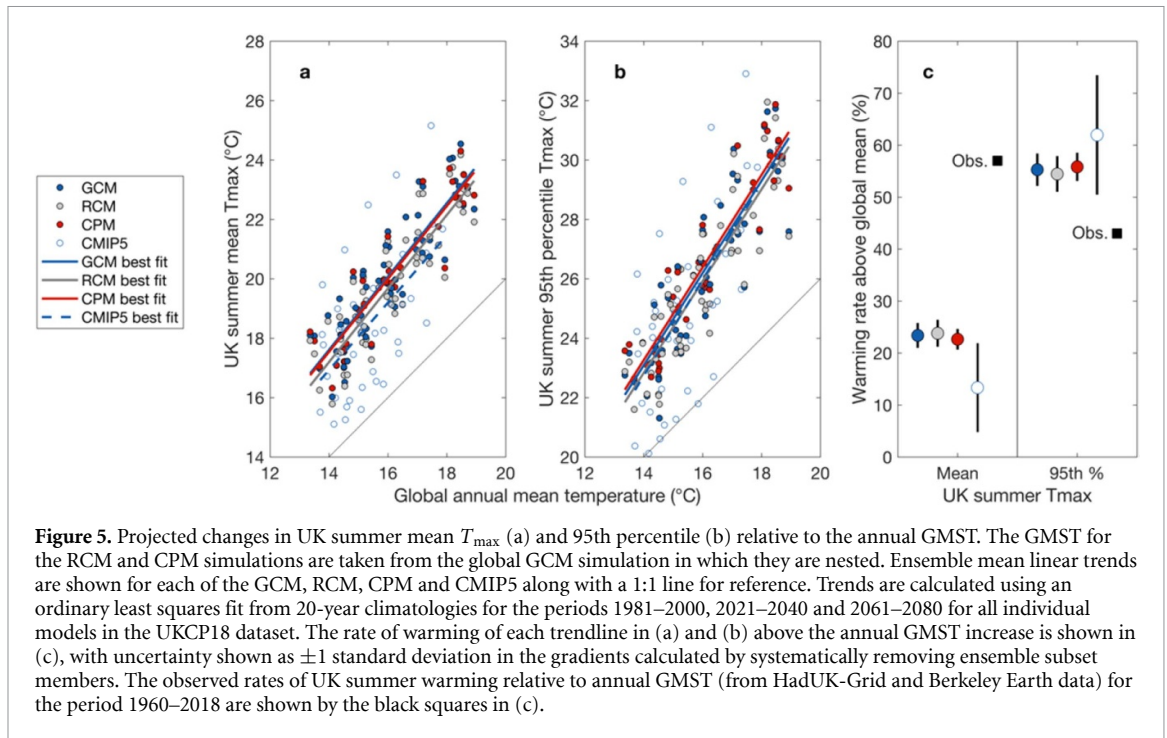
is a small improvement particularly for more intense events, however longer, less intense events (95th percentile for 2–7 d) tend to be underestimated. In general for both T_{max} and sWBGT, the RCM subset has fractionally the lowest error in *HE* and the CMIP5 models have the highest.

3.4. Annual variability and long-term trends
Figure 4 shows the relationship between annual mean and 95th percentile of UK summer T_{max} and sWBGT for each dataset for the period 1981–2000. Details on the linear best fit for each of the datasets shown in figure 4 are listed in supplementary table 5. In terms of T_{max}, shown in figure 4(a), all datasets show a qualitatively similar positive relationship, with the 95th percentile showing amplified interannual

variability compared to the summer mean. Both HadUK-Grid and ERA5 show amplification >60% (e.g. the gradient of the linear fit between the T_{max} mean and 95th percentile is 1.61 for HadUK-Grid). For the GCM, RCM and CPM subsets, the amplification is ∼33% and for the CMIP5 subset the amplification is ∼25%.

In terms of sWBGT, shown in figure 4(b), a similar positive relationship is shown. The amplification of summer 95th percentile sWBGT compared to summer mean sWBGT is again greatest for HadUK-Grid and ERA5 (57%–66%), and again the model subsets substantially underestimate the amplification in summer 95th percentile as a function of summer mean values (ranging from 14%–18%).

Figure 5. Projected changes in UK summer mean T_{max} (a) and 95th percentile (b) relative to the annual GMST. The GMST for the RCM and CPM simulations are taken from the global GCM simulation in which they are nested. Ensemble mean linear trends are shown for each of the GCM, RCM, CPM and CMIP5 along with a 1:1 line for reference. Trends are calculated using an ordinary least squares fit from 20-year climatologies for the periods 1981–2000, 2021–2040 and 2061–2080 for all individual models in the UKCP18 dataset. The rate of warming of each trendline in (a) and (b) above the annual GMST increase is shown in (c), with uncertainty shown as ±1 standard deviation in the gradients calculated by systematically removing ensemble subset members. The observed rates of UK summer warming relative to annual GMST (from HadUK-Grid and Berkeley Earth data) for the period 1960–2018 are shown by the black squares in (c).

Figure 6. Observed trends in UK summer (a) mean and (b) 95th percentile T_{max} vs. annual GMST for the period 1960–2018, with a 20-year moving average. The gradient of the ordinary least squared fitted trendline (red) is shown in each panel; (c) the difference between the UK summer 95th percentile and mean T_{max} overlaid with a 20-year moving average of the North Atlantic Oscillation (NAO) index.

UKCP18 simulations also provide projections under RCP8.5 up to 2080 for all models. The relationship between the projected annual global mean surface temperature (GMST) and both UK summer T_{max} mean and 95th percentile using 20-year climate averages are shown in figure 5. The differences in the rate of UK warming from each subset relative to the GMST are shown in figure 5(c). The magnitude of the long-term change in climate is much smaller than the interannual variability shown in figure 4.

The trendlines in figure 5(a) show that modelled UK summer mean T_{max} increases 13%–24% quicker than the GMST, while modelled UK summer 95th percentile T_{max} increases much quicker than GMST, with model subsets ranging 54%–62% quicker (trendlines in figure 5(b)).

To further assess the robustness of these modelled changes, HadUK-Grid observations of UK summer T_{max} were compared with Berkeley Earth estimates for annual GMST for the period 1960–2018 (Rohde *et al* 2013), as shown in figure 6. Using a 20-year moving climate window, the observations suggest a reasonably strong linear fit between UK summer mean T_{max} and annual GMST (R-squared of 0.94; figure 6(a)), with the UK summer mean warming 57% faster than GMST—considerably more than is found in the model simulations (figure 5(c)).

Observed UK summer 95th percentile was found to warm 43% faster than the GMST (i.e. less than the UK summer mean; figure 6(b)). This is likely related to the observed UK summer T_{max} 95th percentile having strong decadal variability associated with the NAO (figure 6(c)) (Folland *et al* 2009, Sanderson *et al*

2017, National Oceanic and Atmospheric Administration 2020) overprinted on the long-term warming trend. Linear trends presented for the observed rate of UK summer 95th percentile warming in figure 5(c) should therefore be treated as an approximation with high uncertainty.

4. Discussion and conclusions

This study provides a detailed evaluation of heatwaves and heat stress metrics in UKCP18 over the recent past, showing the impact of horizontal resolution on model performance for a range of heat-related variables. GCM (HadGEM3-GC3.05) simulations are found to perform as well or better than CMIP5 models for UK heat-related variables, generally lying within the range of offset between ERA5 and HadUK-Grid observation based estimates. It may be preferable for model simulations to perform better compared to HadUK-Grid or ERA5 depending on what variable (including spatial and temporal extent) is being assessed. For example, T_{max} is derived from hourly data for ERA5 and therefore could underestimate the variability observed in HadUK-Grid, whereas vapour pressure is only available at monthly temporal resolution for HadUK-Grid compared to daily resolution in ERA5.

The increased spatial resolution of the RCM and CPM offer significant potential for improvement of climate projections over the UK for certain applications, for example in capturing variability in coastal and upland regions, as has previously been shown for other high resolution modelling experiments (Qiu *et al* 2020). Additionally, although not assessed here, urban heat island effects are clearly important for UK summer heat extremes (Heaviside *et al* 2016) and could be more appropriately represented at higher resolution.

Generally on a national scale, increasing model spatial resolution yields incremental improvements in model performance. For example, our results indicate that modelled variability improves with increasing resolution, with both a decrease in $RMSE_{SD}$ for UK summer climate variables and a fractional improvement in *HE* in the RCM compared to the GCM. However, increased resolution does not result in major systematic improvements across all aspects of the simulations. There are likely a number of reasons for this.

The majority of large scale variability in the RCM and CPM simulations likely originates from the GCM simulation which is used to drive them, therefore resulting in a similar performance across resolutions when evaluated on a large scale. Additionally, it is important to note that the lower resolution simulations (GCM and CMIP5) were evaluated against HadUK-Grid and ERA5 data that had been regridded to the same lower spatial scale. Therefore,

models were not penalised for smoothing out localised variability that occurs when spatially aggregating data. Future evaluation work should focus on whether higher resolution produces better reproduction of dynamical features that contribute to heatwaves such as blocking events. Kendon *et al* (2019) suggest the RCM and CPM are expected to behave similarly in this regard, however a thorough evaluation of this is beyond the scope of the current study. Soil moisture availability is noted to behave differently between the CPM and RCM (Kendon *et al* 2019) and this can have implications for the amplification of heatwaves (Miralles *et al* 2012, Petch *et al* 2020). Finally, as shown here, it is important to consider whether dominant modes of climate variability such as the NAO are well captured by models when considering UK summer temperature extremes.

Focussing specifically on heatwave events, defined in terms of their intensity and duration, the models show overestimations of past heatwave variability when defined in terms of T_{max}. However, this overestimation is generally reduced when defined in terms of heat stress metrics such as sWBGT, particularly in CMIP5 models. This is consistent with other global studies which show that uncertainty in high percentiles of heat stress is reduced by compensation between the errors of extreme temperature and humidity (Fischer and Knutti 2013). Similar uncertainties in temperature and Humidex extremes over Europe have been previously reported between RCMs and reanalysis (Scoccimarro *et al* 2017). In our study, it is shown that the most extreme events (in terms of intensity and duration) have relatively larger uncertainties, although this could be the result of the relatively small sample size of such events in the historical record. Given the large relative errors for more extreme events, it may be inappropriate to use UKCP18 subsets to analyse future extreme events above a certain magnitude (found to be heatwaves exceeding the 98th percentile for four or more days using the *HE* method presented here). Large ensembles and extreme value statistics could be better suited for assessing more extreme events (Sippel *et al* 2015, Suarez-Gutierrez *et al* 2020).

Previous research using CMIP5 models has shown that extremes in temperature are projected to warm quicker than the annual mean temperature (Seneviratne *et al* 2016). Assessment of T_{max} in UKCP18 showed that for the recent past (1981–2000), the UK summer 95th percentile has amplified interannual variability compared to the UK summer mean T_{max} in good agreement with observational and reanalysis data. The magnitude and impacts of past UK heatwaves have already been shown to have been influenced by greenhouse gas emissions (Mitchell *et al* 2016, Mccarthy *et al* 2019b). When assessing projected long-term UK summer T_{max} in relation to the GMST, in UKCP18 both the summer

mean and 95th percentile temperatures are projected to warm faster than the GMST. This suggests that UK heatwaves could be further amplified with future warming. A process based understanding of the differences between these models and observations should be a priority in future research, and systematic biases should be carefully considered when using the UKCP18 simulations in impacts studies.

Acknowledgments

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: http://data.ceda.acuk/badc/ukcp18/data.

ORCID iDs

A T Kennedy-Asser ⓘ https://orcid.org/0000-0001-5143-8932
D M Mitchell ⓘ https://orcid.org/0000-0002-0117-3486
R F Warren ⓘ https://orcid.org/0000-0002-0122-1599

References

Andrews O, Le Quéré C, Kjellstrom T, Lemke B and Haines A 2018 Implications for workability and survivability in populations exposed to extreme heat under climate change: a modelling study *Lancet Planet. Health* **2** E540–7

Arbuthnott K G and Hajat S 2017 The health effects of hotter summers and heat waves in the population of the United Kingdom: a review of the evidence *Environ. Health* **16** 119

Arnell N W, Lowe J A, Challinor A J and Osborn T J 2019 Global and regional impacts of climate change at different levels of global temperature increase *Clim. Change* **155** 377–91

Arup 2020 *Bristol One City Climate Strategy Preliminary Climate Resilience Assessment* (London: Ove Arup And Partners Ltd) (www.bristolonecity.com/wp-content/uploads/2020/02/Preliminary-climate-resilience-assessment.pdf)

Australian Bureau Of Meteorology 2010 *About the approximation to the WBGT used by the Bureau of Meteorology* (www.bom.gov.au/info/thermal_stress/#wbgt) (Accessed: 4 February 2020)

Buzan J, Oleson K and Huber M 2015 Implementation and comparison of a suite of heat stress metrics within the Community Land Model version 4.5 *Geosci. Model Dev.* **8** 151

Centre for Environmental Data Analysis 2018 CEDA Archive (http://data.ceda.ac.uk/badc/ukcp18/data) (Accessed: 04 November 2020)

Chen X, Li N, Liu J, Zhang Z and Liu Y 2019 Global heat wave hazard considering humidity effects during the 21st century *Int. J. Environ. Res. Public Health* **16** 1513

Coffel E D, Horton R M and De Sherbinin A 2017 Temperature and humidity based projections of a rapid rise in global heat stress exposure during the 21st century *Environ. Res. Lett.* **13** 014001

Costa H, Floater G, Hooyberghs H, Verbeke S and De Ridder K 2016 *Climate Change, Heat Stress and Labour Productivity: A Cost Methodology for City Economies* (London: Centre For Climate Change Economics And Policy) (www.lse.ac.uk/granthaminstitute/wp-content/uploads/2016/07/Working-Paper-248-Costa-et-al.pdf)

Di Napoli C, Barnard C, Prudhomme C, Cloke H L and Pappenberger F 2020 ERA5-HEAT: a global gridded historical dataset of human thermal comfort indices from climate reanalysis *Geosci. Data J.* **00** 1–9

Edey A, Burton R and Gadian A 2020 A case study investigation of summer temperature conditions at two coastal sites in the UK, and analysis of future temperatures and heat wave structures in a warming climate scenario (arXiv:2003.09702)

Fischer E M and Knutti R 2013 Robust projections of combined humidity and temperature extremes *Nat. Clim. Change* **3** 126–30

Folland C K, Knight J, Linderholm H W, Fereday D, Ineson S and Hurrell J W 2009 The Summer North Atlantic Oscillation: past, present, and future *J. Clim.* **22** 1082–103

Green H K, Andrews N, Armstrong B, Bickler G and Pebody R 2016 Mortality during the 2013 heatwave in England–how did it compare to previous heatwaves? A retrospective observational study *Environ. Res.* **147** 343–9

Heaviside C, Vardoulakis S and Cai X-M 2016 Attribution of mortality to the urban heat island during heatwaves in the West Midlands, UK *Environ. Health* **15** S27

Hersbach H *et al* 2020 The ERA5 global reanalysis *Q. J. R. Meteorol. Soc.* **146** 1999–2049

Hollis D, Mccarthy M, Kendon M, Legg T and Simpson I 2019 HadUK-Grid—a new UK dataset of gridded climate observations *Geosci. Data J.* **6** 151–9

ISO 2017 *Ergonomics of the Thermal Environment — Assessment of Heat Stress Using the WBGT (Wet Bulb Globe Temperature) Index* 3 edn (Geneva: International Standards Organization)

Johnson H, Kovats R S, Mcgregor G, Stedman J, Gibbs M, Walton H, Cook L and Black E 2005 The impact of the 2003 heat wave on mortality and hospital admissions in England *Health Stat. Q.* **25** 6–11

Kendon E *et al* 2019 UKCP Convection-permitting model projections: science report Met Office Retrieved from (www.metoffice.gov.uk/pub/data/weather/uk/ukcp18/science-reports/UKCP-Convection-permitting-model-projections-report.pdf)

Kjellstrom T, Freyberg C, Lemke B, Otto M and Briggs D 2018 Estimating population heat exposure and impacts on working people in conjunction with climate change *Int. J. Biometeorol.* **62** 291–306

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

Lo E, Mitchell D, Bohnenstengel S, Collins M, Hawkins E, Hegerl G, Joshi M and Stott P 2020 U.K. Climate Projections: Summer Daytime and Nighttime Urban Heat Island Changes in England's Major Cities *J. Clim.* **33** 9015–30

Lowe J A, Bernie D, Bett P, Bricheno L, Brown S, Calvert D, Clark R, Eagle K, Edwards T and Fosser G 2018 *UKCP18 Science Overview Report* (Exeter: Met Office Hadley Centre)

Masterton J M and Richardson F 1979 *Humidex: A Method of Quantifying Human Discomfort Due to Excessive Heat and Humidity* (Downsview: Environment Canada, Atmospheric Environment)

Matthews T K, Wilby R L and Murphy C 2017 Communicating the deadly consequences of global warming for human heat stress *Proc. Natl Acad. Sci.* **114** 3861–6

Mccarthy M, Armstrong L and Armstrong N 2019a A new heatwave definition for the UK *Weather* **74** 382–7

Mccarthy M, Christidis N, Dunstone N, Fereday D, Kay G, Klein-Tank A, Lowe J, Petch J, Scaife A and Stott P 2019b Drivers of the UK summer heatwave of 2018 *Weather* **74** 390–6

Miralles D G, Van Den Berg M J, Teuling A J and De Jeu R A M 2012 Soil moisture-temperature coupling: a multiscale observational analysis *Geophys. Res. Lett.* **39** L21707

Mitchell D, Heaviside C, Vardoulakis S, Huntingford C, Masato G, Guillod B P, Frumhoff P, Bowery A, Wallom D and Allen M 2016 Attributing human mortality during extreme heat waves to anthropogenic climate change *Environ. Res. Lett.* **11** 074006

Murphy J, Harris G, Sexton D, Kendon E, Bett P, Clark R and Yamazaki K 2019 UKCP18 land projections: science report Met Office (www.metoffice.gov.uk/pub/data/weather/uk/UKCP18/science-reports/UKCP18-land-report.pdf)

National Oceanic and Atmospheric Administration 2020 *North Atlantic Oscillation (NAO)* (National Oceanic and Atmospheric Administration) (Accessed: 7 September 2020) (www.ncdc.noaa.gov/teleconnections/nao/)

NIOSH 1986 *Criteria for a Recommended Standard: Occupational Exposure to Hot Environments: Revised Criteria 1986* (Cincinnati, OH: U.S. Department Of Health And Human Services, Centers For Disease Control, National Institute for Occupational Safety and Health, DHHS (NIOSH))

Perkins S E and Alexander L V 2013 On the measurement of heat waves *J. Clim.* **26** 4500–17

Petch J C, Short C J, Best M J, Mccarthy M, Lewis H W, Vosper S B and Weeks M 2020 Sensitivity of the 2018 UK Summer Heatwave to local sea temperatures and soil moisture *Atmos. Sci. Lett.* **21** E948

Public Health England 2019a *Heatwave Plan For England* (London: PHE Publications)

Public Health England 2019b *PHE Heatwave Mortality Monitoring: Summer 2019* (London: Public Health England)

Qiu L, Im E-S, Hur J and Shim K-M 2020 Added value of very high resolution climate simulations over South Korea using WRF modeling system *Clim. Dyn.* **54** 173–89

Rivington M, Spencer M, Gimona A, Artz R, Wardell-Johnson D and Ball J 2019 *Snow Cover and Climate Change in the Cairngorms National Park: Summary Assessment* (Edinburgh: ClimateXChange) (www.climatexchange.org.uk/media/3900/cxc-snow-cover-and-climate-change-in-the-cairngorms-national-park_1.pdf)

Rohde R, Muller R, Jacobsen R, Muller E, Perlmutter S, Rosenfeld A, Wurtele J, Groom D and Wickham C 2013 A new estimate of the average earth surface land temperature spanning 1753–2011 *Geoinfor. Geostat.: an overview* **1** 1–7

Russo S, Sillmann J and Sterl A 2017 Humid heat waves at different warming levels *Sci. Rep.* **7** 1–7

Sanderson M G, Economou T, Salmon K H and Jones S E 2017 Historical trends and variability in heat waves in the United Kingdom *Atmosphere* **8** 191

Scoccimarro E, Fogli P G and Gualdi S 2017 The role of humidity in determining scenarios of perceived temperature extremes in Europe *Environ. Res. Lett.* **12** 114029

Seneviratne S I, Donat M G, Pitman A J, Knutti R and Wilby R L 2016 Allowable CO_2 emissions based on regional and impact-related climate targets *Nature* **529** 477–83

Sherwood S C 2018 How important is humidity in heat stress? *J. Geophys. Res.: Atmos.* **123** 11808–10

Sherwood S C and Huber M 2010 An adaptability limit to climate change due to heat stress *Proc. Natl Acad. Sci.* **107** 9552

Sippel S, Mitchell D, Black M T, Dittus A J, Harrington L, Schaller N and Otto F E 2015 Combining large model ensembles with extreme value statistics to improve attribution statements of rare events *Weather Clim. Extremes* **9** 25–35

Steadman R G 1979 The assessment of sultriness. Part I: a temperature-humidity index based on human physiology and clothing science *J. Appl. Meteorol.* **18** 861–73

Suarez-Gutierrez L, Müller W A, Li C and Marotzke J 2020 Dynamical and thermodynamical drivers of variability in European summer heat extremes *Clim. Dyn.* **54** 4351–66

Vicedo-Cabrera A M *et al* 2018 Temperature-related mortality impacts under and beyond Paris Agreement climate change scenarios *Clim. Change* **150** 391–402

Zhao Y, Ducharne A, Sultan B, Braconnot P and Vautard R 2015 Estimating heat stress from climate-based indicators: present-day biases and future spreads in the CMIP5 global climate model ensemble *Environ. Res. Lett.* **10** 084013