Understanding Projected Heatwave Impacts on Demographic Groups in Different Climate Change Storylines

Cuiyi Fei, Supervisors: Prof. Rachel White, Prof. Ethan Raker

Abstract. The bias in climate model projections of heatwaves and surface temperature trends is primarily observed at regional spatial scales. However, demographic projections in various SSPs (Shared Socioeconomic Pathways) are typically constrained to human-defined country or provincial levels rather than natural climatological regions, underscoring the importance of county-level analysis for accurate local policy-making. This study uses county-level population projections to examine demographic changes by age, sex, and race in the US affected by biases in heatwaves and surface temperature trends in CMIP6 climate models. Most CMIP6 models tend to underestimate persistent heatwaves in northeastern North America, and surface temperature trends are underestimated in the western United States. Comparing more vulnerable regions to less vulnerable ones in terms of persistent heatwaves and surface temperature trends in summer, we find that the child population and white population exhibit distinct features in most comparisons. The Hispanic population and sex ratios also show differences in some cases, while the Black population, elderly population, and other racial groups appear less sensitive to these vulnerabilities. The population affected by climate model biases on heatwaves and surface temperature trends is highly sensitive to various uncertainties—the internal variability can adjust the total vulnerable population by a factor of 2 to 10. In comparison, uncertainties due to different models are relatively small. Our results suggest the need for detailed climate adaptation storylines that indicate the relative likelihood and magnitude of populations affected by climate model biases on extreme heat-related variables.

1 Introduction

Heatwaves, characterized by prolonged periods of abnormally high surface temperatures, have been increasing in both frequency and intensity due to climate change (Perkins et al., 2012). These extreme events are linked to a variety of disasters, including impacts on agriculture, wildfires, energy demand, ecosystems, and public health (Brás et al., 2021; Shaposhnikov et al., 2014; Larcom et al., 2019; Yin et al., 2023; Kovats and Kristie, 2006). Vulnerable demographic groups are particularly at risk, as heatwaves exacerbate existing health conditions and significantly increase mortality rates among populations less capable of adapting to extreme temperatures (Chambers, 2020; Kravchenko et al., 2013). Given the severity of these impacts, there is an urgent need to estimate the potential consequences of heatwaves on human populations.

However, accurately simulating heatwaves in climate models presents a significant challenge due to inherent biases (Domeisen et al., 2023). These biases are often related to the models' difficulties in accurately representing atmospheric large-scale circulations, such as blocking events, which can lead to underestimating the persistence of heatwaves (Brunner et al., 2018; Masato et al., 2013). Furthermore, the poor simulation of interactions between different components of the Earth system, like air-ocean and air-land interactions, further complicates heatwave predictions (Mueller and Seneviratne, 2012; Wills et al., 2024). Beyond

1

heatwaves, climate models also struggle with simulating surface temperature trends at regional scales, despite their success in capturing global mean temperature trends (Davy and Esau, 2014; Wang et al., 2023). Fortunately, model biases tend to remain relatively constant under global warming, allowing for more reliable projections after bias correction (Krinner and Flanner, 2018).

This research employs the storyline method alongside a simple bias correction to improve the estimation of heatwave impacts and summer surface temperature trends (Shepherd et al., 2018). The uncertainties addressed in this study primarily stem from two sources: model bias and internal variability, as the scenario uncertainty is embedded within different SSPs data already. By quantifying these uncertainties, the study aims to provide a more accurate assessment of the demographic groups most affected by heatwaves and temperature trends, considering model biases.

A novel aspect of this research is the use of population data calculated by Hauer (2019) at the county level under different Shared Socioeconomic Pathways (SSPs). This granular approach offers greater flexibility in assessing the impacts of natural systems, which do not adhere to national or provincial boundaries. Although this population data has been utilized in coastal vulnerability studies, there is currently no research at the county level focusing on the demographic impacts of heatwaves.

This study aims to explore how different demographic groups—defined by age, sex, race, and geographic location—are affected by heatwaves and long-term trends in summer temperatures. Understanding the uncertainties surrounding the impacts of these extreme weather events is crucial for developing targeted public health interventions and informing policy decisions. By analyzing historical data on heatwave occurrences and summer temperature trends, this research seeks to identify the most at-risk demographic groups and examine how these risks have evolved over time.

2 Data and methods

2.1 Climate data

In this research, we analyzed ERA5 reanalysis - an approximation of the observed state of the atmosphere - and historical simulations from five CMIP6 models, as well as Scenario Model Intercomparison Project (ScenarioMIP) projections (SSP1-2.6, SSP2-4.5, SSP3-7.0, SSP5-8.5), with resolutions interpolated to 1°x1° latitude-longitude (Hersbach et al., 2020; Eyring et al., 2016; O'Neill et al., 2016). While this resolution is relatively coarse, it is sufficient for capturing large-scale circulation patterns. The five CMIP6 models used in this study are the Australian Community Climate and Earth System Simulator coupled model (ACCESS-CM2), the Centre National de Recherches Météorologiques Earth System Model (CNRM-ESM2-1), The Flexible Global Ocean-Atmosphere-Land System Model Grid-Point Version 3 (FGOALS-g3), the Korea Meteorological Administration Advanced Community Earth-System model (KACE), and the Meteorological Research Institute Earth System Model Version 2.0 (MRI-ESM2.0). Each selected model includes three ensemble members in historical experiments and in each SSP projection. The original surface temperature data was downloaded through acccmip6 software.

ACCESS-CM2 includes the UK Met Office (UKMO) Unified Model (UM) atmospheric model (v10.6) in the GA7.1 configuration, with a resolution of N96 (1.875°×1.25°), 85 vertical levels; the Community Atmosphere Biosphere Land Exchange land surface model (CABLE2.5) developed by Australian researchers; the Geophysical Fluid Dynamics Laboratory (GFDL)

Modular Ocean Model (MOM5) at 1° resolution; the Los Alamos National Lab Sea Ice Model (LANL CICE5.1) in the UKMO configuration; and the Ocean Atmosphere Sea Ice Soil Model Coupling Toolkit (OASIS-MCT) maintained by the Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS) and the Centre National de la Recherche Scientifique (CNRS) in France (Bi et al., 2020; Walters et al., 2019; Griffies et al., 2012; Hunke et al., 2010; Craig et al., 2017).

65   CNRM-ESM2-1 is the second-generation Earth system model developed by the CNRM/CERFACS modeling group. It includes ARPEGE-Climat v6.3 for the atmosphere, developed at CNRM; Nucleus for European Modelling of the Ocean (NEMO), developed by the NEMO consortium, which includes Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), CNRS, Istituto Nazionale di Geofisica e Vulcanologia (INGV), Mercator-ocean, Met-Office, and National Oceanography Centre (NOC); Global Experimental Leads and Sea Ice for Atmosphere and Ocean (GELATO) for sea ice, developed at CNRM

70 and embedded in NEMO for coupled simulations or in Surface Externalisée (SURFEX) v8.0 for SST-prescribed simulations; Interaction Soil-Biosphere-Atmosphere CNRM version of Total Runoff Integrating Pathways (ISBA-CTRIP) for land surface processes and river routing; FLake for lake thermal processes, developed at IGB-Berlin (Leibniz-Institute of Freshwater Ecology and Inland Fisheries) and revised at CNRM; PISCESv2-gas (Pelagic Interactions Scheme for Carbon and Ecosystem Studies volume 2) for marine biogeochemistry; ISBA-CC (Interaction Soil Biosphere Atmosphere Carbon Cycle) for conti-

75 nental biogeochemistry; TACTIC for aerosols; and REPROBUS for chemistry. Components without specific source attribution are developed by CNRM (Séférian et al., 2019; Madec et al., 2017; Mélia, 2002; Voldoire et al., 2017).

    FGOALS-g3 includes Version 3 of the Grid-Point Atmospheric Model of LASG-IAP (GAMIL3) for the atmosphere, Version 3 of the LASG-IAP Climate System Ocean Model (LICOM3) for the ocean, Version 4 of the LANL CICE for sea ice, and the CAS-Land Surface Model (CAS-LSM) for land surface processes, based on the Community Land Model Version 4.5

80 (CLM4.5). The model also incorporates two couplers: CPL7, developed at the National Center for Atmospheric Research (NCAR), and C-Coupler2 (Community Coupler Version 2), developed at Tsinghua University. Components without source attribution are developed by LAST-IAP (Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics - Institute of Atmospheric Physics) (Li et al., 2020).

    KACE is composed of the UM as the atmospheric component, MOM from NOAA GFDL as the ocean component, LANL

85 CICE as the sea ice component, Joint UK Land Environment Simulator (JULES) as the land component, and the OASIS3-MCT coupler (Lee et al., 2020).

    MRI-ESM2.0 consists of four major component models: an atmospheric general circulation model (AGCM) MRI-AGCM3.5, an aerosol model called Model of Aerosol Species IN the Global Atmosphere (MASINGAR), an atmospheric chemistry model MRI Chemistry Climate Model version 2.1 (MRI-CCM2.1), and an ocean–sea-ice general circulation model MRI Community

90 Ocean Model version 4 (MRI COMv4) (Yukimoto et al., 2019).

    To ensure comparability with ERA5 data from 1940 and historical experiments in CMIP6, the historical period for this research is defined as 1940–2014, and ScenarioMIP time period is 2015-2100. In this study, heatwaves are identified as periods when temperatures exceed the 95th percentile of all summer (June-July-August) days, after applying a linear detrending, for at least three consecutive days. Additionally, the summer seasonal temperature trend per year is calculated.

2.2 Population data and methods

In the U.S., we use county-level population projection data from Hauer (2019). This dataset provides projections by age, sex, and race in five-year intervals for all U.S. counties from 2020 to 2100. The projections are calibrated using historic U.S. census data from 1990 to 2015. Population projections are generated using cohort-change differences (CCDs) through ARIMA models for eighteen five-year age groups (0–85+), two sex groups (Male and Female), and four race groups (White Non-Hispanic, Black Non-Hispanic, Other Non-Hispanic, and Hispanic). These projections are further controlled by the Shared Socioeconomic Pathways (SSPs).

Because sociological county definitions do not align with natural latitude-longitude grids, we interpolate the geographical centroids of the counties to correspond with the climatological data's latitude and longitude grid points. We calculate the distance between the centroids of the counties and the four surrounding lat-lon grid points. Counties closest to a grid point within the threshold are considered within that threshold, and vice versa.

For the demographic analysis, we employ the Welch Two Sample t-test (Welch, 1947). This test is particularly effective in situations where the variances of two populations are unequal, making it suitable for small sample sizes or samples with different variances. The test involves calculating the means and variances of the two samples, then using these results to compute the degrees of freedom and the t-statistic.

3 Results

3.1 Climate model bias patterns

We first examined the extent to which persistent heatwave days can be simulated by historical climate models compared to ERA5 reanalysis. Given that the number of days exceeding the 95th percentile in summer from 1940 to 2014 is constant, the discrepancies between climate model simulations and reanalysis reflect differences in the persistence of these extreme heatwaves. In Figure 1, lighter colors indicate fewer heatwave days simulated by the climate models relative to reanalysis/observation. The results across all climate models are similar: 1) significant biases in heatwave days are observed in northeastern North America and central Asia, 2) heatwaves over land are poorly captured, with simulated heatwave days being at least 40% fewer than in reanalysis/observation, and 3) heatwaves over the ocean are generally well captured, except in FGOALS-g3. Focusing on the US, the lighter color in northeastern regions suggests that people living there in the future may experience more frequent and unexpected heatwaves.

We then examined the uncertainties associated with the bias in persistent heatwave days across climate models. Figure 2a displays the standard deviation of persistent heatwave day biases for each climate model mean, highlighting uncertainties stemming from various model biases. Figure 2b-f show the standard deviations among ensembles of each climate model, reflecting natural variability in the climate system. The standard deviations in each climate model are approximately 5% in North America, indicating that the differences observed between climate models and ERA5 reanalysis in Figure 1 are more than 2 standard deviations from the heatwave day spectrum, suggesting a solid model bias. Notably, the standard deviation
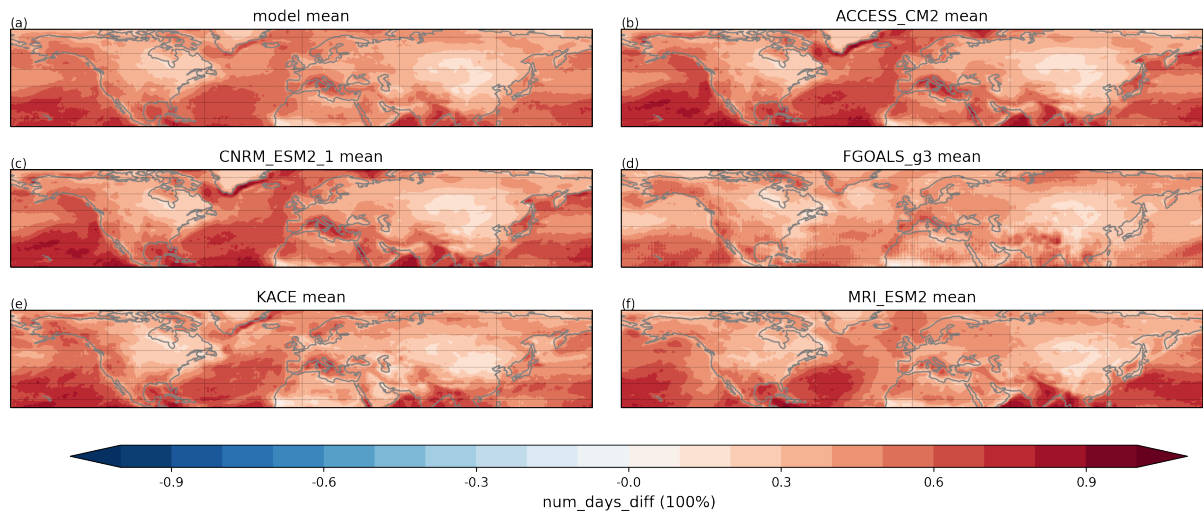
Figure 1. The ratio of simulated persistent heatwave days in climate models relative to ERA5 reanalysis. (a) Average of five models' mean, (b) average of ACCESS-CM2 ensembles, (c) average of CNRM-ESM2-1 ensembles, (d) average of FGOALS-g3 ensembles, (e) average of KACE1-0-G ensembles, and (f) average of MRI-ESM2-1 ensembles.

among different climate models is small over land in North America but quite large along the coast. In North America, the standard deviation among different climate models is mostly under 4% over land, except for two regions in the western and midwestern US. Meanwhile, the standard deviations within each climate model ensemble are around 5%, indicating that natural
130   variability may contribute to greater uncertainty than model biases.

In Figure 3, model bias in surface temperature trends differs from the bias observed in persistent heatwave days. While climate models generally capture global mean temperature trends accurately, these trends vary across different Shared Socioeconomic Pathways (SSPs). To analyze regional temperature trends, we subtract the global mean trend from surface temperature trends at each grid point and compare these regional trends to ERA5 data. Figure 3 illustrates the differences between climate
135   model regional trends and ERA5 regional trends. Common features across all models include: 1) A significant underestimation of land surface temperature in the Arctic region, indicating an Arctic Amplification signal; 2) Underestimation of surface temperature trends in the western US (particularly over the Rockies), the Iranian Plateau, and the Tibetan Plateau, suggesting that the models struggle to represent topographic effects accurately. Consequently, while the west US may not experience the persistent heatwave day bias, residents could face higher summer average temperatures than predicted by climate models. People
140   in different regions in the US may be influenced by different kinds of climate model biases.

Finally, we examined the standard deviations of surface temperature trends among different climate models and their ensembles, as shown in Figure 4. Figure 4a reveals that the standard deviation is less than 0.01 K/year, indicating that 2 standard deviations among different models are still smaller than the 'model bias' over the Rockies, suggesting discrepancies among models. However, the standard deviations within climate model ensembles exceed 0.04 K/year, meaning that 2 standard devia-
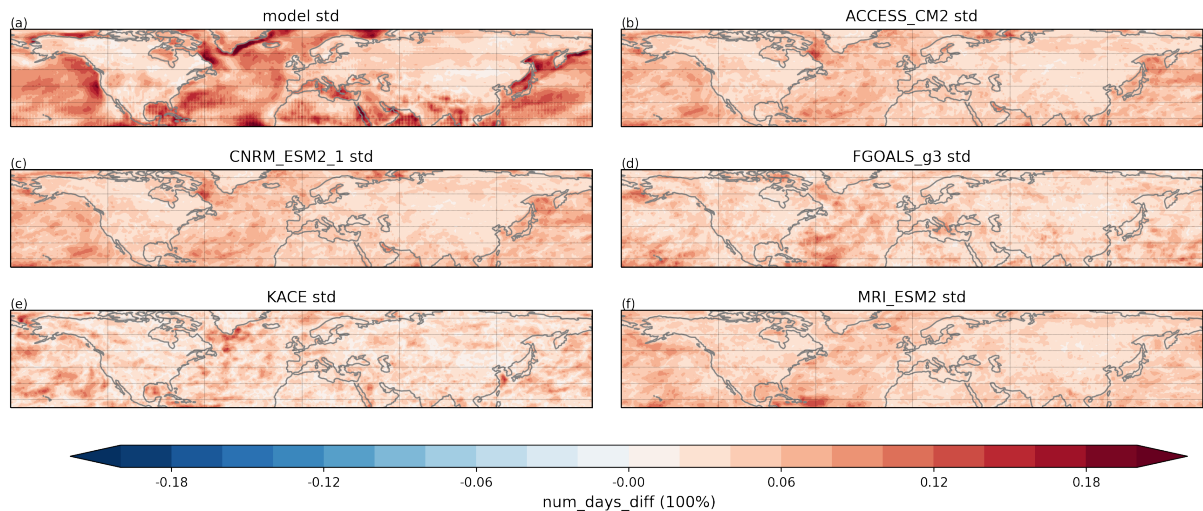
Figure 2. The ratio of simulated persistent heatwave days in climate models relative to ERA5 reanalysis. (a) Standard deviation of five models' mean, (b) standard deviation of ACCESS-CM2 ensembles, (c) standard deviation of CNRM-ESM2-1 ensembles, (d) standard deviation of FGOALS-g3 ensembles, (e) standard deviation of KACE1-0-G ensembles, and (f) standard deviation of MRI-ESM2-1 ensembles.
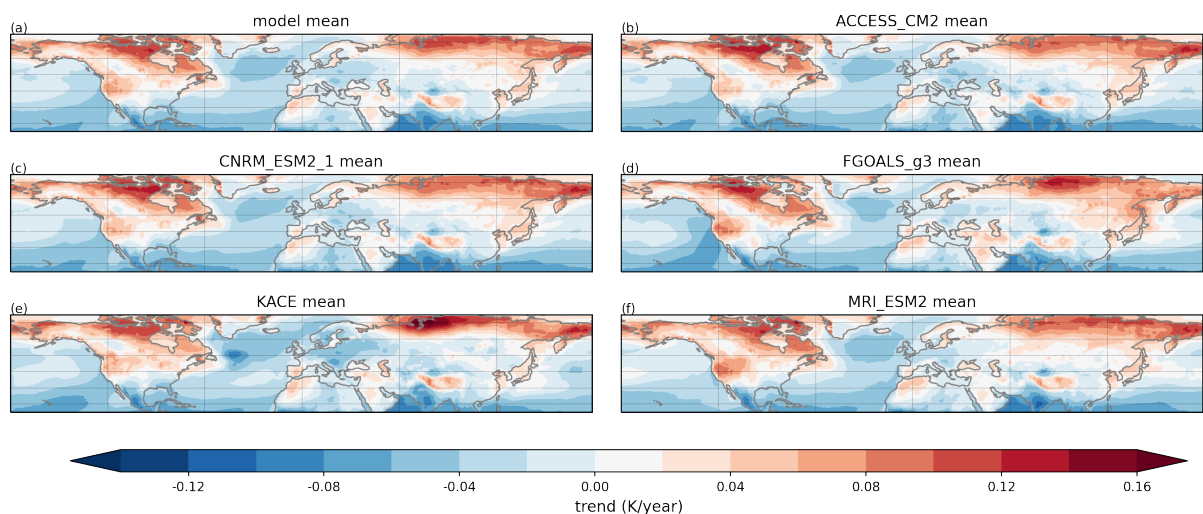

Figure 3. Same as Figure 1, except for surface temperature trends rather than persistent heatwave days.

145   tions can surpass the so-called model bias. This implies that the observed 'model bias' may be due to natural variability alone, meaning that the differences shown in Figure 3 could be attributed to this variability. This does not necessarily indicate that the models are incorrect; rather, even if the models are accurate, we would still face uncertainty about future conditions. While residents in the west US might not be directly affected by model bias, all climate models face challenges in predicting future
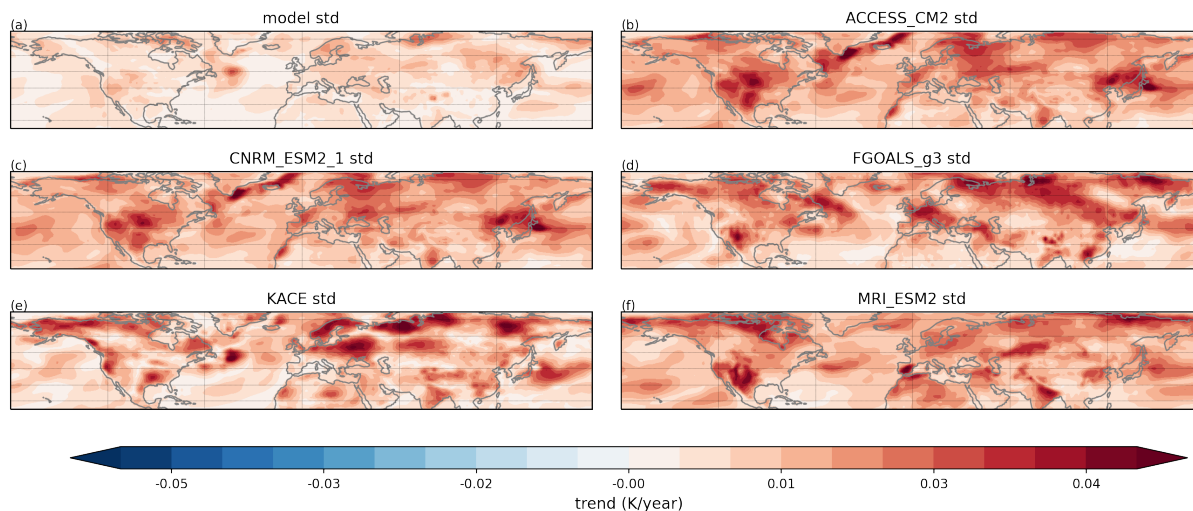
6

Figure 4. Same as Figure 12, except for surface temperature trends rather than persistent heatwave days.

conditions due to this variability. This underscores the importance of employing a storyline approach to prepare for all possible
scenarios.

Since the persistent heatwave days bias and surface temperature trends are calculated over a long time period, the pattern
of climate model bias remains consistent over the years. We also examined these biases across different SSPs and found no
significant differences from the biases observed in historical simulations (about 10% compared with the model discrepancies),
confirming our initial hypothesis. Given that our current metrics are based on relative thresholds, the different SSPs do not
impact the final results. Therefore, we will focus on biases from historical simulations in this report to maintain simplicity.
However, it's important to note that when absolute values, such as absolute heatwave temperature, are considered, different
SSPs will play a crucial role.

3.2 Population maps associated with climate model bias

After assessing the climate model biases and the effects of natural variability, we next connect these findings with population
maps to identify the counties affected by these biases and variabilities.

We first examine the population distribution for 2020 and 2100 in Figures 5 and 6 to understand demographic trends. Since
population projections begin from 2015, differences among various SSPs are minimal. Figure 5 shows that the population is
concentrated along the east and west coasts, with some counties in southern California and Florida having the largest popula-
tions. In contrast, the Midwest and western US have relatively few counties with large populations

The situation changes significantly by 2100, with a marked increase in urbanization over the next 80 years. Most of the
population becomes concentrated in a limited number of counties, with a noticeable trend of faster population growth in the
southern regions (California, Texas, and Florida) compared to the northeast US and the Great Lakes area. It is striking to observe
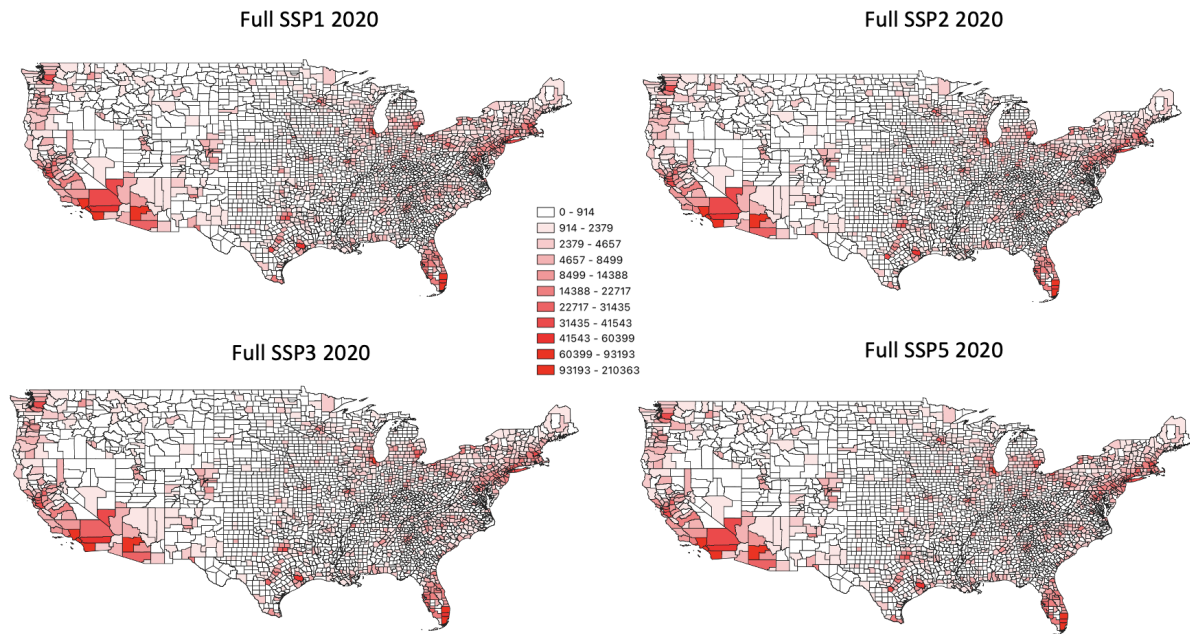
7

Figure 5. The population projection in each county in 2020 in continental US. The population projection starts from 2015.

that, by 2100, some counties could have populations in the millions, compared to only hundreds of thousands in 2020—a tenfold increase! Besides the overall demographic shifts, it's interesting to note the variations in population distribution across
170   different SSPs. In SSP3, which assumes high regional rivalry, the population in the western US mountain regions is significantly lower than in the other SSPs. Conversely, SSP1, SSP2, and SSP5 all indicate that, regardless of the socioeconomic development path, the population will be concentrated in major cities.

An interesting result from comparing the demographic features across different SSPs is that the population characteristics are not significantly different among them. In the analysis below, we focus on SSP5, as it has the largest population and clearly
175   illustrates these features. While there are instances where different SSPs exhibit distinct features, these will be discussed in the next subsection on population characteristics.

First, we focus on the counties affected by the persistent heatwave days bias. We established a series of thresholds based on the bias value in the continental US region, roughly defined as 25-50N, 125-70W over land. This definition might have some inaccuracies, especially around the Great Lakes. We initially set thresholds at 90%, 75%, and 50% to examine the
180   demographic distributions. However, a challenge arose with the first figure in this series: since the persistent heatwave days bias is concentrated in the northeastern US, using the 90% threshold resulted in no US counties being selected (indicating that Canadians might be more affected by this model bias!). As a result, we adjusted the thresholds to 10% and 25% in place of 90% and 75% for the persistent heatwave days analysis. These adjusted thresholds allow for a more meaningful comparison with the counties affected by surface temperature trends, enabling us to identify areas where the distributions do not overlap.
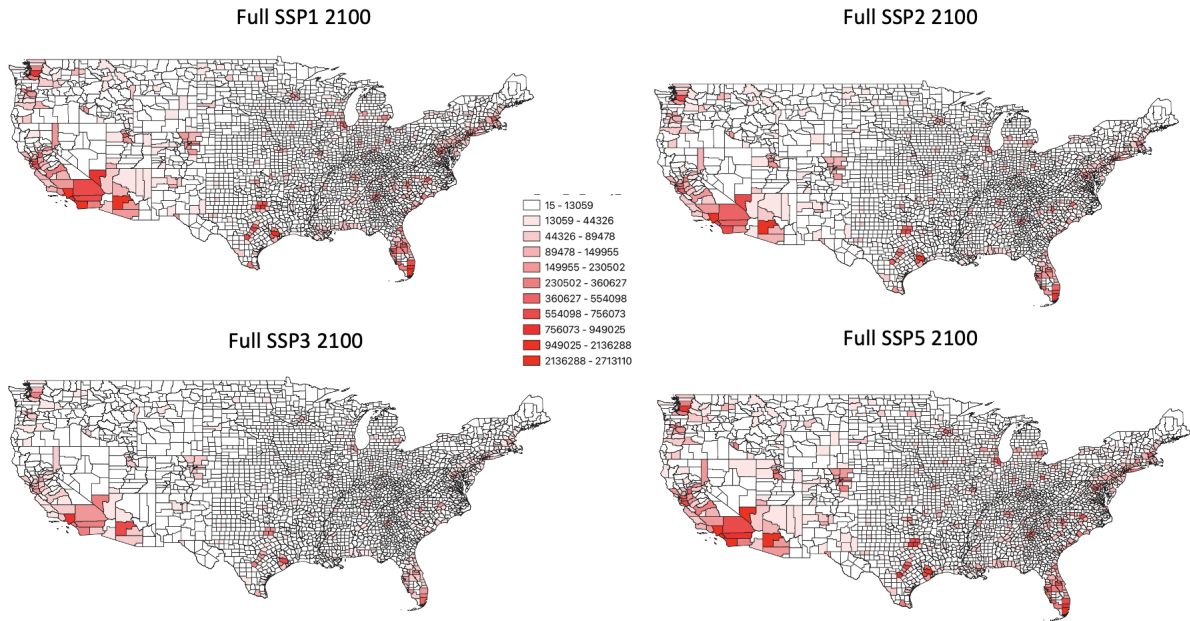
Figure 6. Same as Figure 5, except for 2100 projection.

Starting from Figure 7, we adjusted the legend to ensure that white represents regions not selected by the thresholds, while light pink indicates areas with few people living in the selected counties. Moving from left to right, we see that with the 10% threshold, only a few counties in the southern and southwestern US are not selected; with the 25% threshold, the selected area shrinks northward and eastward, with only some counties in the northwestern US remaining selected; and with the 50% threshold, the selection is limited to the northeastern US and areas around the Great Lakes. The number of counties selected ranges from 2,831 to 695. The standard deviation, reflecting the role of internal variability, can be significant, particularly when one standard deviation is removed. Given the common definition of natural variability in climate science—typically 2 standard deviations—the population affected by this variability could vary by a factor of 2 to 10 (not shown). The number of counties selected with 1 standard deviation ranges from 1,066 to 2,399.

Beyond the regions severely affected by model bias in persistent heatwave days, it's crucial to identify where uncertainties are substantial for informed policy-making. Figure 8 highlights the regions most impacted by model uncertainties and natural variability, using thresholds ranging from 90% to 50%. The first row illustrates the impact of model uncertainties, while the second row shows the influence of natural variability. Model uncertainties are concentrated in southern Florida, parts of the Rockies, and the Midwest. As the threshold is lowered to 75%, the affected areas expand, including the northern boundary of the US, the Great Lakes, the southeastern US, and southern California. At the 50% threshold, most of the US is affected, except for the northeastern US and parts of Arizona, among others. Natural variability impacts start in similar regions—Florida and the western US—but cover a much larger area, including parts of the southern US. As the threshold lowers, the affected regions
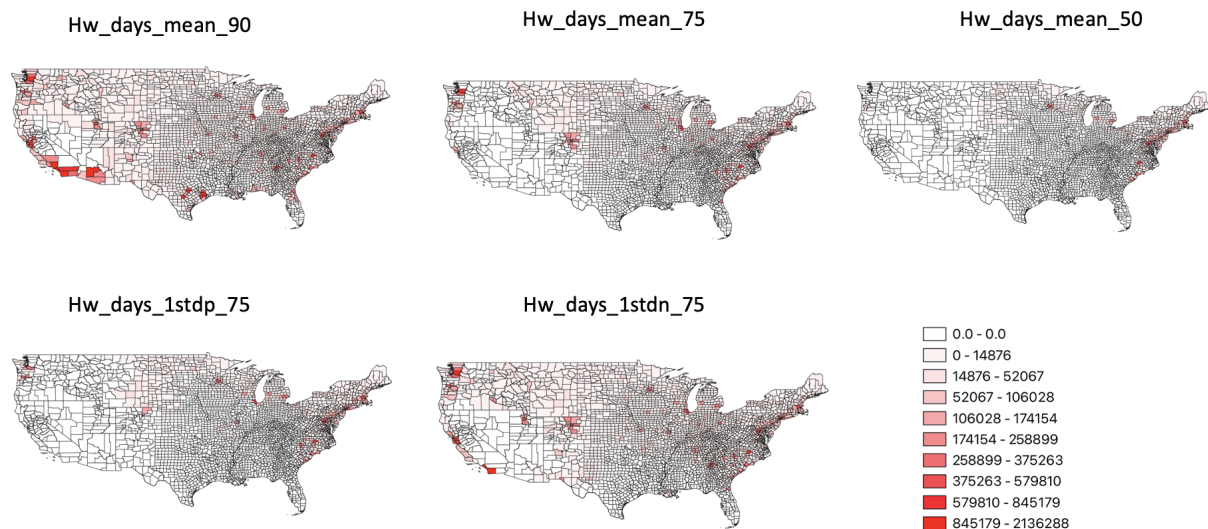
Figure 7. Counties and their populations affected by different thresholds of persistent heatwave days bias: (a) exceeding 10% of the persistent heatwave days bias within the continental US (25-50N, 125-70W over land), (b) exceeding 25%, (c) exceeding 50%, (d) exceeding 25% with 1 standard deviation added, and (e) exceeding 25% with 1 standard deviation removed.

expand further, eventually covering most of the US, including parts of the northeastern US. It's important to note that model uncertainties and natural variability overlap in regions like Florida and the western US, which is not entirely surprising given that model bias can be linked to natural variability, especially decadal variabilities. However, this overlap raises additional

205   concerns for local policy-making and underscores the importance of climate adaptation strategies. The county number affected by the large uncertainties ranges from 291 to 2863 in Figure 8.

     The surface temperature trend 'bias', or the strong effect of natural variability around the Rockies, is much more regionally confined. In Figure 9, the regions strongly affected start in the western US, excluding the southern part, and gradually expand southward and eastward, reaching areas with high population densities like southern California and the northeastern US. The

210   population's sensitivity to the standard deviation is notable—not because of the spatial range covered by the standard deviation, but due to the high population density in these affected counties. The county number ranges from 203 to 979 in Figure 9.

     Unsurprisingly, the uncertainties due to climate models and natural variability overlap more in Figure 10, as model uncertainties already reflect natural variability. Compared to Figure 9, it's important to note that even if some counties in southern California and Texas aren't strongly affected by the bias, natural variability can still play a significant role. These uncertainties

215   extend to most counties when using the 50% threshold. The number of counties affected ranges from 761 to 2,633 in Figure 9, with the variation largely due to the differing county densities between the western and eastern US.
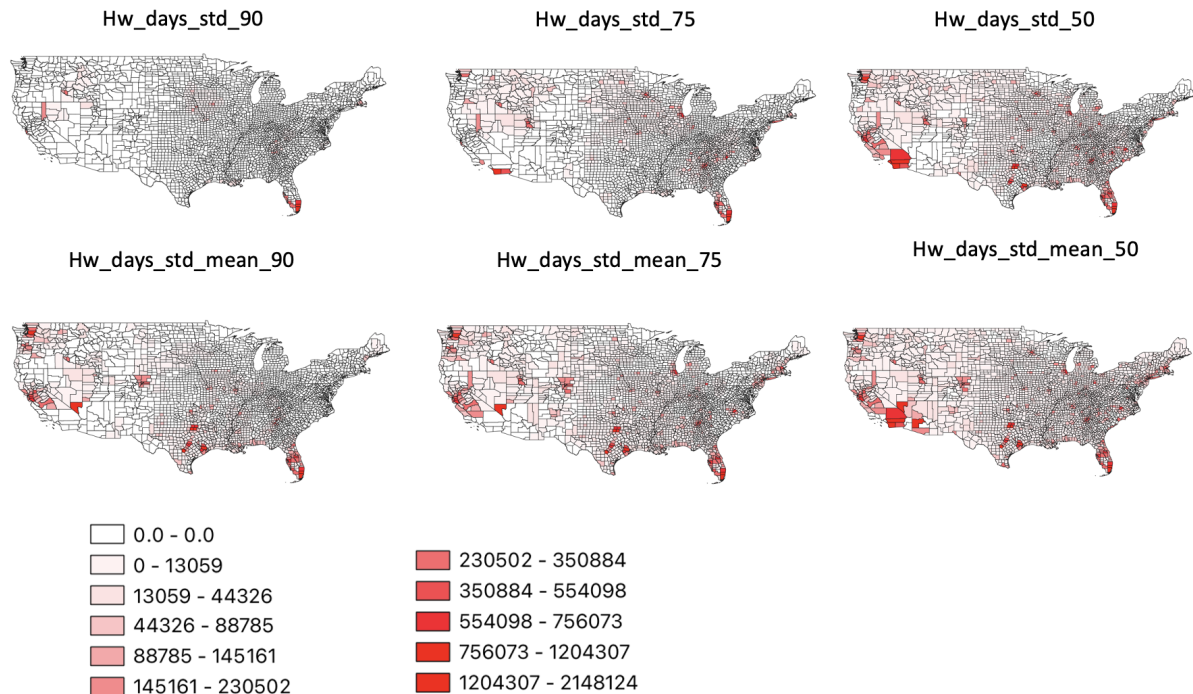
Figure 8. Counties and their populations affected by different thresholds of persistent heatwave days standard deviation: (a) exceeding 90% of the persistent heatwave days bias standard deviation among different climate model means within the continental US (25-50N, 125-70W over land), (b) exceeding 75%, (c) exceeding 50%, (d) exceeding 90% of the mean persistent heatwave days bias standard deviation within each climate model in the continental US, (e) exceeding 75%, and (f) exceeding 50%.

3.3 Population feature comparison

We then want to examine whether the populations in more vulnerable regions—whether to persistent heatwave days or surface temperature trends—differ in characteristics from those in less vulnerable regions. Given the relatively large populations within the 90% and 10% thresholds, we did not include a comparison with the total population, as it may overlap with the 90% threshold population. For the next analysis, we used the 75% threshold (except for the 25% threshold for the mean of persistent heatwave days) and continued to use SSP5 as the example. In the following boxplots, the light peach box indicates the selected regions and the orange box indicates the lefted regions. The box represents the 25th to 75th percentile distribution, with the red line indicating the median value. The minimum and maximum values correspond to 25% minus 1.5 times the Interquartile Range (IQR, 25th to 75th percentile) or 75% plus 1.5 times the IQR.

Figure 11 compares population characteristics between regions where persistent heatwave days exceed the 25th percentile threshold and the remaining regions (shown in the first and third pairs of columns in each subplot), and between regions where surface temperature trends exceed the 75th percentile threshold and the remaining regions (shown in the second and fourth pairs of columns in each subplot). According to the Welch T-test, only in the comparison of the elderly population for surface
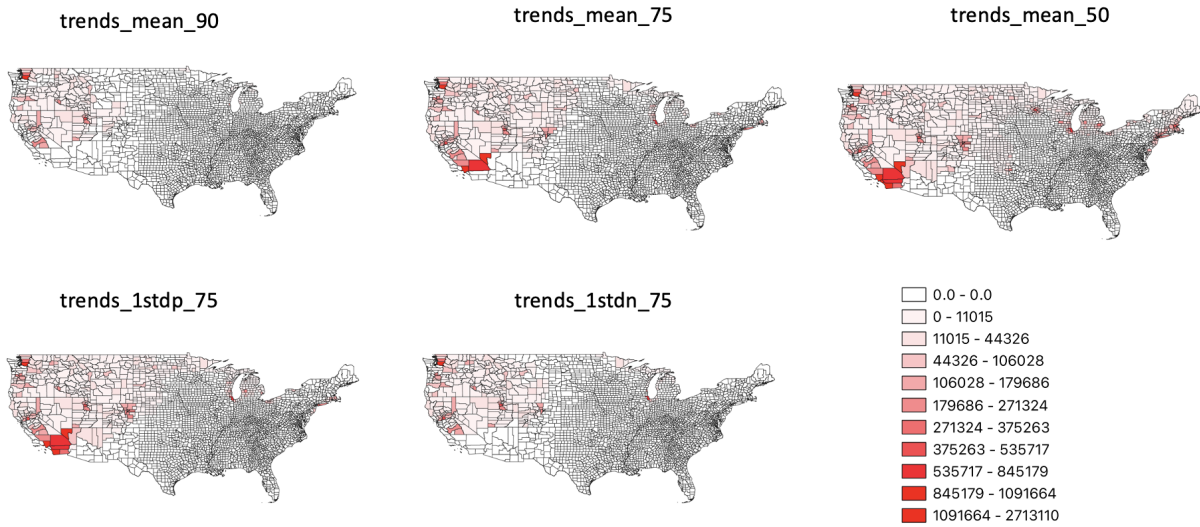
11

Figure 9. Same as Figure 7, except for surface temperature trend mean difference rather than persistent heatwave days mean difference.
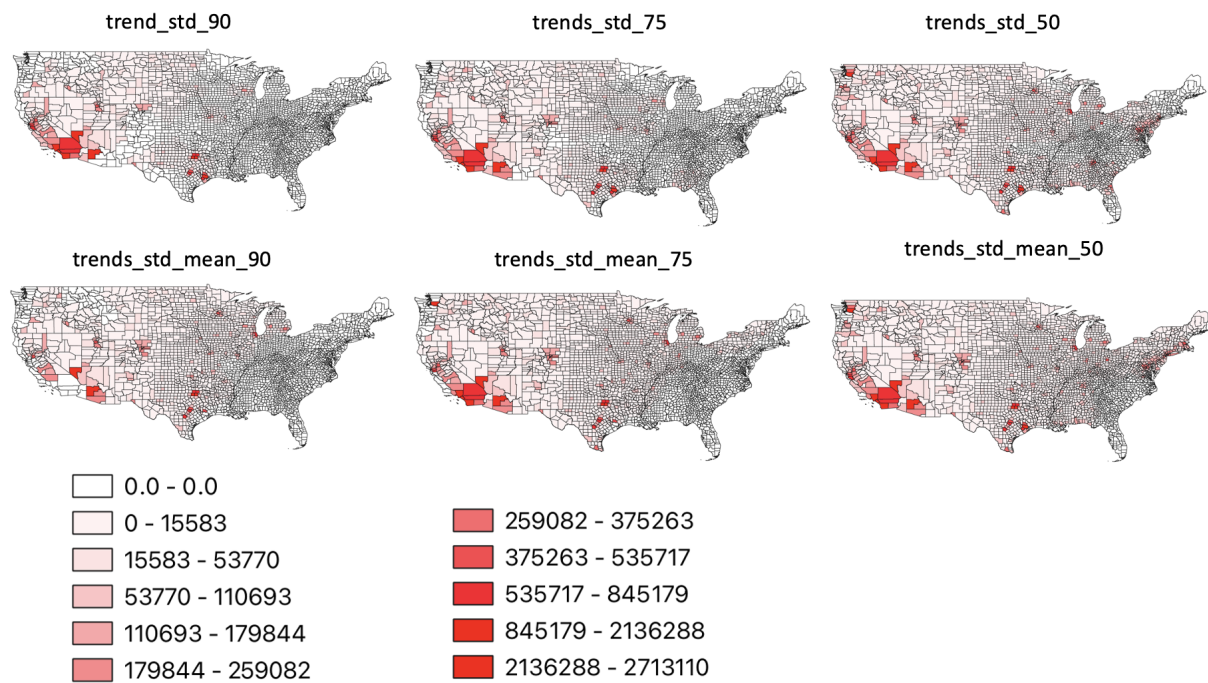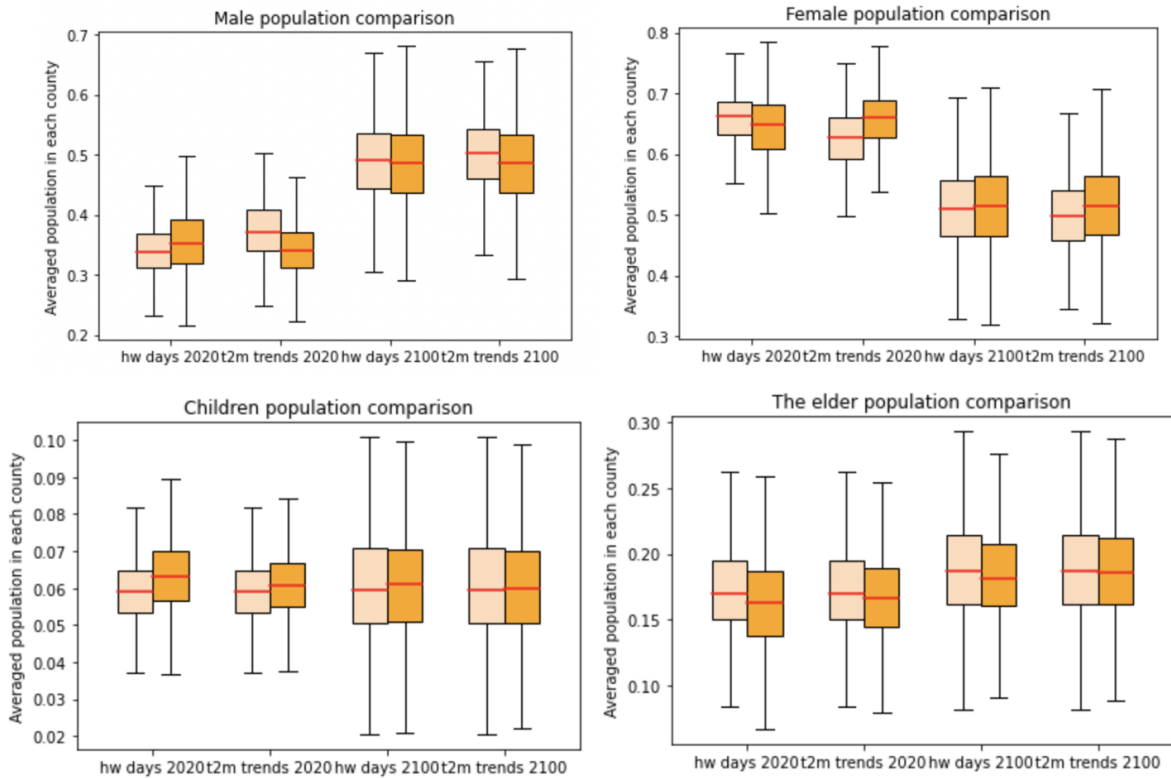

Figure 10. Same as Figure 8, except for surface temperature trend mean standard deviation rather than persistent heatwave days standard deviation.

230 temperature trends in 2100 do the counties above and below the threshold show no significant difference (p-value over 0.7).

Figure 11. Comparison of population characteristics between regions where persistent heatwave days exceed the 25th percentile threshold and the remaining regions, as well as between regions where surface temperature trends exceed the 75th percentile threshold and the remaining regions. Panel (a) compares the male population to the total population ratio for persistent heatwave days in 2020, surface temperature trends in 2020, persistent heatwave days in 2100, and surface temperature trends in 2100. Panel (b) presents the same comparison for the female population, while panel (c) focuses on the child population, and panel (d) on the elderly population.

For the male, female, and child populations in regions affected by persistent heatwave days in 2100, slight differences were observed (p-values between 0.05 and 0.1).

Notably, the ratios of the subset populations are all different in 2020. Regions with heatwave days bias (northeast US) have relatively fewer males and more females, whereas in the western US, where surface temperature trends are prominent, the opposite pattern is observed. This could be related to the industrial structures in different US regions. Interestingly, this sex ratio distribution is expected to reverse by 2100 if immigration is not considered. It is questionable why the male ratio is so far from 0.5 in 2020, prompting further analysis to verify the accuracy of these results.

For both the child and elderly populations, the child population is smaller, and the elderly population is larger in the selected regions compared to the remaining regions in 2020. However, this trend is less pronounced in 2100.

Figure 12, similar to Figure 11, illustrates the comparisons across the four racial groups. A distinct feature is the substantial reduction in the White population across the US, accompanied by an increase in the Black, Hispanic, and other populations, with the Hispanic population experiencing the most significant growth. This trend is widespread and not necessarily linked to heatwaves or surface temperature trends, as the box plots for both groups (above and below the thresholds) show similar patterns.

However, regions above the model bias thresholds can differ notably from those below, especially for the White and Hispanic populations. A consistent feature observed in both 2020 and 2100 is a higher White population and a lower Hispanic population in more vulnerable regions, a difference confirmed as significant by Welch's t-test, though this disparity diminishes considerably by 2100. Additionally, there is a significantly lower Black population in the vulnerable regions, except for those impacted by heatwave day bias in 2100. The "other" population category is less affected by trends in bias, which mainly reflects regional differences between the western US and other areas. However, fewer people of other races are found in regions affected by heatwave day bias (northeast US) in 2020, a trend that reverses by 2100, with more people of other races residing in these vulnerable regions.

In comparing regions where the standard deviations of persistent heatwave days and surface temperature trends among different climate model means exceed the 75th percentile threshold (Figure 13), issues with the sex ratio persist. Notably, the sex ratios for heatwave days and summer temperature trends in 2100 still pass Welch's t-test. The sex ratio in 2020 (with fewer males and more females in regions over the threshold, such as South Florida and certain mountainous areas) reverses in 2100. For child and elderly populations, the difference between selected and non-selected regions in 2100 for both heatwave days bias and surface temperature trend bias is significant: there are more children and fewer elderly in the selected regions, which aligns with the observed population growth in the southern and western U.S. in Figure 6. However, the differences in 2020 are less pronounced, with some not being statistically significant, emphasizing the importance of monitoring the age and sex population in the future.

In Figure 14, the racial population characteristics across the country are examined. For White and Black populations, significant differences between regions with large and small model uncertainties are observed in 2020, but these differences diminish by 2100. Specifically, in 2020, regions with high model uncertainties have a higher proportion of White and a lower proportion of Black populations, but this disparity is reduced in 2100. In contrast, the Hispanic population, which was relatively low in 2020, increases significantly by 2100. Despite this overall increase, regions with high model uncertainties still show a relatively lower Hispanic population. For other racial groups, no significant differences are noted.

Figure 15 compares regions where the average of each climate model's standard deviations for persistent heatwave days and surface temperature trends exceeds the 75th percentile threshold. Aside from a potential bias in sex ratios in 2020, there are no significant differences in sex ratios between regions with high and low uncertainties for persistent heatwave days, indicating similar impacts from large or small natural variability. However, significant differences are observed in sex ratios for regions with high or low uncertainties in surface temperature trends, particularly in the West US. In the West US, where high uncertainties are located, there is a higher male-to-female ratio compared to other regions.
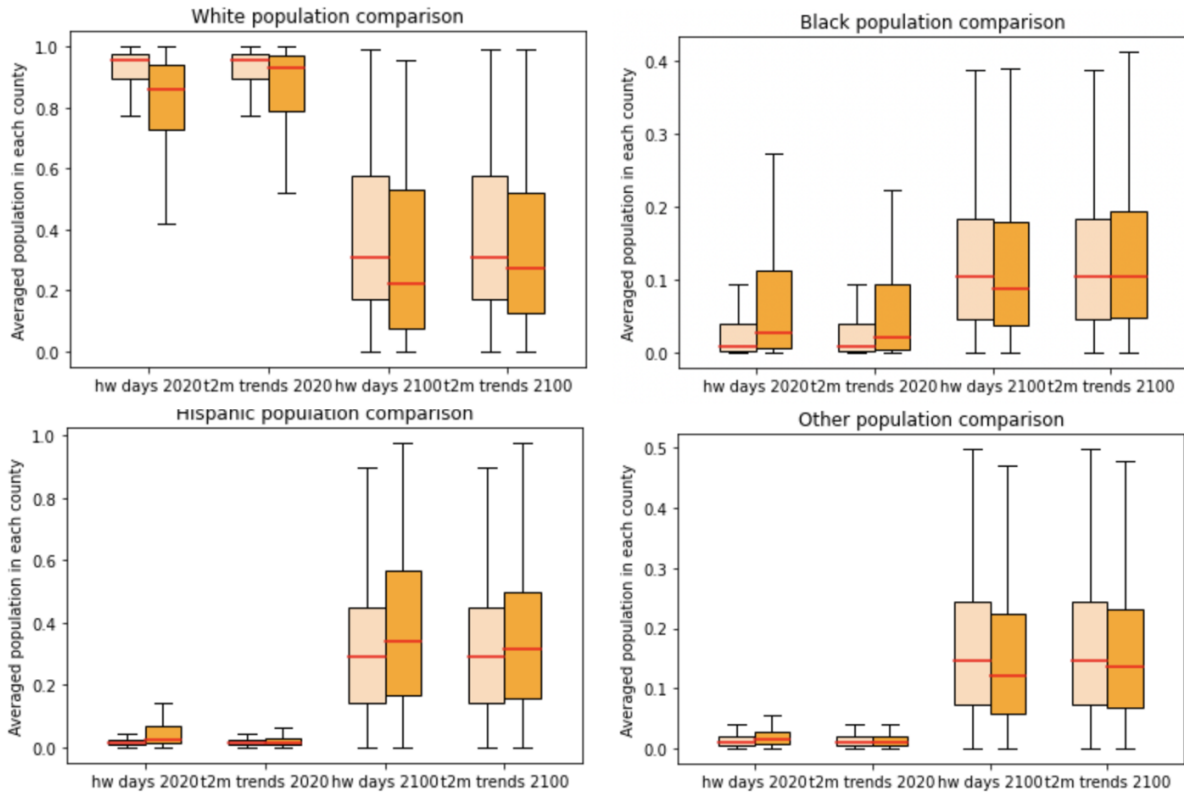
14

Figure 12. The same as Figure 11, except for (a) the White population, (b) the Black population, (c) the Hispanic population and (d) the other races population.

For child populations, regions in the West US with high uncertainties consistently show a larger ratio of children and a smaller ratio of elderly individuals compared to other regions. However, the difference in elderly population ratios for surface temperature trends in 2020 is not strongly significant from other regions (p-value around 0.07).

Finally, in Figure 16, we analyze racial population features in regions where the average of each climate model's standard deviations for persistent heatwave days and surface temperature trends exceeds the 75th percentile threshold. White and Hispanic populations consistently show significant differences between regions with high and low uncertainties. In regions with high uncertainties, particularly in the West US, the Hispanic population is consistently higher in both 2020 and 2100, whereas the White population is higher in 2020 but lower in 2100. For the Black population, there is no significant difference in regions affected by natural variability in heatwave days, but a significant difference is observed in regions affected by natural variability in surface temperature trends, such as the West US. Similarly, other races show a higher population presence in the West US.
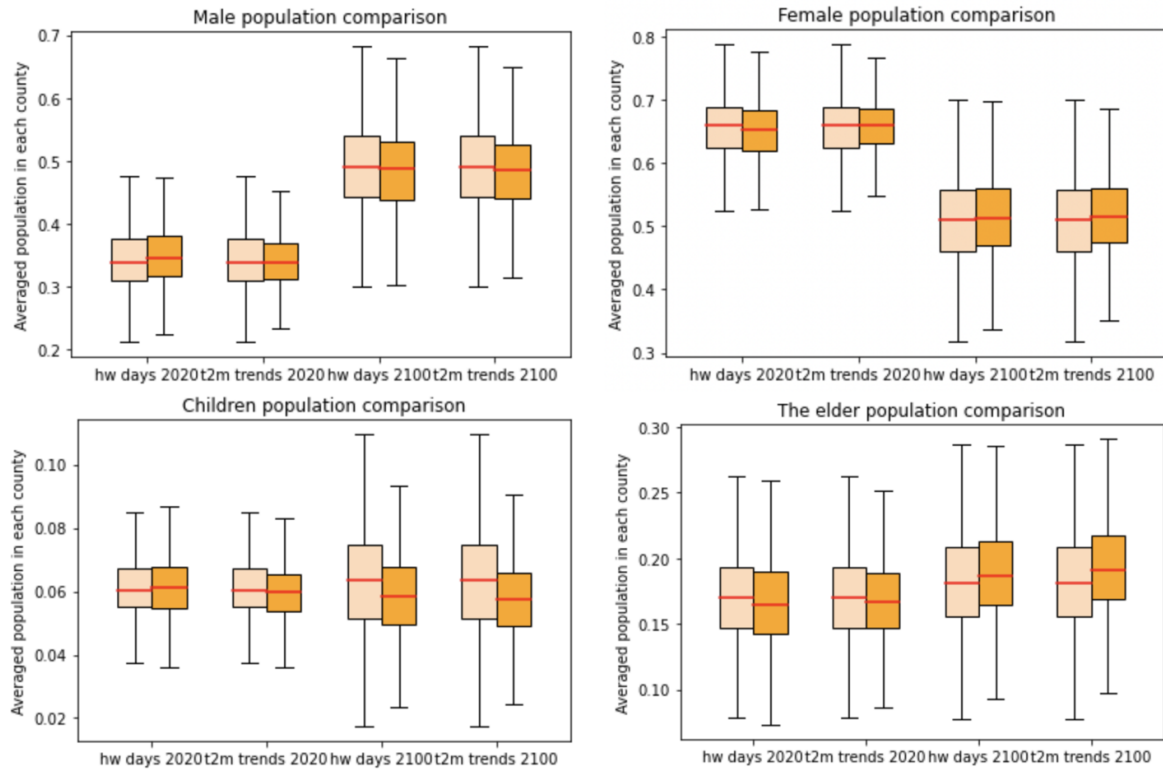
15

Figure 13. The same as Figure 11, but focusing on regions where the standard deviations of persistent heatwave days and surface temperature trends among different climate model means exceed the 75th percentile threshold.

4 Summary and Discussion

In this research, we analyzed two types of surface temperature metric biases in climate models and examined how vulnerable population groups might be impacted by these biases. We found that the selected CMIP6 climate models exhibit similar behaviors: a strong bias in persistent heatwave days in the northeastern U.S. and a bias in surface temperature trends in the western U.S. Both biases are influenced more by natural variability than by uncertainties among different climate models. The bias in persistent heatwaves is consistent and pronounced, while the bias in surface temperature trends is more strongly controlled by natural variabilities.

Overall, the U.S. population shares some common trends: 1) continued urbanization, 2) an increasing ratio of non-White populations, particularly Hispanic, and 3) a growing proportion of the elderly population. The uncertainties from climate model analyses contribute to significant uncertainties in the affected populations as well, due to large variations in the spatial range corresponding to 1 or 2 standard deviations from the thresholds, as well as the impact of counties with extremely high populations and ongoing urbanization.
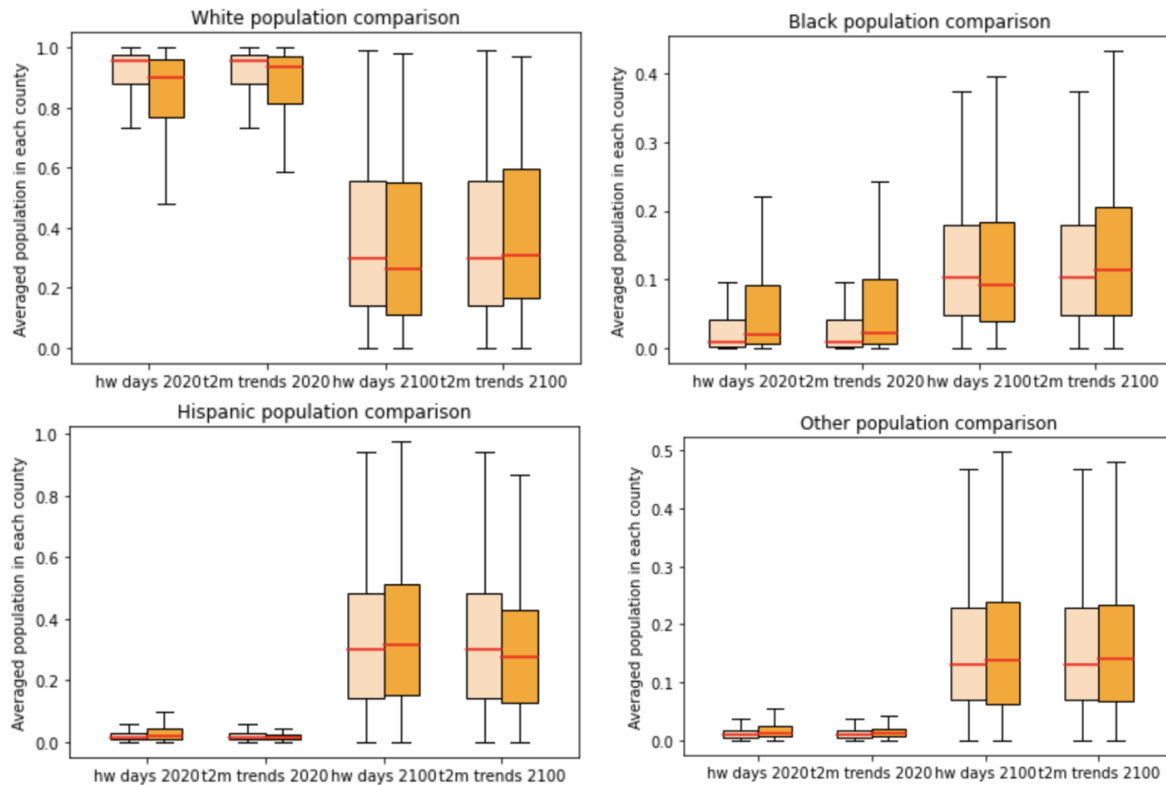
16

Figure 14. The same as Figure 12, but focusing on regions where the standard deviations of persistent heatwave days and surface temperature trends among different climate model means exceed the 75th percentile threshold.

In addition to these general trends, some key observations include: 1) there are more female populations on the East Coast and more male populations on the West Coast, possibly linked to industrial structures. Consequently, regions with surface temperature trends bias tend to have a higher male population, while those suffering from persistent heatwave days tend to have more females. 2) In most cases, regions with a higher elderly population have fewer children, and vice versa, suggesting that policymakers should prepare for more vulnerable populations across different scenarios. 3) For racial demographics, the differences between selected and non-selected regions are relatively small across all racial groups. A significant difference exists between 2020 and 2100, with a notable trend of more White populations in regions affected by both types of model biases. However, by 2100, regions experiencing large natural variability and both types of model biases will see an increase in the Hispanic population and a decrease in the White population.

We have also examined the phase preference of quasi-stationary Rossby waves in the upper troposphere (Fei and White, 2023). These long-lasting atmospheric circulation patterns play a significant role in surface heatwave occurrences. The phase preference feature of quasi-stationary Rossby waves, identified through odds ratios, is associated with heatwave location pref-
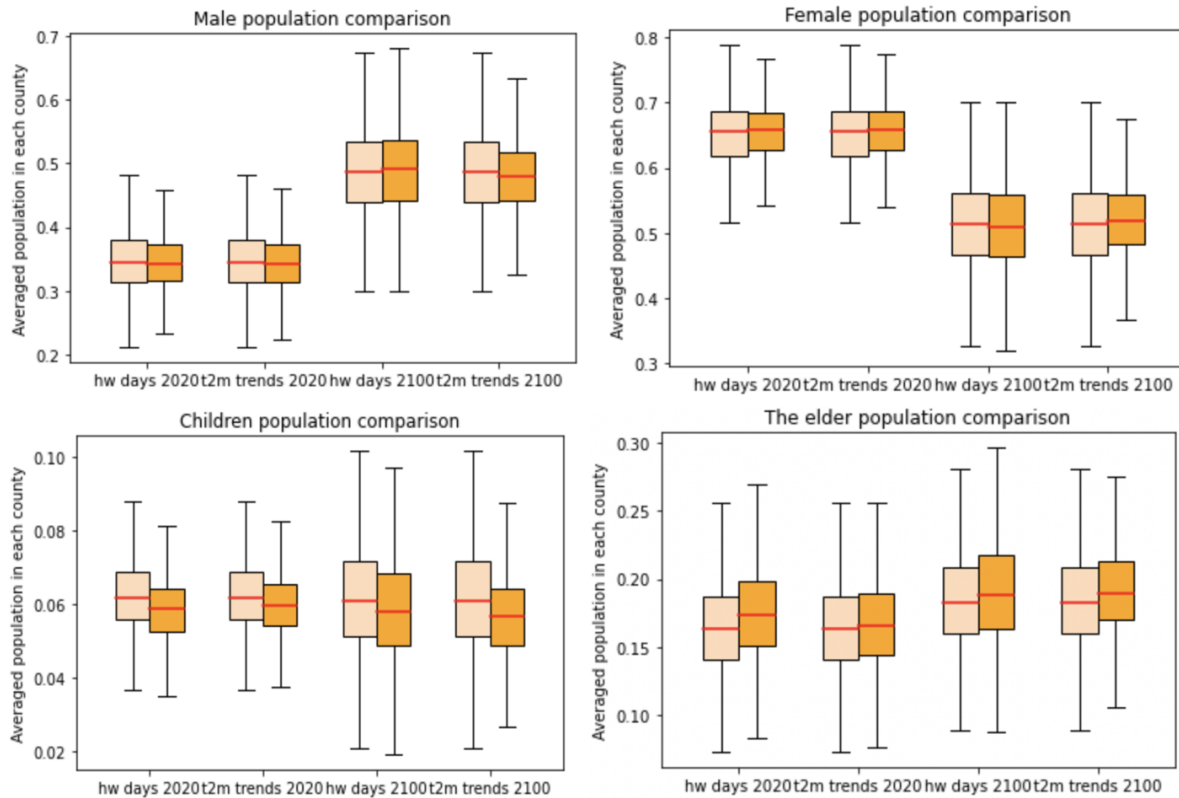
17

Figure 15. The same as Figure 11, but focusing on regions where the average of each climate model's standard deviations for persistent heatwave days and surface temperature trends exceeds the 75th percentile threshold.

erences in the western US. However, this region differs from the northeastern US, where the most significant heatwave days bias from climate models is observed, indicating a less direct relevance to the current research target. Furthermore, the phase

310   preference feature is based on absolute values rather than relative thresholds, suggesting that future analyses should focus on the impact of absolute high temperatures on human populations.

4.1 shortcoming and next steps

There are several shortcomings and ongoing next steps for this research. One clear limitation is the need to investigate the sex ratio at the start of the population projection. Additionally, selecting grid points based on the continental US range is crucial to

315   avoid having zero counties for the 90th percentile threshold of heatwave days mean. Furthermore, the fact that the U.S. is an immigration country has not been accounted for in the population projection dataset, which introduces significant limitations to the projection (Black et al., 2011).
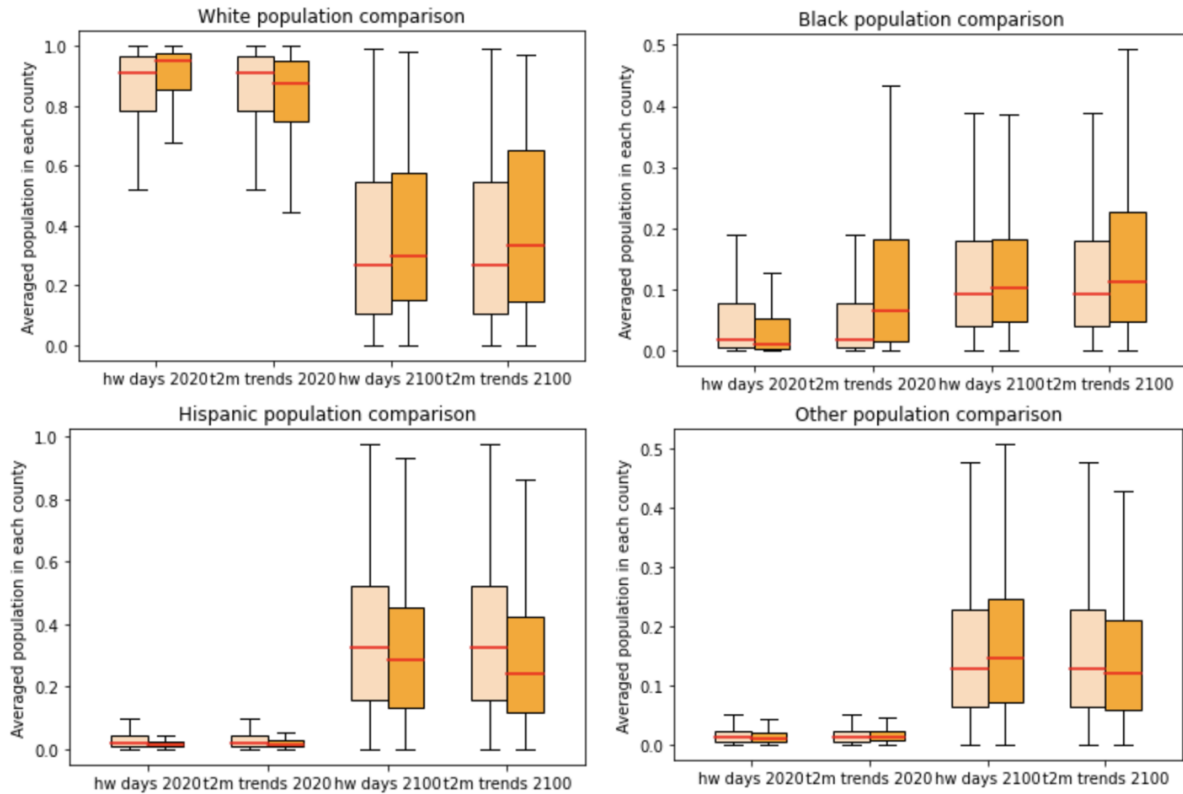
Figure 16. The same as Figure 12, but focusing on regions where the average of each climate model's standard deviations for persistent heatwave days and surface temperature trends exceeds the 75th percentile threshold.

Despite these shortcomings, there are exciting research directions ahead. For instance, we have begun projecting the Canadian population to 2100 at the subdivision level, considering Canada's vast territory. This includes examining populations with different mother tongues, which may indicate indigenous groups speaking less common languages. However, due to the lack of household population data at the subdivision level from the Canadian Census, identifying group quarters with normal family structures is challenging. Consequently, population projections may become less accurate after a few steps.

Another promising avenue is to consider absolute temperature values, particularly in SSP5 scenarios and temperatures exceeding 35°C. Incorporating relative humidity is also important for assessing human comfort (Heo et al., 2019; Raymond et al., 2020). Additionally, exploring demographic differences between 2020 and 2100 could reveal whether changes are due to climate model biases or population projections. Lastly, we plan to complete the analysis for the remaining six climate models using the same methods.

330

References

Bi, D., Dix, M., Marsland, S., O'farrell, S., Sullivan, A., Bodman, R., Law, R., Harman, I., Srbinovsky, J., Rashid, H. A., et al.: Configuration and spin-up of ACCESS-CM2, the new generation Australian community climate and earth system simulator coupled model, Journal of Southern Hemisphere Earth Systems Science, 70, 225–251, 2020.

Black, R., Bennett, S. R., Thomas, S. M., and Beddington, J. R.: Migration as adaptation, Nature, 478, 447–449, 2011.

Brás, T. A., Seixas, J., Carvalhais, N., and Jägermeyr, J.: Severity of drought and heatwave crop losses tripled over the last five decades in Europe, Environmental Research Letters, 16, 065 012, 2021.

Brunner, L., Schaller, N., Anstey, J., Sillmann, J., and Steiner, A. K.: Dependence of present and future European temperature extremes on the location of atmospheric blocking, Geophysical research letters, 45, 6311–6320, 2018.

Chambers, J.: Global and cross-country analysis of exposure of vulnerable populations to heatwaves from 1980 to 2018, Climatic Change, 163, 539–558, 2020.

Craig, A., Valcke, S., and Coquart, L.: Development and performance of a new version of the OASIS coupler, OASIS3-MCT_3. 0, Geoscientific Model Development, 10, 3297–3308, 2017.

Davy, R. and Esau, I.: Global climate models' bias in surface temperature trends and variability, Environmental Research Letters, 9, 114 024, 2014.

Domeisen, D. I., Eltahir, E. A., Fischer, E. M., Knutti, R., Perkins-Kirkpatrick, S. E., Schär, C., Seneviratne, S. I., Weisheimer, A., and Wernli, H.: Prediction and projection of heatwaves, Nature Reviews Earth & Environment, 4, 36–50, 2023.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958, 2016.

Fei, C. and White, R. H.: Large-amplitude quasi-stationary Rossby wave events in ERA5 and the CESM2: features, precursors, and model biases in Northern Hemisphere winter, Journal of the Atmospheric Sciences, 80, 2075–2090, 2023.

Griffies, S. M. et al.: Elements of the modular ocean model (MOM), GFDL Ocean Group Tech. Rep, 7, 47, 2012.

Hauer, M. E.: Population projections for US counties by age, sex, and race controlled to shared socioeconomic pathway, Scientific data, 6, 1–15, 2019.

Heo, S., Bell, M. L., and Lee, J.-T.: Comparison of health risks by heat wave definition: Applicability of wet-bulb globe temperature for heat wave criteria, Environmental research, 168, 158–170, 2019.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, 2020.

Hunke, E. C., Lipscomb, W. H., Turner, A. K., Jeffery, N., and Elliott, S.: Cice: the los alamos sea ice model documentation and software user's manual version 4.1 la-cc-06-012, T-3 Fluid Dynamics Group, Los Alamos National Laboratory, 675, 500, 2010.

Kovats, R. S. and Kristie, L. E.: Heatwaves and public health in Europe, The European Journal of Public Health, 16, 592–599, 2006.

Kravchenko, J., Abernethy, A. P., Fawzy, M., and Lyerly, H. K.: Minimization of heatwave morbidity and mortality, American journal of preventive medicine, 44, 274–282, 2013.

Krinner, G. and Flanner, M. G.: Striking stationarity of large-scale climate model bias patterns under strong climate change, Proceedings of the National Academy of Sciences, 115, 9462–9466, 2018.

Larcom, S., She, P.-W., and van Gevelt, T.: The UK summer heatwave of 2018 and public concern over energy security, Nature Climate Change, 9, 370–373, 2019.

Lee, J., Kim, J., Sun, M.-A., Kim, B.-H., Moon, H., Sung, H. M., Kim, J., and Byun, Y.-H.: Evaluation of the Korea meteorological administration advanced community earth-system model (K-ACE), Asia-Pacific Journal of Atmospheric Sciences, 56, 381–395, 2020.

Li, L., Yu, Y., Tang, Y., Lin, P., Xie, J., Song, M., Dong, L., Zhou, T., Liu, L., Wang, L., et al.: The flexible global ocean-atmosphere-land system model grid-point version 3 (FGOALS-g3): Description and evaluation, Journal of Advances in Modeling Earth Systems, 12, e2019MS002 012, 2020.

Madec, G., Bourdallé-Badie, R., Bouttier, P.-A., Bricaud, C., Bruciaferri, D., Calvert, D., Chanut, J., Clementi, E., Coward, A., Delrosso, D., et al.: NEMO ocean engine, 2017.

Masato, G., Hoskins, B. J., and Woollings, T.: Winter and summer Northern Hemisphere blocking in CMIP5 models, Journal of Climate, 26, 7044–7059, 2013.

Mélia, D. S.: A global coupled sea ice–ocean model, Ocean Modelling, 4, 137–172, 2002.

Mueller, B. and Seneviratne, S. I.: Hot days induced by precipitation deficits at the global scale, Proceedings of the national academy of sciences, 109, 12 398–12 403, 2012.

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., et al.: The scenario model intercomparison project (ScenarioMIP) for CMIP6, Geoscientific Model Development, 9, 3461–3482, 2016.

Perkins, S. E., Alexander, L., and Nairn, J.: Increasing frequency, intensity and duration of observed global heatwaves and warm spells, Geophysical Research Letters, 39, 2012.

Raymond, C., Matthews, T., and Horton, R. M.: The emergence of heat and humidity too severe for human tolerance, Science Advances, 6, eaaw1838, 2020.

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., Decharme, B., Delire, C., Berthet, S., Chevallier, M., et al.: Evaluation of CNRM Earth system model, CNRM-ESM2-1: Role of Earth system processes in present-day and future climate, Journal of Advances in Modeling Earth Systems, 11, 4182–4227, 2019.

Shaposhnikov, D., Revich, B., Bellander, T., Bedada, G. B., Bottai, M., Kharkova, T., Kvasha, E., Lezina, E., Lind, T., Semutnikova, E., et al.: Mortality related to air pollution with the Moscow heat wave and wildfire of 2010, Epidemiology, 25, 359–364, 2014.

Shepherd, T. G., Boyd, E., Calel, R. A., Chapman, S. C., Dessai, S., Dima-West, I. M., Fowler, H. J., James, R., Maraun, D., Martius, O., et al.: Storylines: an alternative approach to representing uncertainty in physical aspects of climate change, Climatic change, 151, 555–571, 2018.

Voldoire, A., Decharme, B., Pianezze, J., Lebeaupin Brossier, C., Sevault, F., Seyfried, L., Garnier, V., Bielli, S., Valcke, S., Alias, A., et al.: SURFEX v8. 0 interface with OASIS3-MCT to couple atmosphere with hydrology, ocean, waves and sea-ice models, from coastal to global scales, Geoscientific Model Development, 10, 4207–4227, 2017.

Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., et al.: The Met Office Unified Model global atmosphere 7.0/7.1 and JULES global land 7.0 configurations, Geoscientific Model Development, 12, 1909–1963, 2019.

Wang, T., Gong, H., Liu, Y., and Lu, J.: Polar and topographic amplifications of intermodel spread of surface temperature in climate models, Journal of Geophysical Research: Atmospheres, 128, e2022JD037 509, 2023.

Welch, B. L.: The generalization of 'STUDENT'S' problem when several different population varlances are involved, Biometrika, 34, 28–35, 1947.

Wills, R. C., Herrington, A. R., Simpson, I. R., and Battisti, D. S.: Resolving weather fronts increases the large-scale circulation response to Gulf Stream SST anomalies in variable-resolution CESM2 simulations, Journal of Advances in Modeling Earth Systems, 16, e2023MS004 123, 2024.

Yin, J., Gentine, P., Slater, L., Gu, L., Pokhrel, Y., Hanasaki, N., Guo, S., Xiong, L., and Schlenker, W.: Future socio-ecosystem productivity threatened by compound drought–heatwave events, Nature Sustainability, 6, 259–272, 2023.

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., et al.: The Meteorological Research Institute Earth System Model version 2.0, MRI-ESM2. 0: Description and basic evaluation of the physical component, Journal of the Meteorological Society of Japan. Ser. II, 97, 931–965, 2019.