Changes in apparent temperature around the Beijing-Tianjin megalopolis under greenhouse gas and stratospheric aerosol injection scenarios

Jun Wang1, John C. Moore1,2*, Liyun Zhao1*

1 College of Global Change and Earth Systems Science, Beijing Normal University, Beijing 100875, China
2 Arctic Center, University of Lapland, Rovaniemi, Finland
Correspondence to: zhaoliyun@bnu.edu.cn, john.moore.bnu@gmail.com

Abstract. We compare apparent temperatures – that is a combination of 2 m air temperature, relative humidity and surface wind speed in four Earth System Models under the modest greenhouse emissions RCP4.5, the “business-as-usual” RCP8.5 and the stratospheric aerosol injection G4 geoengineering scenarios. Apparent temperatures come from both a 10 km resolution dynamically downscaled model (WRF), and a statistically bias corrected (ISIMIP) and downscaled simulation for the greater Beijing region. ISIMIP downscaling method tends to simulate apparent temperatures well at present in all seasons, and WRF produces warmer winters than does ISIMIP. WRF produces warmer winters and cooler summers than does ISIMIP in the future. These differences mean that estimates of numbers of days with extreme apparent temperatures vary systematically with downscaling method, as well as between climate models and scenarios. Air temperature changes dominate differences in apparent temperatures between future scenarios even more than they do at present because the reductions in humidity expected under solar geoengineering are overwhelmed by rising vapor pressure due to rising temperatures and the lower windspeeds expected in the region in all future scenarios. Urban centres see larger rises in extreme apparent temperatures than rural surroundings due to differences in land surface type, and since these are also the most densely populated, health impacts will be dominated by the larger rises in apparent temperatures in these urban areas.

500 character non-technical text
Apparent temperatures that include humidity and wind speed in addition to surface temperature measure human heat stress and comfort. We show that apparent temperatures will reach dangerous levels more commonly in future and rise faster than air temperatures because of water vapor pressure rises and lower expected wind speeds. Solar geoengineering can reduce the frequency of extreme events significantly relative to modest, and especially “business as usual” greenhouse scenarios.
1. Introduction

Global mean surface temperature has increased by 0.92℃ (0.68-1.17℃) during 1880-2012 (IPCC, 2021), which naturally also impacts the human living environment (Kraaijenbrink et al., 2017; Garcia et al., 2018). However, neither land surface temperature nor near-surface air temperature can adequately represent the temperature we experience. Apparent temperature (AP), that is how the temperature feels, is formulated to reflect human thermal comfort and is probably a more important indication of health than daily maximum or minimum temperatures. There are various approaches to estimating how the weather conditions affect comfort, but apparent temperature is governed by air temperature, humidity and wind speed (Steadman 1984; Steadman 1994). These are known empirically to affect human thermal comfort (Jacobs et al., 2013), and thresholds have been designed to indicate danger and health risks under extreme heat events (Ho et al., 2016). Analysis of historical apparent temperatures in China (Wu et al., 2017; Chi et al., 2018; Wang et al., 2019), Australia (Jacobs et al., 2013), and the USA (Grundstein et al., 2011) all find that apparent temperature is increasing faster than air temperature. This is due to both decreasing wind speeds and, especially to increasing vapor pressure (Song et al., 2022).

As the world warms, apparent temperature is expected to rise faster than air temperatures in the future (Li et al., 2018; Song et al., 2022). Hence, humans, and other species, will face more heat-related stress but less cold-related environmental stress in the warmer future (Wang et al., 2018; Zhu et al., 2019). Since most of the population is now urban, the conditions in cities will determine how tolerable are future climates for much of humanity, while the differences in thermal comfort between urbanized and rural regions will be a factor in driving urbanization. Reliable estimates of future urban temperatures and their rural surroundings require methods to improve on standard climate model resolution to adequately represent the different land surface types; especially the rapid and accelerating changes in land cover in the huge urban areas characteristic of sprawling developments in the developing world. This is usually done with either statistical or dynamic downscaling approaches, and in this article we examine both methods.

The focus here is in the differences in apparent temperature that may arise from solar geoengineering (that is reduction in incoming short-wave radiation to offset longwave absorption by greenhouse gases) via stratospheric aerosol injection (SAI), and pure greenhouse gas climates. We use all four climate models that have provided sufficient data from the G4 scenario described by the Geoengineering Model Intercomparison Project (GeoMIP). G4 specifies sulfates as the aerosol, and greenhouse gas emissions from the RCP4.5 scenario (Kravitz et al., 2011). The impacts of G4 on surface temperature and precipitation have been discussed at regional scales (Yu et al., 2015) and both are lowered relative to RCP4.5. Some studies have focused on regional impact of SAI on apparent or wet bulb temperatures: in Europe, (Jones et al., 2018);
East Asia (Kim et al., 2020); and the Maritime Continent (Kuswanto et al., 2021). But none of these studies have considered apparent temperature at scales appropriate for rapidly urbanizing regions such as on the North China Plain.

The greater Beijing megalopolis lies in complex terrain, surrounded by hills and mountains on three sides, and a flat plain to the southeast coast (Fig. 1). Over the period 1978-2008, Beijing experienced an increasing trend of 12.7% or 2.07 days per decade in extreme warm nights (Wang et al., 2013), and urbanization produced an average increase in temperature of approximately 0.60°C. By the end of 2019, the permanent resident population in Beijing exceeded 21 million. Tianjin, 100 km from Beijing, is the fourth largest city in China with a population of about 15 million, and Langfang (population 4 million) is about 50 km from Beijing. Thus, the region contains a comparable urbanized population as the northeast US megalopolis. Since its climate is characterized by hot and moist summer monsoon conditions, the population is at an enhanced risk as urban heat island effects lead to city temperatures warming faster than their rural counterparts.

Accurate meteorological data are crucial in simulating future apparent temperatures because all ESM suffer from bias, and this problem is especially egregious at small scales. A companion paper (Wang et al., 2022 in review) looked at differences between downscaling methods with the same 4 Earth System Models (ESM), domain and scenarios as we use here. In this paper, we use the downscaled data to explore the effect of SAI on apparent temperature over the greater Beijing megalopolis. The paper is organized as follows. The data, method of calculating AP and AP thresholds are briefly described in Section 2. The results from present simulation and future projections on apparent temperature are given in Section 3, along with the impact analysis. Finally, Section 4 discusses and concludes the study.
Figure 1. a, The 10 km WRF domain (red box) nested inside a 30 km resolution WRF domain (large black sector). b, The inner domain topography and major conurbations (red dots), with the urban areas of Beijing and Tianjin enclosed in red curves. Panels c and d show the population density (persons per km²) of Beijing and Tianjin provinces (defined by black borders) in 2010 and the grid cells within the Beijing-Tianjin province (blue boxes) when downscaled by ISIMIP (c) and WRF (d).

2. Data and Methods

2.1 Scenarios, ESM, downscaling methods and bias correction

The scenarios, ESM, downscaling methods and bias correction methods we use here are as described in detail by Wang et al., (in review, 2022), and we just summarize the method briefly here. We use three different scenarios: RCP4.5 and RCP8.5 (Riahi et al., 2011) and the GeoMIP G4 scenario which span a useful range of climate scenarios: RCP4.5 is similar (Vandyck et al., 2016) to the expected trajectory of emissions under the 2015 Paris Climate Accord agreed Nationally Determined Contributions (NDCs); RCP8.5 represents a formerly business-as-usual, no climate mitigation policies, large signal to noise ratio scenario; G4 represents a similar radiative forcing as produced by the 1991 Mount Pinatubo volcanic eruption repeating every 4 years.

Climate forcing comes from 4 ESM: BNU-ESM (Ji et al., 2014), HadGEM2-ES (Collins et al., 2011), MIROC-ESM (Watanabe et al., 2011) and MIROC-ESM-CHEM (Watanabe et al., 2011). We compare dynamical and statistical downscaling methods to convert the ESM data to scales more suited to capturing differences between contrasting rural and urban environments. The observational data set we use to assess the performance of two downscaling methods is the daily ERA5 (Hersbach et al., 2018) reanalysis data with a resolution of 0.25°×0.25° over the
domain in Fig. 1b during 2008-2017. Dynamical downscaling for the 4 ESM datasets was done with WRFv.3.9.1 with a parameter set used for urban China studies (Wang et al., 2012) in two nested domains at 30 and 10 km resolution over 2 time slices (2008-2017 and 2060-2069). We corrected the biases in WRF output using the quantile delta mapping method (QDM; Wilcke et al., 2013) with ERA5 to preserve the mean probability density function of the output over the domain without degrading the WRF spatial pattern. All WRF results presented are after QDM bias correction. Statistical downscaling was done with the trend-preserving statistical bias-correction Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) method (Hempel et al., 2013) for the raw ESM output, producing output matching the mean ERA5 observational data in the reference historical period with the same spatial resolution, while allowing the individual ESM trends in each variable to be preserved.

### 2.2 Apparent temperature

We use the formula proposed in Steadman (1984) to estimate apparent temperature under shade, which has been widely used to study heat waves, heat stress and temperature-related mortality (Perkins and Alexander, 2013; Lyon and Barnston, 2017; Lee and Sheridan, 2018; Zhu et al., 2021):

\[
AP = -2.7 + 1.04 \times T + 2 \times P - 0.65 \times W
\]

(1)

where \(AP\) is the apparent temperature (°C) under shade meaning that radiation is not considered; \(T\) is the 2 m temperature (°C), \(W\) is the wind speed at 10 m above the ground (m/s), and \(P\) is the vapor pressure (kPa) calculated by

\[
P = P_s \times RH
\]

(2)

where \(P_s\) is the saturation vapor pressure (kPa), and \(RH\) is the relative humidity (%). \(P_s\) is calculated using the Clausius–Clapeyron relation:

\[
P_s = \begin{cases} 
0.61078 \times e^{\left(\frac{17.30388 \times T}{T + 259.5}\right)}, & T \geq 0 \\
0.61078 \times e^{\left(\frac{21.735104 \times (T-3)}{T + 259.5}\right)}, & T < 0 
\end{cases}
\]

(3)

To assess the potential risks of heat-related exposure from apparent temperature, we also count the number of days with \(AP > 32°C\) (NdAP_32) in the Beijing-Tianjin province. This threshold does not lead to extreme risk and death, instead it is classified as requiring “extreme caution” by the US National Weather Service. While hotter AP thresholds would give a more direct estimate of health risks, the statistics of these presently rare events mean that detecting differences between scenarios is less reliable than using the cooler NdAP_32 threshold. We presume that similar differences between scenarios would apply for higher thresholds.

### 2.3 Population Data Set

Since health impacts are more important where there are more people, we calculate the NdAP_32 weighted by population (Fig. 1c and 1d). We employ gridded population data (Fu et al., 2014; https://doi.org/10.3974/geodb.2014.01.06.V1) with a spatial resolution of 1×1 km collected in 2010. The population density distribution in Beijing and Tianjin provinces with the ISIMIP and WRF grid cells contained are shown in the Fig. 1c and 1d.
2.4 Determination of each factor’s contribution to change in AP

Equation (1) describes how AP is calculated, and this can be broken down into how much equivalent temperature is produced by each term (Fig. 2), with 2008-2017 as the baseline interval for season-by-season contributors to AP. Across scenario seasonal differences in contributors are then calculated as follows. We use multiple linear regression to reconstruct the relationship between change in AP ($\Delta AP$) and changes in each factor for each future scenario,

$$\Delta AP = \sum \alpha_i X_i + b$$

(4)

where $X_{i(i=1,2,3)}$ are the daily changes of the three meteorological factors between two scenarios: 2 m temperature ($\Delta T$), 2 m relative humidity ($\Delta RH$) and 10 m wind speed ($\Delta W$), $\alpha_i$ are the regression coefficients of the $X_i$ with $\Delta AP$, and $b$ is the intercept, which is a constant. We assume that all three meteorological factors should be included in the regression and we estimate the contributions of each factor to changes of AP as:

$$C_i = \frac{\alpha_i X_i}{\sum \alpha_i X_i}$$

(5)

where $C_{i(i=1,2,3)}$ is the contributions (in units of temperature) from each factor to the changes of the AP, and $\overline{X}_i$ are the mean differences in temperature equivalent due to each factor between two scenarios.

3. Results

3.1 Recent apparent temperatures
Figure 2. Seasonal averaged AP and equivalent temperature of each term in equation 1 for Beijing-Tianjin province (a-c) and Beijing-Tianjin urban areas (d-f) during 2008-2017 from ERA5 (a, d), 4-model ensemble mean after ISIMIP (b, e) and ensemble mean after WRF (c, f). Term 1 is 1.04T, term 2 is 2P and term 3 is -0.65W.

Figure 2 shows the seasonal averaged AP and equivalent temperatures caused by temperature, relative humidity and wind speed in Beijing-Tianjin province and Beijing-Tianjin urban areas during 2008-2017. According to the ERA5 results (Fig. 2a, 2d), AP and the separate 3 terms show similar seasonal patterns over the whole province and just the urban areas. Vapor pressure is higher in summer and wind speed is higher in spring. AP is lower than 2 m temperature in all seasons except summer, and especially lower in winter. AP, temperature, vapor pressure and wind speed are all higher in urban areas than in the surrounding rural in any season. The ISIMIP results (Fig. 2b, 2e), by design, perfectly reproduce the ERA5 seasonal characteristics of AP, temperature, vapor pressure and wind speed. WRF shows a similar pattern with that from ERA5, but for the Beijing-Tianjin province, WRF overestimates both 2 m temperature and AP in winter by 2.1°C and by 2.4°C respectively relative to ERA5 (Fig. 2c). In the Beijing-Tianjin urban areas, WRF overestimates the temperature and AP relative to ERA5 in all seasons, especially in winter (Fig. 2f).
Figure 3. Top row: the spatial distribution of mean apparent temperature from ERA5 (a), 4-model ensemble mean after ISIMIP (b) and ensemble mean after WRF (c) during 2008-2017. Bottom row: the spatial distribution of annual mean number of days with AP > 32°C from ERA5 (d), ISIMIP (e) and WRF (f) during 2008-2017. Fig. S1 and Fig. S2 show the pattern of AP and NdAP_32 for the individual ESM.

We compare the simulations of mean apparent temperature and NdAP_32 from both WRF dynamical downscaling with QDM and from ISIMIP statistical downscaling during 2008-2017 in Fig. 3. Both WRF with QDM and ISIMIP methods produce a pattern of apparent temperature which is close to that from ERA5. The average annual AP from ISIMIP is almost the same as that from ERA5 over the Beijing-Tianjin province (Table 1). While WRF produces warmer apparent temperatures in the city centers of Beijing and Tianjin and lower ones in the high Zhangjiakou mountains than recorded in the lower resolution ERA5 observations. There are also differences between different models after WRF downscaling. For example, apparent temperatures from the two MIROC models from WRF are the warmest. In contrast AP from all 4 ESMs after ISIMIP shows very similar patterns (Fig. S1). Both ISIMIP and WRF appear to overestimate the NdAP_32 in Beijing urban areas and the southerly lowland areas although NdAP_32 is close to zero for all methods in the colder rural areas at relatively high altitude. While some of these differences may be due to the 0.25°×0.25° resolution ERA5, which is coarser than the 10 km WRF simulation, it probably does not account for the broad overestimate across most the North China Plain that is within the WRF and ISIMIP domains. ERA5 gives about 10 NdAP_32 per year in southern Beijing and Tianjin, but there are nearly 15 NdAP_32 from ISIMIP, and over 20 NaAP_32 per year from WRF downscaling in the Beijing-Tianjin urban areas during 2008-2017. NdAP_32 from WRF and ISIMIP downscaling of all ESM is overestimated relative to ERA5. But there are curious differences in ESM under the two downscalings: with ISIMIP, HadGEM2-ES and BNU-ESM have more NdAP_32 than the two MIROC models, while the reverse occurs with WRF (Fig. S2).
Table 1. The annual mean apparent temperature and population weighted NdAP\_32 in Beijing-Tianjin province and Beijing-Tianjin urban areas (Fig. 1b) from ERA5, ISIMIP and WRF during 2008-2017.

<table>
<thead>
<tr>
<th>Data Sources</th>
<th>Provinces AP (°C)</th>
<th>Urban AP (°C)</th>
<th>Population weighted for province (Fig. 1c, 1d) NdAP_32 (day yr^-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WRF ISIMIP</td>
<td>WRF ISIMIP</td>
<td>WRF ISIMIP</td>
</tr>
<tr>
<td>MIROC-ESM</td>
<td>10.5 9.6</td>
<td>13.6 11.4</td>
<td>22.2 10.1</td>
</tr>
<tr>
<td>MIROC-ESM-CHEM</td>
<td>10.5 9.6</td>
<td>13.6 11.4</td>
<td>21.9 11.0</td>
</tr>
<tr>
<td>HadGEM2-ES</td>
<td>9.5 9.6</td>
<td>12.0 11.4</td>
<td>12.3 11.1</td>
</tr>
<tr>
<td>BNU-ESM</td>
<td>9.4 9.7</td>
<td>11.8 11.5</td>
<td>10.2 12.7</td>
</tr>
<tr>
<td>ERA5</td>
<td>9.6 11.4</td>
<td></td>
<td>7.7</td>
</tr>
</tbody>
</table>

The Taylor diagram of the daily mean apparent temperature in Beijing-Tianjin province and Beijing-Tianjin urban areas from 2008-2017 for the 4 ESMs shows that all models under both downscaling methods have correlation coefficients with ERA5 > 0.85. Although AP over the both whole Beijing-Tianjin province and the urban areas are overestimated by WRF, it performs slightly better than ISIMIP on the Taylor plot relative to ERA5 (Fig. S3). Under the ISIMIP method, MIROC-ESM, MIROC-ESM-CHEM and HadGEM2-ES show little differences in correlation or errors while the performance of BNU-ESM is slightly worse. Under WRF simulations, MIROC-ESM and MIROC-ESM-CHEM have larger correlation coefficients and smaller errors than HadGEM2-ES and BNU-ESM.

Figure 4 shows the probability density functions (pdf) of daily AP from the four ESMs under ISIMIP and WRF in Beijing-Tianjin province and Beijing-Tianjin urban areas during 2008-2017. ISIMIP overestimates the probability of extreme cold AP relative to ERA5 (especially BNU-ESM), although all ESM reproduce the ERA5 pdf well at high AP. WRF can reproduce the ERA5 distribution of AP better than ISIMIP, but high AP is overestimated relative to ERA5 and the urban areas perform less well than the whole Beijing-Tianjin province. In urban areas all ESMs driving WRF tend to underestimate the probability of lower AP and to overestimate the probability of higher AP, especially the two MIROC models (Fig. 4d). Fig. S4 displays the annual cycle of monthly AP, with ISIMIP proving excellent by design, at reproducing the monthly AP. While under WRF downscaling AP shows more across model differences, especially during summer and with greater spread for the urban areas.
Figure 4. The probability density function (pdf) for daily apparent temperature under ISIMIP (a, c) and WRF (b, d) results in Beijing-Tianjin province (a, b) and Beijing-Tianjin urban areas (c, d) during 2008-2017.

3.2 2060s apparent temperatures

3.2.1 Changes of apparent temperature

Figure 5. Spatial pattern of ensemble mean apparent temperature difference under different scenarios over 2060-2069: G4-2010s (left column), G4-RCP4.5 (middle column) and G4-RCP8.5 (right column) based on ISIMIP and WRF methods. 2010s refers to the 2008-2017 period. Stippling indicates grid...
points where differences or changes are not significant at the 95% level according to the Wilcoxon signed rank test.

Figure 5 shows the ISIMIP and WRF ensemble mean changes in the annual mean AP under G4 during 2060-2069 relative to the past and the two future RCP scenarios. ISIMIP-downscaled AP (Fig. 5a-5c) shows significant anomalies (p<0.05) across the whole domain, even for the relatively small differences in G4-RCP4.5. There are no models with obvious regional differences in AP anomalies (Fig. S6). G4 is about 2°C warmer than the 2008-2017 period and about 1°C colder than RCP4.5 and 3°C colder than RCP8.5. WRF downscaling (Fig. 5d-5f) anomalies are similar but the warming under G4 relative to the 2010s is smaller and the coolings relative to both RCP scenarios are a little smaller than those under ISIMIP. Individual ESM driven ISIMIP results are in Fig. S6 and WRF results in Fig. S7. For both ISIMIP and WRF downscaling the MIROC models show stronger warming than the other two models between G4 and the 2010s. WRF-downscaled AP driven by HadGEM2-ES exhibits the strongest cooling (> 1.5°C for G4-RCP4.5 and 3°C for G4-RCP4.5). AP changes, whether across all province or just urban areas, are essentially the same (Table 2), which is consistent with patterns in figure 5. The ensemble mean differences in AP between G4 and RCP scenarios calculated both using ISIMIP and WRF downscaling are small, however ensemble mean AP differences between G4 and the 2010s over urban areas are 1.0°C under WRF and 2.0°C, under ISIMIP.

**Table 2.** Difference of apparent temperature between the G4 and other scenarios for the Beijing-Tianjin province and Beijing-Tianjin urban areas as defined in Fig. 1b during 2060-2069. Bold indicates the differences or changes are significant at the 5% level according to the Wilcoxon signed rank test.

<table>
<thead>
<tr>
<th>Model</th>
<th>G4-2010s</th>
<th>G4-RCP4.5</th>
<th>G4-RCP8.5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WRF</td>
<td>ISIMIP</td>
<td>WRF</td>
</tr>
<tr>
<td>MIROC-ESM</td>
<td>0.9</td>
<td>1.5</td>
<td>2.2</td>
</tr>
<tr>
<td>MIROC-ESM-CHEM</td>
<td>0.9</td>
<td>1.5</td>
<td>2.9</td>
</tr>
<tr>
<td>HadGEM2-ES</td>
<td>1.1</td>
<td>1.0</td>
<td>1.8</td>
</tr>
<tr>
<td>BNU-ESM</td>
<td>1.2</td>
<td>1.1</td>
<td>1.2</td>
</tr>
<tr>
<td>Ensemble</td>
<td>1.0</td>
<td>1.3</td>
<td>2.0</td>
</tr>
</tbody>
</table>
Figure 6. The seasonal changes of AP (\(\Delta AP\)) and the seasonal contribution of climatic factors to \(\Delta AP\) for Beijing and Tianjin urban areas under ISIMIP and WRF between G4 and 2010s (a), G4 and 2010s (b), G4 and RCP4.5 (c) and G4 and RCP8.5 (d) based on ensemble mean results. Bold tabulated numbers and ** above the columns indicate differences are significant at the 95% under the Wilcoxon test.

We show the seasonal contribution of temperature, humidity and wind to differences in AP between G4, the 2010s, RCP4.5 and RCP8.5 from ISIMIP and WRF downscaling over Beijing-Tianjin urban areas in Fig. 6. Undoubtedly, temperature makes the biggest contribution to \(\Delta AP\) between different scenarios, and \(\Delta AP\) is smaller under WRF than under ISIMIP. The projected differences in scenario temperatures explain more than 90% of the \(\Delta AP\) differences. There are striking differences between WRF and ISIMIP in the seasonal contribution of humidity to \(\Delta AP\) for both G4 and RCP4.5 relative to the 2010s (Fig. 6a, 6b). Under WRF, summer differences in humidity makes a negative contribution to \(\Delta AP\) for G4 while under RCP4.5 humidity makes only a slightly negative but non-significant contribution, but the summer \(\Delta AP\) is much lower than in other seasons. Wind increases \(\Delta AP\) under both G4 and RCP4.5 relative to the 2010s. Fig. 6c and 6d show that \(\Delta AP\) under G4 compared with RCP4.5 and RCP8.5 is significantly affected by humidity in summer. The negative contributions from humidity under WRF amount to 6-9%, but under ISIMIP the contributions are much smaller, and even acts to reduce differences in \(\Delta AP\) between G4 and RCP4.5. Changes in wind are insignificant for \(\Delta AP\) between G4 and RCP4.5 under ISIMIP, but with WRF changes in wind are generally significant and amount to over 3% in summer. In contrast, the seasonal contribution of wind is about 2.5-4.7% under ISIMIP to differences between G4 and RCP8.5 but close to 0.
Figure 7. The change of apparent temperature based on air temperature under three scenarios (G4, RCP4.5 and RCP8.5) in four ESMs under ISIMIP (left column) and WRF (right column) for urban areas relative to the 2010s.

A useful measure of heat impacts that may be missed if considering only at air temperatures is the seasonality of the differences between AP and air temperature under WRF.
The four model ensemble annual mean \( \Delta \) (AP-T) under ISIMIP is projected to rise by 0.4°C, 0.5°C and 0.9°C under G4, RCP4.5 and RCP8.5, relative to the 2010s. Under WRF, \( \Delta \) (AP-T) is much smaller than under ISIMIP but still rising faster than air temperatures: by 0.2°C, 0.3°C and 0.5°C under G4, RCP4.5 and RCP8.5 relative to the 2010s, respectively. In general, the largest anomalies in \( \Delta \) (AP-T) are in summer under both WRF and ISIMIP downscaling, but the two MIROC models under WRF have small or even negative \( \Delta \) (AP-T) in summer with WRF.

### 3.2.2 Changes of the number of days with AP>32°C

![Figure 8](https://doi.org/10.5194/esd-2022-47)

**Figure 8.** Ensemble mean differences in annual number of days with AP > 32°C (NdAP_32) between scenarios for 2060-2069: G4-2010s (left column), G4-RCP4.5 (second column) and G4-RCP8.5 (right column) based on ISIMIP method and WRF. 2010s means the results simulated during 2008-2017. Stippling indicates grid points where differences or changes are not significant at the 5% level according to the Wilcoxon signed rank test. Corresponding ISIMIP results for each ESM are in Fig. S8, and WRF results in Fig. S9.

The NdAP_32 anomalies in Figure 8 show that ISIMIP projects an increase of about 20 days per year with AP>32 °C for the southeast of Beijing province and 10 days in the western areas of Beijing under G4 relative to the 2010s. NdAP_32 is about 10 days fewer under G4 than RCP4.5 with no clear spatial differences. G4 has about 35 fewer NdAP_32 days in the southern part of the domain and 20 fewer days in the western domain than the RCP8.5 scenario. In contrast WRF suggests that most areas do not show any significant difference between G4 and the 2010s, while the anomalies relative to RCP4.5 are similar as ISIMIP, although differences are less significant. G4-RCP8.5 anomalies with WRF are less significant and smaller than with ISIMIP. The urban areas show larger decreases in NdAP_32 than the more rural areas, even in the low altitude plain. Individual ESM show almost no statistically significant differences between G4 and RCP4.5 (Fig. S8 and S9), but the differences seen in Fig. 8 are significant because of the larger sample size in the significance test.
All ESMs with ISIMIP show more NdAP$_{32}$ in the urban areas under G4 than the 2010s, while two MIROC models driving WRF show fewer NdAP$_{32}$ in Beijing-Tianjin urban areas (Fig. S8, S9).

The pdf of daily apparent temperature in summer over Beijing-Tianjin urban areas (Fig. 9) shifts rightwards for G4, RCP4.5 and RCP8.5 during the 2060s relative to the 2010s. Figure 9 shows that by the 2060s, the dangerous threshold of AP$>39$ is crossed frequently under RCP8.5 with both WRF and ISIMIP downscaling, but for the RCP4.5 and G4 scenarios these events are much rarer. ISIMIP results tend to show higher probability tails (extreme events) than under WRF simulations.
Population weighted NdAP_32 in the 2060s for Beijing-Tianjin province is shown in Table 3. ISIMIP downscaling suggests ensemble mean rises in NdAP_32 of 22.4 days per year under G4 relative to the 2010s, but that G4 has 8.6 and 33.5 days per year fewer than RCP4.5 and RCP8.5, respectively. NdAP_32 from WRF under G4 is reduced by 19.6 days per year relative to RCP8.5, and by 6.3 days relative to RCP4.5 (Table 3).

Table 3. Difference of population weighted NdAP_32 between the G4 and other scenarios for Beijing-Tianjin province (Fig. 1c, 1d) during 2060-2069. Bold indicates the changes are significant at the 5% level according to the Wilcoxon signed rank test. (Units: day y$^{-1}$).

<table>
<thead>
<tr>
<th>Beijing-Tianjin province</th>
<th>G4-2010s</th>
<th>G4-RCP4.5</th>
<th>G4-RCP8.5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ISIMIP</td>
<td>WRF</td>
<td>ISIMIP</td>
</tr>
<tr>
<td>MIROC-ESM</td>
<td>18.6</td>
<td>-8.1</td>
<td>-17.0</td>
</tr>
<tr>
<td>MIROC-ESM-CHEM</td>
<td>28.7</td>
<td>-10.2</td>
<td>3.9</td>
</tr>
<tr>
<td>HadGEM2-ES</td>
<td>25.7</td>
<td>9.4</td>
<td>-12.5</td>
</tr>
<tr>
<td>BNU-ESM</td>
<td>16.4</td>
<td>13.6</td>
<td>-8.6</td>
</tr>
<tr>
<td>Ensemble</td>
<td>22.4±2.9</td>
<td>1.2±6.0</td>
<td>-8.6±4.5</td>
</tr>
</tbody>
</table>

4. Discussion and Conclusion

Our study on thermal comfort under geoengineering scenarios for the Beijing megalopolis may be useful across the developing world which is expected to suffer disproportionate climate impact damages relative the global mean, while also undergoing rapid urbanization. Assessing health impacts and mortality due to heat stress under greenhouse gas scenarios should consider urbanization and the change to concrete surfaces from vegetation that leads to differences in heat capacities, rates of evapotranspiration, and hence humidity and apparent temperature. These require downscaled analyses, accurate meteorological and high-resolution land surface datasets.

In our analysis we assumed the urban area did not change over time, and also that population remains distributed as in the recent past. This may be reasonable in the highly developed and relatively mature greater Beijing-Tianjin region but should be considered in rapidly urbanizing regions elsewhere. But there certainly will be changes over time in the radiative cooling from surface pollution sources. PM2.5 is a health issue in many developing regions (Ran et al., 2022), but as wealth increases efforts to curb air pollution generally clean the air. This has clear health benefits, but also removes aerosols from the troposphere that cool the surface. The urban areas that have higher apparent temperatures at present are also the areas with greatest aerosol load and hence greatest cooling. Once that is removed direct radiation, air temperatures and apparent temperatures will all rise – by several degrees (Wang et al., 2016). So a future more comprehensive health impact study would include both the negative health impacts of aerosol pollution and the potential cooling effects those
Both ISIMIP and WRF can reproduce the observed (ERA5) spatial patterns and seasonal variabilities of apparent temperature in the region around Beijing. WRF shows warm biases in AP during all months relative to ERA5 due to warmer temperatures in urban areas, with the exception of driving from the BNU-ESM and HadGEM2-ES in summer (Fig. S5). Both ISIMIP and WRF tend to overestimate population weighted NdAP_32 by 46% and 116%, respectively. These large discrepancies are due to relatively small overestimates of the likelihood of the tails of the probability distributions which leads to a dramatic increase in the frequency of extreme climate events (Dimri et al., 2018; Huang et al., 2021). AP is about 1.5°C warmer that 2 m temperature over the Beijing and Tianjin urban areas in summer due to higher vapor pressures amplifying warmer urban temperatures, and this is despite humidity being lower over the cities. Under high humidity conditions, a slight increase in temperature will cause a large increase in heat stress (Li et al., 2018; Luo and Lau, 2019). AP is nearly 4°C colder than 2 m temperature in winter due to wind speed (Fig. 2d). Differences between AP and 2 m temperature (AP-T) during summer are greater in urban areas than neighboring rural areas.

The apparent temperatures in Beijing Tianjin urban areas under G4 in the 2060s are simulated to be 1°C and 2.5°C lower than RCP4.5 and RCP8.5, although AP would be higher than in the recent past. The cooling effect of G4 relative to RCP4.5 and RCP8.5 is greatest under HadGEM2-ES (Fig. S6, S7), due to the ESM having largest temperature differences between scenarios (Wang et al., 2022 in review). WRF downscaling produces reduced seasonality in AP compared with ISIMIP, and WRF produces relatively cooler summers and warmer winters than ISIMIP, and so much less differences in apparent temperature ranges (Fig. 10). Differences in AP between G4 and the RCP scenarios are mainly driven by temperature. In all scenarios and downscalings AP rises faster than the temperature due to decreased wind speeds in the future (Li et al., 2018; Zhu et al., 2021) but mainly because of rises in vapor pressure driven by rising temperatures. This effect occurs despite the general drying expected under solar geoengineering (Bala et al., 2008; Yu et al., 2015).

The NdAP_32 under G4 is projected to decrease by 8.6 days per year by ISIMIP and 6.3 days per year by WRF relative to RCP4.5 for Beijing-Tianjin Province. Much larger reductions in NdAP_32 of 33.5 days per year (ISIMIP) and 19.6 days per year (WRF) are projected relative to RCP8.5. Differences between scenarios in frequency of dangerously hot days are far larger using ISIMIP statistical downscaling than using WRF. This is another impact of the reduced seasonality of WRF compared with ISIMIP (Fig. 10).

The higher resolution WRF simulation produces a much larger range of apparent temperatures across the domain than ERA5 and ISIMIP downscaling. This increased variability makes reaching a statistical significance threshold more challenging for
WRF than ISIMIP results. Despite this, the ESM-driven differences in WRF output are less than from ISIMIP, reflecting the physically based processes in the dynamic WRF simulation. This reduces the impact of differences in ESM forcing at the domain boundaries with WRF compared with the statistical bias correction and downscaling methods. Although there are some uncertainties between models and downscaling methods, G4 SAI can not only reduce the mean apparent temperature but also decrease the probability of PDF tails (extreme events) in summer.

Figure 10. Seasonal cycles of apparent temperature from MIROC-ESM, MIROC-ESM-CHEM, HadGEM2-ES and BNU-ESM under G4, RCP4.5 and RCP8.5 in Beijing-Tianjin urban areas during 2060s based on ISIMIP (red) and WRF (black) methods.
Code and data availability

All ESM data used in this work are available from the Earth System Grid Federation (WCRP, 2021; https://esgf-node.llnl.gov/projects/cmip6, last access: 14 July 2021). The WRF and ISIMIP bias-corrected and downscaled results are available for the authors on request. WRF and ISIMIP codes are freely available at the references cited in the methods sections.

Supplement link

The link to the supplement will be included by Copernicus.

Author contribution

JCM and LZ designed the experiments, JW performed the simulations. All the authors contribute to the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Disclaimer

Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Special issue statement:

This article is part of the special issue “Resolving uncertainties in solar geoengineering through multi-model and large-ensemble simulations (ACP/ESD inter-journal SI)”. It is not associated with a conference.

Acknowledgements

We thank the climate modeling groups for participating in the Geoengineering Model Intercomparison Project and their model development teams; the CLIVAR/WCRP Working Group on Coupled Modeling for endorsing the GeoMIP; and the scientists managing the earth system grid data nodes who have assisted with making GeoMIP output available. This research was funded by the National Key Science Program for
Global Change Research (2015CB953602).

References


Vandyck, T., Keramidas, K., Saveyn, B., et al.: A global stocktake of the Paris pledges: Implications for energy systems and economy, Global Environmental Change, 41,


