Detecting Extreme Temperature Events Using Gaussian Mixture Models

Aytaç Paçal[1,2], Birgit Hassler[1], Katja Weigel[2,1], M. Levent Kurnaz [4], Michael F. Wehner[3], Veronika Eyring[1,2]

[1]Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany
[2]University of Bremen, Institute of Environmental Physics (IUP), Bremen, Germany
[3]Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[4]Center for Climate Change and Policy Studies, Boğaziçi University, İstanbul, Turkey

Key Points:

- Extreme temperature events are detected with Gaussian Mixture Models to follow a multimodal rather than a unimodal distribution.
- 10-year temperature extremes will occur 15 times more frequently under 3.0°C future warming.
- Cold days are getting warmer faster than hot days in high latitudes, whereas it is the opposite in low latitudes.

Corresponding author: Aytaç Paçal, `aytac.pacal@dlr.de`

Abstract

Extreme temperature events have traditionally been detected assuming a unimodal distribution of temperature data. We found that surface temperature data can be described more accurately with a multimodal rather than a unimodal distribution. Here, we applied Gaussian Mixture Models (GMM) to daily near-surface maximum air temperature data from the historical and future Coupled Model Intercomparison Project Phase 6 (CMIP6) simulations for 46 land regions defined by the Intergovernmental Panel on Climate Change (IPCC). Using the multimodal distribution, we found that temperature extremes, defined based on daily data in the warmest mode of the GMM distributions, are getting more frequent in all regions. Globally, a 10-year extreme temperature event relative to 1980-2010 conditions will occur 15 times more frequently in the future under 3.0°C of Global Warming Levels (GWL). The frequency increase can be even higher in tropical regions, such that 10-year extreme temperature events will occur almost twice a week. Additionally, we analysed the change in future temperature distributions under different GWL and found that the hot temperatures are increasing faster than cold temperatures in low latitudes, while the cold temperatures are increasing faster than the hot temperatures in high latitudes. The smallest changes in temperature distribution can be found in tropical regions, where the annual temperature range is small. Our method captures the differences in geographical regions and shows that the frequency of extreme events will be even higher than reported in previous studies.

Plain Language Summary

Extreme temperature events are unusual weather conditions with exceptionally low or high temperatures. Traditionally, the temperature range was determined by assuming a single distribution, which describes the frequency of temperatures at a given climate using their mean and variability. This single distribution was then used to detect extreme weather events. In this study, we found that temperature data from reanalyses and climate models can be more accurately described using a mixture of multiple Gaussian distributions. We used the information from this mixture of Gaussians to determine the cold and hot extremes of the distributions. We analysed their change in a future climate and found that hot temperature extremes are getting more frequent in all analyzed regions at a rate that is even higher than found in previous studies. For example, a global 10-year event will occur 15 times more frequently under 3.0°C of global warming. Furthermore, our results show that the temperatures of hot days will increase faster than the temperature of cold days in equatorial regions, while the opposite will occur in polar regions. Extreme hot temperatures will be the new normal in highly populated regions such as the Mediterranean basin.

1 Introduction

Increasing levels of atmospheric carbon dioxide (CO_2) concentration unequivocally transformed the earth's climate (IPCC, 2021). This surplus of CO_2 in the atmosphere contributes to the greenhouse effect, and by increasing the mean and the variability of global temperatures, it amplifies the risk of high-impact temperature extremes (Baker et al., 2018). The effects of anthropogenic global warming led to the emergence of heat extremes that would not have occurred previously (Robinson et al., 2021). This means that unprecedented heat extremes like the 2010 Russian heatwave or the 2021 western North America heatwave would have likely not happened without the warming effect (Rahmstorf & Coumou, 2011; Christidis et al., 2015; Thompson et al., 2022). The latter was found to be a remarkable four standard deviations away from the mean (Thompson et al., 2022). The Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report (AR6) concluded that human influence on the climate system is unequivocal (Eyring et al., 2021) and *virtually certain* to be the main driver of the changes in hot and cold extremes (Seneviratne et al., 2021). It introduced more frequent and intense hot extremes since the 1950s on land areas while a decrease in cold extremes is observed (IPCC, 2021). Several studies found that the duration, frequency, and intensity of extreme events will increase, and extreme events will be introduced at new locations (Seneviratne et al., 2012; Rahmstorf & Coumou, 2011; Kharin et al., 2013; Sillmann, Kharin, Zhang, et al., 2013; Sillmann, Kharin, Zwiers, et al., 2013; Pfleiderer et al., 2019; Perkins-Kirkpatrick & Lewis, 2020; Vogel, Hauser, & Seneviratne, 2020; Raymond et al., 2020; Seneviratne et al., 2021; Mallick et al., 2022). As the number of occurrences of heat extremes like the 2003 European heatwave and their duration increase, the socio-economic burden of climate change poses a threat to societies (Meehl & Tebaldi, 2004; Robine et al., 2008; García-León et al., 2021; Demiroglu et al., 2020; Perera et al., 2020; Seneviratne et al., 2021).

The warming of the climate causes different changes in different regions. Tropics, polar regions and the Middle East and North Africa (MENA) region, are hot spots of notable climate trend shifts (Hao et al., 2018; Y. Zhang et al., 2022). Iyakaremye et al. (2022) have shown that an abrupt shift in the daily maximum temperatures occurred in Africa in the last two decades compared to the previous 20 years, which introduced more frequent and intense hot days. Moreover, climate model projections also show a 1.6°C increase in the annual maximum of daily maximum temperature over Africa in the future, despite a projected 1.5°C global warming level (Iyakaremye et al., 2021). By the end of the century, the frequency and intensity of heatwaves will highly increase in the MENA region under a business-as-usual pathway scenario, which will affect about half of the MENA population (Lelieveld et al., 2016; Zittis et al., 2021; Ozturk et al., 2021). The number of occurrences of exceptionally hot summers, which have 2-4°C hotter temperatures than the long-term average, has also increased from a single event between 1951 and 1980 to five events between 2001 and 2010 in Central and Eastern Europe, where the 2010 heatwave was the hottest and longest event with the largest geographical extent that ever occurred over Europe (Twardosz & Kossowska-Cezak, 2013; Guerreiro et al., 2018). Similarly, other studies also found that the temperature extremes in Europe will increase 20-fold at the end of the century, compared to 1961-1990 (Nikulin et al., 2011; Schär et al., 2004; Barriopedro et al., 2011). Over the Americas, the dry and hot extremes showed an increase both in frequency and spatial scope over the past 122 years (Alizadeh et al., 2020; Cai et al., 2014).

Correctly characterizing the temperature distributions to analyze extreme events is a still-continuing issue as extremes are by definition rare events, and several studies showed that the assumption of distributions or a stationary climate often underestimates the observed heat records (Benestad, 2004; Schär et al., 2004; Anderson & Kostinski, 2010; Fischer & Schär, 2010; Barriopedro et al., 2011; C. Li et al., 2019; Loikith & Neelin, 2019). Thompson et al. (2022) characterized extreme events by calculating a daily extreme index which is the difference between the daily maximum temperature and mean daily maximum temperature divided by the standard deviation. With the assumption of a normal distribution, they found that the 2021 North American heatwave was one of the most extreme events with 4 standard deviations from the mean. Moreover, the authors projected that 20% of the weather risk attribution forecast regions (Stone, 2019) will experience extreme events that are four standard deviations from the means in the future. Other studies found that hot summers will be the norm, i.e. mean temperatures exceed the temperature of the historically hottest summer, within the next 1-2 decades (Mueller et al., 2016; Lewis et al., 2017; Vogel, Hauser, & Seneviratne, 2020; Vogel, Zscheischler, et al., 2020).

Common indices to monitor and analyze climate extremes that are used in the climate community at the moment, such as ETCCDI (the Expert Team on Climate Change Detection and Indices), are mostly based on daily mean near-surface air temperature or daily maximum near-surface air temperature (X. Zhang et al., 2011; Alexander et al., 2006). Two standard approaches to detect extreme events are the percentile-over-threshold

(POT) and the block maxima method. The block maxima method groups data into an equal length of blocks, e.g. month, season, or year, and use the maximum temperature value of each block to fit the data. The POT method defines a threshold, e.g. percentiles, and uses all temperature values above this threshold in the analysis. Choosing the percentiles for defining extremes is not trivial as the temperature thresholds have a strong seasonality and temporal dependence (Huang et al., 2016). The block maxima method is more commonly used in climate studies because of its simplicity with monthly, seasonal or annual block periods for fitting generalized extreme value (GEV) distribution to temperature and precipitation extremes (Kharin et al., 2013; Wang et al., 2016; Paciorek et al., 2018; Wehner et al., 2018; C. Li et al., 2021; IPCC, 2021). The block maxima method, however, does not use all available data, as calculating a single maximum value from a block period throws out the rest of the data. To be approximated by the GEV distribution, the blocks are assumed to be long enough and "max-stable", which means that if you take the maximum of a group of values selected from a specific GEV distribution, the result will be GEV distributed with the same shape parameter (Huang et al., 2016; Ben Alaya et al., 2020). However, these assumptions might not be valid for all possible use cases or all possible variables. For example, GEV is not the best fit for shorter block lengths as the fit improves with increasing block size (Ben Alaya et al., 2020; Wang et al., 2016). Ben Alaya et al. (2020) argued that the identically distributed random variables assumption of extreme value theory might be problematic for extreme precipitation events. They considered a mixture of GEV distributions to fit precipitation data to demonstrate that the mixture distribution could be a potential explanation for the instability of annual maxima. Kollu et al. (2012) tested wind speed characteristics using mixture probability distribution functions (PDF). They found that conventional PDFs are inadequate to describe wind speed distributions compared to the mixture distributions that they used in the study. A mixture of Gaussians was used by Shin et al. (2022) to describe the distribution of the daily thermal comfort index in South Korea, an index that has a strong seasonality. Ice surface temperature data follows a clear multimodal distribution, according to Clarkson et al. (2022). They also found that a unimodal distribution fit is particularly poor at modelling the tail probabilities. Probability distributions with one and two components are called unimodal and bimodal, respectively, whereas distributions with multiple (two or more) components are called multimodal distributions.

The temperature distributions are expected to move towards warmer temperatures and to change their shape with changing means and standard deviations (IPCC, 2021). Also, the assumption of distribution might not be correct for all geographical regions as daily weather variables show a distinct non-Gaussianity (E. M. Volodin & Yurova, 2012; Perron & Sura, 2013; Kodra & Ganguly, 2014; Sardeshmukh et al., 2015; Linz et al., 2018; Tamarin-Brodsky et al., 2019). Furthermore, several studies found that daily mean, daily maximum and real forecast data of 2m temperatures show bimodal features (Grace, 1995; Wilks, 2002; Donat & Alexander, 2012; Cho & Jeong, 2016; Bertossa et al., 2021). These changes, shifts and bimodalities in the temperature distributions affect the probabilities in the tails. As extreme events are rare events that lie in the tails of a distribution, correctly describing the tails is very important for extreme event detection. Even though the block maxima method is widely used in studies which used block sizes large enough to converge asymptotically to GEV distributions, a GEV distribution is not well suited to describe extreme value data when the bimodality is apparent or block sizes are short (Sardeshmukh et al., 2015; Wang et al., 2016; Knoben et al., 2019; Ben Alaya et al., 2020). Therefore, the properties of the entire probability distribution, i.e. mean, standard deviation and shape, are needed to get the tail properties right (Sardeshmukh et al., 2015). A distribution can be described by not only the mean and the standard deviation, but also skewness and kurtosis. Donat and Alexander (2012) found that daily minimum and maximum temperatures have significantly shifted towards higher values and skewed towards the hotter part of the distribution. They highlighted that the changes in extremes are related not only to the means but also to other parameters of the daily temperature distribution. Sardeshmukh and Sura (2009) found a parabolic relationship between kurtosis and skewness that cause the non-Gaussianity of the observed daily weather anomalies. Similarly, Tamarin-Brodsky et al. (2022) used a mixture model with three Gaussians to describe the PDF of near-surface atmospheric temperature to analyze the relationship between kurtosis and skewness, as they are important to explain how the tails of the distribution change. They found that two- and three-Gaussian models are useful to explain the relationship between kurtosis and skewness.

In the study presented here, our approach is to utilize the entire temperature distribution to detect extreme events. We implemented Gaussian Mixture Models (GMM), which describe the probability distribution function of data points as a mixture of Gaussian distributions. We determined the number of Gaussian components in the temperature distribution of each grid cell of 46 land regions defined by the Intergovernmental Panel on Climate Change (IPCC) using daily near-surface maximum air temperature data from the historical and future Coupled Model Intercomparison Project Phase 6 (CMIP6) simulations. This choice was supported by previous studies which found distinct bimodality in daily weather variables (Grace, 1995; Wilks, 2002; Donat & Alexander, 2012; Cho & Jeong, 2016; Bertossa et al., 2021) and was verified by applying the same analysis to the European Centre for Medium-Range Weather Forecasts Reanalysis 5th Generation (ECMWF-ERA5) data for the same historical time period (1980-2010). The parameters from the determined distribution components, namely means, standard deviations and weights, were used to calculate the change in the return period of extreme temperature events between the historical and future periods determined by using global warming levels (GWL). The return period of an event describes the average time between the occurrences of a certain event of a defined size. In this study, we analysed 1-year, 5-year, 10-year and 20-year events, where an n-year event means that the event in question would occur once in every n years. We only calculated return periods equal to or less than the available future data period to prevent overestimating the return periods of extreme events, since GMM distributions are not bounded. Section 2 presents the climate data and warming levels used in this study, as well as the analyzed regions, and explains the methodology of detecting extreme event return periods by using GMM. Section 3 shows our results obtained using the GMM method for all analyzed IPCC land regions, and section 4 finalizes the paper with a summary and discussion.

2 Data and Methodology

2.1 Climate Data

For this study, we used daily near-surface maximum temperatures from the Coupled Model Intercomparison Project Phase 6 (CMIP6), and for which both the historical simulations and the simulations for Shared Socioeconomic Pathways (SSPs) 1-2.6, 2-4.5, 3-7.0 and 5-8.5 scenarios were available (O'Neill et al., 2014; Eyring et al., 2016; O'Neill et al., 2016). Additionally, ECMWF-ERA5 dataset was included for the 31-year time period (1980-2010). Table 1 shows the list of models and their resolutions. The 31-year time period from 1980 to 2010 from historical simulations is used as the base to calculate the return values of extreme temperature events, i.e. 1-year, 5-year, 10-year and 20-year events. The GWL, as introduced in the IPCC AR6 report, are used to assess the changes in future climate in line with the warming levels defined in the Paris Agreement which are compared to the pre-industrial period (IPCC, 2021). The future period for each model is defined as a 20-year period when the central year of the future 20-year running global daily near-surface temperature mean of that model first exceeds a GWL of 1.5°C, 2°C, 3°C, and 4°C between 2015 and 2100, relative to the global daily near-surface temperature mean of the 1850-1900 base period (IPCC, 2021; Hauser et al., 2022). As some datasets did not exceed certain warming levels, they were excluded from the analysis (e.g NOR-ESM2-MM was not used in calculations for 4°C warming under SPP5-8.5, as it

226 did not exceed this level). Figure 1 shows the historical and future GWL periods for each CMIP6 model used in this study.

Table 1. Reanalysis data and CMIP6 models used in this study to detect extreme temperature events. Climate models with spatial resolutions ranging from 50 to 500 km were used in the analyses. The first available ensemble members were chosen. The Renalysis dataset that has a resolution of 25km was regridded to 100 km and used for evaluating modality.

Model	Variant	Resolution	Reference
ECMWF-ERA5	Reanalysis	25 km	(Hersbach et al., 2020)
ACCESS-CM2	r1i1p1f1	250 km	(Dix et al., 2019)
ACCESS-ESM1-5	r1i1p1f1	250 km	(Ziehn et al., 2019)
AWI-CM-1-1-MR	r1i1p1f1	100 km	(Semmler et al., 2018)
BCC-CSM2-MR	r1i1p1f1	100 km	(Wu et al., 2018)
CanESM5	r1i1p1f1	500 km	(Swart et al., 2019)
CNRM-CM6-1	r1i1p1f2	250 km	(Voldoire, 2018)
CNRM-CM6-1-HR	r1i1p1f2	50 km	(Voldoire, 2019)
CNRM-ESM2-1	r1i1p1f2	250 km	(Seferian, 2018)
EC-Earth3	r1i1p1f1	100 km	((EC-Earth), 2019a)
EC-Earth3-CC	r1i1p1f1	100 km	((EC-Earth), 2021)
EC-Earth3-Veg	r1i1p1f1	100 km	((EC-Earth), 2019b)
EC-Earth3-Veg-LR	r1i1p1f1	250 km	((EC-Earth), 2020)
FGOALS-g3	r1i1p1f1	250 km	(L. Li, 2019)
GFDL-ESM4	r1i1p1f1	100 km	(Krasting et al., 2018)
HadGEM3-GC31-LL	r1i1p1f3	250 km	(Ridley et al., 2019a)
HadGEM3-GC31-MM	r1i1p1f3	100 km	(Ridley et al., 2019b)
INM-CM4-8	r1i1p1f1	100 km	(von et al., 2019)
INM-CM5-0	r1i1p1f1	100 km	(E. Volodin et al., 2019)
IPSL-CM6A-LR	r1i1p1f1	250 km	(Boucher et al., 2018)
KACE-1-0-G	r1i1p1f1	250 km	(Byun et al., 2019)
MIROC6	r1i1p1f1	250 km	(Tatebe & Watanabe, 2018)
MIROC-ES2L	r1i1p1f2	500 km	(Hajima et al., 2019)
MPI-ESM1-2-HR	r1i1p1f1	100 km	(Jungclaus et al., 2019)
MPI-ESM1-2-LR	r1i1p1f1	250 km	(Wieners et al., 2019)
MRI-ESM2-0	r1i1p1f1	100 km	(Yukimoto et al., 2019)
NESM3	r1i1p1f1	250 km	(Cao & Wang, 2019)
NorESM2-LM	r1i1p1f1	250 km	(Seland et al., 2019)
NorESM2-MM	r1i1p1f1	100 km	(Bentsen et al., 2019)
UKESM1-0-LL	r1i1p1f2	250 km	(Tang et al., 2019)

227

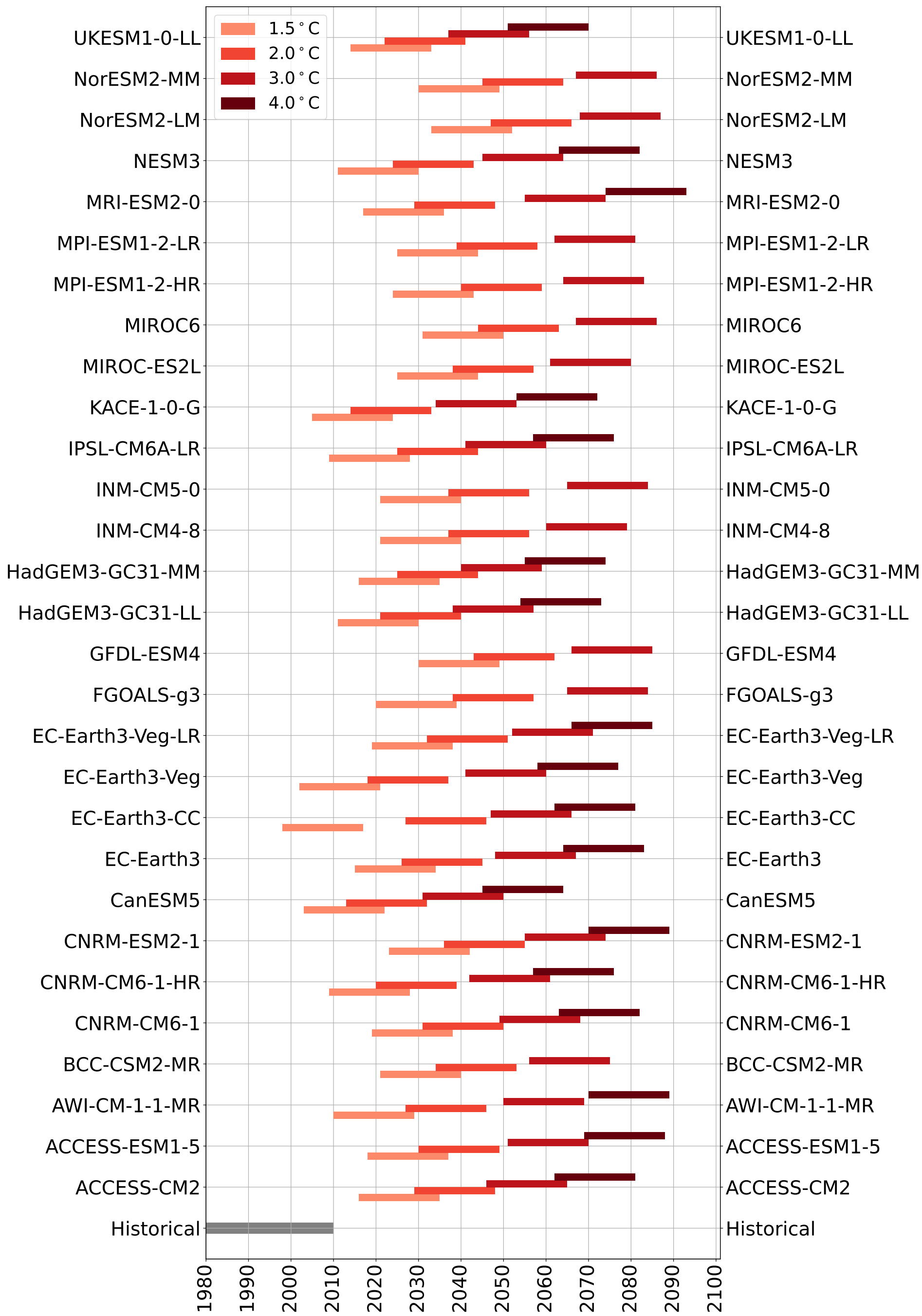


Figure 1. Future periods of the CMIP6 models when the central year of the 20-year running window exceeds global warming levels relative to 1850-1900 base for the SSP5-8.5 scenario (Hauser et al., 2022). The colors in the graph go from light to dark, each color representing a different level of warming 1.5°C, 2°C, 3°C, and 4°C. These levels are expected to be exceeded around 2026, 2040, 2060, and 2070 respectively. 31-year historical base period indicated in gray. Note that different models have different time periods when they exceed the GWL. Future periods for other SSP scenarios are presented in the Supplementary Material Figure S2 to S4.

We extracted daily maximum near-surface air temperature for 31-year historical and 20-year future periods under GWL for each SSP individually for 46 IPCC land regions that are shown in Figure 2 (Iturbide et al., 2020). All data extraction and preprocessing in this study were performed by using the Earth System Model Evaluation Tool (ESMValTool) version 2.5.0, which is an open-source software package for analysing and evaluating model simulations (Eyring et al., 2020; Lauer et al., 2020; Righi et al., 2020; Weigel et al., 2021). We extracted the daily maximum near-surface air temperature from each model for each region using shapefiles provided by IPCC (Iturbide et al., 2020), converted units from Kelvin to Celsius, and created a single spatiotemporal Network Common Data Form (NetCDF) file for each region. The data were then ready to be used in the diagnostic script written in Python where the extreme events and their return periods were analyzed.

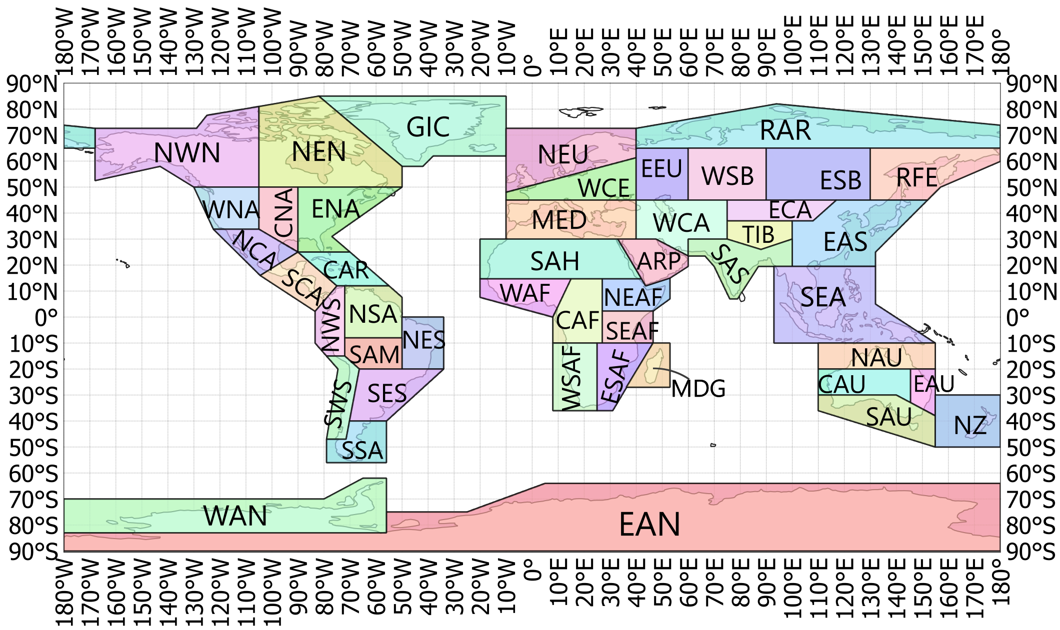


Figure 2. 46 land regions defined by IPCC are used in the study. See Supplementary Material Table S2 for region definitions.

2.2 Return Period Analyses

For the return period calculation of extreme temperature events, i.e. 1-year, 5-year, 10-year and 20-year events, we defined a temperature threshold for an event by calculating the standard deviation distance of the event temperature from the mean temperature in the past, i.e. how many standard deviations away the event temperature *was* from the mean. We then applied this temperature threshold value to the future period but calculated its standard deviation distance from the mean using the parameters from the future distribution, i.e. how many standard deviations away the event temperature *will be* from the mean. To test the underlying distribution shape of the daily near-surface maximum temperature distribution, we first analyzed data from individual grid cells of each climate model. We found that daily maximum near-surface air temperature data in climate grid cells usually do not follow a unimodal distribution, but rather follow a bimodal distribution, a probability distribution composed of two components.

To calculate the return periods of extreme events, we modelled the temperature data from a grid cell as mixtures of Gaussian distributions, rather than a single Gaussian distribution.GMM is a probabilistic model that describes the data points in a population as a mixture of Gaussian distributions with unknown parameters which are the mean, standard deviation and weight of each Gaussian component, five parameters in total for a bimodal distribution. An example goodness-of-fit test for normal distribution, GEV distribution with different shape parameters and GMM distributions on the daily maximum temperature data from a random grid cell is presented in the Supplementary Material Section 1. We used an unsupervised machine-learning package, the "GaussianMixture" package from open-sourced machine-learning library Scikit-learn, to compute the unknown parameters of the Gaussian components in a mixture that generates all observed data points (Pedregosa et al., 2011). We applied this package to the daily maximum near-surface air temperature data in each grid cell of the CMIP6 models. The "GaussianMixture" package first randomly assigns values to component parameters and then uses the expectation-maximization algorithm (EM) to converge their values. EM fits GMM to data by alternating between two steps, Expectation (E) and Maximization (M). In the E step, it randomly assumes components and calculates the probability of each point to be generated by that component. In the M step, it tweaks parameters to maximize the likelihood found in the first step. It also uses the Bayesian Information Criteria (BIC) score, which is used to estimate the goodness-of-fit of a distribution, and which accounts for both the likelihood function and the number of parameters. For a Gaussian mixture with K components, μ_k is the mean, σ_k is the standard deviation, and ω_k is the weight of k^{th} component. Then, the probability distribution function of the GMM with K components would be:

$$p(x) = \sum_{i=1}^{K} \omega_i \mathcal{N}(x \mid \mu_i, \sigma_i) \tag{1}$$

$$\mathcal{N}(x \mid \mu_i, \sigma_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right) \tag{2}$$

$$\sum_{i=1}^{K} \omega_i = 1 \tag{3}$$

In our analysis, we disregarded three or more Gaussian components. This choice was supported by the value of the BIC score, and the fact that increasing the number of components tends to cause overfitting, even though BIC scores penalise adding more parameters. Furthermore, we used the gradient of BIC scores rather than using the lowest score. We selected the number of Gaussian components where the highest gradient change occurs in the BIC scores as the best fit. To further prevent overfitting, we also applied the following unimodality test after estimating the BIC scores: If the BIC score returned a bimodal distribution, then the parameters of the mixture distribution components were used for the unimodality test. As shown in Equation 4, if the difference between the means of Gaussian components was less than or equal to twice the minimum of standard deviations, then unimodal distribution was assumed, otherwise, the bimodal distribution fit for the data was kept. After all these tests and checks, the majority of grid cells showed a clear bimodal distribution. For a bimodal distribution, hereafter we referred to the right (left) Gaussian component as "hot (cold) Gaussian" as shown in Figure 3).

$$|\mu_1 - \mu_2| \leq 2\min(\sigma_1, \sigma_2) \tag{4}$$

First, we grouped grid cells of a region depending on their modality, either unimodal or bimodal, for each CMIP6 model, and calculated the percentages of grid modalities among all grid cells of a region for each CMIP6 model. We then determined the multi-model mean percentages of grid cell modalities of a region as shown in Figure 4. Additionally, we calculated the global multi-model mean percentage of grid cell modalities using all regions and CMIP6 models. We found that globally 88.78% of all grid cells follow a bimodal distribution in the historical period as shown in the white box in the upper centre part of Figure 4. Furthermore, we analysed the ECMWF-ERA5 dataset for

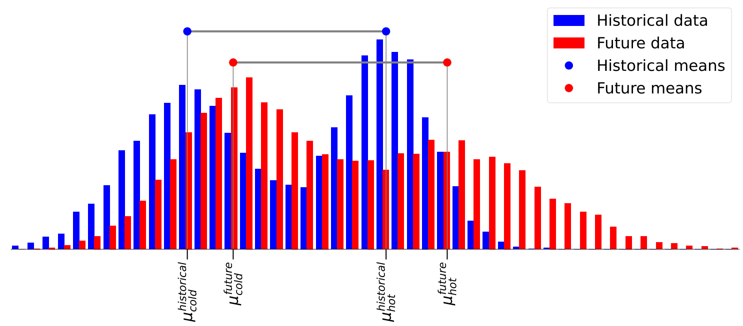


Figure 3. Exemplary bimodal distributions from a hypothetical grid cell for the historical (blue) and future (red) simulations. The parameters of each Gaussian component, means (blue and red dots), standard deviations and weights, determine the distribution shape, and are used in the analyses.

the same historical time period (1980-2010) to confirm whether bimodality is also found in data other than model simulations. We regridded the ECMWF-ERA5 data from a 25-km grid to a 100-km grid using the nearest neighbour method to have a similar resolution as many CMIP6 datasets. ECMWF-ERA5 reanalysis dataset shows similar results to the CMIP6 models: Globally 87.68% of all grid cells in the ECMWF-ERA5 reanalysis dataset follow a bimodal distribution as shown in the white box in the upper centre part of Figure 5, while only 12.32% of them follow a unimodal distribution.

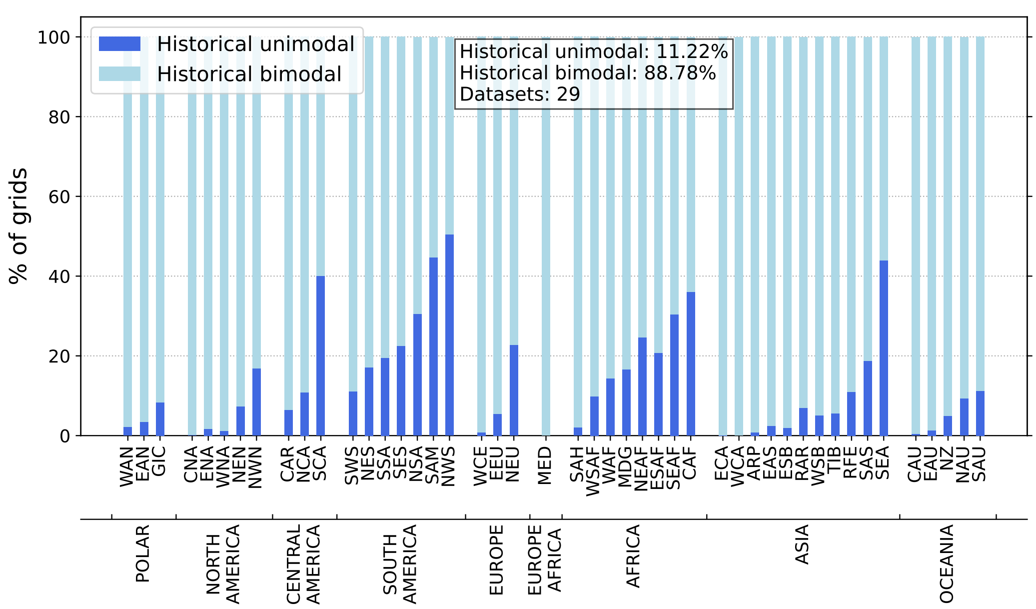


Figure 4. Multi-model mean percentages of grid modalities for the historical period in study regions grouped by continents. Dark and light blue bars show the percentage of grid cells with unimodal or bimodal distribution, respectively, for the historical period of 29 CMIP6 simulations.

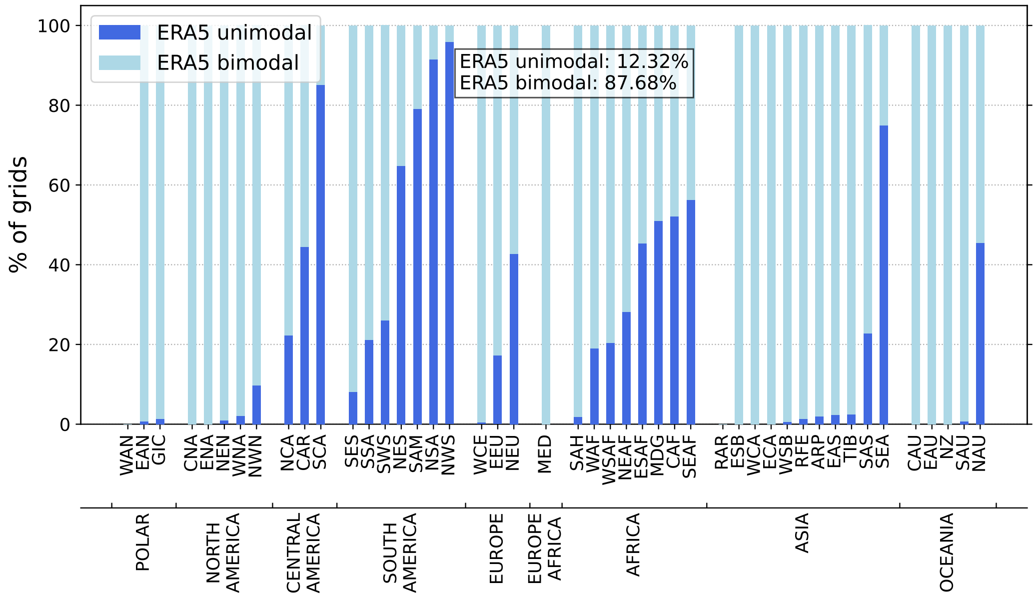


Figure 5. Same as Figure 4 but for ECMWF-ERA5 reanalysis dataset.

Then, the parameters of the hot Gaussian component, $\mu_{hot}^{historical}$, $\sigma_{hot}^{historical}$ and $\omega_{hot}^{historical}$, were used to calculate the change in return periods. We only analysed 1-year, 5-year, 10-year and 20-year events, as GMM are unbounded. One should be careful while calculating the return periods using GMM, as the unbounded tails of Gaussian component could overestimate the probabilities of longer return periods. Therefore, return periods equal to or less than the analysis period were calculated using GMM. The change in return periods is calculated first in each grid cell of a region and then averaged together to produce regional results for each CMIP6 simulation.

For normally distributed data, the expected percentage of the population inside the $\mu \pm d\sigma$ range is defined as

$$E(\mu \pm d\sigma) = \text{erf}\left(\frac{d}{\sqrt{2}}\right) \tag{5}$$

where erf is the error function and d is the standard deviation distance. The approximate expected frequency, f, outside this range is then defined as the return period of an extreme.

$$1\ in\ \frac{1}{1 - \text{erf}\left(\frac{d}{\sqrt{2}}\right)}\ days \tag{6}$$

An $n-year$ event with this definition refers then to a temperature event occurring once in every n "$year$", where "$year$" is defined as the number of days covered by the hot Gaussian component. For example, we can assume that a symmetrical bimodal distribution results in 180 days of cold weather and 180 days of hot weather in a normal 365-day calendar year. For such a symmetric case, a 10-year event would then be a temperature event that occurs once in every 1800 days (10 years$\times 180\frac{days}{year}$). Since we cannot assume a symmetric distribution for grid cells of each model, we calculated the number of days covered by the hot Gaussian component using the component weights and dataset size.

Let $\mathcal{D}$ denote the number of days in L years. Then, a "*year*" in the historical period, $\mid \mathcal{N}(\mu_{hot}^{historical}, \sigma_{hot}^{historical}) \mid$ is defined as

$$\mid \mathcal{N}(\mu_{hot}^{historical}, \sigma_{hot}^{historical}) \mid = \frac{\omega_{hot}^{historical}\mathcal{D}}{L} \tag{7}$$

where $\mu_{hot}^{historical}$ is the mean, $\sigma_{hot}^{historical}$ is the standard deviation and $\omega_{hot}^{historical}$ is the weight of hot Gaussian component. The expected frequency of n-year events in the historical period, $f_n^{historical}$, is then calculated by using the length of a year,

$$f_n^{historical} = n \times \mid \mathcal{N}(\mu_{hot}^{historical}, \sigma_{hot}^{historical}) \mid \quad n = 1, 5, 10, 20 \tag{8}$$

The standard deviation distance of range, $d_n^{historical}$, for an extreme event in the historical period can be calculated by using Equation 6,

$$d_n^{historical} = \text{erf}^{-1}\left(1 - \frac{1}{f_n^{historical}}\right)\sqrt{2} \tag{9}$$

where erf^{-1} is inverse error function. Now, we can calculate a temperature threshold, $\tau_n^{historical}$, for an *n-year* event in the historical period.

$$\tau_n^{historical} = \mu_{hot}^{historical} + d_n^{historical}\sigma_{hot}^{historical} \tag{10}$$

Using this temperature threshold from the historical period, we calculate the standard deviation distance of the temperature threshold of *n-year* event in the future, d_n^{future}, by using the mean μ_{hot}^{future}, and standard deviation σ_{hot}^{future} from the hot Gaussian component of the future distribution.

$$d_n^{future} = \frac{\tau_n^{historical} - \mu_{hot}^{future}}{\sigma_{hot}^{future}} \tag{11}$$

$$f_{\dot{n}}^{future} = \frac{1}{1 - \text{erf}\left(\frac{d_n^{future}}{\sqrt{2}}\right)} \tag{12}$$

$$\tag{13}$$

Finally, the new value of the return period in the future $\dot{n}$, i.e. $\dot{n}$-year event, is calculated by using Equation 8

$$\dot{n} = \frac{f_{\dot{n}}^{future}}{|\mathcal{N}(\mu_{hot}^{future}, \sigma_{hot}^{future})|} \tag{14}$$

where $\mathcal{N}(\mu_{hot}^{future}, \sigma_{hot}^{future})|$ is length of a "*year*" in the future period.

With this method, we can also analyse if and how much the Gaussian components will shift in the future relative to the historical period. We defined ΔT, as the difference in differences between the means of cold and hot Gaussian components as shown in Equation 15:

$$\Delta T = \delta T_{cold} - \delta T_{hot} \tag{15}$$

$$\delta T_{cold} = \mu_{cold}^{future} - \mu_{cold}^{historical} \tag{16}$$

$$\delta T_{hot} = \mu_{hot}^{future} - \mu_{hot}^{historical} \tag{17}$$

In Figure 3 this change in hot and cold Gaussian means is schematically illustrated. Assuming the future means of Gaussian components are higher than the historical periods, δT_{cold} and δT_{hot} will always be positive. Therefore, a negative ΔT means that the peaks are diverging in the future: the hot Gaussian moves toward warmer temperatures faster than the cold Gaussian, which increases the frequency of hot extremes and induces an overall warmer climate. A positive ΔT means that the peaks are converging: the cold Gaussian moves closer to the hot Gaussian, which increases the number of days with warmer temperatures in the colder mode.

3 Results

First, we checked the change in the percentage of modalities from the present to the future time periods. For this, we analyzed the modality of the temperature data from each individual grid cell of an IPCC land region by counting the number of grid cells with each modality. We found that the percentages of grid cells with bimodal distributions stay almost the same under different warming levels. As some of the CMIP6 datasets do not exceed certain warming levels, the number of datasets are not identical for the historical and future period and therefore affect the change in percentages. We analysed modalities of grid cells under different GWL for all SSP scenarios but we only present SSP5-8.5 results here, as the SSP5-8.5 scenario had data from 29 CMIP6 models and the GWL are scenario independent. Globally, almost 90% of all grid cells follow a bimodal distribution as shown in Figure 4 for the historical period, Figure 5 for the reanalysis data and Figure 6 for GWL 3.0°C for different regions grouped by continents (See Supplementary Material Figure S5 to S7 for other warming levels). East Antarctica (EAN) region is not included in the analysis because it is composed of many grid cells near the pole, causing numerical problems. Global averages and the number of datasets are shown in the white box in the upper centre part of each figure. In the historical period, the grid cells in tropical and sub-tropical regions have slightly higher percentages of unimodal distributions compared to higher latitude regions. However, regions still mostly follow a bimodal distribution as shown in Figure 4. The multi-model mean percentage of unimodal distributions does not exceed 50% of grid cells in any of the regions, except in N.W.South-America (NWS) and South-American-Monsoon (SAM) regions where 52.5% and 51.2% of the grid cells follow a unimodal distribution, respectively, in the historical period. The higher percentage of unimodal distributions in lower latitudes is consistent with tropical climate features, where hot temperatures are observed all year round and the annual temperature range is small (Richter, 2016; Beck et al., 2018). This climate type is therefore expected to likely experience a temperature distribution close to a single Gaussian. All grid cells (99.9%) in CMIP6 models follow a bimodal distribution in the Mediterranean (MED) region in the historical period and under all future periods. In polar regions, more than 90% of the grid cells follow a bimodal distribution in the historical period. The percentage of grid cells with unimodal distributions in polar regions slightly increases under future global warming levels.

As previously mentioned in Section 2.2, large values of ΔT (see Equation 15) will cause the temperature distribution to change its modality for future GWL periods with respect to the historical base period of 1980-2010. We analysed all regional grids for all CMIP6 models for the modality changes under GWL 1.5°C, 2°C, 3°C, and 4°C. Figure 7 shows the percentage of changes in grid cell distribution modalities under GWL3.0°C. Globally, the percentage of grids changing from a unimodal (bimodal) distribution in the historical period to a bimodal (unimodal) distribution in the future periods is between 2.98% (2.41%) and 5.99% (3.99%) for different scenarios and GWL as shown in Table 2. The change from unimodal to bimodal distribution in the future period is most prevalent in regions where the highest percentage of unimodality was observed in the historical period, as shown in Figure 4. This suggests that regions that were previously characterized by more consistent temperatures (as indicated by a unimodal temperature distribution) may experience more variability in temperature in the future. As our analysis uses the mean and standard deviation of the same component from the historical and future daily maximum temperature distributions, we only used the grid cells which have the same modality in the historical and future periods. We disregarded the grid cells with changing modalities, i.e. unimodal to bimodal or vice versa, as this will affect the mean and standard deviation, and hence the return period analysis.

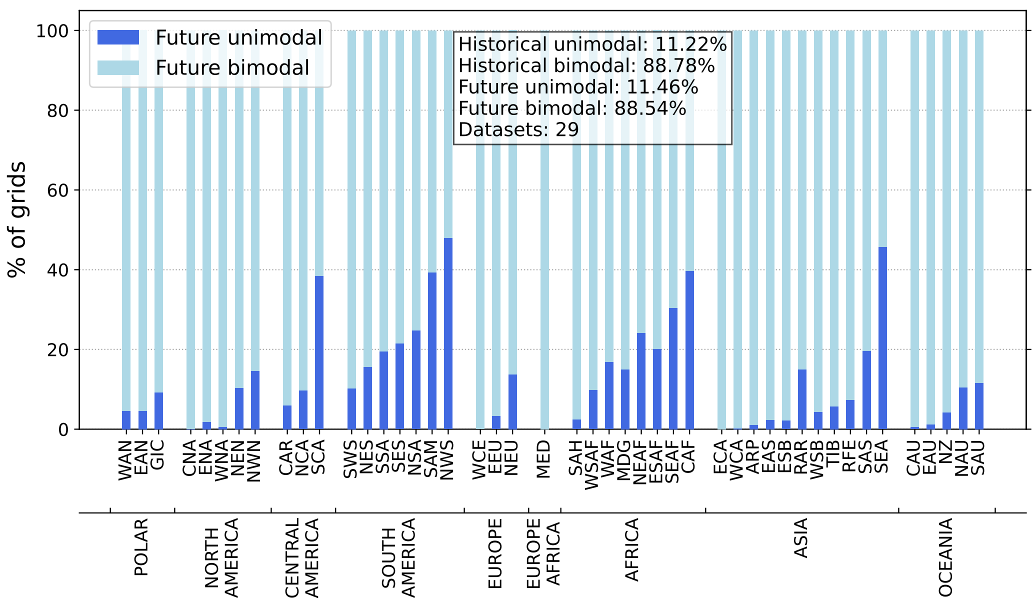


Figure 6. Same as Figure 4 but for future SSP5-8.5 scenario under GWL 3.0°C.

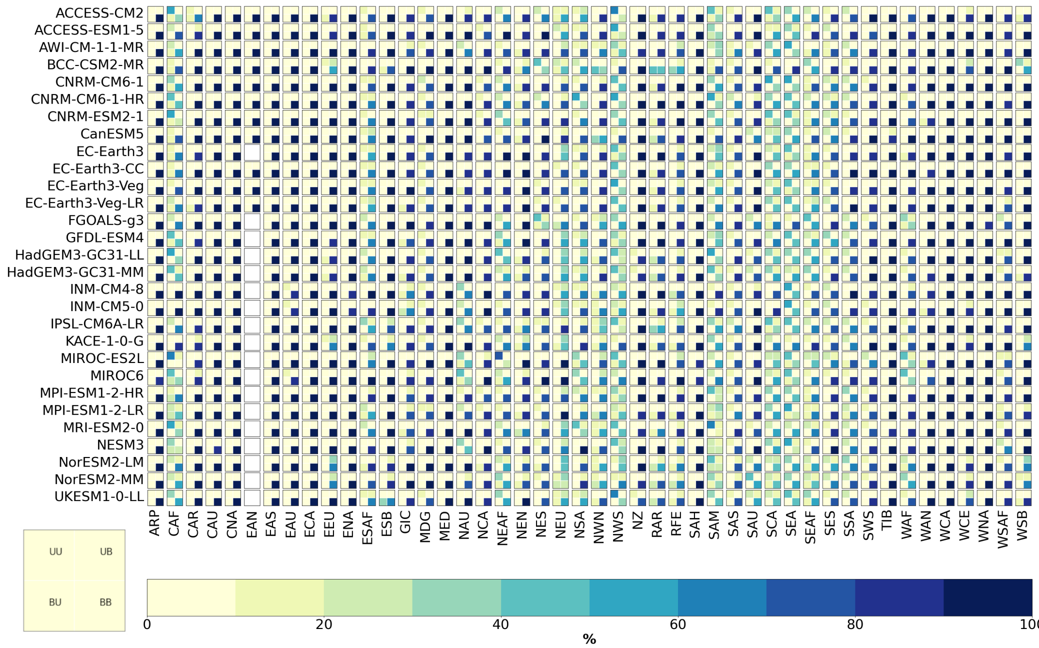


Figure 7. Percentage of changes in grid cell modalities relative to 1980-2010 distribution shape for SSP5-8.5 under GWL3.0°C. Each cell represents a region of a CMIP6 model and is divided into 4 quadrants. Each quadrant of squares, q_{ij}, uses index notation, where i represents the modality in the historical period and j represents the modality in the future period, 1 for a unimodal distribution and 2 for a bimodal distribution. The top-left quadrant, q_{11}, shows the percentage of grid cells with unimodal distribution both in the historical and the future periods, i.e. unimodal to unimodal (UU). The top-right quadrant, q_{12}, shows the percentage of grid cells that change from unimodal distribution in the historical period to bimodal distribution in the future (UB). The bottom-left quadrant, q_{21}, shows bimodal to unimodal (BU). The bottom-right quadrant, q_{22}, shows bimodal to bimodal distribution (BB). The colour of the quadrants shows the percentage of grid cells. For several models, East Antarctica (EAN) region is not included in the analysis because it is composed of many grid cells near the pole, causing numerical problems.

We also analysed the movements of the Gaussian components relative to each other using the ΔT definition from Equation 15 in grid cells with a bimodal distribution. Figure 8 shows the ΔT results for all analysed regions for SSP5-8.5 under 3.0°C warming (see Supplementary Material Figure S8 to S10 for other warming levels). Changes in distribution peaks are smaller for the lower warming levels. This is consistent with the fact that the time periods for exceeding warming levels are very close to the historical period as shown in Figure 1. For the future 3.0°C warming scenario, we observed that the mean temperatures are increasing in all regions. Temperature distributions for the European regions have negative ΔT values, -0.46 degrees on average. This will cause already bimodal peaks in the historical period to separate further from each other in the future, while the whole distribution moves towards higher temperatures. Divergence of peaks will result in more extreme hot temperatures in Europe, as the hot Gaussian moves faster. This result is in agreement with findings from the IPCC AR6 report, in which temperatures in Europe are reported to increase faster than the rest of the globe (IPCC, 2021). Polar regions, Northern America and parts of Northern Asia have positive ΔT values, i.e. converging peaks in grid cells with bimodal distributions. The distribution shape shifts to warmer temperatures and approaches a unimodal distribution as the cold Gaussian part of the distribution moves toward the warmer temperatures faster than the

Table 2. Global average percentage of grid cells with varying distribution modality between the historical and future periods.

EXP	GWL	Unimodal→Unimodal	Unimodal→Bimodal	Bimodal→Unimodal	Bimodal→Bimodal
SSP1-2.6	1.5°C	11.22%	2.98%	2.41%	83.40%
SSP1-2.6	2.0°C	10.29%	3.65%	2.56%	83.51%
SSP2-4.5	1.5°C	10.96%	2.98%	2.43%	83.63%
SSP2-4.5	2.0°C	10.24%	3.70%	2.84%	83.21%
SSP2-4.5	3.0°C	8.77%	4.75%	3.18%	83.29%
SSP2-4.5	4.0°C	7.17%	5.99%	3.75%	83.08%
SSP3-7.0	1.5°C	10.79%	3.12%	2.54%	83.55%
SSP3-7.0	2.0°C	10.15%	3.76%	3.01%	83.08%
SSP3-7.0	3.0°C	8.92%	4.79%	3.56%	82.73%
SSP3-7.0	4.0°C	7.81%	5.50%	3.89%	82.80%
SSP5-8.5	1.5°C	11.03%	3.04%	2.47%	83.46%
SSP5-8.5	2.0°C	10.32%	3.75%	3.17%	82.76%
SSP5-8.5	3.0°C	9.16%	4.91%	3.90%	82.03%
SSP5-8.5	4.0°C	8.20%	5.57%	3.99%	82.24%

hot Gaussian part. This convergence is also consistent with the slight increase in the percentage of unimodal distribution in polar regions as shown in Figure 6. This will cause polar regions to have more days with warmer temperatures also in the colder mode while also having an overall warmer climate. The convergence of peaks in three polar regions (EAN, WAN, GIC) and three northern regions (RAR, NEN and NWN) becomes clear when the regions are sorted by the mean temperature of cold Gaussian component as shown in Figure 9. High ΔT values in polar regions are also supported by previous studies reporting that Arctic regions are warming faster than the global average (Taylor et al., 2022). The lowest ΔT values are in MED and SAM regions, -0.97 and -1.25 degrees respectively, which will cause both bimodal peaks to diverge from each other while both are moving towards warmer temperatures. Regions in Oceania, Central- and parts of South-America have ΔT values close to zero, i.e. the cold and hot Gaussian peaks shifts toward the warmer temperatures at the same rate. This will cause these regions to have warmer cold and hot periods under future global warming levels compared to the historical period. When all regions are considered, we observe that the extreme temperature events will increase everywhere, as the mean temperatures increase in all regions compared to the historical distributions. The fact that the peaks are converging only in cold climate regions while diverging in other regions shows that shifts in the Gaussian components with respect to each other are essential for extreme temperature event analyses as these changes affect the overall distribution shape and extent.

After analysing the distribution shapes and peak movements, we calculated the return periods -the average time between the occurrences of a certain event- of 1-year, 5-year, 10-year and 20-year events using only the grid cells with constant modalities, i.e. unimodal or bimodal both for the historical and future periods, as described in Equation 14. As we did not analyse the extremes in blocks such as season or a full year, but the overall extremes in the region's temperature distribution, we used the number of data points under the Gaussian component as the length of a *year* defined in Equation 7. For example, globally a 10-*year* event was a temperature event that occurs once in every 1880 days (10 years$\times 188 \frac{days}{year}$) (for bimodal distributions) in the historical period, but it will occur once in every 564, 311, 122 and 72 days under GWL 1.5°CC, 2.0°C, 3.0°C and 4.0°C scenarios, as shown in the plot showing global results in Figure 10, respectively. In other words, historical 10-*year* events will be 3-*year*, 1.65-*year*, 0.65-*year* and 0.3-*year* events under the future GWL 1.5°C, 2.0°C, 3.0°C and 4.0°C scenarios, respectively. After calculating the frequency of extreme events using the temperature distributions in each grid cell individually for an IPCC land region, we averaged the results for the whole region for a single model. The global map with box plots in Figure 10 shows multi-model 10-

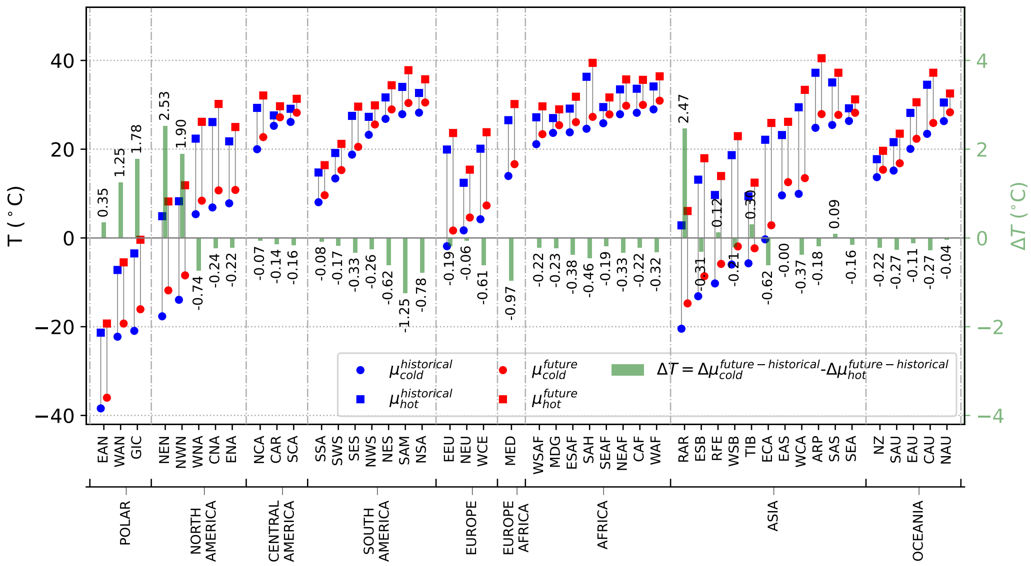


Figure 8. Multi-model peak mean change of region temperature distributions from bimodal grid cells for SSP5-8.5 under GWL3.0°C. Blue and red dots are peak means for the historical and future periods, respectively, and are plotted on the left y-axis. Green bars describe the change in the peak mean temperature, ΔT, and are plotted on the right y-axis. The upward shift in both blue and red dots represents overall warming (see Supplementary Material Figure S8 to S10 for other warming levels).

year event frequencies of each region for SSP5-8.5 scenario under different GWL, where blue, green, pink and red boxes represent 1.5°C, 2.0°C, 3.0°C and 4.0°C, respectively. Results for 1-year, 5-year, and 20-year events are left out for simplicity and presented in the Supplementary Material Figure S15 to S28. The length of a *"year"* in each region that is used for return period calculations, i.e. the number of days in 10 years, is shown on the top right corner of each sub-plot in Figure 10.

As shown in Figure 10, return periods of extreme temperature events are getting shorter for all regions under all GWL scenarios as the median of each box is smaller than the base period. The frequency of extreme events is higher in lower latitudes compared to higher latitudes. For example, the return periods are getting prominently shorter in regions around the equator -where a higher percentage of unimodal grid cells was observed- compared to the other regions. Furthermore, CMIP6 models show narrower boxes and shorter whiskers in lower latitudes compared to wider boxes and longer whiskers in higher latitudes for all analyzed GWL. Among all analysed regions, the Caribbean (CAR) region has the highest increase in the frequency of a 10-year event, from once in 1910 days for the historical period to once in every 93.0, 24.9, 4.2 and 1.8 days under GWL 1.5, 2, 3, and 4°C, respectively. Regions around the equator (namely CAR, NSA, NWS, NES, SEA, SCA, SAM, MDG, WAF, and SEAF regions) are the top 10 regions with the highest increase in the frequency of extreme events under all GWL. The frequency of a temperature event equivalent to a 10-year event (historically once in every 1610 days) in the Mediterranean (MED) region increases to once in 367.7, 184.2, 62.8 and 27.1 days in the future under GWL 1.5, 2, 3, and 4°C, respectively. Within the European continent, the West&Central Europe (WCE) region has a higher increase in the frequency of extreme events compared to the Eastern Europe (EEU) and the North Eastern Europe (NEU) regions, where the latter two regions are among the regions with the least increase in ex-

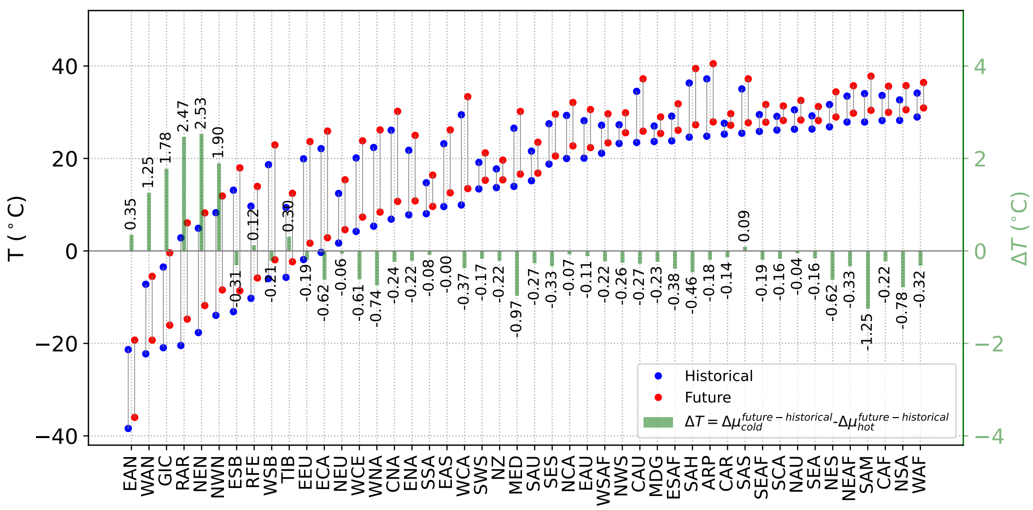


Figure 9. Multi-model peak mean change of region temperature distributions sorted by cold Gaussian mean temperatures for SSP5-8.5 under GWL 3.0°C. Blue and red dots are peak means for the historical and future periods, respectively, and are plotted on the left y-axis. Green bars describe the change in the peak mean temperatures, ΔT, and are plotted on the right y-axis. The colder regions, three polar regions (EAN, WAN, GIC) and three northern regions (RAR, NEN and NWN), have positive ΔT values and their absolute values are higher than the other regions. The upward shift in blue dots shows that the temperature of cold days is getting warmer and this increase is faster in polar regions compared to the rest of the world (see Supplementary Material Figure S11 to S13 for other warming levels).

treme temperature event frequency. The smallest increase in the frequency of hot extremes is observed in the Western Antarctica (WAN) region, where the return periods of 10-year events will decrease from once in 1830 days to once in 1062.12, 844.7, 541.9 and 339.8 days under GWL 1.5, 2, 3, and 4°C, respectively. High latitude regions, such as WAN, NEU, EAN, NWN, ESB, GIC, RAR, SSA, TIB, and NEN regions are the 10 regions with the smallest decrease in return periods of extreme hot temperature events. Some of these regions are polar regions with positive ΔT values as shown in Figure 9. This will cause more days with warmer temperatures in the colder mode of these regions while having an increase in hot extremes.

4 Summary and Discussion

Detection of extreme events is important to mitigate their impact on natural and anthropogenic systems. Future projections suggest that the mean and standard deviations of maximum surface temperature will increase. This change in the shape of maximum surface temperature distributions increases the intensity and frequency of extreme events in the future. However, not only the shift to warmer temperatures but also the modality of temperature distribution affects the parameters of the entire distribution which is important to calculate the return periods as shown in this study.

GMM are a promising method for calculating the return periods of extreme events, and additionally determining the shape of the entire distribution for daily maximum temperature data. Some studies used seasonal periods to analyse extreme events (Walt & Fitchett, 2021; Prodhomme et al., 2022), however, onsets and length of seasons are pre-

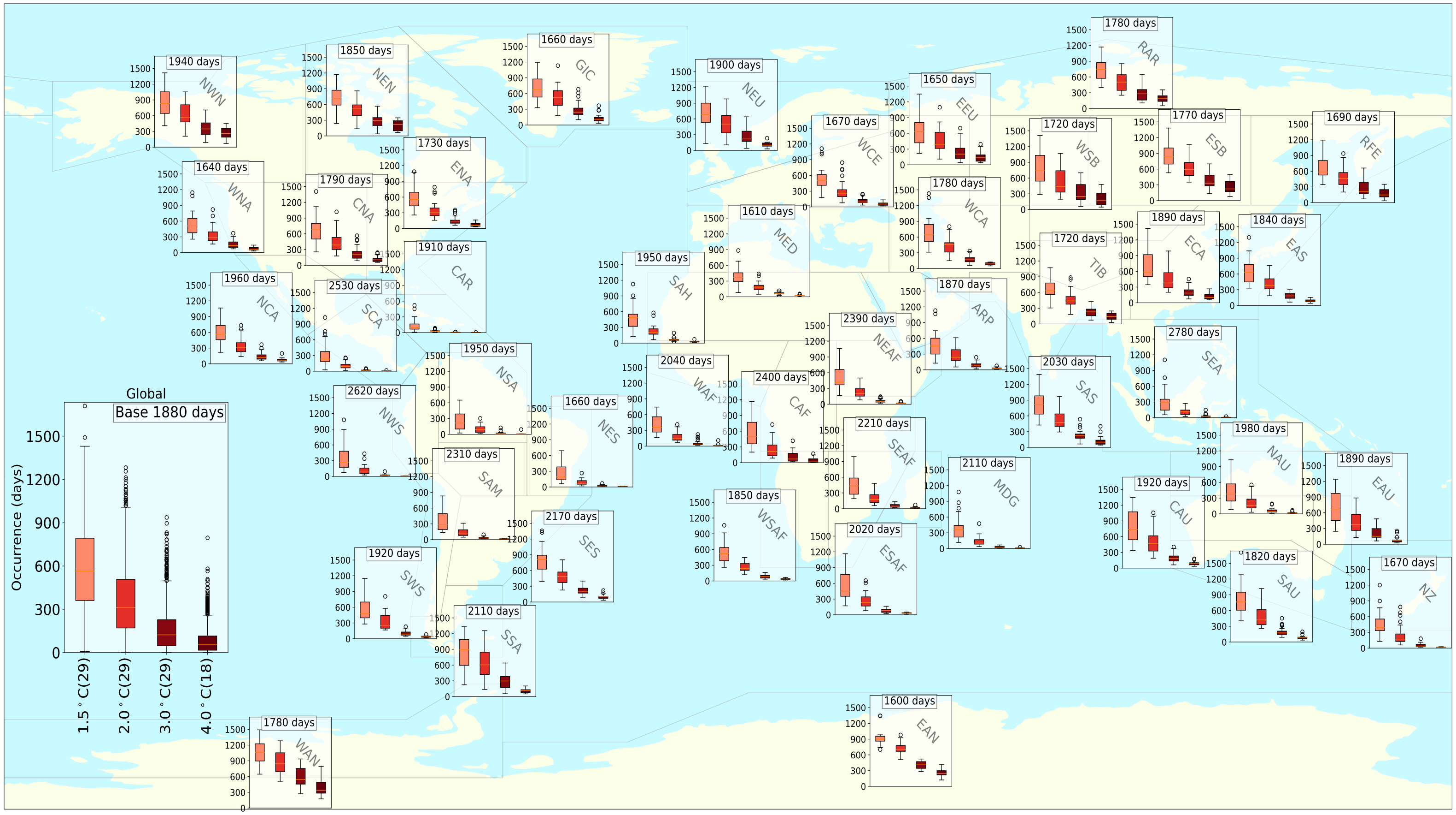


Figure 10. Multi-model median of event frequencies for 10-year hot temperature events compared to the 1980-2010 period under GWL 1.5, 2, 3 and 4°C relative to 1850-1900 baseline for SSP5-8.5 scenario. Blue, green, pink and red boxes represent 1.5°C, 2.0°C, 3.0°C and 4.0°C, respectively. The orange lines inside the boxes show the CMIP6 multi-model median, and the boxes extend between the first quartile (Q1) to the third quartile (Q3) of the data, i.e. inter-quartile range (IQR). The vertical lines, i.e. whiskers, stretch out 1.5 IQR from the box. The circles represent the models outside of the interquartile range, i.e. outliers. The length of the hot period used for return period calculations, i.e. number of days in 10 years, is shown in the top right corner of each plot. The global return periods are shown on the left. The more outlier points in the global box plot are because of the differences between regional return periods (See Supplementary Material Figure S15 to S28 for other return periods).

dicted to change with climate change (Wang et al., 2021). Therefore, the definition of current seasonal periods will not necessarily be valid for future climates. One can utilize GMM to determine the hot Gaussian component of a region to define the length of the analysis period instead of using fixed seasonal definitions. Furthermore, one loses most of the data with current extreme event indices. GEV distributions are a better fit for longer block sizes than for shorter blocks like daily data. If the available dataset is short, the longer block sizes will produce fewer data which can increase the variability in parameter estimation (Huang et al., 2016; Wang et al., 2016). If there is more than one extremely hot day in the block (month, season or year), e.g. several consecutive days, block maxima methods consider only the hottest, and hence only one day in a block, while GMM considers all days hotter than the threshold. Assuming that a heat wave lasts usually days to a few weeks, a substantial number of hot days might not be seen by block maxima methods as long as they fall into the same block. However, since the Gaussian components of GMM are not bounded, it is important to only calculate the return periods of extreme events equal to or less than the study period when applying GMM. Additionally, we only used grid cells which have the same number of Gaussian components in their temperature distribution, i.e. unimodal or bimodal distribution, both for the historical and future periods. Grid cells with changing distribution shapes, e.g. transforming from a bimodal distribution in the historical period to a unimodal distribution in the future or vice versa, were found in less than 10% of the grid cell for each GWL as shown in Table 2, and were disregarded in the analysis as calculating the temperature thresholds becomes problematic with the abrupt change in means and standard deviations. Finally, GMM can provide information on different climate features in different regions such as cold and hot periods, and their changes.

For the first time, the IPCC AR6 Report includes a new dedicated chapter on weather and climate extreme events (IPCC, 2021). This emphasizes the importance of robust methods of extreme event detection to be able to mitigate the impact of such events. IPCC AR6 reports that the return periods of 10-year events will increase around the world, with the highest changes projected to happen in some mid-latitude and semi-arid regions. Our findings are in agreement with these results. Furthermore, IPCC AR6 projects the warming rate in mid-latitudes to be higher than the average global warming rate. This will introduce the highest increase in the temperature of the hottest days. For example, almost all grid cells in the Mediterranean region follow a bimodal distribution, and the peaks of bimodal distribution will diverge in the future. This might explain why the Mediterranean region is identified as one of the most responsive regions to climate change and a hot spot of climate extremes (IPCC, 2021). Similarly, Arctic regions are projected to have the highest increase in temperature of the coldest days (IPCC, 2021; C. Li et al., 2021). Our results are also consistent with these increases as shown in Figure 8, where diverging peaks in mid-latitudes will shift the hot Gaussian part of temperature distributions to the higher temperature ranges. This shift in the Gaussian components of temperature distribution will cause those land regions to have warmer temperature extremes and can explain the higher average warming rate than the global average. Likewise, converging peaks in polar regions as shown in Figure 8 will move the cold Gaussian part toward warmer temperatures, thereby introducing higher warming on the coldest days.

According to our analyses, 10-year events will increase almost 3-fold under GWL 1.5°C compared to the historical period for all SSP scenarios as shown in Figure 11 when looking at the whole globe. This means a temperature event that occurs once in every 10 years (1880 days) will occur 3.3 times in every 10 years under GWL 1.5°C. 10-year extreme temperature events will become even more frequent globally under GWL 2°C, 3°C and 4°C; 6.0, 15.3, and 32.7 times in every 10 years, respectively. In other words, current 10-year events will be 3.0-year, 1.65-year, 0.65-year and 0.3-year events in the future under GWL 1.5°C, 2°C, 3°C and 4°C, respectively. Our results show a higher increase compared to the IPCC AR6 report, where the frequency of 10-year events is projected to increase approximately 3, 4, 5.5 and 9-fold under GWL 1.5°C, 2°C, 3°C and

4°C, respectively (IPCC, 2021), using a block maxima method for determining the extreme events. The higher increase in our method compared to IPCC AR6 can most likely be explained by the fact that GMM considers all days hotter than the threshold, while the block maxima method only uses the maximum of a block. Another important point deduced from the analyses of different regions for several CMIP6 models is that the ensemble of analyzed CMIP6 models shows coherent results for regions as shown in the regional box plots in Figure 10. Most of the individual model results fall within the first and third quartile, and only a few models fall outside this range. The higher number of outlier points in the global box plot in Figure 10, and also shown for different SSP scenarios in Figure 11, are caused by the differences between regional return periods. All SSP scenarios show similar results with each other as the return periods are calculated for GWL which have the same forcing on climate.

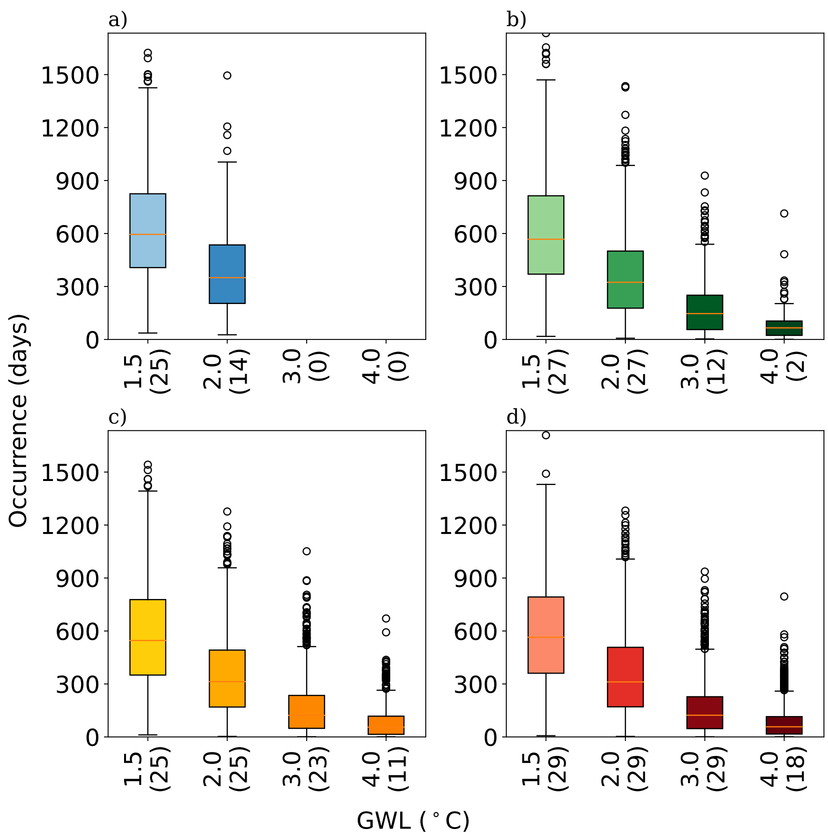


Figure 11. Global multi-model median of event frequencies for 10-year temperature events under 1.5, 2, 3 and 4°C warming levels for a) SSP1-2.6, b) SSP2-4.5, c) SSP3-7.0 and d) SSP5-8.5 scenarios. The orange lines inside the boxes show the CMIP6 multi-model median, and the boxes extend between the first quartile (Q1) to the third quartile (Q3) of the data, i.e. interquartile range (IQR). The vertical lines, i.e. whiskers, stretch out 1.5 IQR from the box. The circles represent the models outside of the interquartile range, i.e. outliers. The length of the hot period used for return period calculations, i.e. number of days in 10 years, is shown in the top right corner of each plot. The number of datasets is given in parentheses. All plots show similar results for different SSP scenarios as the GWL are scenario-independent.

Return periods of extreme events become shorter in every region, which means that the frequency of extreme temperature events increases. This will become larger with increasing global warming levels. Some climate models have already exceeded GWL 1.5°C with respect to the 1850-1900 period as shown in Figure 1. This fact further emphasises the importance of robust methods to detect extreme events. Even though there is a delay in taking the necessary precautions to reduce the speed of the warming of the climate, as time goes by, tomorrow's projections become today's reality.

Code and data availability

The recipes to extract regional data from CMIP6 models using ESMValTool, python scripts to analyse extreme events and to produce all figures of this manuscript are accessible in the following GitHub repository: `https://github.com/EyringMLClimateGroup/pacal23jgr_GaussianMixtureModels_Extremes`. The regional output files amount to hundreds of GB.

The latest release of ESMValTool is publicly at `https://github.com/ESMValGroup/ESMValTool` (Andela et al., 2022).

Acknowledgments

Funding for this study was provided by the European Research Council (ERC) Synergy Grant "Understanding and Modelling the Earth System with Machine Learning (USMILE)" under the Horizon 2020 research and innovation programme (Grant Agreement No. 855187). This work used resources of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project ID bd1083. We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modelling groups for producing and making available their model outputs, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies that support CMIP6 and ESGF. We would like to extend our sincere gratitude to Dr. Pauline Bonnet for her valuable comments and suggestions to improve the manuscript.

References

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., . . . Vazquez-Aguirre, J. L. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research: Atmospheres*, *111*(D5). doi: 10.1029/2005JD006290

Alizadeh, M. R., Adamowski, J., Nikoo, M. R., AghaKouchak, A., Dennison, P., & Sadegh, M. (2020). A century of observations reveals increasing likelihood of continental-scale compound dry-hot extremes. *Science Advances*, *6*(39), eaaz4571. doi: 10.1126/sciadv.aaz4571

Andela, B., Broetz, B., de Mora, L., Drost, N., Eyring, V., Koldunov, N., . . . Winterstein, F. (2022). *ESMValTool* [software]. Retrieved from `https://github.com/ESMValGroup/ESMValTool/` doi: 10.5281/zenodo.3401363

Anderson, A., & Kostinski, A. (2010). Reversible record breaking and variability: Temperature distributions across the globe. *Journal of Applied Meteorology and Climatology*, *49*(8), 1681–1691. doi: 10.1175/2010JAMC2407.1

Baker, H. S., Millar, R. J., Karoly, D. J., Beyerle, U., Guillod, B. P., Mitchell, D., . . . Allen, M. R. (2018). Higher co2 concentrations increase extreme event risk in a 1.5°c world. *Nature Climate Change*, *8*(7), 604–608. doi: 10.1038/s41558-018-0190-1

Barriopedro, D., Fischer, E. M., Luterbacher, J., Trigo, R. M., & García-Herrera, R. (2011). The hot summer of 2010: Redrawing the temperature record map of europe. *Science*, *332*(6026), 220-224. doi: 10.1126/science.1201224

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., &

Wood, E. F. (2018). Present and future köppen-geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*(1), 180214. doi: 10.1038/sdata.2018.214
Ben Alaya, M. A., Zwiers, F., & Zhang, X. (2020). An evaluation of block-maximum-based estimation of very long return period precipitation extremes with a large ensemble climate simulation. *Journal of Climate*, *33*(16), 6957–6970. doi: 10.1175/JCLI-D-19-0011.1
Benestad, R. (2004). Record-values, nonstationarity tests and extreme value distributions. *Global and Planetary Change*, *44*(1–4), 11–26. doi: 10.1016/j.gloplacha.2004.06.002
Bentsen, M., Oliviè, D. J. L., Seland, y., Toniazzo, T., Gjermundsen, A., Graff, L. S., ... Schulz, M. (2019). *Ncc noresm2-mm model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.8040
Bertossa, C., Hitchcock, P., DeGaetano, A., & Plougonven, R. (2021). Bimodality in ensemble forecasts of 2 m temperature: identification. *Weather and Climate Dynamics*, *2*(4), 1209–1224. Retrieved from `https://wcd.copernicus.org/articles/2/1209/2021/` doi: 10.5194/wcd-2-1209-2021
Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., ... Cheruy, F. (2018). *Ipsl ipsl-cm6a-lr model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.5195
Byun, Y.-H., Lim, Y.-J., Sung, H. M., Kim, J., Sun, M., & Kim, B.-H. (2019). *Nims-kma kace1.0-g model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.8378
Cai, W., Borlace, S., Lengaigne, M., van Rensch, P., Collins, M., Vecchi, G., ... Jin, F.-F. (2014). Increasing frequency of extreme el niño events due to greenhouse warming. *Nature Climate Change*, *4*(22), 111–116. doi: 10.1038/nclimate2100
Cao, J., & Wang, B. (2019). *Nuist nesmv3 model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.8769
Cho, H., & Jeong, S. (2016). Estimation and comparative analysis on the distribution functions of air and water temperatures in korean coastal seas. *Journal of Korean Society of Coastal and Ocean Engineers*, *28*, 171-176. doi: 10.9765/KSCOE.2016.28.3.171
Christidis, N., Jones, G. S., & Stott, P. A. (2015). Dramatically increasing chance of extremely hot summers since the 2003 european heatwave. *Nature Climate Change*, *5*(11), 46–50. doi: 10.1038/nclimate2468
Clarkson, D., Eastoe, E., & Leeson, A. (2022). Melt probabilities and surface temperature trends on the greenland ice sheet using a gaussian mixture model. *The Cryosphere*, *16*(5), 1597–1607. Retrieved from `https://tc.copernicus.org/articles/16/1597/2022/` doi: 10.5194/tc-16-1597-2022
Demiroglu, O. C., Saygili-Araci, F. S., Pacal, A., Hall, C. M., & Kurnaz, M. L. (2020). Future holiday climate index (hci) performance of urban and beach destinations in the mediterranean. *Atmosphere*, *11*(9), 911. doi: 10.3390/atmos11090911
Dix, M., Bi, D., Dobrohotoff, P., Fiedler, R., Harman, I., Law, R., ... Yang, R. (2019). *Csiro-arccss access-cm2 model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.4271
Donat, M. G., & Alexander, L. V. (2012). The shifting probability distribution of global daytime and night-time temperatures. *Geophysical Research Letters*, *39*(14). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012GL052459` doi: https://doi.org/10.1029/2012GL052459
(EC-Earth), E.-E. C. (2019a). *Ec-earth-consortium ec-earth3 model output prepared*

for cmip6 cmip historical [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.4700
(EC-Earth), E.-E. C. (2019b). *Ec-earth-consortium ec-earth3-veg model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.4706
(EC-Earth), E.-E. C. (2020). *Ec-earth-consortium ec-earth3-veg-lr model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.4707
(EC-Earth), E.-E. C. (2021). *Ec-earth-consortium ec-earth-3-cc model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.4702
Eyring, V., Bock, L., Lauer, A., Righi, M., Schlund, M., Andela, B., . . . Zimmermann, K. (2020). Earth system model evaluation tool (esmvaltool) v2.0 – an extended set of large-scale diagnostics for quasi-operational and comprehensive evaluation of earth system models in cmip. *Geoscientific Model Development*, *13*(7), 3383–3438. doi: 10.5194/gmd-13-3383-2020
Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. doi: 10.5194/gmd-9-1937-2016
Eyring, V., Gillett, N. P., Achuta Rao, K. M., Barimalala, R., Barreiro Parrillo, M., Bellouin, N., . . . Sun, Y. (2021). Human influence on the climate system [Book Section]. 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. 3). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.
Fischer, E. M., & Schär, C. (2010). Consistent geographical patterns of changes in high-impact european heatwaves. *Nature Geoscience*, *3*(6), 398-403. doi: 10.1038/ngeo866
García-León, D., Casanueva, A., Standardi, G., Burgstall, A., Flouris, A. D., & Nybo, L. (2021). Current and projected regional economic impacts of heatwaves in europe. *Nature Communications*, *12*(1), 5807. doi: 10.1038/s41467-021-26050-z
Grace, W. (1995). A binormal model of frequency distributions of daily maximum temperature. *Australian meteorological magazine*, *42*, 151-161.
Guerreiro, S. B., Dawson, R. J., Kilsby, C., Lewis, E., & Ford, A. (2018). Future heat-waves, droughts and floods in 571 european cities. *Environmental Research Letters*, *13*(3), 034009. doi: 10.1088/1748-9326/aaaad3
Hajima, T., Abe, M., Arakawa, O., Suzuki, T., Komuro, Y., Ogura, T., . . . Tachiiri, K. (2019). *Miroc miroc-es2l model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.5602
Hao, Z., Hao, F., Singh, V. P., & Zhang, X. (2018). Changes in the severity of compound drought and hot extremes over global land areas. *Environmental Research Letters*, *13*(12), 124022. doi: 10.1088/1748-9326/aaee96
Hauser, M., Engelbrecht, F., & Fischer, E. M. (2022). *Transient global warming levels for CMIP5 and CMIP6* [dataset]. Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.7390473` doi: 10.5281/zenodo.7390473
Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., . . . Thépaut, J.-N. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999-2049. doi: 10.1002/qj.3803
Huang, W. K., Stein, M. L., McInerney, D. J., Sun, S., & Moyer, E. J. (2016). Estimating changes in temperature extremes from millennial-scale climate simulations using generalized extreme value (gev) distributions. *Advances in Statistical Climatology, Meteorology and Oceanography*, *2*(1), 79–103. doi:

10.5194/ascmo-2-79-2016

IPCC. (2021). *Climate change 2021: The physical science basis* (V. Masson-Delmotte et al., Eds.). Cambridge University Press (In Press).

Iturbide, M., Gutiérrez, J. M., Alves, L. M., Bedia, J., Cerezo-Mota, R., Cimadevilla, E., ... Vera, C. S. (2020). An update of ipcc climate reference regions for subcontinental analysis of climate model data: definition and aggregated datasets. *Earth System Science Data*, *12*(4), 2959–2970. doi: 10.5194/essd-12-2959-2020

Iyakaremye, V., Zeng, G., Ullah, I., Gahigi, A., Mumo, R., & Ayugi, B. (2022). Recent observed changes in extreme high-temperature events and associated meteorological conditions over africa. *International Journal of Climatology*, joc.7485. doi: 10.1002/joc.7485

Iyakaremye, V., Zeng, G., & Zhang, G. (2021). Changes in extreme temperature events over africa under 1.5 and 2.0c global warming scenarios. *International Journal of Climatology*, *41*(2), 1506–1524. doi: 10.1002/joc.6868

Jungclaus, J., Bittner, M., Wieners, K.-H., Wachsmann, F., Schupfner, M., Legutke, S., ... Roeckner, E. (2019). *Mpi-m mpi-esm1.2-hr model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.6594

Kharin, V. V., Zwiers, F. W., Zhang, X., & Wehner, M. (2013). Changes in temperature and precipitation extremes in the cmip5 ensemble. *Climatic Change*, *119*(2), 345–357. doi: 10.1007/s10584-013-0705-8

Knoben, W. J. M., Woods, R. A., & Freer, J. E. (2019). Global bimodal precipitation seasonality: A systematic overview. *International Journal of Climatology*, *39*(1), 558-567. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.5786` doi: https://doi.org/10.1002/joc.5786

Kodra, E., & Ganguly, A. R. (2014). Asymmetry of projected increases in extreme temperature distributions. *Scientific Reports*, *4*(1), 5884. Retrieved 2023-01-08, from `https://www.nature.com/articles/srep05884` (Number: 1 Publisher: Nature Publishing Group) doi: 10.1038/srep05884

Kollu, R., Rayapudi, S. R., Narasimham, S., & Pakkurthi, K. M. (2012). Mixture probability distribution functions to model wind speed distributions. *International Journal of Energy and Environmental Engineering*, *3*(1), 27. doi: 10.1186/2251-6832-3-27

Krasting, J. P., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., ... Zhao, M. (2018). *Noaa-gfdl gfdl-esm4 model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.8597

Lauer, A., Eyring, V., Bellprat, O., Bock, L., Gier, B. K., Hunter, A., ... Zechlau, S. (2020). Earth system model evaluation tool (esmvaltool) v2.0 – diagnostics for emergent constraints and future projections from earth system models in cmip. *Geoscientific Model Development*, *13*(9), 4205–4228. doi: 10.5194/gmd-13-4205-2020

Lelieveld, J., Proestos, Y., Hadjinicolaou, P., Tanarhte, M., Tyrlis, E., & Zittis, G. (2016). Strongly increasing heat extremes in the middle east and north africa (mena) in the 21st century. *Climatic Change*, *137*(1–2), 245–260. doi: 10.1007/s10584-016-1665-6

Lewis, S. C., King, A. D., & Perkins-Kirkpatrick, S. E. (2017). Defining a new normal for extremes in a warming world. *Bulletin of the American Meteorological Society*, *98*(6), 1139–1151. doi: 10.1175/BAMS-D-16-0183.1

Li, C., Zwiers, F., Zhang, X., & Li, G. (2019). How much information is required to well constrain local estimates of future precipitation extremes? *Earth's Future*, *7*(1), 11–24. doi: 10.1029/2018EF001001

Li, C., Zwiers, F., Zhang, X., Li, G., Sun, Y., & Wehner, M. (2021). Changes in annual extremes of daily temperature and precipitation in cmip6 models. *Journal*

of Climate, *34*(9), 3441 - 3460. doi: 10.1175/JCLI-D-19-1013.1
Li, L. (2019). *Cas fgoals-g3 model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.3356
Linz, M., Chen, G., & Hu, Z. (2018). Large-scale atmospheric control on non-gaussian tails of midlatitude temperature distributions. *Geophysical Research Letters*, *45*(17), 9141–9149. doi: 10.1029/2018GL079324
Loikith, P. C., & Neelin, J. D. (2019). Non-gaussian cold-side temperature distribution tails and associated synoptic meteorology. *Journal of Climate*, *32*(23), 8399–8414. doi: 10.1175/JCLI-D-19-0344.1
Mallick, J., Islam, A. R. M. T., Ghose, B., Islam, H. M. T., Rana, Y., Hu, Z., ... Ismail, Z. B. (2022). Spatiotemporal trends of temperature extremes in bangladesh under changing climate using multi-statistical techniques. *Theoretical and Applied Climatology*, *147*(1–2), 307–324. doi: 10.1007/s00704-021-03828-1
Meehl, G. A., & Tebaldi, C. (2004). More intense, more frequent, and longer lasting heat waves in the 21st century. *Science*, *305*(5686), 994–997. doi: 10.1126/science.1098704
Mueller, B., Zhang, X., & Zwiers, F. W. (2016). Historically hottest summers projected to be the norm for more than half of the world's population within 20 years. *Environmental Research Letters*, *11*(4), 044011. doi: 10.1088/1748-9326/11/4/044011
Nikulin, G., Kjellström, E., Hansson, U., Strandberg, G., & Ullerstig, A. (2011). Evaluation and future projections of temperature, precipitation and wind extremes over europe in an ensemble of regional climate simulations. *Tellus A*, *63*(1), 41–55. doi: 10.1111/j.1600-0870.2010.00466.x
O'Neill, B. C., Kriegler, E., Riahi, K., Ebi, K. L., Hallegatte, S., Carter, T. R., ... van Vuuren, D. P. (2014). A new scenario framework for climate change research: the concept of shared socioeconomic pathways. *Climatic Change*, *122*(3), 387-400. doi: 10.1007/s10584-013-0905-2
O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., ... Sanderson, B. M. (2016). The scenario model intercomparison project (scenariomip) for cmip6. *Geoscientific Model Development*, *9*(9), 3461–3482. doi: 10.5194/gmd-9-3461-2016
Ozturk, T., Saygili-Araci, F. S., & Kurnaz, M. L. (2021). Projected changes in extreme temperature and precipitation indices over cordex-mena domain. *Atmosphere*, *12*(55), 622. doi: 10.3390/atmos12050622
Paciorek, C. J., Stone, D. A., & Wehner, M. F. (2018). Quantifying statistical uncertainty in the attribution of human influence on severe weather. *Weather and Climate Extremes*, *20*, 69-80. Retrieved from `https://www.sciencedirect.com/science/article/pii/S2212094717300841` doi: https://doi.org/10.1016/j.wace.2018.01.002
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.
Perera, A. T. D., Nik, V. M., Chen, D., Scartezzini, J.-L., & Hong, T. (2020). Quantifying the impacts of climate change and extreme climate events on energy systems. *Nature Energy*, *5*(2), 150-159. doi: 10.1038/s41560-020-0558-0
Perkins-Kirkpatrick, S. E., & Lewis, S. C. (2020). Increasing trends in regional heatwaves. *Nature Communications*, *11*(1), 3357. doi: 10.1038/s41467-020-16970-7
Perron, M., & Sura, P. (2013). Climatology of non-gaussian atmospheric statistics. *Journal of Climate*, *26*(3), 1063–1083. doi: 10.1175/JCLI-D-11-00504.1
Pfleiderer, P., Schleussner, C.-F., Kornhuber, K., & Coumou, D. (2019, August). Summer weather becomes more persistent in a 2 °c world. *Nature Climate Change*, *9*(9), 666–671. Retrieved from `https://doi.org/10.1038/`

s41558-019-0555-0 doi: 10.1038/s41558-019-0555-0

Prodhomme, C., Materia, S., Ardilouze, C., White, R. H., Batté, L., Guemas, V., ... García-Serrano, J. (2022). Seasonal prediction of European summer heatwaves. *Climate Dynamics*, *58*(7-8), 2149–2166. Retrieved 2022-05-02, from https://link.springer.com/10.1007/s00382-021-05828-3 (3 citations (Semantic Scholar/DOI) [2022-05-02]) doi: 10.1007/s00382-021-05828-3

Rahmstorf, S., & Coumou, D. (2011). Increase of extreme events in a warming world. *Proceedings of the National Academy of Sciences*, *108*(44), 17905-17909. doi: 10.1073/pnas.1101766108

Raymond, C., Matthews, T., & Horton, R. M. (2020). The emergence of heat and humidity too severe for human tolerance. *Science Advances*, *6*(19). doi: 10.1126/sciadv.aaw1838

Richter, M. (2016). Temperatures in the tropics. In L. Pancel & M. Köhl (Eds.), *Tropical forestry handbook* (pp. 343–361). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-54601-3_38

Ridley, J., Menary, M., Kuhlbrodt, T., Andrews, M., & Andrews, T. (2019a). *Mohc hadgem3-gc31-ll model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.6109

Ridley, J., Menary, M., Kuhlbrodt, T., Andrews, M., & Andrews, T. (2019b). *Mohc hadgem3-gc31-mm model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.6112

Righi, M., Andela, B., Eyring, V., Lauer, A., Predoi, V., Schlund, M., ... Zimmermann, K. (2020). Earth system model evaluation tool (esmvaltool) v2.0 – technical overview. *Geoscientific Model Development*, *13*(3), 1179–1199. doi: 10.5194/gmd-13-1179-2020

Robine, J.-M., Cheung, S. L. K., Le Roy, S., Van Oyen, H., Griffiths, C., Michel, J.-P., & Herrmann, F. R. (2008). Death toll exceeded 70,000 in europe during the summer of 2003. *Comptes Rendus Biologies*, *331*(2), 171-178. doi: 10.1016/j.crvi.2007.12.001

Robinson, A., Lehmann, J., Barriopedro, D., Rahmstorf, S., & Coumou, D. (2021). Increasing heat and rainfall extremes now far outside the historical climate. *npj Climate and Atmospheric Science*, *4*(11), 1–4. doi: 10.1038/s41612-021-00202-w

Sardeshmukh, P. D., Compo, G. P., & Penland, C. (2015). Need for caution in interpreting extreme weather statistics. *Journal of Climate*, *28*(23), 9166–9187. doi: 10.1175/JCLI-D-15-0020.1

Sardeshmukh, P. D., & Sura, P. (2009). Reconciling non-gaussian climate statistics with linear dynamics. *Journal of Climate*, *22*(5), 1193 - 1207. Retrieved from https://journals.ametsoc.org/view/journals/clim/22/5/2008jcli2358.1.xml doi: 10.1175/2008JCLI2358.1

Schär, C., Vidale, P. L., Lüthi, D., Frei, C., Häberli, C., Liniger, M. A., & Appenzeller, C. (2004). The role of increasing temperature variability in european summer heatwaves. *Nature*, *427*(6972), 332-336. doi: 10.1038/nature02300

Seferian, R. (2018). *Cnrm-cerfacs cnrm-esm2-1 model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.4068

Seland, y., Bentsen, M., Oliviè, D. J. L., Toniazzo, T., Gjermundsen, A., Graff, L. S., ... Schulz, M. (2019). *Ncc noresm2-lm model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.8036

Semmler, T., Danilov, S., Rackow, T., Sidorenko, D., Barbi, D., Hegewald, J., ... Jung, T. (2018). *Awi awi-cm1.1mr model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.2686

Seneviratne, S., Nicholls, N., Easterling, D., Goodess, C., Kanae, S., Kossin, J., ...

Zhang, X. (2012). Changes in climate extremes and their impacts on the natural physical environment. In C. Field et al. (Eds.), *Managing the risks of extreme events and disasters to advance climate change adaptation* (p. 109-230). Cambridge, UK, and New York, NY, USA: Cambridge University Press.

Seneviratne, S., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., ... Zhou, B. (2021). Weather and climate extreme events in a changing climate [Book Section]. 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, United Kingdom and New York, NY, USA: Cambridge University Press.

Shin, J.-Y., Kang, M., & Kim, K. R. (2022). Outdoor thermal stress changes in south korea: Increasing inter-annual variability induced by different trends of heat and cold stresses. *Science of The Total Environment*, *805*, 150132. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0048969721052074` doi: https://doi.org/10.1016/j.scitotenv.2021.150132

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. *Journal of Geophysical Research: Atmospheres*, *118*(4), 1716–1733. doi: 10.1002/jgrd.50203

Sillmann, J., Kharin, V. V., Zwiers, F. W., Zhang, X., & Bronaugh, D. (2013). Climate extremes indices in the cmip5 multimodel ensemble: Part 2. future climate projections. *Journal of Geophysical Research: Atmospheres*, *118*(6), 2473–2493. doi: 10.1002/jgrd.50188

Stone, D. A. (2019). A hierarchical collection of political/economic regions for analysis of climate extremes. *Climatic Change*, *155*(4), 639–656. doi: 10.1007/s10584-019-02479-6

Swart, N. C., Cole, J. N., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., ... Sigmond, M. (2019). *Cccma canesm5 model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.3610

Tamarin-Brodsky, T., Hodges, K., Hoskins, B. J., & Shepherd, T. G. (2019). A dynamical perspective on atmospheric temperature variability and its response to climate change. *Journal of Climate*, *32*(6), 1707–1724. doi: 10.1175/JCLI-D-18-0462.1

Tamarin-Brodsky, T., Hodges, K., Hoskins, B. J., & Shepherd, T. G. (2022). A simple model for interpreting temperature variability and its higher-order changes. *Journal of Climate*, *35*(1), 387–403. doi: 10.1175/JCLI-D-21-0310.1

Tang, Y., Rumbold, S., Ellis, R., Kelley, D., Mulcahy, J., Sellar, A., ... Jones, C. (2019). *Mohc ukesm1.0-ll model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.6113

Tatebe, H., & Watanabe, M. (2018). *Miroc miroc6 model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.5603

Taylor, P., Boeke, R., Boisvert, L., Feldl, N., Henry, M., Huang, Y., ... Tan, I. (2022). Process drivers, inter-model spread, and the path forward: A review of amplified arctic warming. *Frontiers in Earth Science*, *9*. doi: 10.3389/feart.2021.758361

Thompson, V., Kennedy-Asser, A. T., Vosper, E., Lo, Y. T. E., Huntingford, C., Andrews, O., ... Mitchell, D. (2022). The 2021 western north america heat wave among the most extreme events ever recorded globally. *Science Advances*, *8*(18), eabm6860. doi: 10.1126/sciadv.abm6860

Twardosz, R., & Kossowska-Cezak, U. (2013). Exceptionally hot summers in central and eastern europe (1951–2010). *Theoretical and Applied Climatology*, *112*(3), 617–628. doi: 10.1007/s00704-012-0757-0

Vogel, M. M., Hauser, M., & Seneviratne, S. I. (2020). Projected changes in hot, dry and wet extreme events' clusters in CMIP6 multi-model ensemble. *Environmental Research Letters*, *15*(9), 094021. doi: 10.1088/1748-9326/ab90a7

Vogel, M. M., Zscheischler, J., Fischer, E. M., & Seneviratne, S. I. (2020). Development of future heatwaves for different hazard thresholds. *Journal of Geophysical Research: Atmospheres*, *125*(9), e2019JD032070. doi: 10.1029/2019JD032070

Voldoire, A. (2018). *Cmip6 simulations of the cnrm-cerfacs based on cnrm-cm6-1 model for cmip experiment historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.4066

Voldoire, A. (2019). *Cnrm-cerfacs cnrm-cm6-1-hr model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.4067

Volodin, E., Mortikov, E., Gritsun, A., Lykossov, V., Galin, V., Diansky, N., ... Emelina, S. (2019). *Inm inm-cm5-0 model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.5070

Volodin, E. M., & Yurova, A. Y. (2012). Summer temperature standard deviation, skewness and strong positive temperature anomalies in the present day climate and under global warming conditions. *Climate Dynamics*, *40*(5-6), 1387–1398. Retrieved from https://doi.org/10.1007/s00382-012-1447-4 doi: 10.1007/s00382-012-1447-4

von, E., Mortikov, E., Gritsun, A., Lykossov, V., Galin, V., Diansky, N., ... Emelina, S. (2019). *Inm inm-cm4-8 model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.5069

Walt, A. J., & Fitchett, J. M. (2021). Trend analysis of cold extremes in south africa: 1960–2016. *International Journal of Climatology*, *41*(3), 2060–2081. doi: 10.1002/joc.6947

Wang, J., Guan, Y., Wu, L., Guan, X., Cai, W., Huang, J., ... Zhang, B. (2021). Changing lengths of the four seasons by global warming. *Geophysical Research Letters*, *48*(6), e2020GL091753. doi: 10.1029/2020GL091753

Wang, J., Han, Y., Stein, M. L., Kotamarthi, V. R., & Huang, W. K. (2016). Evaluation of dynamically downscaled extreme temperature using a spatially-aggregated generalized extreme value (GEV) model. *Climate Dynamics*, *47*(9-10), 2833–2849. Retrieved from https://doi.org/10.1007/s00382-016-3000-3 doi: 10.1007/s00382-016-3000-3

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. Retrieved from https://www.sciencedirect.com/science/article/pii/S2212094717301159 doi: https://doi.org/10.1016/j.wace.2018.03.001

Weigel, K., Bock, L., Gier, B. K., Lauer, A., Righi, M., Schlund, M., ... Eyring, V. (2021). Earth system model evaluation tool (esmvaltool) v2.0 – diagnostics for extreme events, regional and impact evaluation, and analysis of earth system models in cmip. *Geoscientific Model Development*, *14*(6), 3159–3184. doi: 10.5194/gmd-14-3159-2021

Wieners, K.-H., Giorgetta, M., Jungclaus, J., Reick, C., Esch, M., Bittner, M., ... Roeckner, E. (2019). *Mpi-m mpi-esm1.2-lr model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.6595

Wilks, D. S. (2002). Smoothing forecast ensembles with fitted probability distributions. *Quarterly Journal of the Royal Meteorological Society*, *128*(586), 2821-2836. Retrieved from https://rmets.onlinelibrary.wiley.com/doi/abs/10

.1256/qj.01.215 doi: https://doi.org/10.1256/qj.01.215

Wu, T., Chu, M., Dong, M., Fang, Y., Jie, W., Li, J., … Zhang, Y. (2018). *Bcc bcc-csm2mr model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.2948

Yukimoto, S., Koshiro, T., Kawai, H., Oshima, N., Yoshida, K., Urakawa, S., … Adachi, Y. (2019). *Mri mri-esm2.0 model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.6842

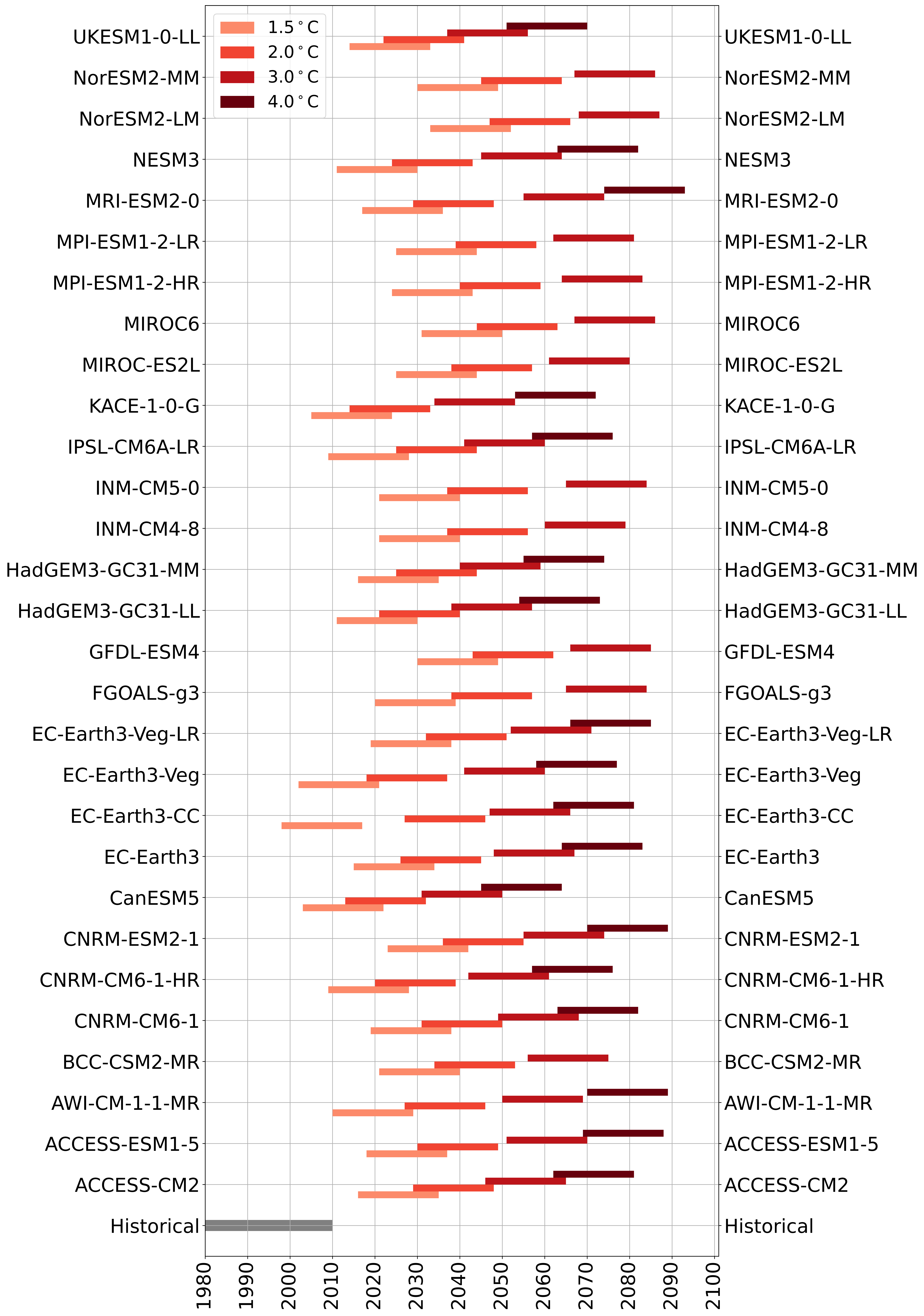
Zhang, X., Alexander, L., Hegerl, G. C., Jones, P., Tank, A. K., Peterson, T. C., … Zwiers, F. W. (2011). Indices for monitoring changes in extremes based on daily temperature and precipitation data. *WIREs Climate Change*, *2*(6), 851-870. doi: 10.1002/wcc.147

Zhang, Y., Li, Q., Ge, Y., Du, X., & Wang, H. (2022). Growing prevalence of heat over cold extremes with overall milder extremes and multiple successive events. *Communications Earth and Environment*, *3*(1), 73. doi: 10.1038/s43247-022-00404-x

Ziehn, T., Chamberlain, M., Lenton, A., Law, R., Bodman, R., Dix, M., … Druken, K. (2019). *Csiro access-esm1.5 model output prepared for cmip6 cmip historical* [dataset]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.4272

Zittis, G., Hadjinicolaou, P., Almazroui, M., Bucchignani, E., Driouech, F., El Rhaz, K., … Lelieveld, J. (2021). Business-as-usual will lead to super and ultra-extreme heatwaves in the middle east and north africa. *npj Climate and Atmospheric Science*, *4*(1), 20. doi: 10.1038/s41612-021-00178-7

Figure 1.



1.5 °C
2.0 °C
3.0 °C
4.0 °C
UKESM1-0-LL
NorESM2-MM
NorESM2-LM
NESM3
MRI-ESM2-0
MPI-ESM1-2-LR
MPI-ESM1-2-HR
MIROC6
MIROC-ES2L
KACE-1-0-G
IPSL-CM6A-LR
INM-CM5-0
INM-CM4-8
HadGEM3-GC31-MM
HadGEM3-GC31-LL
GFDL-ESM4
FGOALS-g3
EC-Earth3-Veg-LR
EC-Earth3-Veg
EC-Earth3-CC
EC-Earth3
CanESM5
CNRM-ESM2-1
CNRM-CM6-1-HR
CNRM-CM6-1
BCC-CSM2-MR
AWI-CM-1-1-MR
ACCESS-ESM1-5
ACCESS-CM2
Historical
1980
1990
2000
2010
2020
2030
2040
2050
2060
2070
2080
2090
2100

Figure 2.

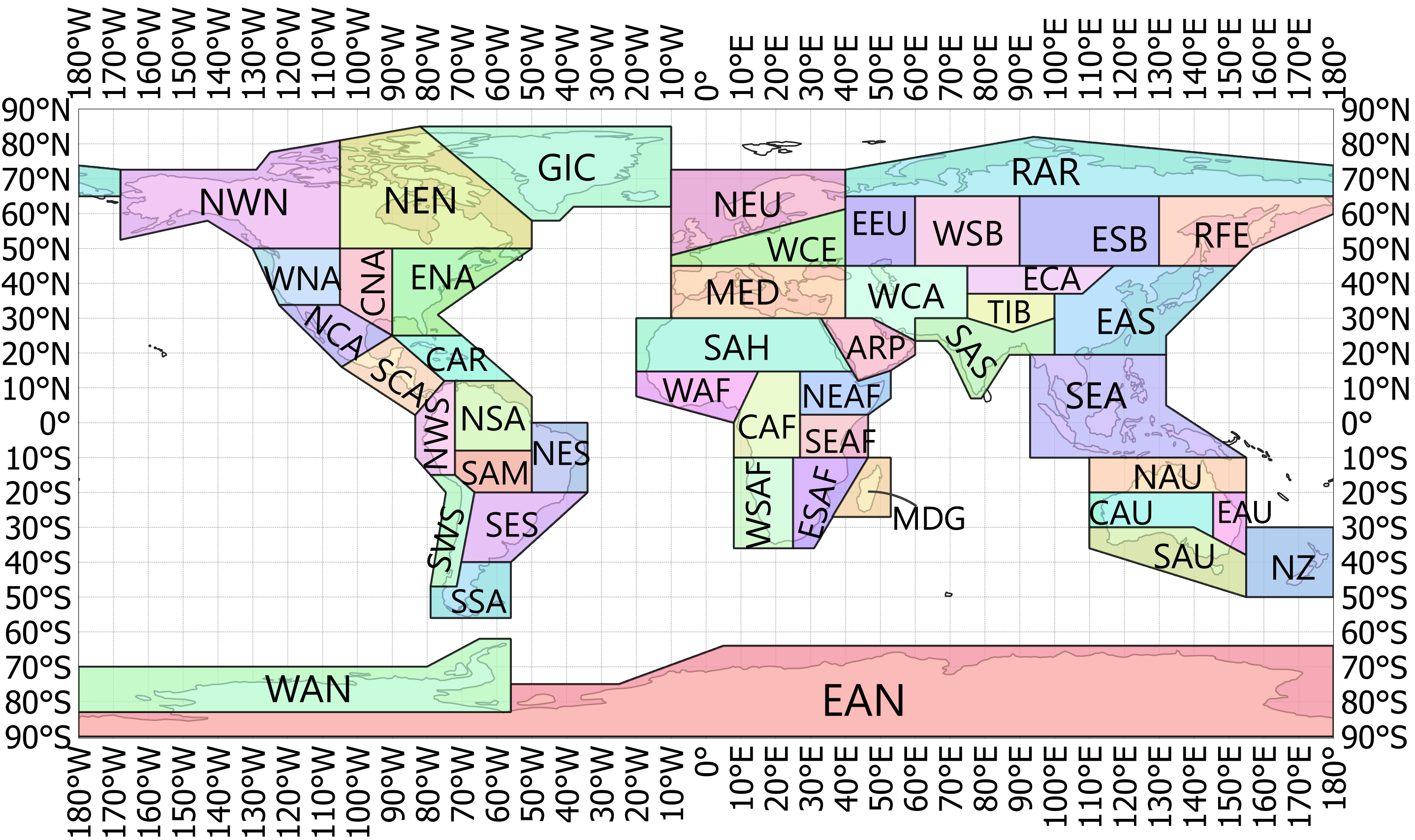

GIC
NWN
NEN
WNA
CNA
ENA
NCA
SCA
CAR
NWS
NSA
NES
SAM
SWS
SES
SSA
NEU
WCE
MED
SAH
WAF
CAF
NEAF
SEAF
WSAF
ESAF
MDG
ARP
RAR
EEU
WSB
ESB
RFE
WCA
ECA
TIB
EAS
SAS
SEA
NAU
CAU
EAU
SAU
NZ
WAN
EAN

Figure 3.

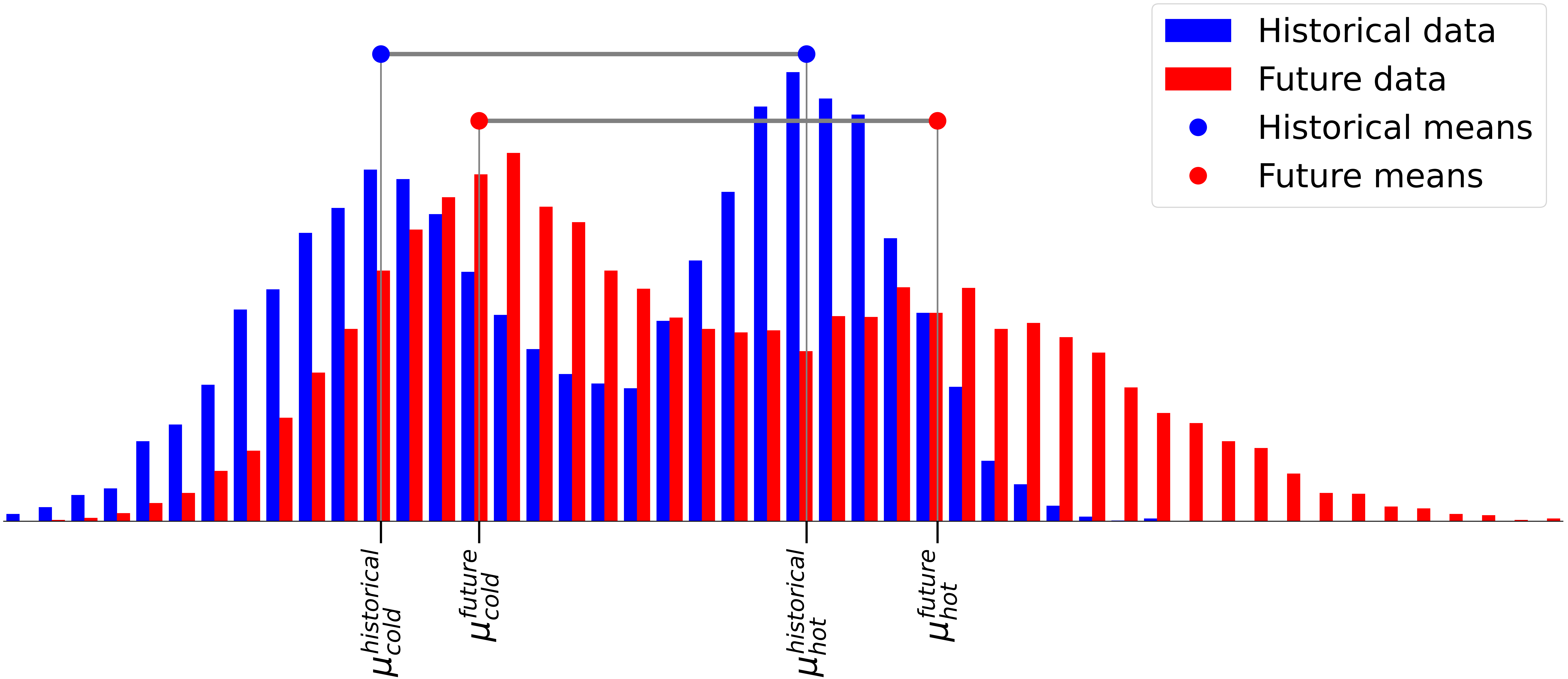
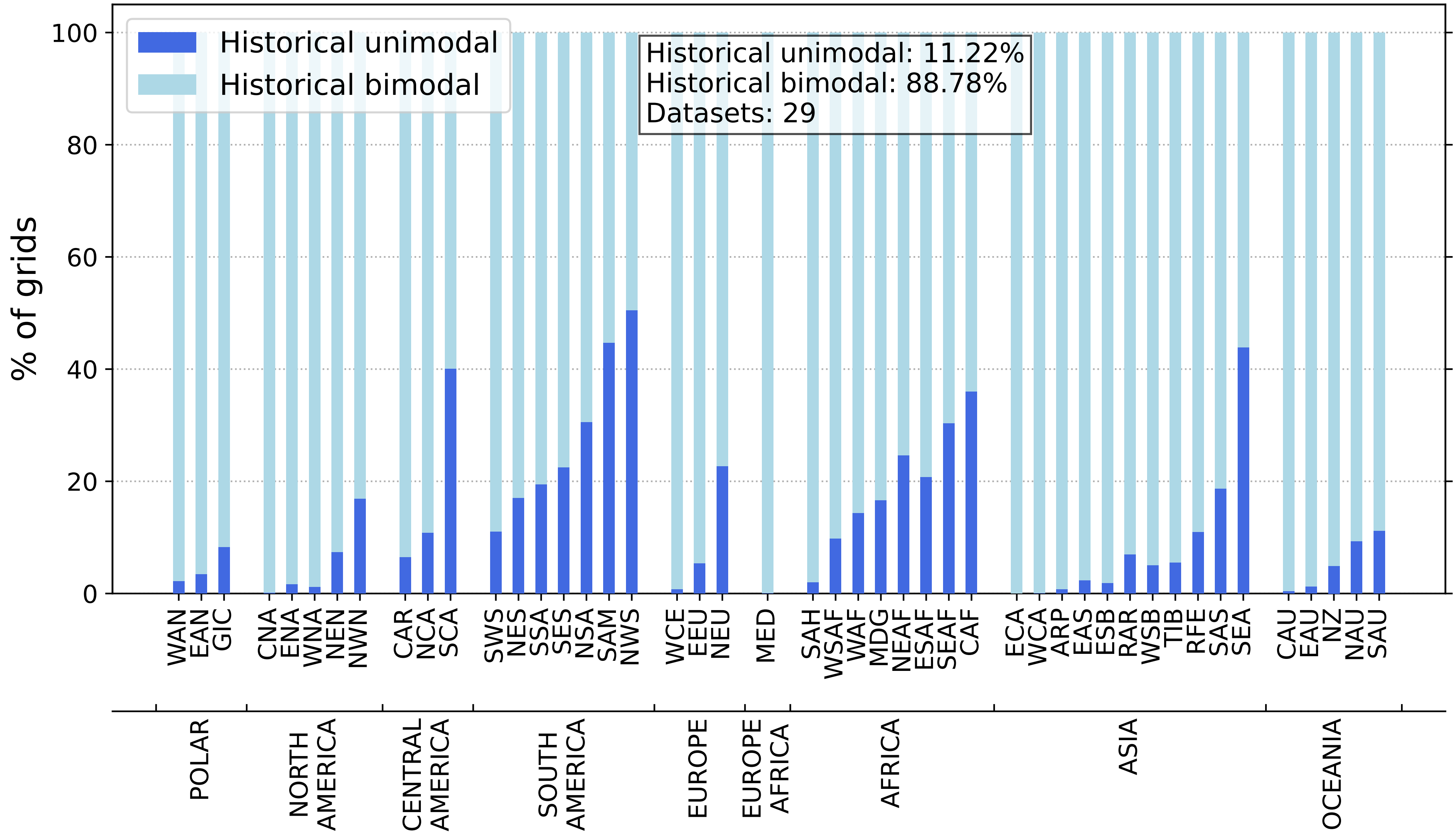

Historical data
Future data
Historical means
Future means
$\mu_{cold}^{historical}$
μ_{cold}^{future}
$\mu_{hot}^{historical}$
μ_{hot}^{future}

Figure 4.



Historical unimodal
Historical bimodal
Historical unimodal: 11.22%
Historical bimodal: 88.78%
Datasets: 29
% of grids
0
20
40
60
80
100
WAN EAN GIC CNA ENA WNA NEN NWN CAR NCA SCA SWS NES SSA SES NSA SAM NWS WCE EEU NEU MED SAH WSAF WAF MDG NEAF ESAF SEAF CAF ECA WCA ARP EAS ESB RAR WSB TIB RFE SAS SEA CAU EAU NZ NAU SAU
POLAR
NORTH AMERICA
CENTRAL AMERICA
SOUTH AMERICA
EUROPE
EUROPE AFRICA
AFRICA
ASIA
OCEANIA

Figure 5.

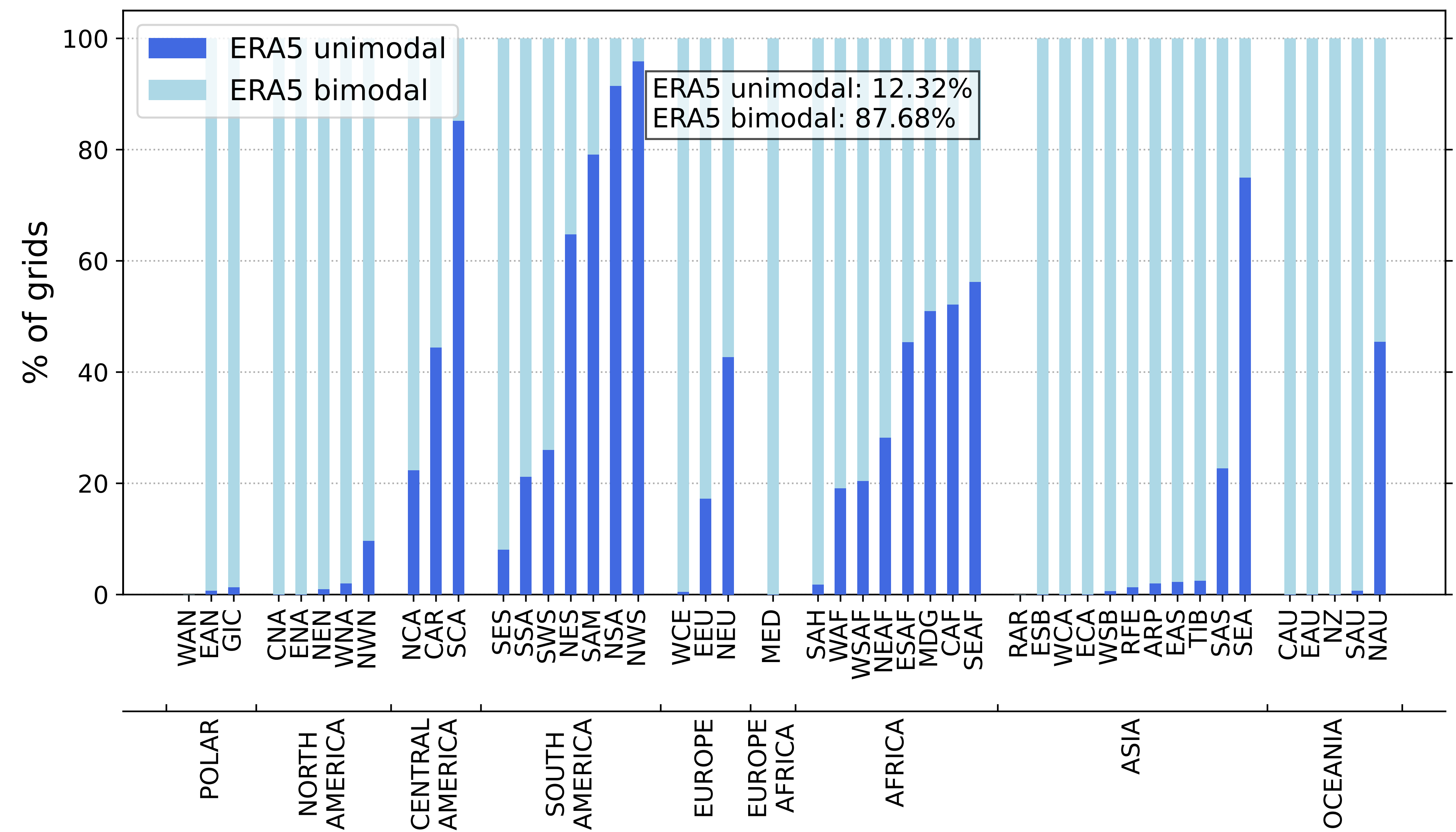
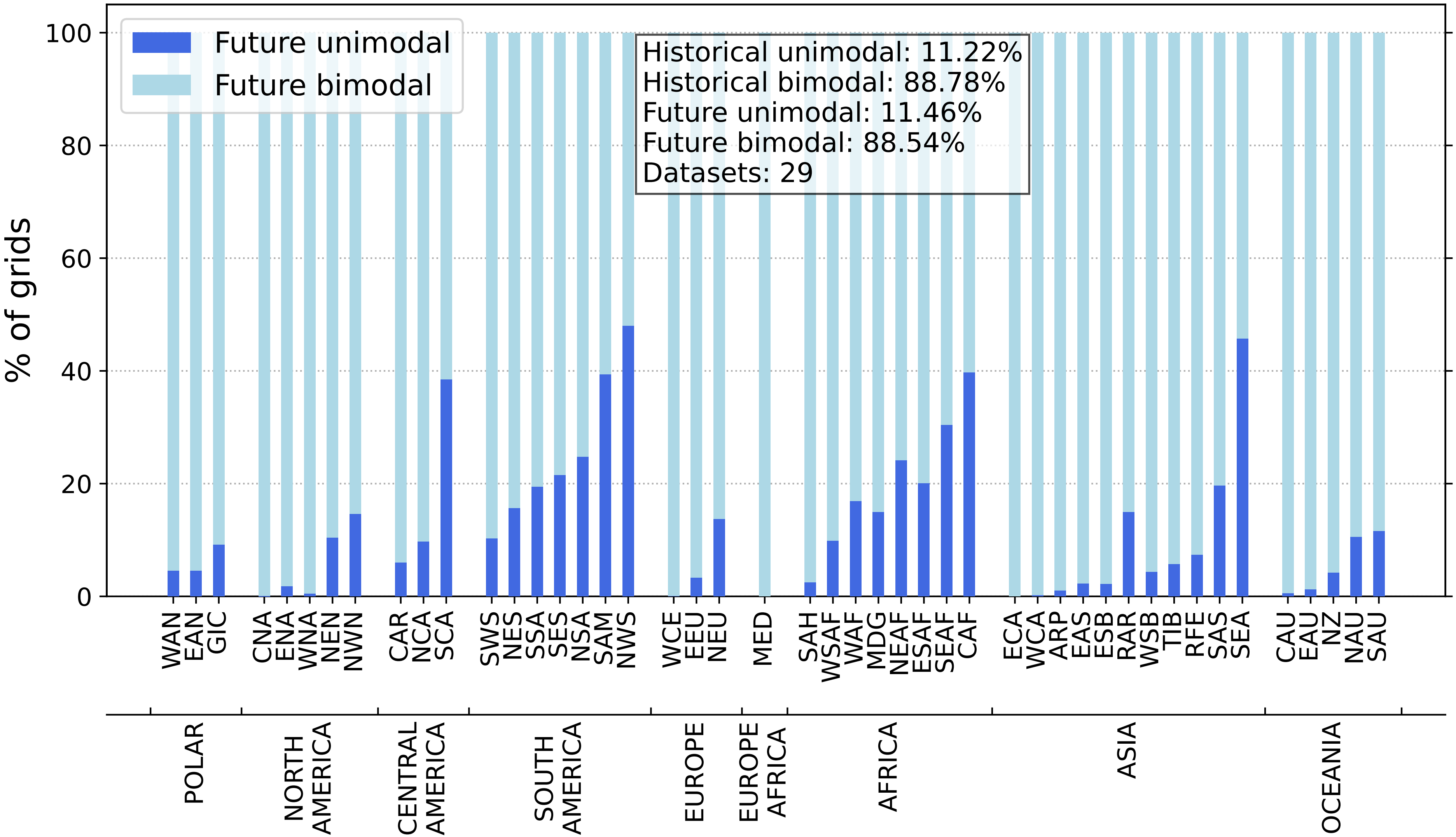

ERA5 unimodal
ERA5 bimodal
ERA5 unimodal: 12.32%
ERA5 bimodal: 87.68%
% of grids
0
20
40
60
80
100
WAN
EAN
GIC
CNA
ENA
NEN
WNA
NWN
NCA
CAR
SCA
SES
SSA
SWS
NES
SAM
NSA
NWS
WCE
EEU
NEU
MED
SAH
WAF
WSAF
NEAF
ESAF
MDG
CAF
SEAF
RAR
ESB
WCA
ECA
WSB
RFE
ARP
EAS
TIB
SAS
SEA
CAU
EAU
NZ
SAU
NAU
POLAR
NORTH AMERICA
CENTRAL AMERICA
SOUTH AMERICA
EUROPE
EUROPE AFRICA
AFRICA
ASIA
OCEANIA

Figure 6.

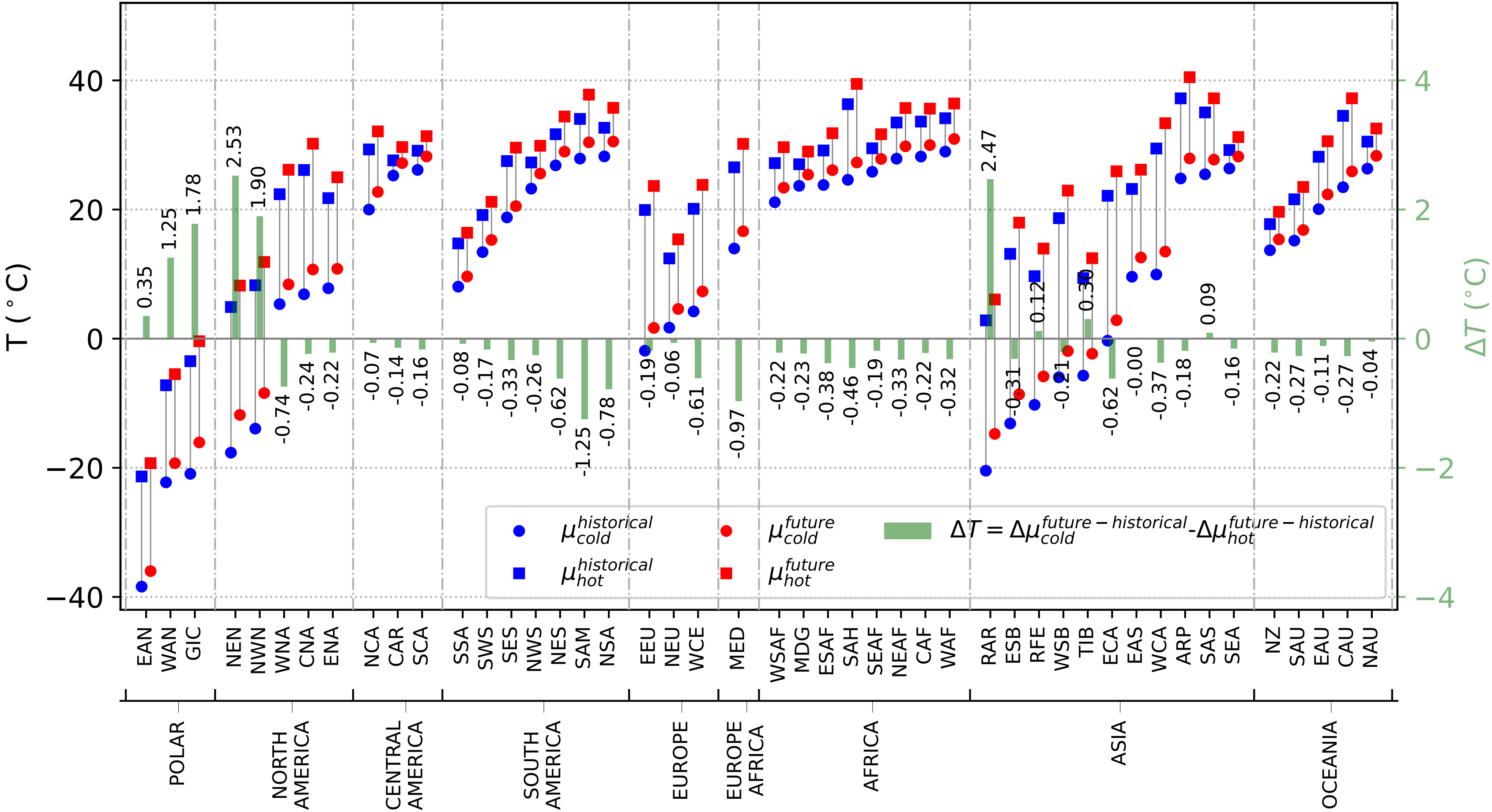


Future unimodal
Future bimodal
Historical unimodal: 11.22%
Historical bimodal: 88.78%
Future unimodal: 11.46%
Future bimodal: 88.54%
Datasets: 29
% of grids
0
20
40
60
80
100
WAN
EAN
GIC
CNA
ENA
WNA
NEN
NWN
CAR
NCA
SCA
SWS
NES
SSA
SES
NSA
SAM
NWS
WCE
EEU
NEU
MED
SAH
WSAF
WAF
MDG
NEAF
ESAF
SEAF
CAF
ECA
WCA
ARP
EAS
ESB
RAR
WSB
TIB
RFE
SAS
SEA
CAU
EAU
NZ
NAU
SAU
POLAR
NORTH AMERICA
CENTRAL AMERICA
SOUTH AMERICA
EUROPE
EUROPE AFRICA
AFRICA
ASIA
OCEANIA

Figure 7.

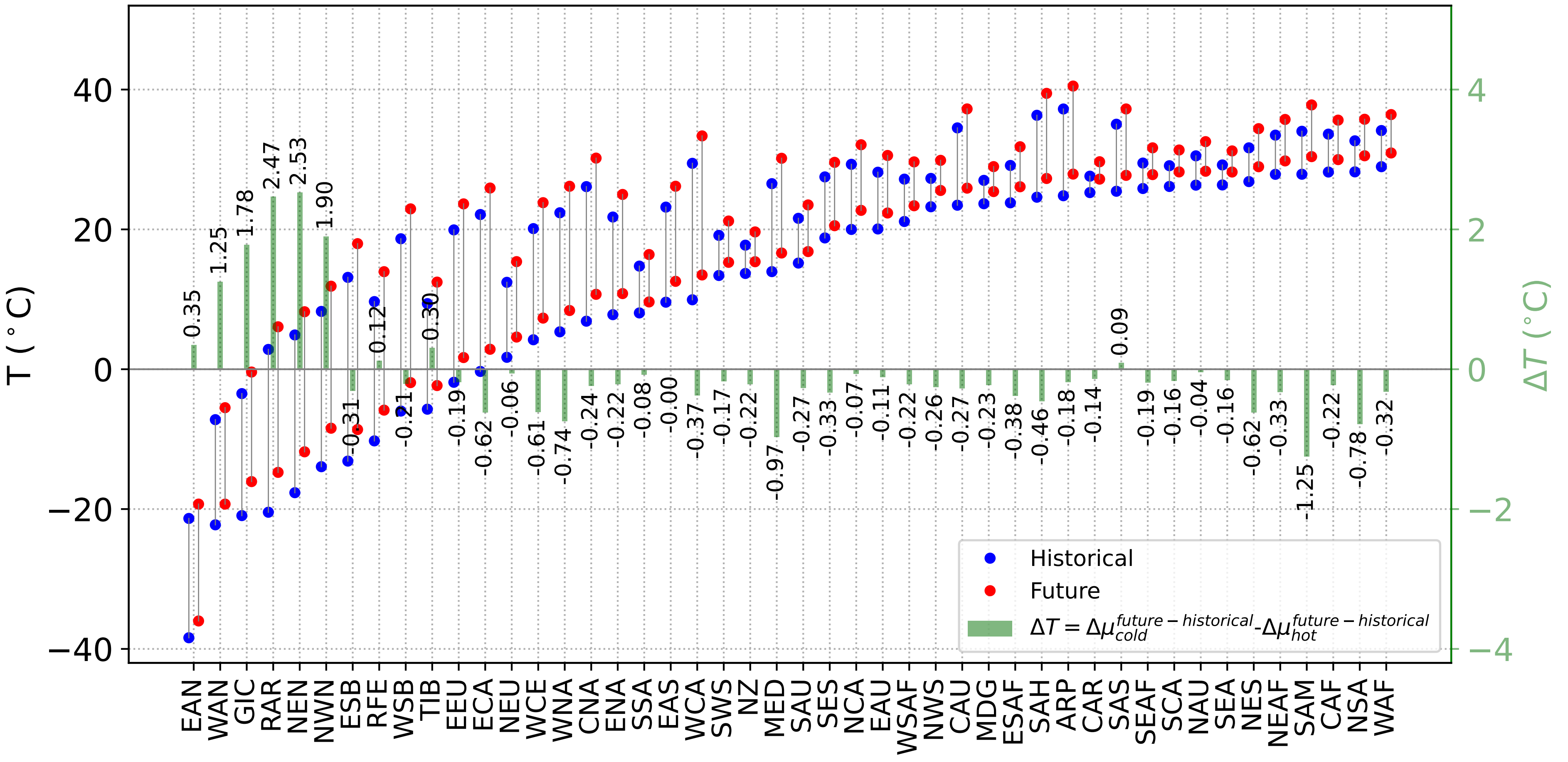

ACCESS-CM2
ACCESS-ESM1-5
AWI-CM-1-1-MR
BCC-CSM2-MR
CNRM-CM6-1
CNRM-CM6-1-HR
CNRM-ESM2-1
CanESM5
EC-Earth3
EC-Earth3-CC
EC-Earth3-Veg
EC-Earth3-Veg-LR
FGOALS-g3
GFDL-ESM4
HadGEM3-GC31-LL
HadGEM3-GC31-MM
INM-CM4-8
INM-CM5-0
IPSL-CM6A-LR
KACE-1-0-G
MIROC-ES2L
MIROC6
MPI-ESM1-2-HR
MPI-ESM1-2-LR
MRI-ESM2-0
NESM3
NorESM2-LM
NorESM2-MM
UKESM1-0-LL
ARP CAF CAR CAU CNA EAN EAS EAU ECA EEU ENA ESAF ESB GIC MDG MED NAU NCA NEAF NEN NES NEU NSA NWN NWS NZ RAR RFE SAH SAM SAS SAU SCA SEA SEAF SES SSA SWS TIB WAF WAN WCA WCE WNA WSAF WSB
UU
UB
BU
BB
0
20
40
60
80
100
%

Figure 8.



T (°C)
ΔT (°C)
$\mu_{cold}^{historical}$
$\mu_{hot}^{historical}$
μ_{cold}^{future}
μ_{hot}^{future}
$\Delta T = \Delta\mu_{cold}^{future-historical}$-$\Delta\mu_{hot}^{future-historical}$
EAN WAN GIC NEN NWN WNA CNA ENA NCA CAR SCA SSA SWS SES NWS NES SAM NSA EEU NEU WCE MED WSAF MDG ESAF SAH SEAF NEAF CAF WAF RAR ESB RFE WSB TIB ECA EAS WCA ARP SAS SEA NZ SAU EAU CAU NAU
POLAR
NORTH AMERICA
CENTRAL AMERICA
SOUTH AMERICA
EUROPE
EUROPE AFRICA
AFRICA
ASIA
OCEANIA
0.35 1.25 1.78 2.53 1.90 -0.74 -0.24 -0.22 -0.07 -0.14 -0.16 -0.08 -0.17 -0.33 -0.26 -0.62 -1.25 -0.78 -0.19 -0.06 -0.61 -0.97 -0.22 -0.23 -0.38 -0.46 -0.19 -0.33 -0.22 -0.32 2.47 -0.31 0.12 -0.21 0.30 -0.62 -0.00 -0.37 -0.18 0.09 -0.16 -0.22 -0.27 -0.11 -0.27 -0.04

Figure 9.

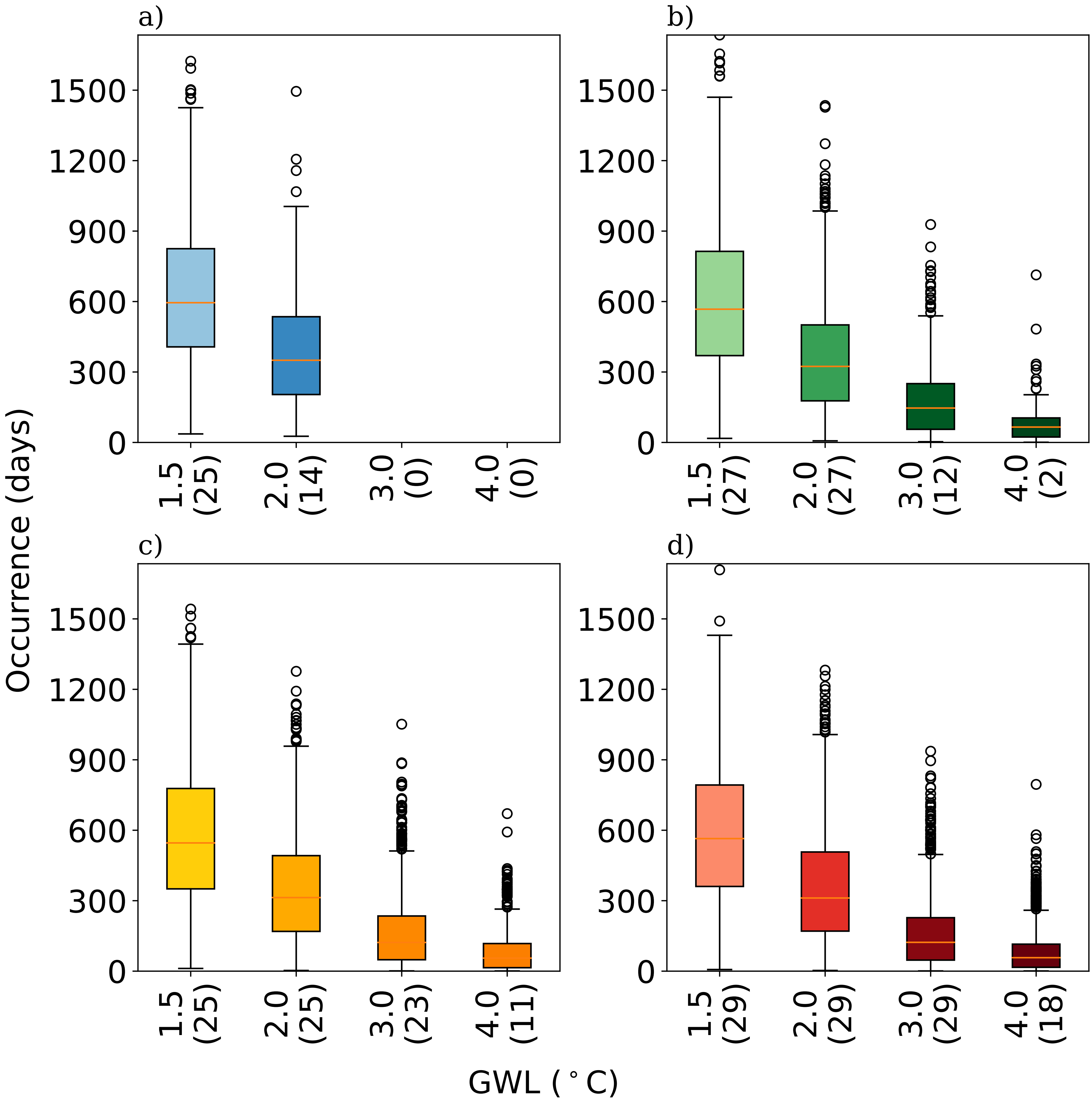


T ($^\circ$C)
ΔT ($^\circ$C)
Historical
Future
$\Delta T = \Delta\mu_{cold}^{future-historical}$-$\Delta\mu_{hot}^{future-historical}$
EAN
WAN
GIC
RAR
NEN
NWN
ESB
RFE
WSB
TIB
EEU
ECA
NEU
WCE
WNA
CNA
ENA
SSA
EAS
WCA
SWS
NZ
MED
SAU
SES
NCA
EAU
WSAF
NWS
CAU
MDG
ESAF
SAH
ARP
CAR
SAS
SEAF
SCA
NAU
SEA
NES
NEAF
SAM
CAF
NSA
WAF

Figure 10.

Global
Base 1880 days
Occurrence (days)
1.5°C(29)
2.0°C(29)
3.0°C(29)
4.0°C(18)
1940 days NWN
1850 days NEN
1660 days GIC
1900 days NEU
1650 days EEU
1670 days WCE
1720 days WSB
1780 days RAR
1770 days ESB
1690 days RFE
1640 days WNA
1790 days CNA
1730 days ENA
1910 days CAR
1960 days NCA
2530 days SCA
1610 days MED
1780 days WCA
1950 days SAH
1870 days ARP
1720 days TIB
1890 days ECA
1840 days EAS
1950 days NSA
2620 days NWS
2310 days SAM
1660 days NES
2040 days WAF
2400 days CAF
2390 days NEAF
2030 days SAS
2780 days SEA
2210 days SEAF
2110 days MDG
1850 days WSAF
2020 days ESAF
1980 days NAU
1890 days EAU
1920 days CAU
1820 days SAU
1670 days NZ
1920 days SWS
2170 days SES
2110 days SSA
1780 days WAN
1600 days EAN

Figure 11.



a)
b)
c)
d)
Occurrence (days)
0
300
600
900
1200
1500
1.5 (25)
2.0 (14)
3.0 (0)
4.0 (0)
1.5 (27)
2.0 (27)
3.0 (12)
4.0 (2)
1.5 (25)
2.0 (25)
3.0 (23)
4.0 (11)
1.5 (29)
2.0 (29)
3.0 (29)
4.0 (18)
GWL (°C)