Journal of Geophysical Research - Atmospheres

Supporting Information for

Characterizing Changes in Eastern U.S. Pollution Events in a Warming World

A. M. Fiore[1,2,*], G.P. Milly[2], Laurel Quiñones[3†], Jared Bowden[4], Sarah E. Hancock[5], Erik Helstrom[2,‡] Jean-François Lamarque[6], Jordan Schnell[7&], Jason West[8], and Yangyang Xu[9]

[1]Department of Earth and Environmental Science, Columbia University, Palisades, NY, USA, [2]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA, [3]Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, USA, [4]Department of Applied Ecology, North Carolina State University, Raleigh, NC, USA, [5]Department of Computer Science, Columbia University, New York, NY, USA, [6]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA, [7]Department of Earth and Planetary Sciences and Institute for Sustainability and Energy at Northwestern University, Evanston, Illinois, USA, [8]Department of Environmental Sciences & Engineering, University of North Carolina, Chapel Hill, NC, [9]Department of Atmospheric Sciences, Texas A&M University, College Station, TX

Corresponding author: Arlene M. Fiore (amfiore@ldeo.columbia.edu)

*now at Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

†now at Department of Mechanical Engineering, Columbia University, New York, NY, USA

‡now at Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

&now at Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder; NOAA Global Systems Laboratory, Boulder, CO, USA

Contents of this file

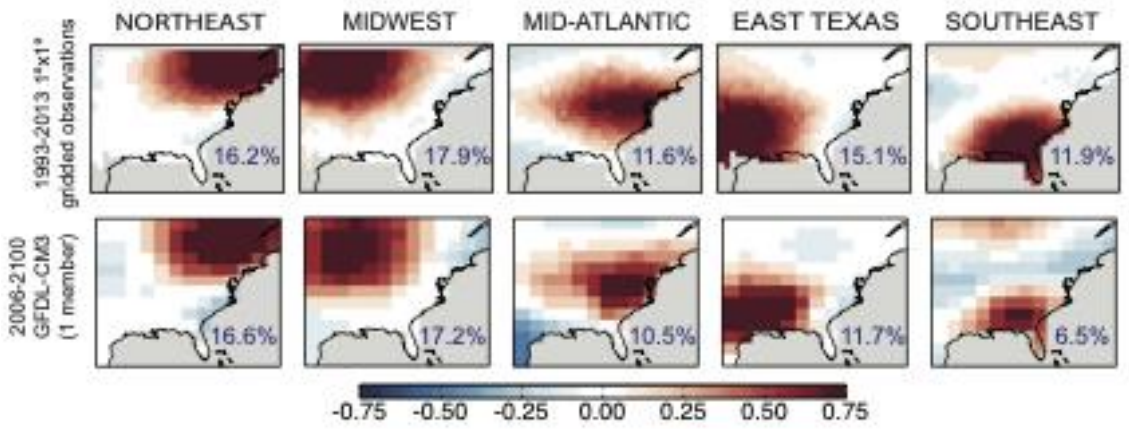

Figure S1a. Regions emerging from an EOF analysis on standardized anomalies of summertime daily maximum 8-hour average (MDA8) O_3 over the EUS. Shown are the EOF pattern loadings derived from (top) gridded observations, (bottom) one of three ensemble members in the GFDL-CM3 chemistry-climate model. Blue text indicates the total variance explained by each EOF.

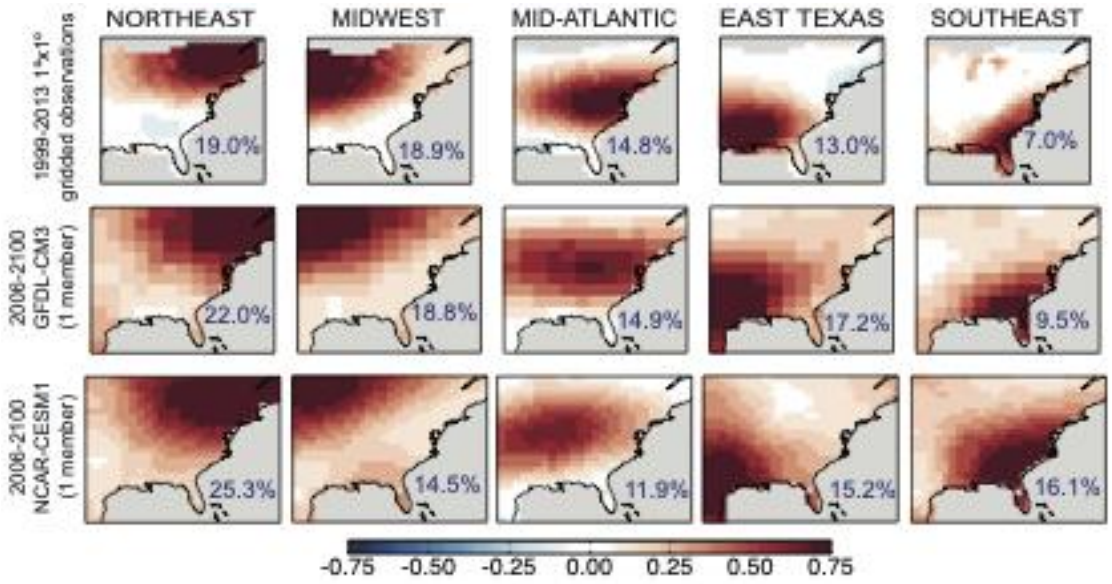

Figure S1b. Regions emerging from an EOF analysis on standardized anomalies of daily temperature over the EUS during summer (June-July-August). Shown are the EOF pattern loadings derived from (top) gridded observations of daily T_{max}, (middle row) daily T_{max} simulated by one of three ensemble members in the GFDL-CM3 chemistry-climate

model, (bottom) daily mean surface temperature simulated by one of 12 NCAR-CESM1 ensemble members. Blue text indicates the total variance explained by each EOF.

.

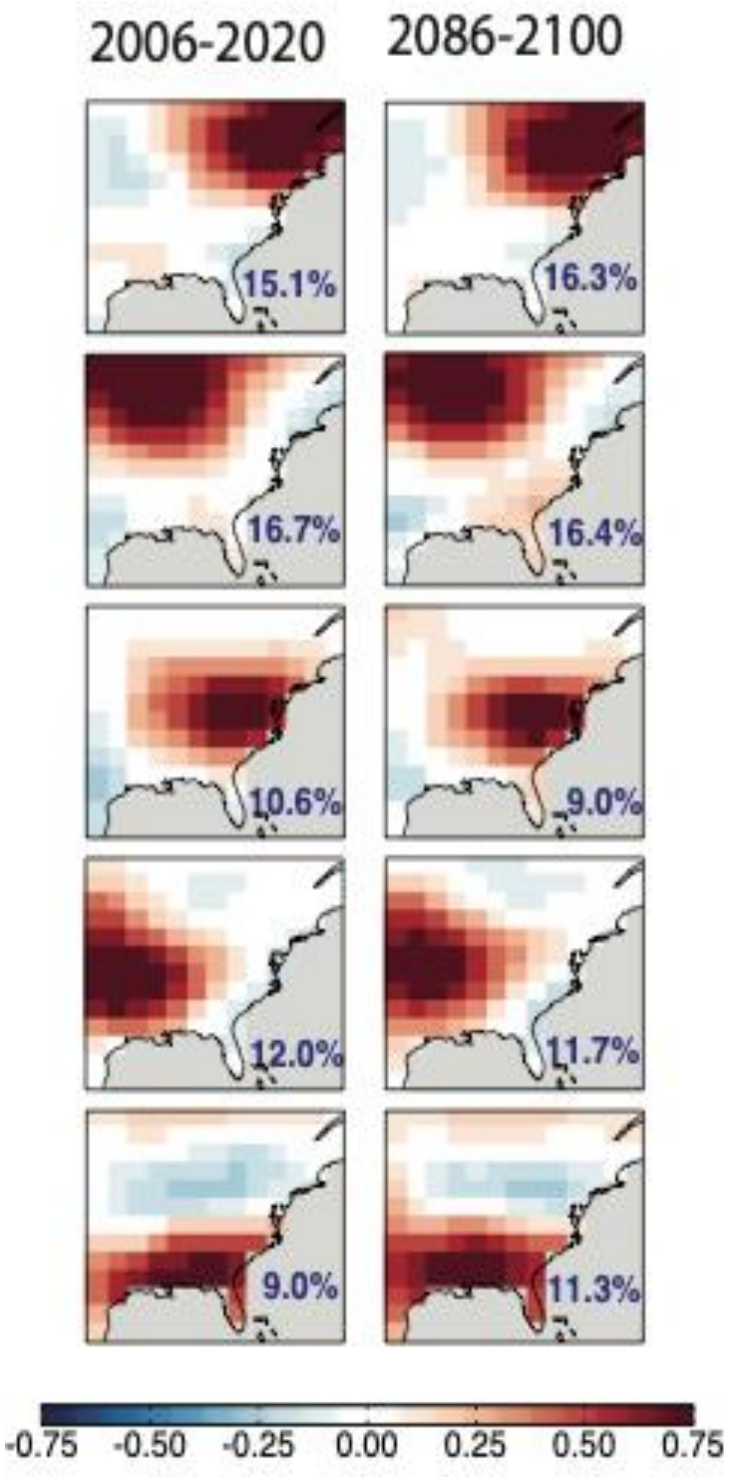

Figure S2. EOFs derived from daily mean $PM_{2.5}$ at the end of the century versus the beginning of the century in the GFDL-CM3 simulation (ensemble member 1).

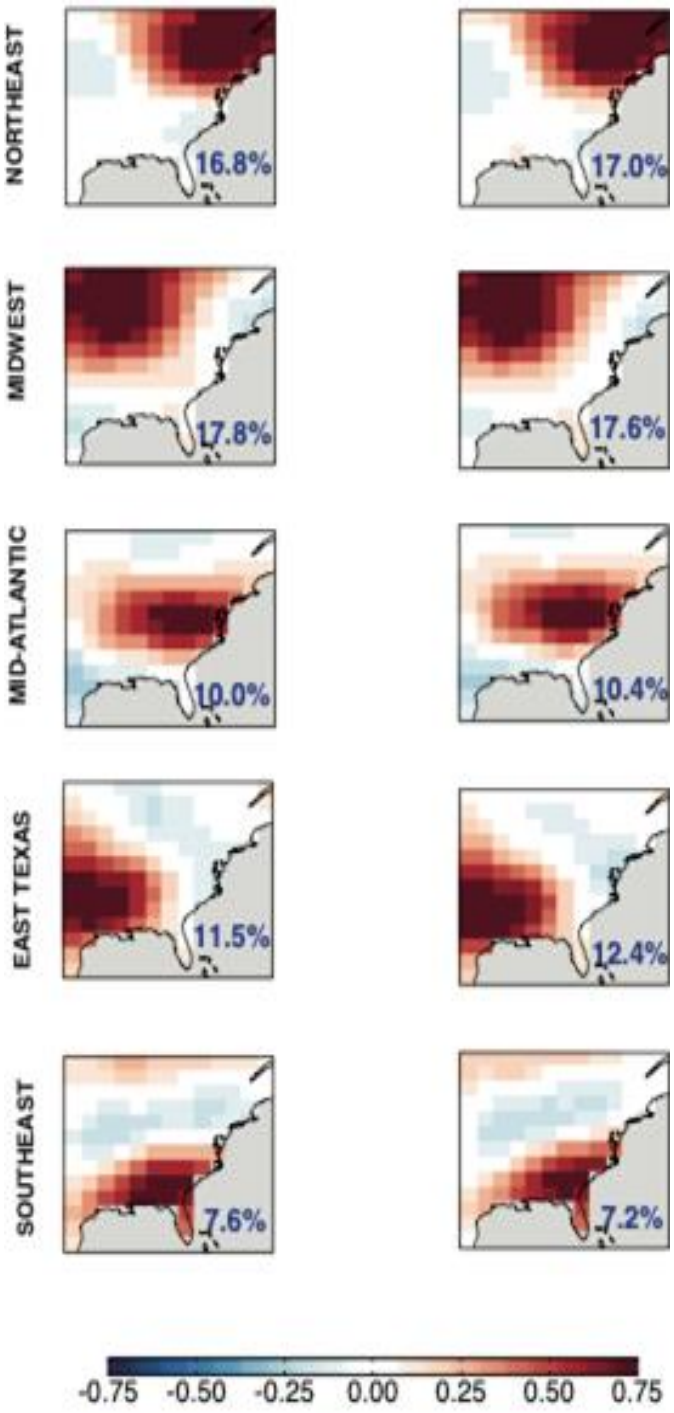

Figure S3a. EOFs derived from daily mean $PM_{2.5}$ in the two GFDL-CM3 ensemble members not shown in Figure 1. Blue text shows the percentage variance explained by each EOF.

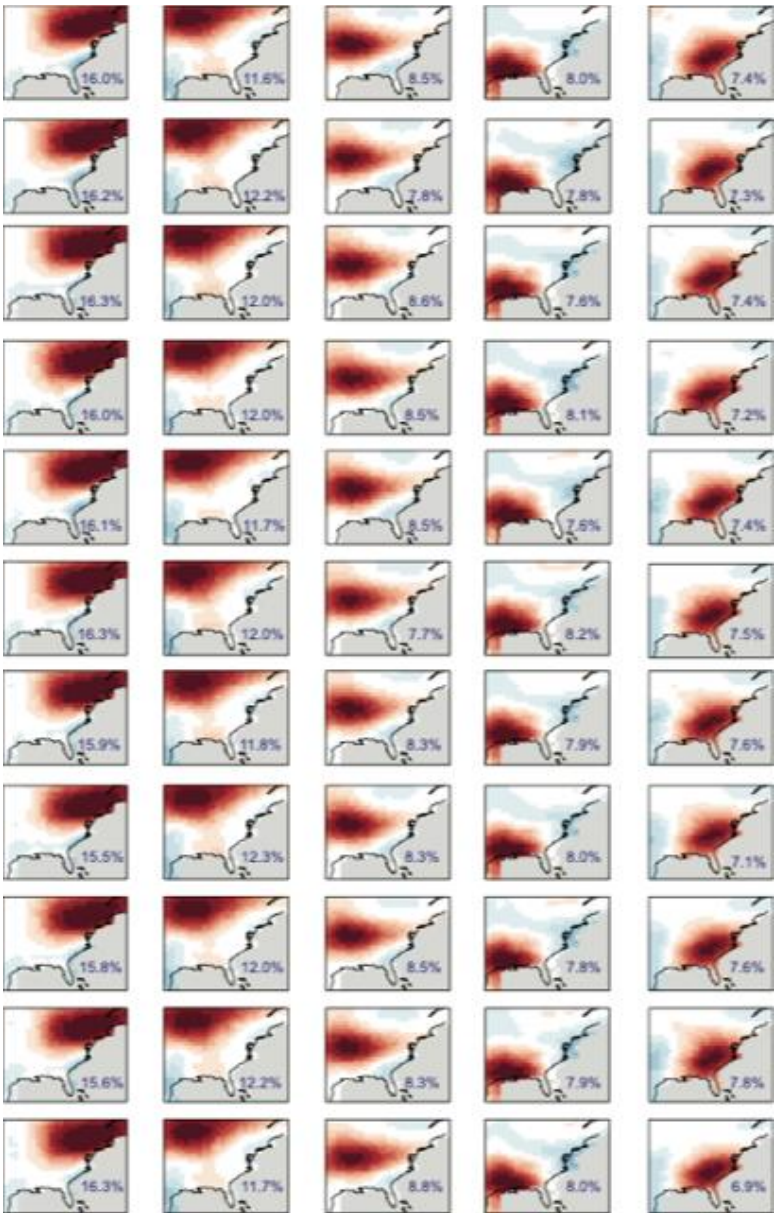

Figure S3b. EOFs derived from daily mean $PM_{2.5}$ in the 11 NCAR-CESM ensemble members not shown in Figure 1. Blue text shows the percentage variance explained by each EOF.

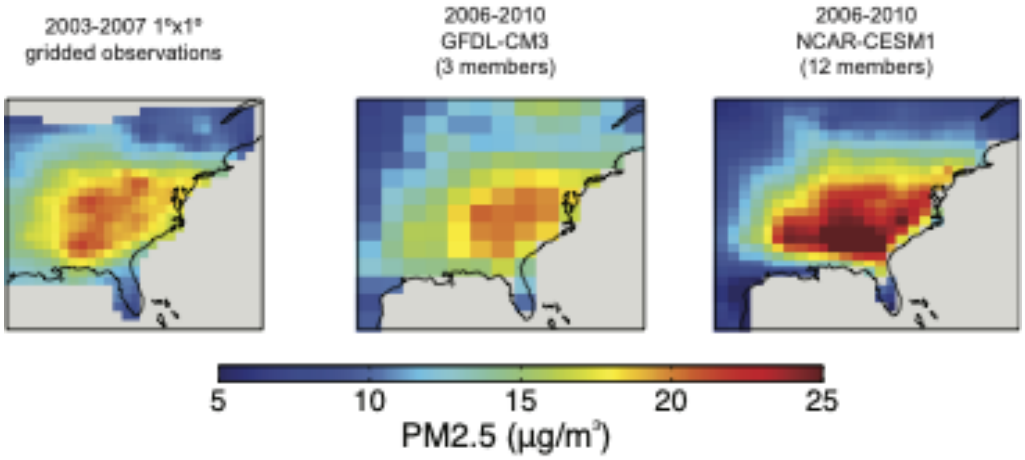

Figure S4a. Model evaluation of summertime ensemble mean $PM_{2.5.}$ with the gridded observational dataset used to derive the EOFs from daily data in Figure 1. Observations are averaged from 2003-2007, centered around 2005, the year for which emissions are perpetually repeated in the model, to avoid strong influence of trends driven by

anthropogenic emissions. By selecting the first five simulation years (2006-2010) for this comparison, we also minimize the influence of climate change.

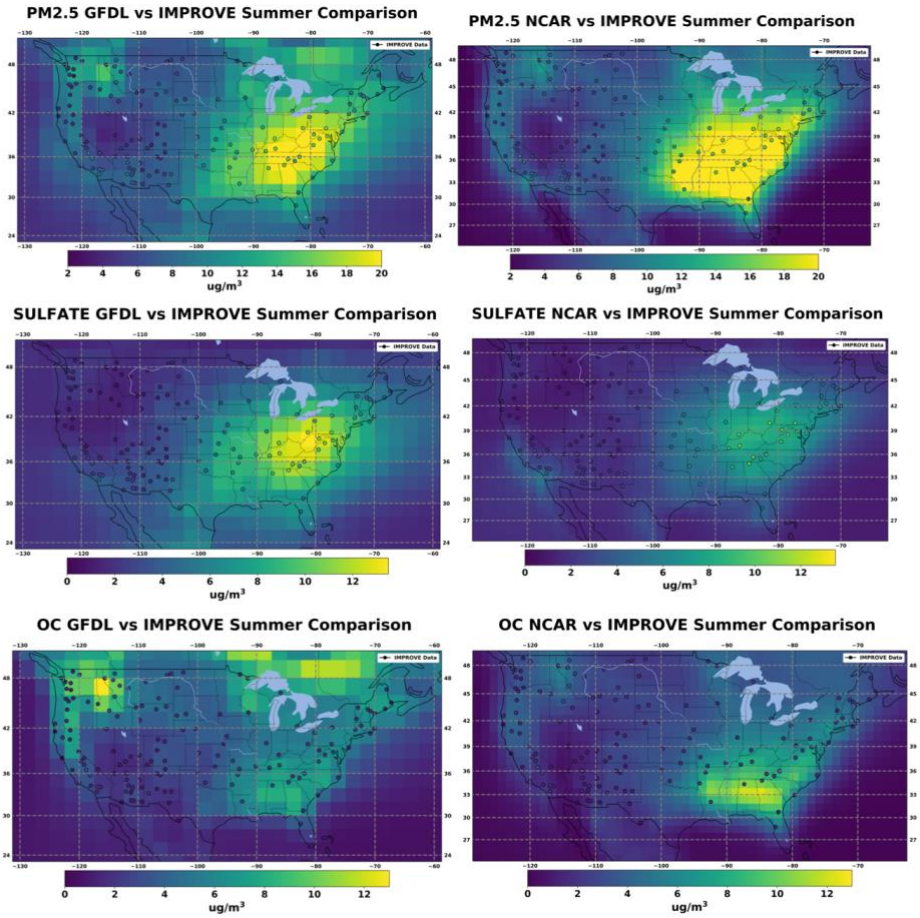

Figure S4b. Model evaluation of summertime ensemble mean $PM_{2.5.}$ with measurements from the IMPROVE network. Both models overestimate observed EUS $PM_{2.5}$ in summer (June-July-August) but differ in their simulation of individual components. Shown are summertime, ensemble mean surface PM2.5 (top), sulfate (middle), and organic carbon (bottom) in the GFDL-CM3 (left) and NCAR-CESM1 (right) chemistry-climate models averaged over 2006-2010 in the RCP8.5_WMGG scenario (Section 2). Filled circles show observations at the IMPROVE network, averaged over 2003-2007. The observed dataset centers around 2005, the year for which emissions are perpetually repeated in the model, to avoid strong influence of trends driven by anthropogenic emissions. By selecting the first five simulation years (2006-2010), we also minimize the influence of climate change.

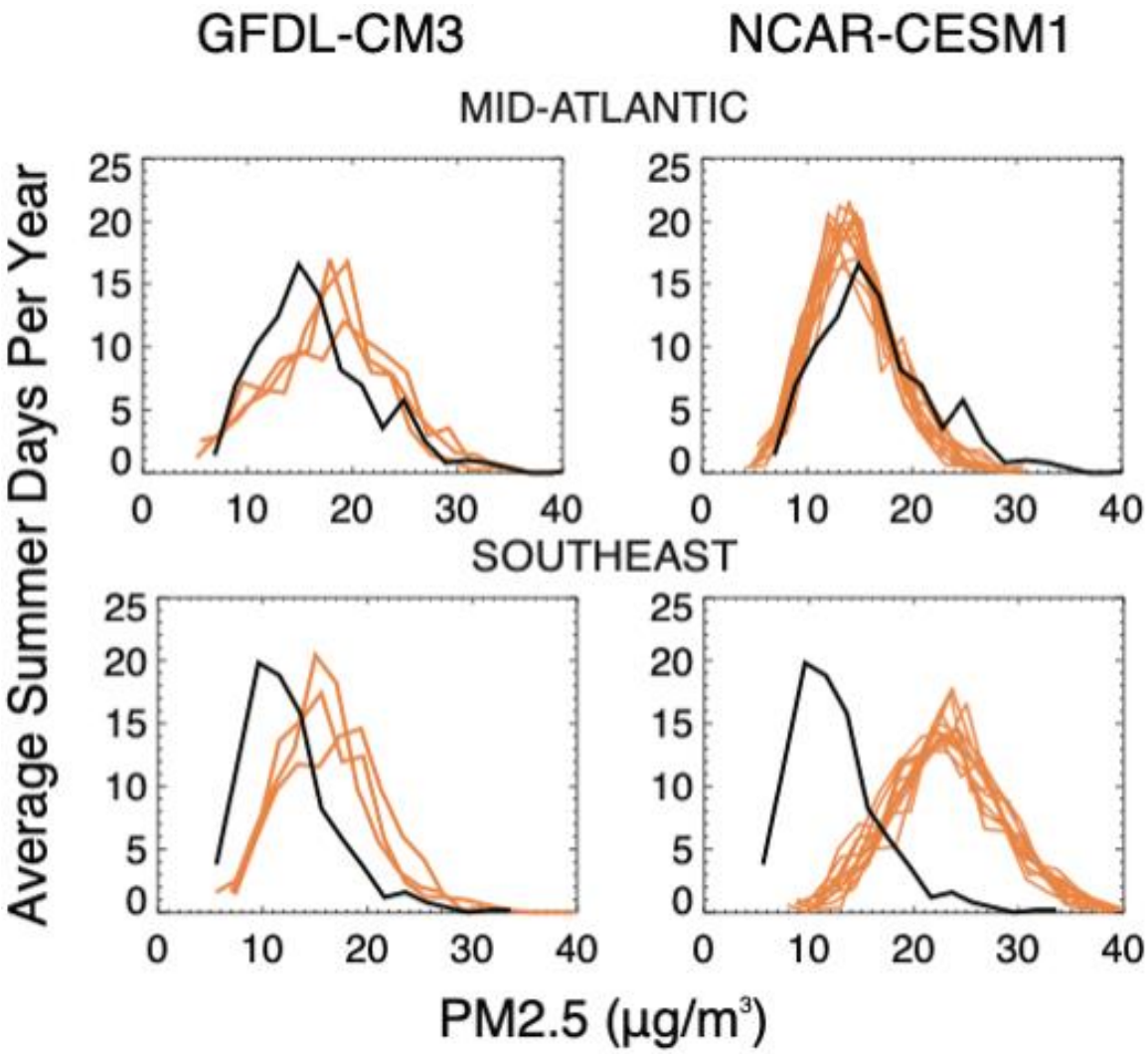

Figure S5. Average number of summer days with daily $PM_{2.5}$ falling within 2 µg m^{-3} concentration bins, regionally averaged (where EOF loading in Figure 1 exceeds > 0.5) in the observations (black) for the years 2003-2007 and in the individual ensemble members (orange) for GFDL-CM3 (left) and NCAR-CESM1 (right) for model years 2006-2010. Note that the mid-Atlantic EOFs derived from CESM1 and observations differ.

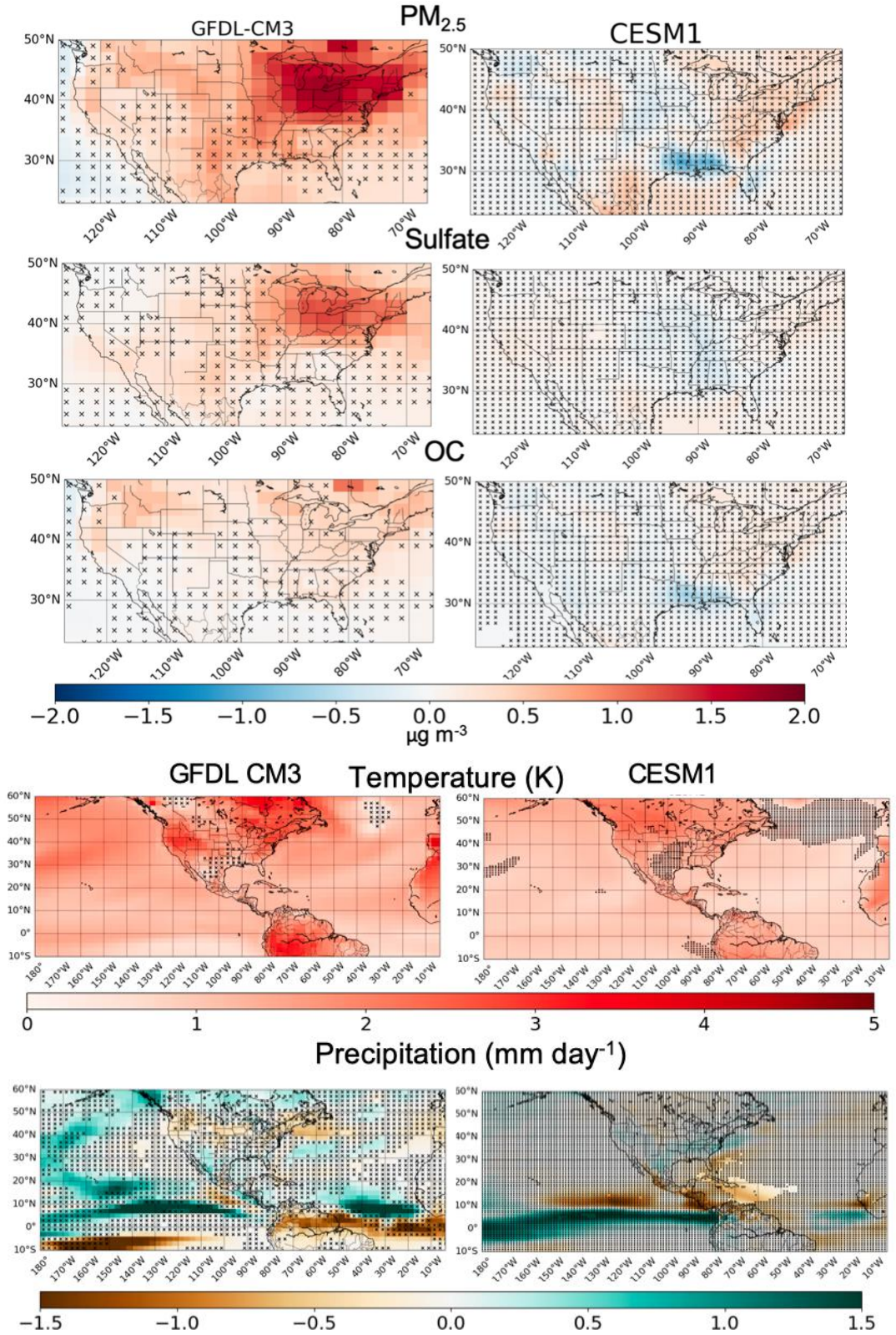

Figure S6. Change in summertime (June-July-August) $PM_{2.5}$, sulfate, organic carbon (OC), daily 2m air temperature (max for GFDL-CM3; mean for CESM1), and precipitation from

2006-2025 to 2041-2060 in the GFDL-CM3 (left; 3 ensemble members) and CESM1 (right; 12 ensemble members) under the RCP8.5_WMGG scenario. Grid cells marked with 'x' indicate that the ensemble mean change is smaller than the range of the changes simulated by individual ensemble members.

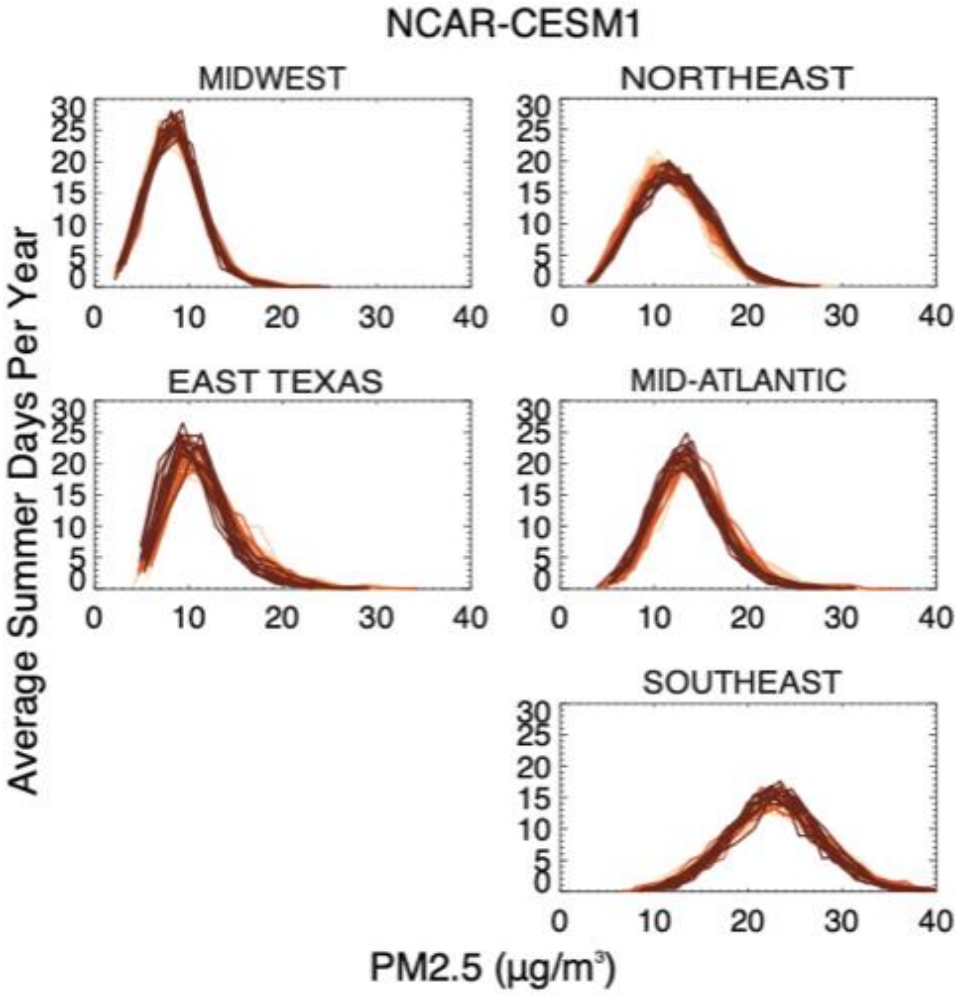

Figure S7. Little detectable change in the surface $PM_{2.5}$ distributions under the RCP8.5_WMGG scenario in the 12-member NCAR-CESM1 ensemble. Average number of summer days with daily $PM_{2.5}$ falling within 2 µg m^{-3} concentration bins, regionally averaged (where EOF loading > 0.5 in Figure 1) in each NCAR-CESM1 ensemble member for the years 2006-2015 (light), 2051-2060 (darker) and 2091-2100 (darkest).

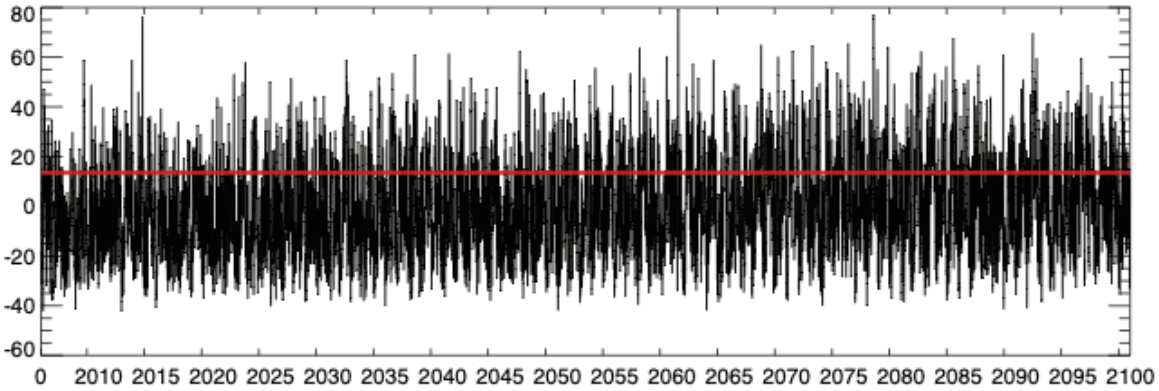

Figure S8. Principal component accompanying the Northeast EOF derived from the GFDL-CM3 model for the first of three ensemble members under the RCP8.5_WMGG scenario from 2006-2100. The red line indicates the 75th percentile.

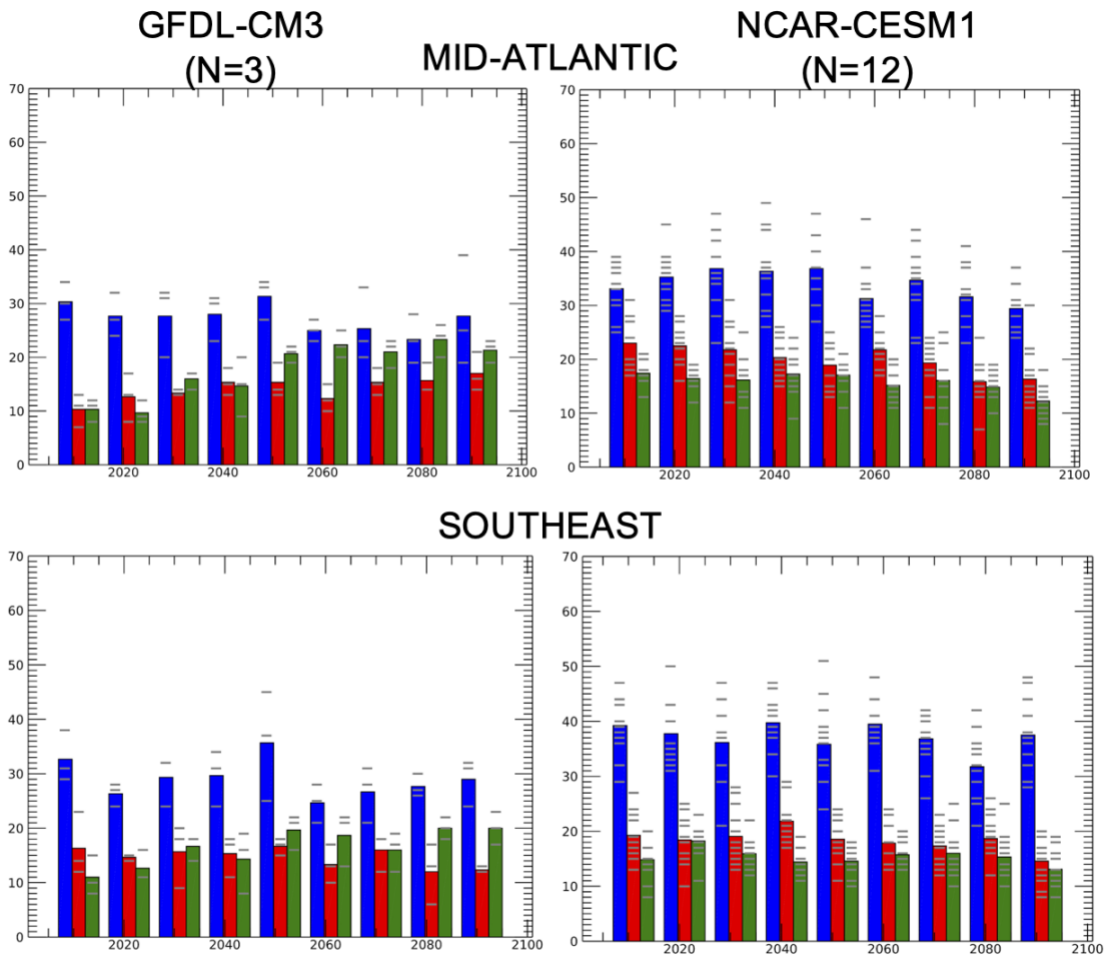

Figure S9. As in Figure 7 in the main text, but for the Mid-Atlantic and Southeast.

EOF	$PM_{2.5}$	MDA8 O_3	T_{max}
1	0.371	0.322	0.386
2	0.172	0.202	0.165
3	0.123	0.123	0.076
4	0.06	0.062	0.058
5	0.047	0.056	0.043
6	0.036	0.031	0.036
7	0.025	0.025	0.025
8	0.02	0.024	0.019
9	0.015	0.016	0.018
10	0.012	0.012	0.016

Table S1. Fraction of total variance explained by the first 10 raw EOFs (empirical orthogonal functions) over the EUS derived from the observational datasets of Schnell and Prather (2017).

EOF	Z1 PM	Z3 PM	Z5 PM	Z1 O3	Z3 O3	Z5 O3	Z1 T	Z3 T	Z5 T
1	0.251	0.243	0.254	0.226	0.226	0.235	0.582	0.601	0.605
2	0.158	0.162	0.163	0.158	0.155	0.154	0.117	0.105	0.111
3	0.102	0.102	0.1	0.1	0.1	0.098	0.055	0.052	0.054
4	0.074	0.074	0.073	0.085	0.085	0.084	0.039	0.039	0.037
5	0.058	0.057	0.055	0.056	0.055	0.055	0.033	0.032	0.03
6	0.045	0.048	0.046	0.039	0.039	0.039	0.021	0.019	0.02
7	0.037	0.038	0.037	0.034	0.034	0.034	0.016	0.016	0.014
8	0.028	0.028	0.028	0.028	0.028	0.028	0.013	0.012	0.012
9	0.024	0.024	0.025	0.026	0.025	0.026	0.011	0.011	0.01
10	0.022	0.023	0.022	0.02	0.02	0.02	0.009	0.01	0.01

Table S2. Fraction of total variance explained by the first 10 raw EOFs over the EUS derived from the GFDL-CM3 model in each individual ensemble member (denoted Z1, Z3, Z5) for surface $PM_{2.5}$ (PM), MDA8 O_3 (O_3) and daily maximum temperature (T) for the simulated years 2006-2100 under the RCP8.5_WMGG scenario. Ensemble member labels follow GFDL internal naming conventions.

EOF	E16	E17	E18	E19	E20	E21	E22	E25	E26	E27	E28	E30
1	0.22	0.22	0.22	0.22	0.22	0.22	0.22	0.22	0.21	0.22	0.22	0.22
2	0.11	0.10	0.10	0.11	0.10	0.10	0.10	0.10	0.11	0.10	0.11	0.10
3	0.07	0.08	0.07	0.07	0.07	0.08	0.08	0.08	0.08	0.08	0.08	0.08
4	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
5	0.04	0.05	0.05	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.05
6	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
7	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
8	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
9	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
10	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02

Table S3a. Fraction of total variance explained by the first 10 raw EUS EOFs derived from surface $PM_{2.5}$ in the 12 individual NCAR-CESM1 ensemble members. The ensemble number follows NCAR internal naming conventions.

EOF	E16	E17	E18	E19	E20	E21	E22	E25	E26	E27	E28	E30
1	0.59	0.58	0.58	0.59	0.56	0.57	0.57	0.58	0.57	0.57	0.57	0.57
2	0.11	0.12	0.11	0.11	0.12	0.11	0.11	0.11	0.12	0.12	0.11	0.12
3	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
4	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
5	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
6	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
7	0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02
8	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
9	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
10	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table S3b. As for Table S3a but for daily mean surface temperature (T).

EOF	$PM_{2.5}$	O_3	T_{max}
Northeast	0.18 (r=0.6)	0.13 (r=0.5)	0.73 (r=0.9)
Midwest	0.19 (r=0.5)	0.03 (r=0.2)	0.72 (r=0.9)
Mid-Atlantic	0.17 (r=0.4)	0.21 (r=0.6)	0.64 (r=0.8)
Texas-Gulf	0.08 (r=0.2)	-0.1 (r=0.3)	0.63 (r=0.8)
Southeast	0.08 (r=0.2)	-0.1 (r=0.1)	0.70 (r=0.8)

Table S4. Regression statistics (slopes in days per year and correlation (r)) for GFDL-CM3 ensemble mean trends in 21st century upper quartile events.

Region defined by EOF	Northeast	Midwest	East Texas	Southeast	Mid-Atlantic
GFDL-CM3 (3 ensemble members)					
Change from 2006-2015 to 2051-2060					
min	1.6	1.8	-0.1	0.6	1.2
mean	2.3	2.1	1.5	1.3	1.9
max	2.7	2.4	2.6	2.5	2.4
Change from 2006-2015 to 2091-2100					
min	3.4	2.3	0.1	-0.2	1.3
mean	3.7	2.8	0.8	0.9	2.2
max	4.0	3.3	1.7	2.6	2.9
CESM1 (12 ensemble members)					
Change from 2006-2015 to 2051-2060					
min	-0.1	-0.3	-2.2	-0.6	-1.1
mean	0.5	0.1	-0.4	-0.1	-0.4
max	1.1	0.3	0.5	0.9	0.6
3 lowest members	0.1	-0.1	-1.5	-0.5	-1.0
3 highest members	1.0	0.3	0.4	0.6	0.3
Change from 2006-2015 to 2091-2100					
min	0.3	-0.4	-2.4	-1.6	-1.8
mean	1.0	0.0	-1.6	-0.5	-1.1
max	1.5	0.5	0.0	1.1	-0.1
3 lowest members	0.5	-0.3	-2.3	-1.4	-1.6
3 highest members	1.4	0.4	-0.7	0.7	-0.7

Table S5. Changes in 75th percentile values of $PM_{2.5}$ concentrations ($\mu g\ m^{-3}$) within EOF-defined regions (loading > 0.5 in Figure 1) from the beginning (2006-2015) to the middle (2051-2060) or end (2091-2100) of the 21st century in the GFDL-CM3 and CESM1 models.

	T and PM: ENS MIN			T and PM: ENS MEAN			T and PM: ENS MAX		
REGION	Lag -1	Lag 0	Lag +1	Lag -1	Lag 0	Lag +1	Lag -1	Lag 0	Lag +1
Northeast	0.50	**0.55**	0.45	0.53	**0.58**	0.48	0.55	**0.60**	0.50
Mid-Atlantic	0.13	0.10	0.04	0.18	0.16	0.08	0.23	0.21	0.13
Upper Midwest	0.42	0.44	0.36	0.45	**0.46**	0.37	0.47	**0.49**	0.40
East Texas	-0.42	**-0.45**	-0.43	-0.39	-0.41	-0.39	-0.36	-0.38	-0.36
Southeast	0.15	0.18	0.16	0.24	0.26	0.24	0.27	0.30	0.27

Table S6. Ensemble minimum, mean, and maximum correlation coefficients (r) between principal component time series for temperature and PM simulated by the 12 NCAR-CESM1 ensemble members (T is daily mean temperature; PM is daily mean $PM_{2.5}$) within each region on the same day (Lag 0) or with temperature lagging (Lag -1) or leading (Lag +1) by a day relative to PM (*e.g.,* Lag -1 indicates that T lags PM by 1 day). Correlations are taken for each individual ensemble member prior to averaging. The strongest correlation for each pair of variables is shown in bold where $r \geq |0.45|$.

		O_3			PM		
Region 1	Region 2	Lag 0	Lag -1	Lag -2	Lag 0	Lag -1	Lag -2
Northeast	Mid-Atlantic	**0.50**	0.39	0.22	**0.53**	0.49	0.38
Northeast	Upper Midwest	0.44	**0.69**	0.52	0.51	**0.71**	0.61
Mid-Atlantic	Upper Midwest	0.58	**0.67**	0.54	0.67	**0.69**	0.57
Upper Midwest	East Texas	0.20	0.29	0.24	0.39	**0.48**	0.43
Mid-Atlantic	East Texas	0.04	0.16	0.18	0.38	**0.46**	0.44
Southeast	East Texas	**0.46**	0.38	0.20	**0.62**	0.56	0.44

Table S7. Ensemble mean correlation coefficients (r) between principal component time series in Region 1 versus Region 2 for O_3 or PM (O_3 is MDA8 O_3; PM is daily mean $PM_{2.5}$) simulated by the GFDL-CM3 model on the same day (Lag 0) or with Region 1 lagging by one (Lag -1) or two (Lag -2) days relative to Region 2. Correlations are taken for each

individual ensemble member prior to averaging. The strongest correlation for each pair of regions is shown in bold where $r \geq 0.4$.