[Geophysical Research Letter]

Supporting Information for

[Strong dependence of U.S. summertime air quality on the decadal variability of Atlantic sea surface temperatures]

[L. Shen[1], L. J. Mickley[1], E. M. Leibensperger[2], M. Li[3]]

[[1]John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
[2]Center for Earth and Environmental Science, SUNY Plattsburgh, Plattsburgh, NY 12901, USA
[3]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA]

Contents of this file

Text S1. Statistical methods to calculate PM$_{2.5}$ concentrations and MDA8 ozone

For PM$_{2.5}$, we first construct the synoptic circulation factors across the eastern United States through the use of singular value decompositions (SVDs) of the spatial correlations between PM$_{2.5}$ in each grid box and five meteorological variables in gridboxes in the surrounding region. The five variables are surface temperature, relative humidity, precipitation, and east-west and north-south wind speed. The SVD method effectively compresses the information from a multi-dimensional matrix of these four variables into a set of scalars that represent the oscillation of the PM$_{2.5}$-related synoptic patterns (Shen et al., 2017a). We next develop a multiple linear regression model to correlate observed JJA monthly mean PM$_{2.5}$ concentrations and the local values of the same five meteorological variables and the two most important synoptic factors in each gridbox, diagnosed using SVD. The model is of the form

$$Y = \sum_{k=1}^{n} \alpha_k X_k + \sum_{n=1}^{2} \beta_n S_n + b \qquad (S1)$$

where Y is the JJA monthly mean PM$_{2.5}$ concentration for 1999-2015, X is a local meteorological variable, S is a synoptic circulation factor, α and β are coefficients, and b is the intercept. In order to avoid over-fitting, we use leave-one-out cross-validation, yielding the best variable combinations for each gridbox. Using leave-one-out cross validation, we find that the coefficients of determination (R^2) between observed and predicted PM$_{2.5}$ are 0.3-0.6 for PM$_{2.5}$ from 1999-2015 (Figure S4b).

For MDA8 ozone, we follow a similar approach. Previous studies have diagnosed the strong dependence of surface ozone on mid-tropospheric wind fields (Shen et al., 2015; Porter et al., 2015), and so we construct the ozone-related synoptic patterns using six meteorological variables, including surface temperature, relative humidity, and east-west and north-south wind speed on the surface, as well as east-west and north-south wind speed at 500 hPa. Using leave-one-out cross validation, R^2 between observed and predicted JJA MDA8 ozone range from 0.4-0.7 (Figure S4a).

We use the hybrid extreme value theory (Hybrid-EVT) model to simulate ozone extremes (> 70 ppbv) conditionally on daily mean temperature [Shen et al., 2016]. For most observations sites, the point process (PP) model, which formulates the Poison process limit of extreme ozone distribution above a threshold, can simulate well the ozone high tails. At sites where ozone suppression occurs, the PP model may fail to capture the ozone-temperature relationship, and so we refit the distribution using a logistic regression and a Generalized Pareto Distribution (GPD) model [Shen et al., 2016]. Unlike the PP model the GPD model only accounts for the ozone distribution above a threshold but not the probability exceeding this threshold [Cole, 2001]. Thus our final Hybrid-EVT model consists of both PP and GDP models, and their parameters are conditioned on daily temperature.

Text S2. Health risk calculation

The health impacts of ozone and PM$_{2.5}$ changes are calculated following a previous study (Shindell et al., 2016). Changes in premature deaths are calculated as

$$\Delta M = M_b \times P \times AF \tag{S2}$$

$$AF = 1 - \exp(-\beta\Delta C) \tag{S3}$$

$$RR = \exp(\beta\Delta C) \tag{S4}$$

where P is the adult population (age $\geq$ 30) in the United States, M_b is the baseline mortality rate for the adult population based on estimates from World Health Organization (WHO 2011), ΔM is the number of premature deaths due to changes in PM$_{2.5}$/ozone, AF is the attributable fraction of deaths due to diseases related to PM$_{2.5}$/ozone, β is the slope of concentration-response function (CRF), ΔC is the change of PM$_{2.5}$/ozone concentration, and RR is the relative risk. The long-term RR per 10 μg m^{-3} increase in annual mean PM$_{2.5}$ is 1.14 (95% confidence interval (CI): 1.04—1.23) for lung cancer and 1.09 (95% CI: 1.03—1.16) for cardiovascular and respiratory diseases (Pope et al., 2002). We increase the CRF slopes by 80% based on expert elicitation (Anenberg et al., 2012), which means the CRF slopes (β) are ln(1.14/10)×1.8 for lung cancer and ln(1.09/10)×1.8 for cardiovascular and respiratory diseases (Pan et al., 2004; Anenberg et al., 2012). The JJA changes of PM$_{2.5}$ concentrations estimated in this study are scaled by 0.25 to represent their annual mean changes. The long-term RRs from respiratory disease is 1.04 (95% CI: 1.010-1.067) per 10 ppbv increase in the maximum 6-month average of 1-h daily maximum ozone (Jerrett et al., 2009), which corresponds to a β of ln(1.04/10). The JJA changes in ozone concentrations are scaled by 0.5 to estimate the maximum 6-month average changes. Using the hourly ozone observations for 1990-2015, we find that the standard deviation of detrended JJA seasonal mean 1-h daily maximum ozone is 10% greater than the MDA8 ozone, so we increase the ΔC of MDA8 ozone by 10% when calculating the health risk. The average health risk as well as its uncertainty is calculated from 10,000 bootstrap simulations.

Table S1. Datasets used in this study.

Dataset	Description
Observations and datasets derived from observations (1895-2015)	Temperature: NCDC[*], GHCN[†], Delaware[‡], MLOST[§] and CRUTEM4[‖] Precipitation: NCDC[*], GPCC[¶], GHCN[†], Delaware[‡] Sea level pressure: HadSLP2[#] Drought: 8 different drought indices in NCDC[*] Ozone and $PM_{2.5}$: U.S. EPA Air Quality System (AQS)[**]
NOAA-20CR 56-member ensemble, v2c, 1895-2014 (Compo et al., 2011)	Results from climate model simulation Boundary conditions: SSTs and sea ice extent Assimilation: surface pressure
ERA-20CM 10-member ensemble, 1900-2010 (Hersbach et al., 2015)	Results from climate model simulation Boundary conditions: SSTs and sea ice extent, forcings from CMIP5 Assimilation: none
GISS ModelE2 21-year equilibrium simulation (Schmidt et al., 2014)	Results from climate model simulation Boundary conditions: prescribed SSTs and sea ice extent in warm and cold AMO phases. Assimilation: none

[*] Historical temperature, precipitation, and drought indices for U.S. climate divisions from National Climatic Data Center (*Vose et al., 2014*). The drought indices include Palmer Drought Severity Index (PDSI), Palmer Hydrological Drought Index (PHDI), Palmer "Z" Index (ZNDX), Modified Palmer Drought Severity Index (PMDI), and Standardized Precipitation Index over 3, 6, 9 and 12 months (SPxx).

[†] Gridded Historical Climatology Network (GHCN, v2 for precipitation and v3 for temperature) provided by NOAA/OAR/ESRL Physical Sciences Division.

[‡] University of Delaware air temperature and precipitation dataset (Willmott et al., 2011).

[§] Merged Land-Ocean Surface Temperature Analysis (MLOST) (Smith et al., 2008; Vose et al., 2012).

[‖] Jones (CRU) Air Temperature Anomalies Version 4 (CRUTEM4) (Jones et al., 2012).

[¶] Global Precipitation Climatology Centre (GPCC) (Schneider et al., 2011).

[#] Hadley Centre sea level pressure dataset (HadSLP2) (Allan et al., 2006).

[**] Site measurements of ozone (1980-2015) are interpolated onto a 2.5°×2.5° latitude-by-longitude grid resolution using spatial averaging (Shen et al., 2015). $PM_{2.5}$ site measurements, which are less dense spatially than ozone and have a shorter record (1999-2015), are interpolated onto the same resolution but using inverse distance weighting (Shen et al., 2017a).

Table S2. Models from the Coupled Model Intercomparison Project Phase 5 (CMIP5) used for this study.

| Model Name	Institute
ACCESS1.0	Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia
ACCESS1.3	CSIRO and BOM, Australia
BCC-CSM1-1	Beijing Climate Center, China Meteorological Administration
BNU-ESM	College of Global Change and Earth System Science, Beijing Normal University
CanESM2	Canadian Centre for Climate Modelling and Analysis
CCSM4	National Center for Atmospheric Research
CESM1-CAM5	Community Earth System Model Contributors
CMCC-CM	Centro Euro-Mediterraneo per i Cambiamenti Climatici
CMCC-CMS	Centro Euro-Mediterraneo per i Cambiamenti Climatici
CNRM-CM5	Centre National de Recherches Météorologiques / Centre Européen de Recherche et Formation Avancée en Calcul Scientifique
CSIRO-MK3-6-0	Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence
EC-EARTH	EC-EARTH consortium
FIO-ESM	The First Institute of Oceanography, SOA, China
FGOALS-g2	LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences
GFDL-CM3	NOAA Geophysical Fluid Dynamics Laboratory
GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory
GFDL-ESM2G	NOAA Geophysical Fluid Dynamics Laboratory
GISS-E2-H	NASA Goddard Institute for Space Studies
GISS-E2-R	NASA Goddard Institute for Space Studies
HadGEM2-AO	Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)
HadGEM2-CC	Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)
HadGEM2-ES	Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)
INMCM4	Institute for Numerical Mathematics
IPSL-CM5A-LR	Institut Pierre-Simon Laplace
IPSL-CM5A-MR	Institut Pierre-Simon Laplace
IPSL-CM5B-LR	Institut Pierre-Simon Laplace
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies
MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies

|	
MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology
MPI-ESM-LR	Max-Planck-Institut für Meteorologie (Max Planck Institute for Meteorology)
MPI-ESM-MR	Max-Planck-Institut für Meteorologie (Max Planck Institute for Meteorology)
MRI-CGCM3	Meteorological Research Institute
NORESM1-M	Norwegian Climate Centre
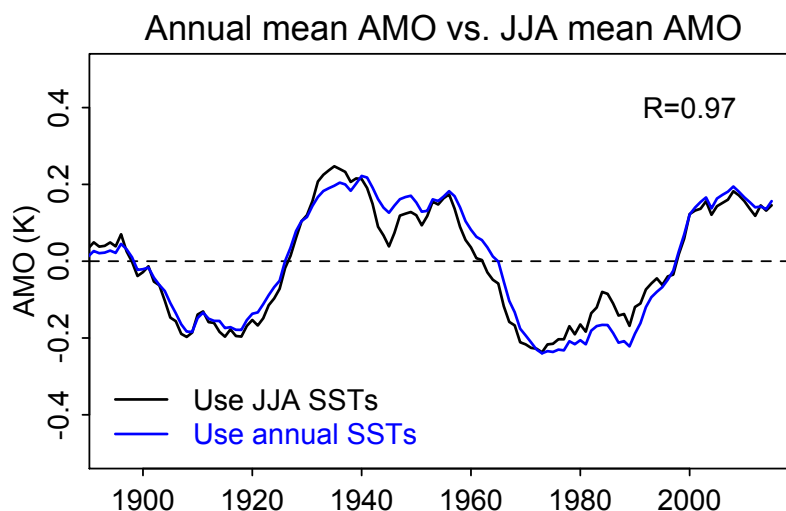
Figure S1. Timeseries of the 11-year running mean values of JJA AMO (black) and annual mean AMO (blue). Correlation between the two timeseries is shown inset. We use JJA AMO in this study.
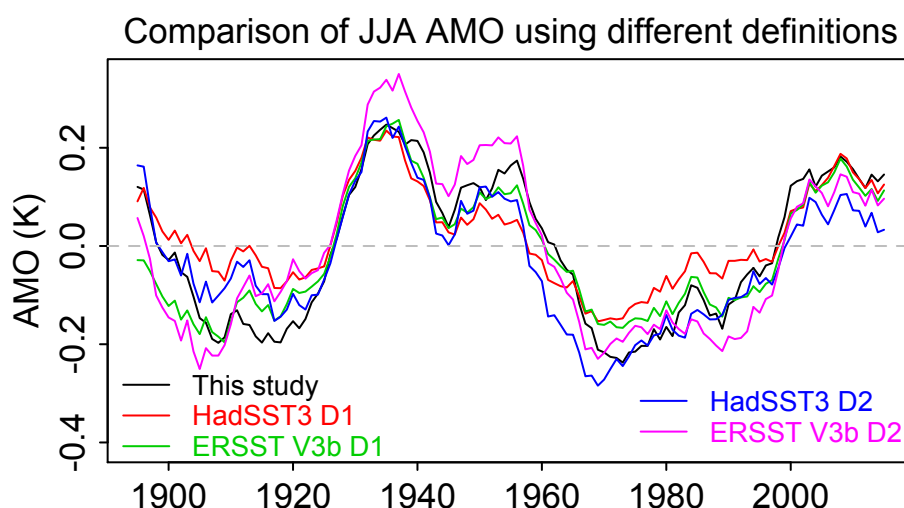
Figure S2. Timeseries of the 11-year running mean values of JJA AMO in this study (black) and those using two other definitions (D1 and D2) with different SST datasets. D1 refers to SSTs averaged over 25°N-60°N, 7°W-75°W minus regression on global mean temperatures (Van Oldenborgh et al., 2009). D2 refers to SSTs averaged over 0°N-60°N, 0°W-80°W minus global mean temperatures from 60°S-60°N (Trenberth and Shea, 2006). These indices can be obtained from Royal Netherlands Meteorological Institute (https://climexp.knmi.nl/selectindex.cgi?id=someone@somewhere). The SST datasets used here are Hadley Centre Sea Ice and Sea Surface Temperature (HadISST, Kennedy et al., 2011) v3 and Extended Reconstructed Sea Surface Temperature (ERSST v3b) (Smith et al., 2008).
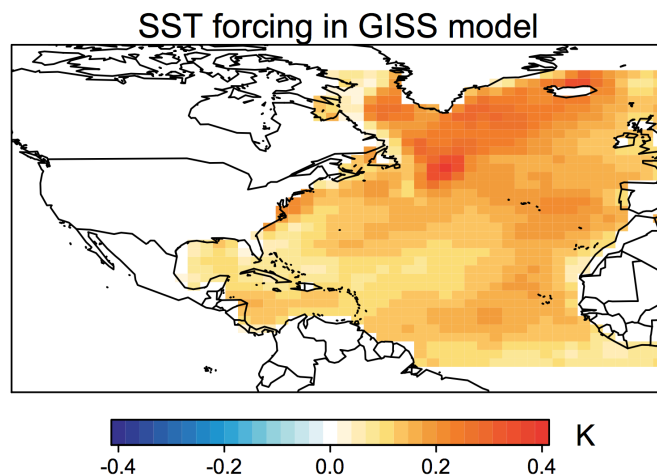
Figure S3. Map of perturbed SSTs used as boundary conditions in the GISS ModelE2 for AMO simulations. The plot shows the SST response to a positive unit change in AMO (see Methods).
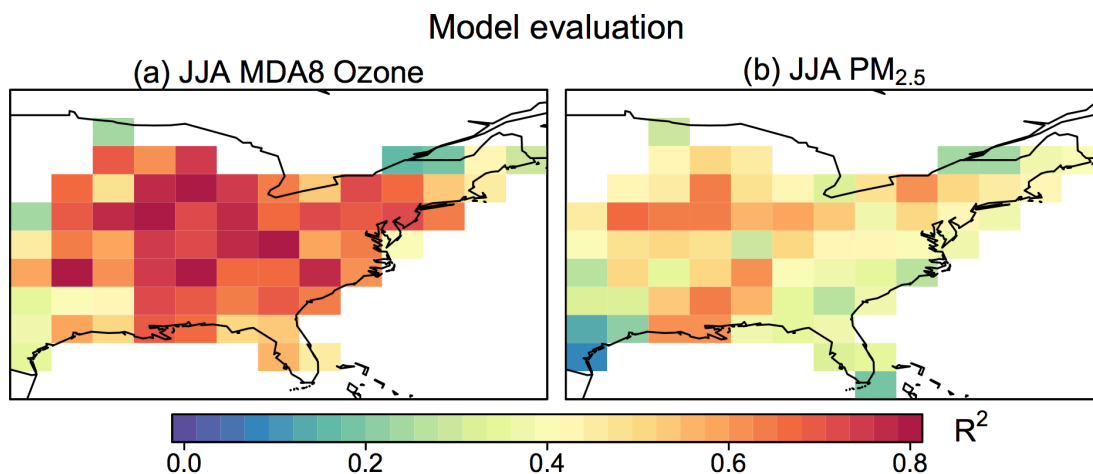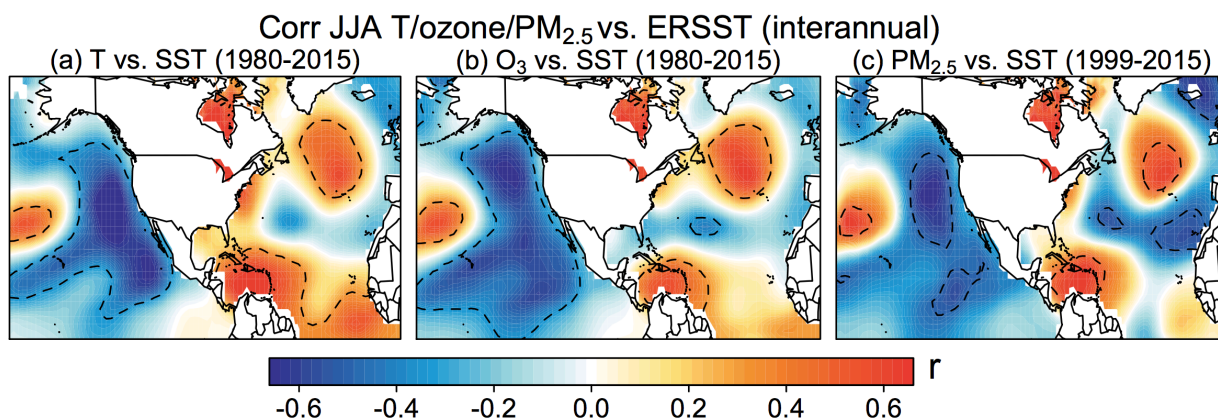

Figure S4. (a) Maps of coefficients of determination (R^2) between observed and predicted JJA seasonal mean MDA8 ozone concentrations for 1980-2015. (b) Same as (a), but using JJA monthly mean PM$_{2.5}$ concentrations for 1999-2015. When averaged across the eastern United States, the R^2 is 0.60 for ozone and 0.41 for PM$_{2.5}$.
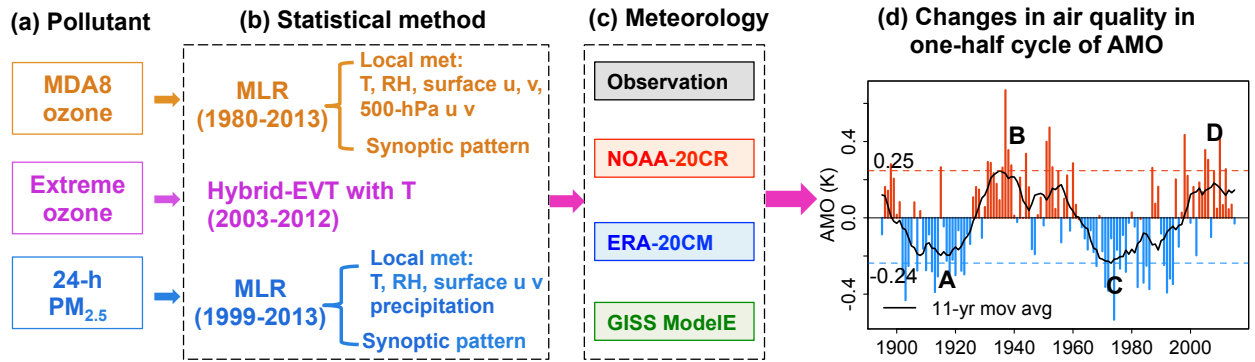
**Figure S5.** Methodology used to estimate the effects of AMO on MDA8 ozone, ozone episodes (> 70 ppbv), and JJA mean PM$_{2.5}$ in one-half AMO cycle from the cold to warm phase. To predict air quality using observed meteorology, we use MLOST temperatures, east-west and north-south windspeeds derived from HadSLP2 sea level pressures, and GPCC precipitation. The changes of air quality in one-half AMO cycle refer to the average changes from the peak in a cold phase (A, C) to the peak in a warm phase (B, D), as described in Methods. Dashed horizontal lines in panel (d) refer to the maximum (+0.25 K) and minimum (-0.24 K) 11-yr running mean AMO indices for the entire time frame. MLR refers to multivariate linear regression and Hybrid-EVT to a hybrid extreme value theory model. Details of the models and datasets are described in Table S1. Bars in panel (d) show the JJA index for positive (red) and negative (blue) AMO; black curve signifies the 11-year moving average of AMO.
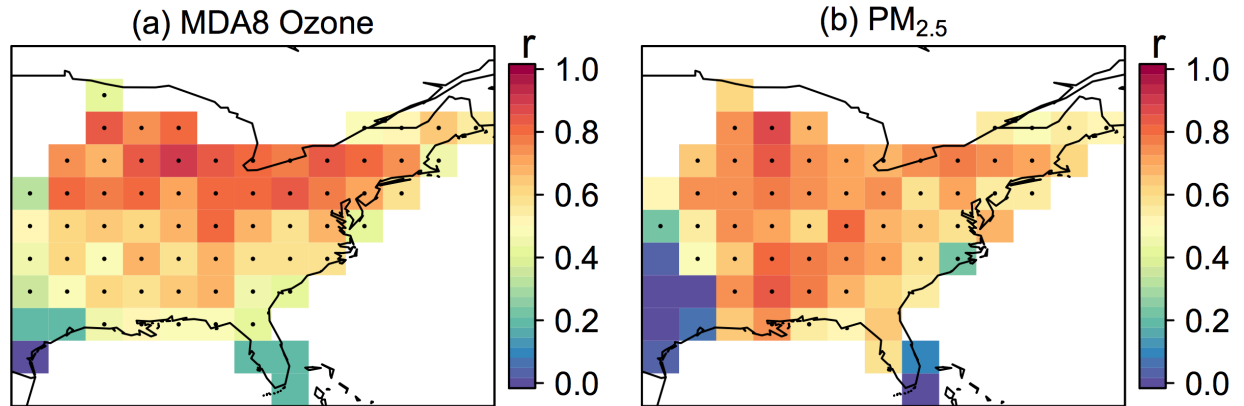

Figure S6. (a) Correlation coefficient *r* of JJA MDA8 ozone and NCEP surface air temperatures from 1980 to 2015. (b) Same as (a), but for PM$_{2.5}$ concentrations from 1999 to 2015. In all panels, gridboxes with statistically significant ($p < 0.05$) correlations are stippled. All data are detrended by subtracting the 7-year moving averages.

Corr JJA T/ozone/PM$_{2.5}$ vs. ERSST (interannual)

(a) T vs. SST (1980-2015) (b) O$_3$ vs. SST (1980-2015) (c) PM$_{2.5}$ vs. SST (1999-2015)

Figure S7. Similar to Figure 1c-e, but using based on NOAA Extended Reconstructed Sea Surface Temperature (ERSST v4) instead of HadISST.
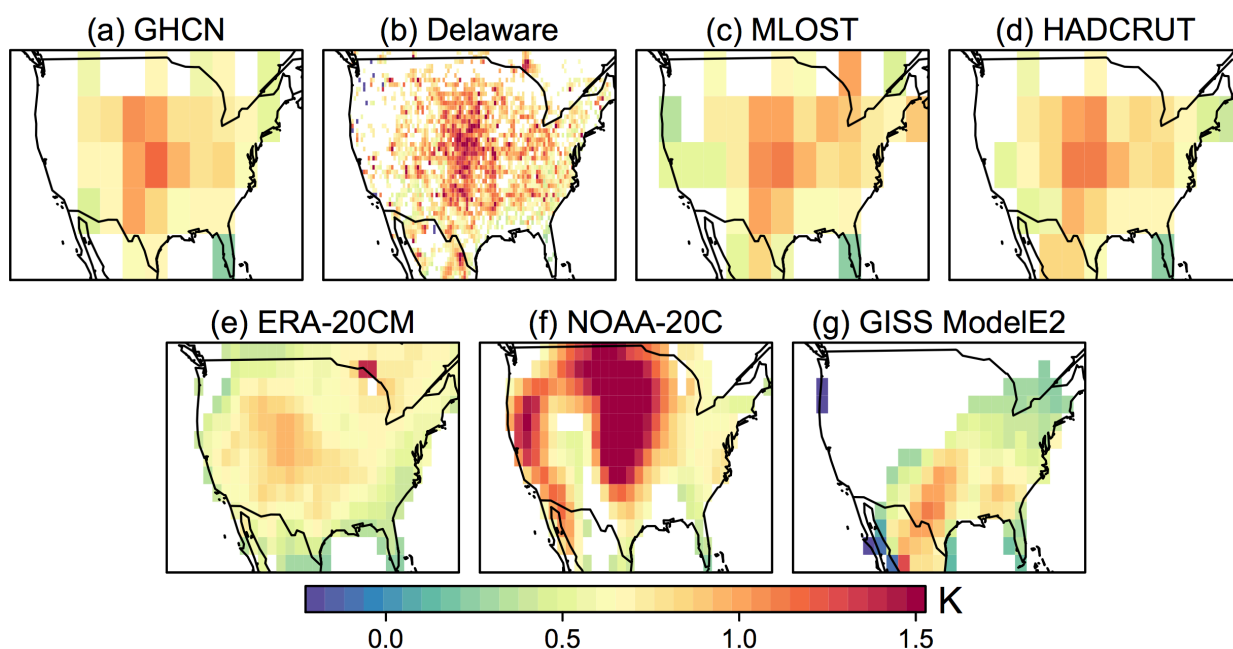

Figure S8. Change of mean JJA surface air temperatures in one-half AMO cycle inferred from different datasets: (a) GHCN, (b) Delaware, (c) MLOST, (d) HADCRUT, (e) ERA-20CM, (f) NOAA-20CR, and (g) GISS ModelE2. See Table S1 and Methods for more details.
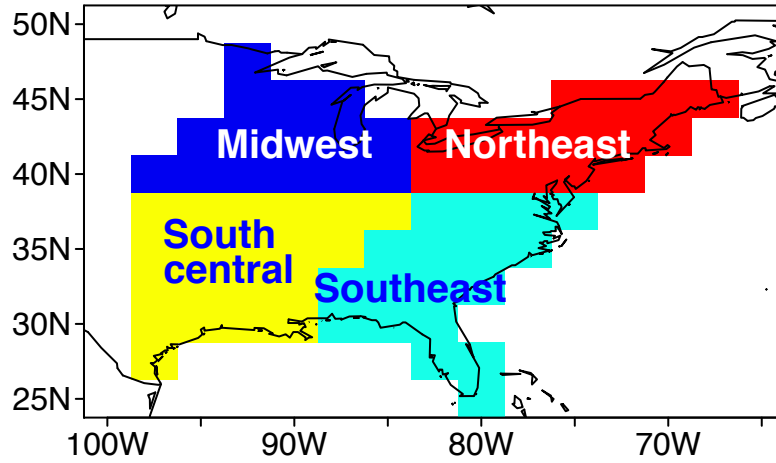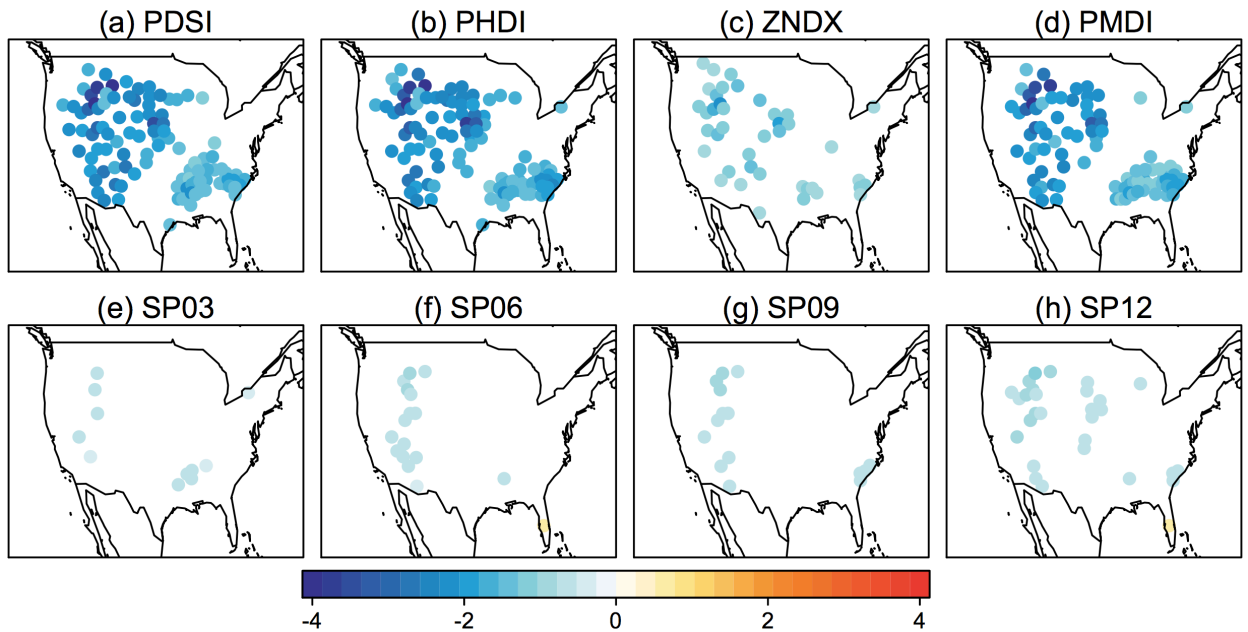
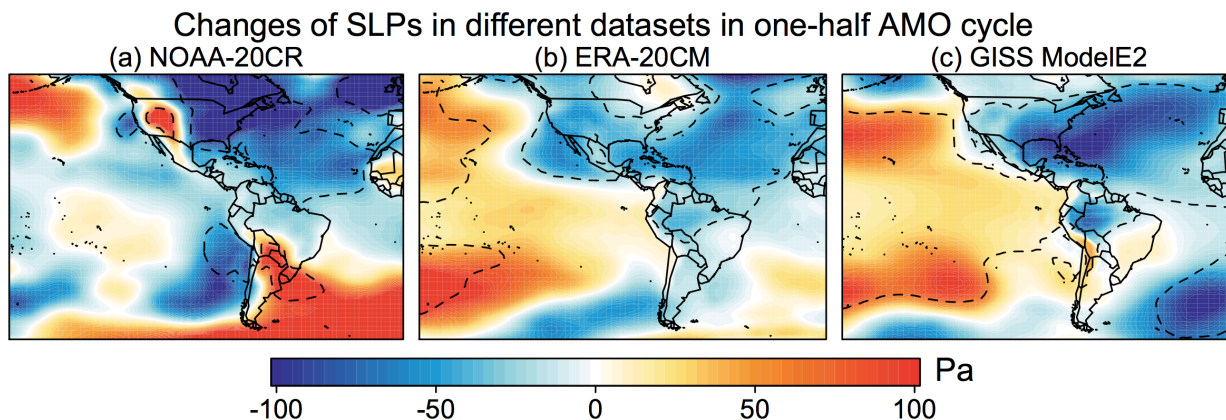**Figure S9**. The four U.S. regions used in this study.
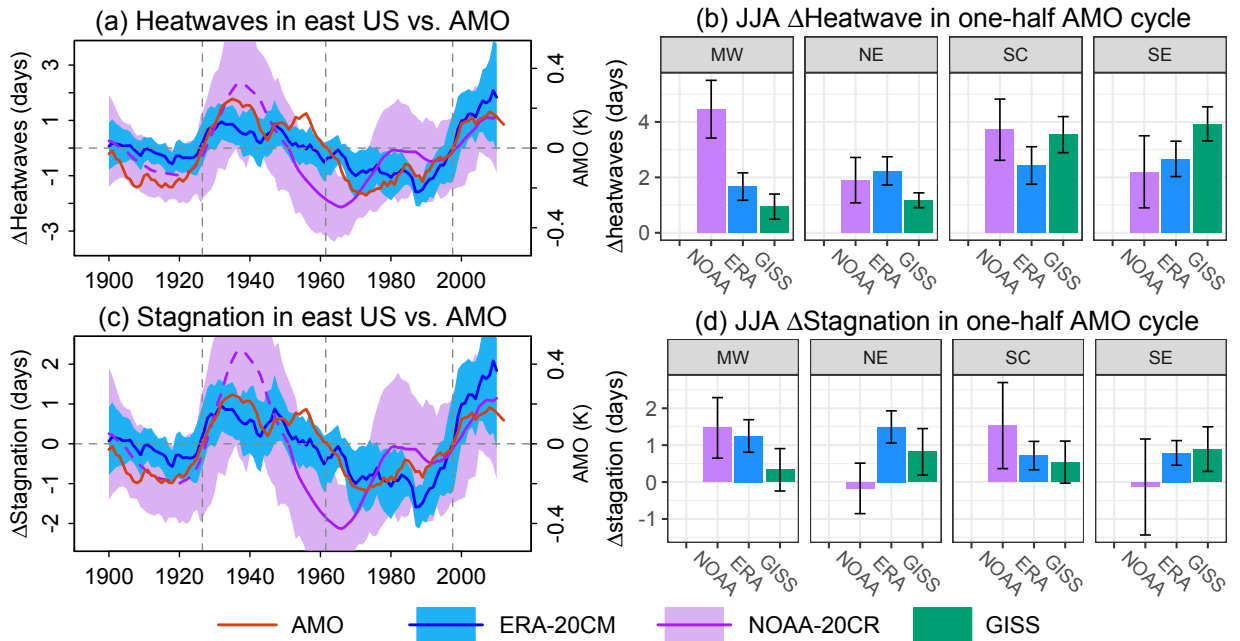
Figure S10. Similar as Figure 2c-d, but for heatwaves and number of stagnant days in the summertime.
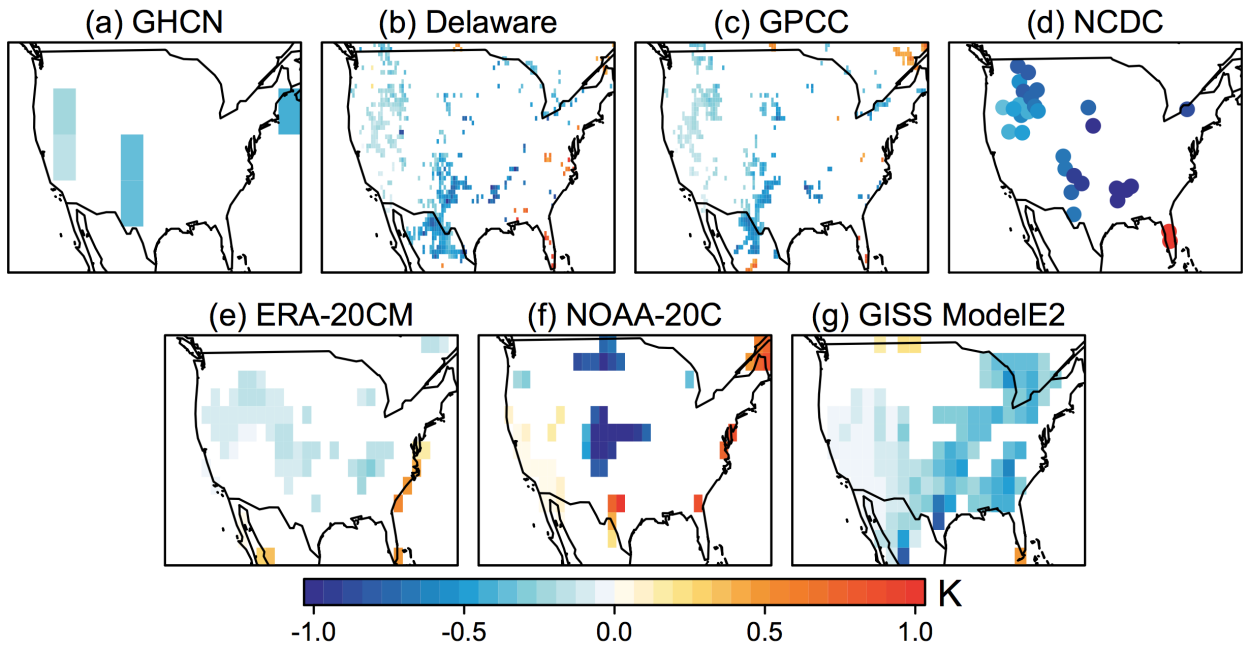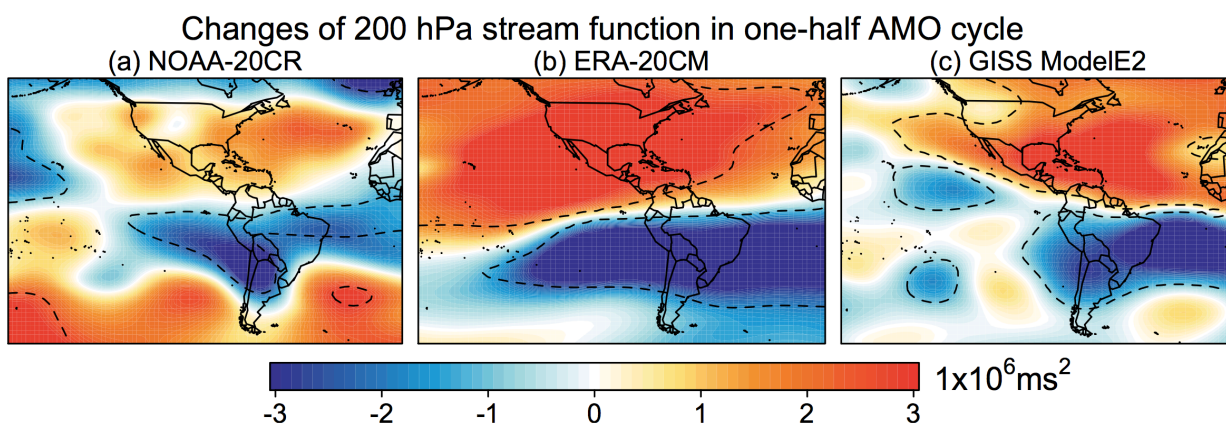
**Figure S11**. Same as Figure S8, but for mean JJA precipitation.

Figure S12. Same as Figure S8, but for JJA drought indices (Table S1).

Changes of SLPs in different datasets in one-half AMO cycle

(a) NOAA-20CR   (b) ERA-20CM   (c) GISS ModelE2


Pa

-100   -50   0   50   100

Figure S13. Changes of mean JJA SLPs in one-half AMO cycle from the cold to warm phase, as simulated by (a) NOAA-20CR, (b) ERA-20CM, and (c) GISS ModelE2. See Table S1 for descriptions of the model setups.

Changes of 200 hPa stream function in one-half AMO cycle

(a) NOAA-20CR   (b) ERA-20CM   (c) GISS ModelE2

$1 \times 10^{6} ms^{2}$

-3   -2   -1   0   1   2   3

Figure S14. Same as Figure S13, but for the mean JJA 200 hPa stream function.
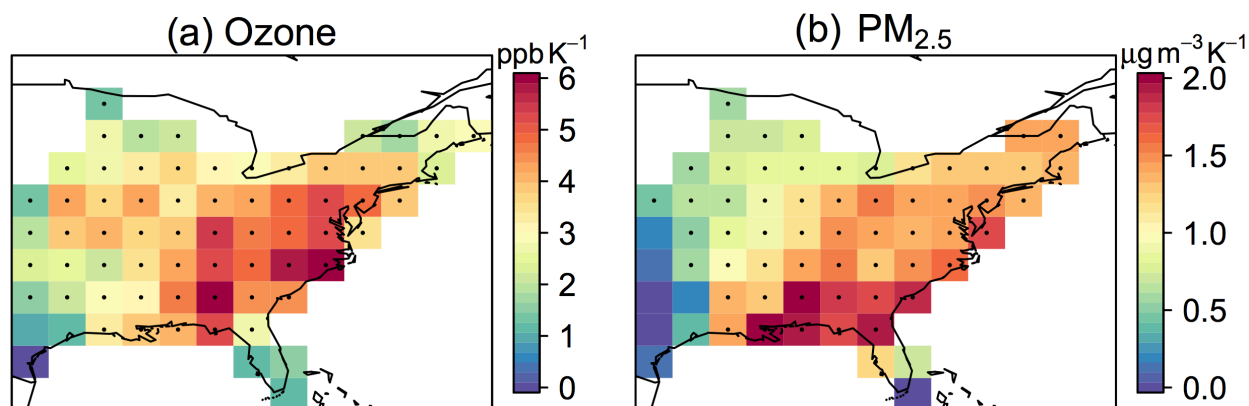
Figure S15. (a) Slopes of JJA seasonal mean MDA8 ozone with surface air temperatures in the eastern United States for 1980-2015. (b) Same as (a), but using monthly mean $PM_{2.5}$ concentrations for 1999-2015. In both panels, gridboxes with statistically significant ($p < 0.05$) correlations are stippled. All data are detrended by subtracting the 7-year moving averages.
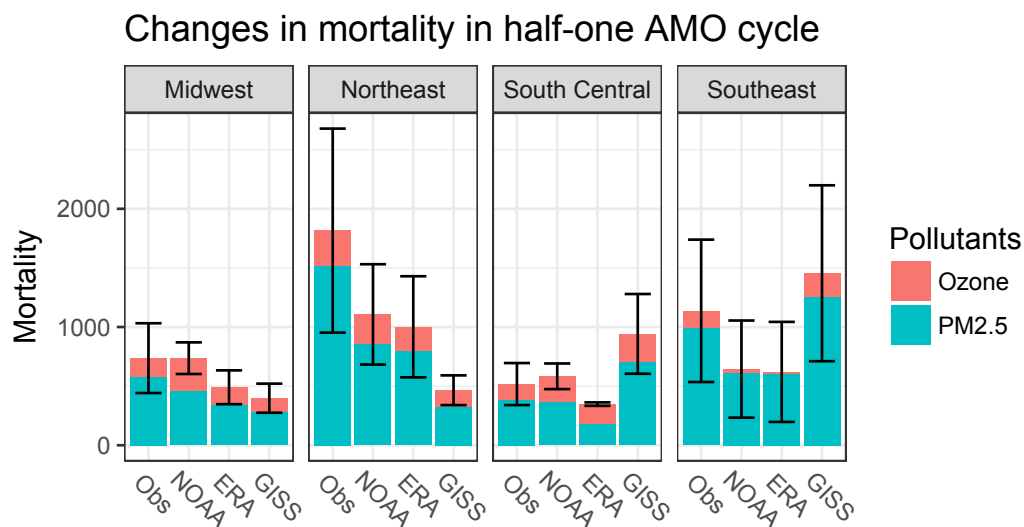
Figure S16. Excess mortality rates per summer due to air quality degradation in each U.S. region over one-half cycle of AMO from the cold to warm phase. Error bars denote one standard deviation around the median across the ensemble of model results. "Obs" signifies results obtained from observed meteorology, including MLOST temperatures, derived east-west and north-south windspeeds from HadSLP2 sea level pressures, and GPCC precipitation. Other datasets are defined in Table S1.
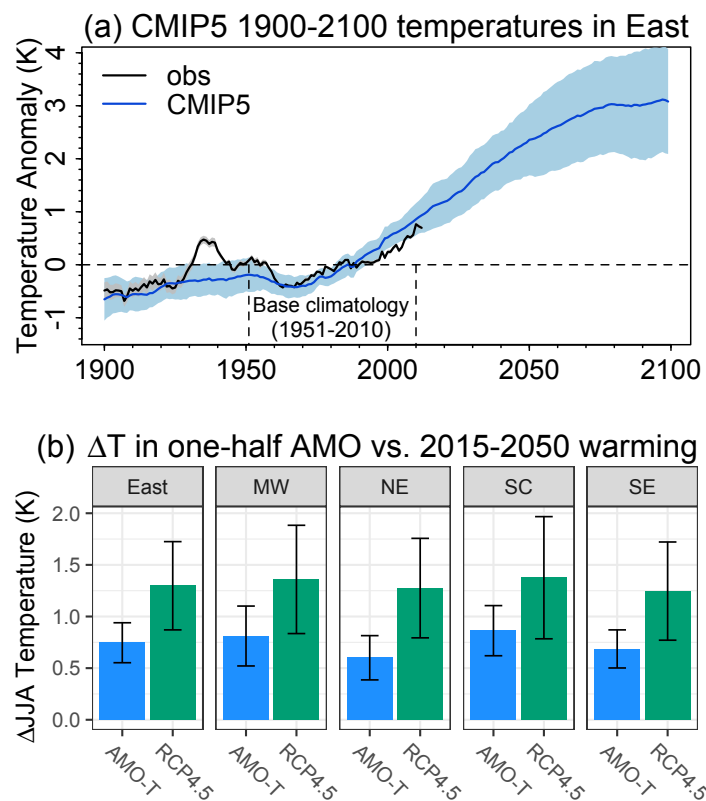
Figure S17. (a) Simulated 11-year running mean JJA temperature in the eastern United States from 1900 to 2100, relative to 1950-2010 time periods. The projected temperature from 2006 to 2100 follows the RCP4.5 scenario, as inferred from an ensemble of 33 CMIP5 models (Table S2). The observed 1900-2015 temperature anomalies are from NCDC and are shown as black line. (b) Increases of JJA temperature in the East in one-half AMO cycle (~35 years) and in CMIP5 projections in the East, Midwest (MW), Northeast (NE), South Central (SC), and Southeast (SE) over future decades (mean 2045-2055 vs. mean 2010-2020). The future climate projections are calculated as the median of an ensemble of 33 CMIP5 models. We use the first ensemble member for each model in Panel (a-b).