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

Arlene M. Fiore[1], George P. Milly[1], Laurel Quiñones[2], Jared Bowden[3], Sarah E. Hancock[2], Erik Helstrom[4], Jean-Francois Lamarque[5], Jordan Schnell[6], J. Jason West[7], and Yangyang Xu[8]

[1]Lamont-Doherty Earth Observatory of Columbia University
[2]Columbia University
[3]North Carolina State University
[4]MIT
[5]NCAR
[6]Cooperative Institute for Research in Environmental Sciences at the University of Colorado Boulder NOAA/Global Systems Laboratory
[7]University of North Carolina at Chapel Hill
[8]Texas A&M University

November 22, 2022

Abstract

Risk assessments of air pollution impacts on human health and ecosystems would ideally consider a broad set of climate and emission scenarios and the role of natural internal climate variability within a single scenario. We analyze initial condition chemistry-climate ensembles to gauge the significance of greenhouse-gas-induced air pollution changes relative to internal climate variability, and response differences in two models. To quantify the effects of climate change on the frequency and duration of summertime regional-scale pollution episodes over the Eastern United States (EUS), we apply an Empirical Orthogonal Function (EOF) analysis to a 3-member GFDL-CM3 ensemble with prognostic ozone and aerosols and a 12-member NCAR-CESM1 ensemble with prognostic aerosols under a 21st century RCP8.5 scenario with air pollutant emissions frozen in 2005. Correlations between GFDL-CM3 principal components for ozone, $PM_{2.5}$ and temperature represent spatiotemporal relationships discerned previously from observational analysis. Over the Northeast region, both models simulate summertime surface temperature increases of over 5 °C from 2006-2025 to 2081-2100 and $PM_{2.5}$ of up to 1-4 µg m^{-3}. The ensemble average decadal incidence of upper quartile Northeast $PM_{2.5}$ events lasting at least five days doubles in GFDL-CM3 and increases >50% in NCAR-CESM1. In other EUS regions, inter-model differences in $PM_{2.5}$ responses to climate change cannot be explained by internal climate variability. Our EOF-based approach anticipates future opportunities to data-mine initial condition chemistry-climate model ensembles for probabilistic assessments of changing frequency and duration of regional-scale pollution and heat events while obviating the need to bias-correct concentration-based thresholds separately in individual models.

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, USA, [9]Department of Atmospheric Sciences, Texas A&M University, College Station, TX, USA

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

Key Points:

- Frequency and duration of Northeast US pollution events increase along with heat events under a high-warming scenario.
- EOF approach enables rapid assessment of regional-scale changes in pollution events without needing to bias correct models individually.
- Larger uncertainty in EUS $PM_{2.5}$ from different model responses to climate change than from climate variability.

Abstract

Risk assessments of air pollution impacts on human health and ecosystems would ideally consider a broad set of climate and emission scenarios and the role of natural internal climate variability within a single scenario. We analyze initial condition chemistry-climate ensembles to gauge the significance of greenhouse-gas-induced air pollution changes relative to internal climate variability, and response differences in two models. To quantify the effects of climate change on the frequency and duration of summertime regional-scale pollution episodes over the Eastern United States (EUS), we apply an Empirical Orthogonal Function (EOF) analysis to a 3-member GFDL-CM3 ensemble with prognostic ozone and aerosols and a 12-member NCAR-CESM1 ensemble with prognostic aerosols under a 21st century RCP8.5 scenario with air pollutant emissions frozen in 2005. Correlations between GFDL-CM3 principal components for ozone, $PM_{2.5}$ and temperature represent spatiotemporal relationships discerned previously from observational analysis. Over the Northeast region, both models simulate summertime surface temperature increases of over 5 °C from 2006-2025 to 2081-2100 and $PM_{2.5}$ of up to 1-4 μg m^{-3}. The ensemble average decadal incidence of upper quartile Northeast $PM_{2.5}$ events lasting at least five days doubles in GFDL-CM3 and increases >50% in NCAR-CESM1. In other EUS regions, inter-model differences in $PM_{2.5}$ responses to climate change cannot be explained by internal climate variability. Our EOF-based approach anticipates future opportunities to data-mine initial condition chemistry-climate model ensembles for probabilistic assessments of changing frequency and duration of regional-scale pollution and heat events while obviating the need to bias-correct concentration-based thresholds separately in individual models.

Plain Language Summary

Prior studies conclude climate change will worsen air quality in some polluted regions but typically neglected the role of climate variability. Uncertainty also arises from differences in climate model responses to a given anthropogenic forcing scenario. Differentiating the relative contributions of these uncertainties (structural versus stochastic) to inter-model differences in projected air pollution responses to climate change is becoming possible with initial-condition climate model ensembles. We analyze day-by-day variations in pollutant levels over five eastern U.S. region to quantify changes in frequency and duration of regional-scale high pollution and heat events with small initial-condition ensembles from two different models. Under a 21st century climate change scenario in which air pollutant emissions are fixed at 2005 levels, our analysis shows longer-lasting and more frequent Northeast U.S. $PM_{2.5}$ (and heat) episodes, which could exacerbate public health burdens. Projecting changes in other EUS regions is limited by inter-model differences that exceed the uncertainty attributable to climate variability. While our ensembles are small relative to those generated most recently with physical climate models, our findings add to a growing recognition that climate variability complicates the detection and attribution of observed and simulated air pollution trends under climate change scenarios.

1 Introduction

High ground-level concentrations of the top two U.S. air pollutants, fine particles ($PM_{2.5}$) and ozone (O_3) sometimes co-occur along with high temperatures across the eastern U.S.A. (EUS) during summer, with >50% same-day coincidence of at least two of these extremes in the Northeast (Schnell & Prather, 2017) and generally about one-third coincidence in the highest O_3 and temperature events (Phalitnonkiat et al., 2018). Air pollution health burdens in other mid-latitude regions have also been found to increase during heat waves (Filleul et al., 2006; García-Herrera et al., 2010; Shaposhnikov et al., 2014), although it is unknown if prolonged versus intermittent exposure to high pollution events elicit different human health responses. Future increases in intensity and frequency of heat stress events are expected (Coffel et al., 2017), raising the possibility that climate change will also exacerbate air pollution and associated adverse health outcomes. Here, we describe an approach to characterize changes in frequency and duration of high pollution and heat events in simulations of 21st century climate change, with a primary focus on $PM_{2.5}$, available from two models, and a secondary focus on the co-occurrence of high $PM_{2.5}$, O_3, and temperature events.

Prior studies identified changes in the severity, duration and spatial extent of U.S. air pollution events under future climate scenarios (Mickley et al., 2004; Rieder et al., 2015; Schnell et al., 2016; S. Wu et al., 2008). Compound extreme weather events such as simultaneous occurrence of air stagnation and heat waves, which are likely to affect air pollution, are projected to increase by mid-to-late century (J. Zhang et al., 2018). Xu et al. (2020) showed a ten-fold increase in the co-occurrence of heatwaves and high $PM_{2.5}$ events by mid-21st century. Air pollution has long been observed to co-vary with meteorology on hourly to interannual time scales (*e.g.* Camalier et al., 2007; Dawson et al., 2013; Kerr et al., 2019; Leibensperger et al., 2008; Lin et al., 2001; Logan, 1989; Rao et al., 1995; Tai et al., 2010; Vukovich, 1995), with an emphasis on air stagnation, temperature inversions, heat waves, and wildfires responding to heat and drought as drivers of the most extreme pollution events (Hong et al., 2019; Horton et al., 2012; Horton et al., 2014; Hou & Wu, 2016; Konovalov et al., 2011; Porter & Heald, 2019; Porter et al., 2015; Shen et al., 2016; Spracklen et al., 2009; Sun et al., 2017; Wang & Angell, 1999). Other work indicates that local observed meteorology-pollutant relationships are strongly shaped by the underlying atmospheric dynamics that control synoptic transport (Barnes & Fiore, 2013; Kerr et al., 2020; Kerr et al., 2019; Oswald et al., 2015; Previdi & Fiore, 2019; Sun et al., 2019; Tai et al., 2012). Overall, a wide range of modeling systems project that climate change will degrade air quality in some currently polluted U.S. regions, although models disagree as to the regional extent and magnitude of projected air pollution changes (*e.g.,* Fiore et al., 2015; Fu & Tian, 2019; Jacob & Winner, 2009; Kirtman et al., 2013; Nolte et al., 2018; Schnell et al., 2016; Weaver et al., 2009).

Some of the inter-model disagreement in the published literature likely reflects a lack of separation of forced climate change (*i.e.*, "signal" due to rising greenhouse gases plus aerosols) from internal variability (*i.e.,* climate "noise" due to natural processes within the climate system) (Deser et al., 2020; East & Garcia-Menendez, 2020; Garcia-Menendez et al., 2017). Computational limitations restricted the length and number of simulations for most prior model projections of future changes in air pollution (Fiore et al., 2015; Fiore et al., 2012; Jacob & Winner, 2009; Weaver et al., 2009). Prior analysis of initial condition ensembles within a single climate model has demonstrated a major role for internal climate variability, measured by the inter-ensemble range, in shaping the near-term regional meteorological trends (Deser et al., 2012ab) to which air pollution will respond. Each ensemble member is one possible future air pollution response to the same forcing scenario, such that with sufficiently large ensembles, statistics can be developed to quantify the probability of 'rare' events in the observed record. Extracting signals of climate change is particularly challenging for extreme quantities. Advances in computational power now permit large ensemble simulations with physical climate models (Deser et al., 2012ab; Deser et al., 2013; Kay et al., 2015), where each ensemble member has different initial conditions but otherwise is forced by the same greenhouse gas and aerosol emission scenarios. The range across individual ensemble members offers a measure of the noise associated with internal climate variability, while the ensemble mean provides an estimate of the forced signal.

Schnell et al. (2014; 2015) have previously concluded that coarse resolution global models capture the observed spatial extent and timing of large-scale O_3 episodes, providing a strong basis for our analysis of air pollution simulated by global climate models. Challenges to quantifying simulated changes in high pollution events include selecting an appropriate threshold and accounting for model biases that may require adjusting the model threshold to ensure a similar frequency of high events as observed. Separate adjustments may be needed not only within each individual model (e.g., as in Horton et al. (2012), but also each region of interest (Schnell et al., 2015; Turnock et al., 2020).

Here, we examine changes in the frequency and duration of high pollution events over five distinct EUS regions. We expand upon Eder et al. (1993), who first applied Empirical Orthogonal Function (EOF) analysis to identify EUS regions in which ground-level ozone is high or low simultaneously across the region. This statistical approach avoids the pervasive problem of identifying relevant model thresholds in the presence of model biases by instead targeting model skill at representing the underlying patterns of spatiotemporal variability. We probe the role of natural climate variability, which arises internally within the climate system, as represented by two chemistry-climate models with interactive aerosol simulations. We also consider co-variations in high $PM_{2.5}$, O_3, and temperature events in one model with full tropospheric chemistry, and compare to observed relationships. The approach described below can be applied to rapidly gauge changing air pollution events as simulated by future large initial condition climate model ensembles that include full tropospheric (gas-phase plus aerosol) chemistry.

2 Data and Methods

2.1 Models and Observations

Our analysis centers on an existing 3-member ensemble generated with the GFDL-CM3 chemistry-climate model to project air pollution during the 21st century under a high warming scenario. We refer to this scenario as "RCP8.5_WMGG" in which Well-Mixed Greenhouse Gases (WMGG) follow the RCP8.5 scenario. Both particulate matter (PM) and ozone precursor emissions are held fixed at 2005 levels as described by Clifton et al. (2014). The simulated warming is less than in the standard RCP8.5 scenario in which aerosol and precursor emissions decline. The GFDL-CM3 model includes fully coupled ocean-atmosphere-sea ice-dynamic vegetation land models, and stratospheric and tropospheric gas-phase chemistry and aerosols (Austin et al., 2013; Donner et al., 2011; Naik et al., 2013). The native model resolution is a c48 cubed sphere which is post-processed to a 2°x2° horizontal grid. All RCP8.5_WMGG ensemble members are identical except for their initial conditions, which are taken from the final day of a corresponding transient 1860-2005 historical simulation. Each historical ensemble member was launched using initial conditions sampled at 50-year intervals in a "pre-industrial control" simulation that perpetually repeats 1860 greenhouse gas, aerosol, air pollutant emissions and other forcings. RCP8.5_WMGG simulations use the same monthly-varying dry deposition and isoprene, soil NO_x and biomass burning emissions every year. Diurnal cycles are imposed for isoprene emissions and ozone dry deposition. Wet deposition and sources of lightning NO_x, dimethyl sulfide (DMS), marine organic aerosol, sea salt and dust are coupled to the simulated meteorology and thus respond to changes in climate (Naik et al., 2013). The simulations neglect feedbacks to air pollution through wildland fires (Abatzoglou & Williams, 2016; Spracklen et al., 2009) as well as changes in terrestrial biogenic emissions or dry deposition (Andersson & Engardt, 2010). These idealized simulations enable us to isolate the influence of rising well-mixed greenhouse gases on pollution events, mainly by changing the meteorology.

Hourly surface ozone, daily maximum temperature at 2m reference height (T_{max}), daily surface $PM_{2.5}$, and monthly chemical components of $PM_{2.5}$ were archived from the lowermost atmospheric layer of all GFDL-CM3 simulations. The $PM_{2.5}$ diagnostic includes sulfate (assumed to be ammonium sulfate), carbonaceous aerosol (organic matter, black carbon, and secondary organic aerosol), the smallest size bin (of five) for dust, and the smallest two size bins (of five) for sea salt. We calculate maximum daily 8-hour average (MDA8) ozone from the hourly ozone fields.

We draw on a 12-member ensemble with the CESM1 climate model generated at NCAR to provide additional context for the changes in high-$PM_{2.5}$ and temperature events diagnosed with the GFDL-CM3 ensemble. As described by Xu and Lamarque (2018), this coupled atmosphere-ocean-sea ice-land model at 1°x1° horizontal resolution includes an interactive aerosol scheme with three internally mixed modes (Ghan et al., 2012; Liu et al., 2012). As for GFDL-CM3, the NCAR-CESM1 simulations hold aerosol and precursor emissions fixed at 2005 levels, as well as the oxidant fields used to drive secondary aerosol formation, but greenhouse gas concentrations rise along the RCP8.5 scenario from 2006 to 2100. In contrast to GFDL-CM3, the CESM1 configuration does not include the fully interactive tropospheric chemistry needed to simulate changes in oxidants. Each NCAR-CESM1 ensemble member is configured identically except for a tiny perturbation ($O(10^{-14})$ K) imposed in the atmospheric temperature initial condition fields (Kay et al., 2015; Xu & Lamarque, 2018). Dust and sea salt emissions respond to meteorology and land surface conditions, while biogenic VOC emissions are held constant (Lamarque et al., 2011). $PM_{2.5}$ is defined as the sum of daily mean sulfate, dust, black carbon, and primary and secondary organic aerosol in the Aitken and accumulation mode in the lowermost atmospheric layer, which we convert from the native model mass mixing ratio (kg/kg) to mass density (ug/m^3). We also use daily mean temperatures at the surface and at 2m reference height from these simulations.

To evaluate simulated EUS spatiotemporal patterns in air pollution, we use observations of near-surface daily mean $PM_{2.5}$ and MDA8 ozone measured at U.S. and Canadian ground-based networks that were optimally interpolated to a 1ºx1º grid over the EUS (Schnell et al., 2014; Schnell & Prather, 2017). These gridded datasets are available for 1999-2013 and 1993-2013 for $PM_{2.5}$ and ozone, respectively. We also use the 1ºx1º temperature fields that Schnell and Prather (2017) regridded from the 0.5°×0.5° European Centre for Medium-Range Weather Forecasting (ECMWF) Interim reanalysis maximum daily 6-hourly temperatures sampled at 2 m reference height.

2.2 Empirical Orthogonal Function (EOF) analysis

We analyze daily $PM_{2.5}$, ozone, and temperature data during summer (June-July-August). We focus on summer, the season when ozone is highest, because we are interested in co-occurrence of ozone and $PM_{2.5}$, which we examine in the GFDL-CM3 model (Section 5). Before conducting Empirical Orthogonal Function (EOF) analysis, we standardize all data, separately for each grid cell, by removing the mean of the entire time series and dividing by the standard deviation. The EOFs are the eigenvectors of the covariance matrix derived from the data matrix (dimensioned space by time). Each EOF is a spatial loading pattern for a mode of spatiotemporal variability that identifies where air pollution or temperature varies coherently (*i.e.*, polluted/clean air and hot/cold temperatures occur across the region indicated by the EOF at the same time). For some of the analysis below we use the EOFs to define a regional mask where the EOF loading exceeds 0.5.

The first five EOFs derived from the $PM_{2.5}$ observations capture 77% of the variance in daily summertime $PM_{2.5}$ concentrations over the EUS (Figure 1). The EOFs derived from the observed MDA8 O_3 and daily maximum temperature datasets capture 77% and 73% of the total variance, respectively. We select five EOFs for Varimax rotation after considering a change point for the amount of variance explained by each successive EOF derived from observations (Wilks, 1995). Table S1 lists the variance explained by the first ten EOFs. We retain the first five EOFs across all variables. The EOF analysis thus reduces the dataset size for further temporal analysis, from the number of individual grid cells (424, 447, and 424 from the gridded $PM_{2.5}$, ozone, and temperature observations, respectively) to five EUS regions. Below we refer to the EOFs by the region names shown in Figure 1.

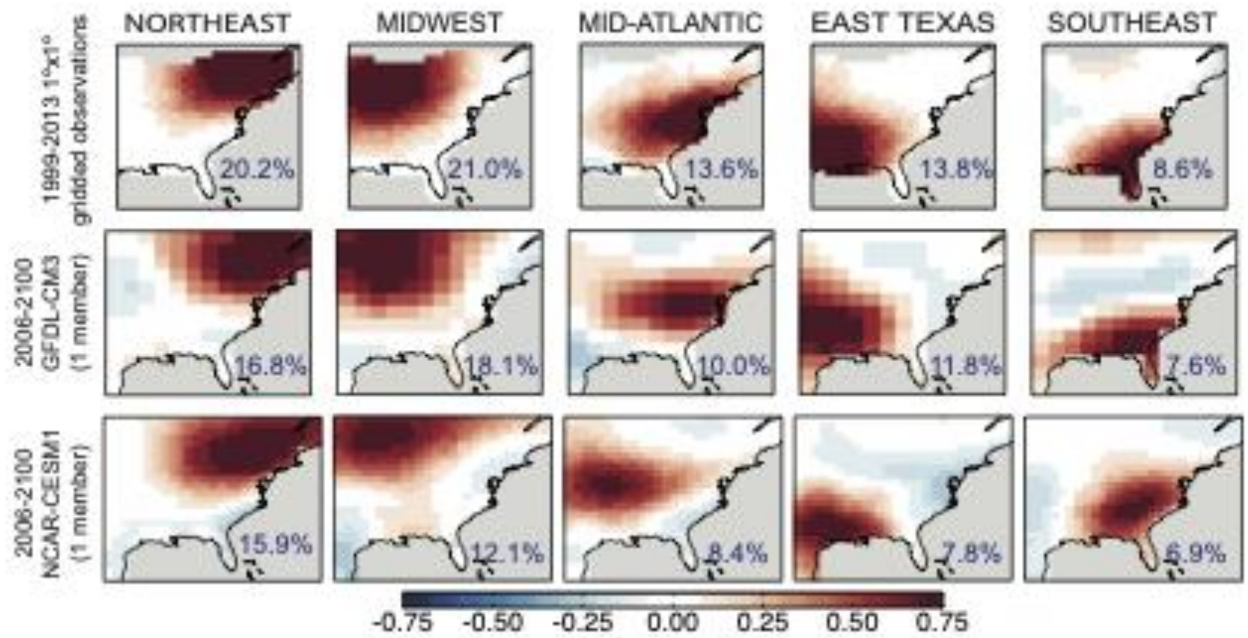

Figure 1. Regions emerging from an EOF analysis on standardized anomalies of summertime daily surface $PM_{2.5}$ over the EUS. Shown are the EOF pattern loadings derived from (top) gridded observations, (middle) one of three ensemble members in the GFDL-CM3 chemistry-climate model, and (bottom) one of 12 NCAR-CESM1 ensemble members. Blue text indicates the total variance explained by each EOF.

Prior analysis of summertime daily ground-level ozone over the EUS for earlier time periods revealed similar EOFs to those in Figure 1 (Eder et al., 1993; Fiore et al., 2003; Lehman et al., 2004). The EOFs derived from summertime ground-level MDA8 ozone observations (Figure S1a) spatially correlate with those for $PM_{2.5}$ (r =0.93-0.99 highest in the Northeast). EOFs for daily T_{max} (Figure S1b) also correlate with those for $PM_{2.5}$ (r = 0.85-0.95, highest in the Northeast and Upper Midwest).

We apply a parallel analysis to the model data; see Tables S2-S3ab for variance explained by the first 10 raw EOFs in the GFDL-CM3 ($PM_{2.5}$, O_3 and temperature) and CESM1 ($PM_{2.5}$ and temperature only) models, respectively (Figures 1, S1ab). The spatial dimension decreases from 113 grid cells in the GFDL-CM3 model and 465 from CESM1 to five regions. We confirm in the GFDL model that the EOFs change little under the 21st century climate change scenario, by conducting the EOF analysis separately on the simulated daily $PM_{2.5}$ for 2006-2025 versus 2086-2100 (Figure S2). We also find that the EOFs are robust across ensemble members (Figures S3ab).

Each EOF is accompanied by a principal component (PC) time series spanning summer days in all years. By definition, the PCs are uncorrelated and combine linearly to explain the largest possible variance captured by the reduced version of the overall dataset. The PC represents how strongly expressed a particular EOF is on each summer day. We orient each PC such that high pollution or temperature values are positive. These time series are the foci for our analysis of changes in the frequency and duration of regional-scale high-pollution events.

We illustrate how the PC can be used to quantify the number of summertime regional-scale pollution events for the Northeast (Figure 2). We consider the observational period during which numerous studies have documented decreasing EUS air pollution in response to emission control programs implemented in the 1990s and 2000s (*e.g.,* Boys et al., 2014; Cooper et al., 2012; Frost et al., 2006; Murphy et al., 2011). For example, 60% decreases in sulfur dioxide emissions from 1990 to 2010 have been linked to 45% lower sulfate aerosol (Skyllakou et al., 2021). Summertime ozone decreases have been attributed to NO_x and VOC emissions reductions of 40% and 14%, respectively, from 2002-2011 (Simon et al., 2015). We define events in the upper quartile (75th percentile; red line in Figure 2) as 'high'. To quantify changes in observed high $PM_{2.5}$ and O_3 events, we count the number of days on which the PC exceeds this threshold. From 1999-2006 to 2007-2013 (time periods separated by the blue dashed vertical line in Figure 2), the number of observed days with high pollution over the Northeast drops: from 265 to 80 days for $PM_{2.5}$ and from 243 to 102 days for MDA8 O_3. This EOF analysis thus enables us to diagnose changes in the frequency of regional-scale high pollution events, without defining an event locally at each monitor or model grid cell relative to a specific concentration threshold.

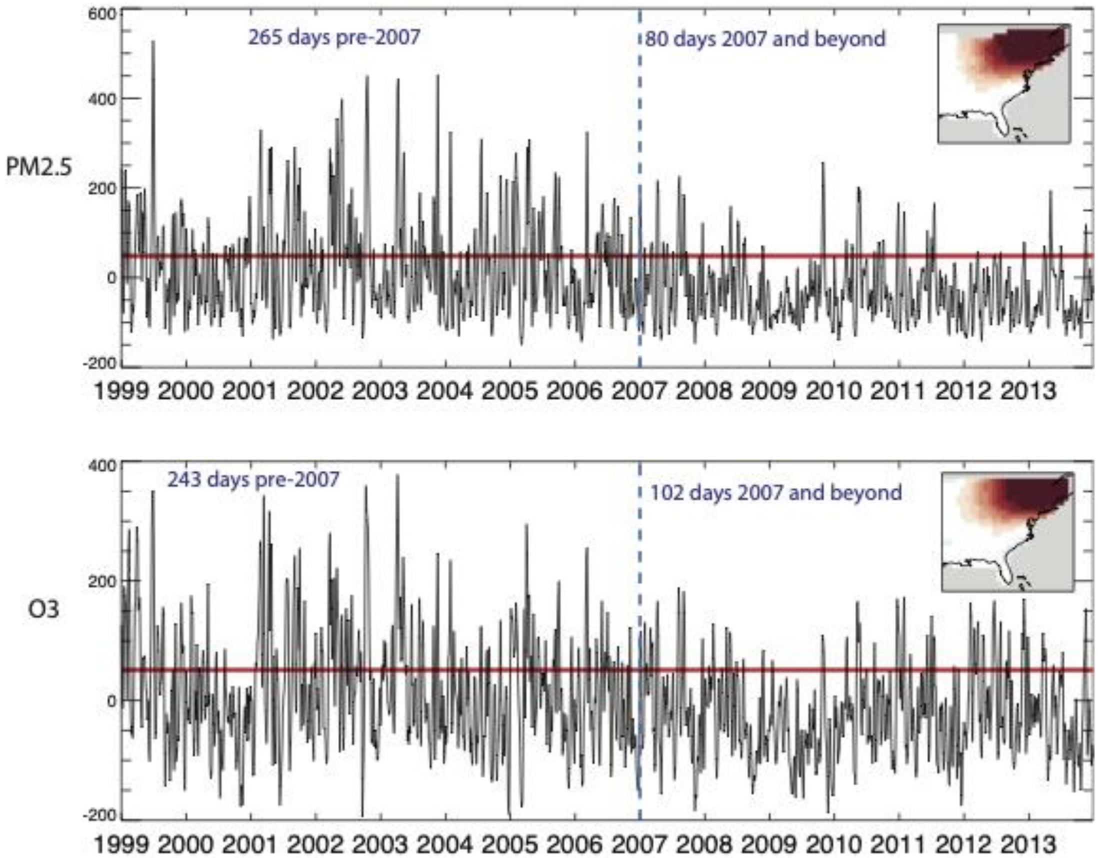

Figure 2. Proof-of-concept demonstration of EOF analysis to track high pollution events. Northeast principal components derived from observed summertime (top) daily mean $PM_{2.5}$ and (bottom) MDA8 ozone from 1999-2013. Shown are the 75th percentile thresholds (red lines) used to define and count the number of high regional-scale pollution events.

Our analysis does not focus on the magnitude of the pollution levels during these events. Rather, our primary interest is to define changes in event frequency and duration, and co-occurrence of high $PM_{2.5}$, O_3, and temperature events under the RCP8.5_WMGG climate change scenario for the 21st century. In any case, the largest, longest-lasting pollution episodes – especially those that are coincident (*i.e.*, high heat, high O_3,, and high $PM_{2.5}$) – typically have the highest pollution levels (Schnell and Prather, 2017).

3 Model Evaluation

Typical approaches evaluating models with observations at specific locations and times are problematic for our study. First, these free-running, fully-coupled chemistry-climate models generate their own weather and thus cannot reproduce the climate variability present in the real atmosphere that is stochastic but imprinted on the air pollutant measurements (*e.g.* year-to-year variations). These measurements are available for a limited number of years when considering variability but can be helpful for evaluating mean errors. Second, the simulations cannot capture observed trends due to changing anthropogenic emissions (since 2005) because they hold air pollutant emissions constant at 2005 levels. In light of these challenges, we evaluate three aspects of the simulations: (1) simulated multi-year summertime average $PM_{2.5}$ and the dominant chemical components (sulfate and organic carbon versus observations, (2) EOFs derived from modeled versus observed daily $PM_{2.5}$, and (3) probability distributions of regionally averaged daily $PM_{2.5}$ derived from the same datasets as in (2). Application of the EOF analysis does not require exact space-time matching, and is ideally suited to evaluate spatiotemporal patterns in climate models that generate their own weather and thus cannot be expected to reproduce observations at a particular location and time. This spatiotemporal evaluation, however, requires extensive observational networks with data of sufficient length and quality, such as are available over the EUS. Section 5 additionally compares observed and simulated cross-correlative relationships between regions and variables.

Summertime mean $PM_{2.5}$ and its major components. The summertime ensemble mean $PM_{2.5}$ simulated by both models reflects the observed spatial pattern of summertime ensemble mean $PM_{2.5}$ in the gridded observations. The NCAR-CESM1 simulation is biased high over the Southeast (Figure S4a). Comparison with the IMPROVE network (Solomon et al., 2014) suggests that both tend to models overestimate EUS $PM_{2.5}$ at these rural sites (Figure S4b), although we note that the comparison with the gridded observations at spatial scales similar to the horizontal resolution of the models is most pertinent here. We also evaluate chemical composition at the IMPROVE sites, which reveals that CESM1 simulates excessive organic carbon, although sulfate tends to be biased low. GFDL-CM3 has a slight tendency to overestimate both species at the IMPROVE sites.

EOFs derived from summertime daily $PM_{2.5}$. The regional patterns that emerge from the EOF analysis applied to daily surface $PM_{2.5}$ are similar in the observations from 1999-2013 and from each of the three GFDL-CM3 model ensemble members for the 2006-2100 period (Figure 1 and S3a). The CM3-derived EOFs capture less overall variance (64-65%; range is over ensemble members) than the observation-derived EOFs (77%). The overall similarity of the patterns implies that this model captures the underlying dynamical and chemical processes that shape the observed spatiotemporal variability. Figure 1 also shows EOFs derived from summertime daily $PM_{2.5}$ simulated by one NCAR-CESM1 ensemble member (Figure S3b displays other ensemble members). The $PM_{2.5}$ EOFs derived from CESM1 capture 54% of the overall variance in the modeled dataset, and four of the EOFs correspond to those derived from observations (Figure 1). Rather than a coastal mid-Atlantic EOF, CESM1 highlights a spatial mode of variability centered over Missouri and Kansas. The spatial error in this pattern as compared to the observation-derived EOF may reflect shortcomings in the geographical placement of the Atlantic or Pacific subtropical high pressure systems and the Great Plains Low Level Jet, and their accompanying precipitation patterns (Bowden et al., 2013; Li et al., 2013; Schmidt & Grise, 2019; Tang et al., 2017). The Northeast EOF, where the two models agree most in their projected changes, serves as a major focus of our analysis and is similarly well captured by both CESM1 and GFDL-CM3.

Probability distributions of daily regional averaged $PM_{2.5}$ in summer. From Figure 1, we selects grid cells in which the EOF loading exceeds 0.5 to define a regional mask separately for each model and the observations. We apply this mask to calculate a daily regional mean $PM_{2.5}$ for the 2006-2010 and 2003-2007 time periods for the models and observations, respectively, and sort the data into 2 μg m^{-3} concentration bins. The mis-match of time periods reflects a compromise to align the model, with constant year 2005 emissions, and the observations, with a 5-year period intended to minimize influences from both emission trends and weather fluctuations. Figure 3 shows the distribution of the average number of days each summer as a function of regional mean $PM_{2.5}$ concentrations for the Northeast, Upper Midwest, and East Texas regions (Figure S5 shows the mid-Atlantic and Southeast). The individual GFDL-CM3 ensemble members fall near or span the observed frequency for the $PM_{2.5}$ bins > 22 μg m^{-3} (Figure 3). This high tail is most relevant to understanding how high $PM_{2.5}$ events will change as the planet warms, and is generally better captured by GFDL-CM3 than CESM1, except over the East Texas region where GFDL-CM3 captures the mode but underestimates the frequency of the highest $PM_{2.5}$ concentrations (>17 μg m^{-3}; Figure 3). While the mode over East Texas is underestimated by CESM1, some ensemble members simulate observed $PM_{2.5}$ levels >26 μg m^{-3} as observed. The GFDL-CM3 distributions over the Northeast, mid-Atlantic and Southeast reflect the mean positive bias evident from Figures S4ab. In CESM1, the positive bias is even higher over the Northeast (and Southeast), with little similarity to the observed distribution shape. Despite mis-placing the Mid-Atlantic EOF relative to observations, CESM1 captures the mode with a slight underestimate, but misses the high tail of the distribution (Figure S5). Below we analyze more deeply high events in GFDL-CM3, which enables us to examine high events of surface O_3 alongside $PM_{2.5}$ and temperature. The NCAR-CESM1 simulations provide a broader context on inter-model differences and on climate variability as measured by the range across ensemble members.

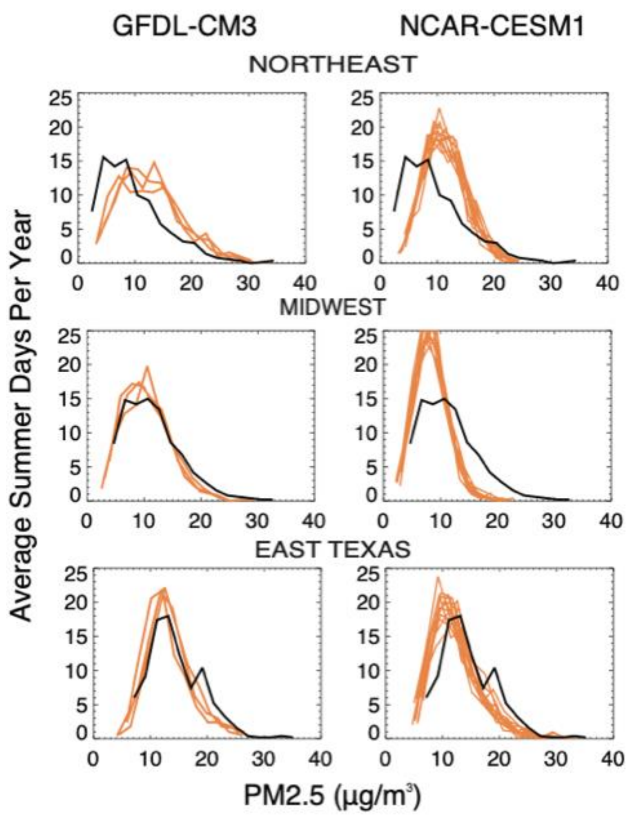

Figure 3. Distributions of the number of summer days with regionally averaged daily $PM_{2.5}$ falling within 2 μg m^{-3} concentration bins. Averages are taken over the regions where the EOF loading in Figure 1 exceeds > 0.5 in the observations (black) for the years 2003-2007 and in the

individual (orange) GFDL-CM3 (left) NCAR-CESM1 (right) ensemble members over the Northeast (top), Midwest (middle), and East Texas (bottom) for model years 2006-2010.

4 21st Century Changes in Summertime $PM_{2.5}$

4.1 Mean values, composition, and probability density functions

Summertime mean $PM_{2.5}$ increases across the contiguous U.S.A. during the 21st century in the GFDL-CM3 ensemble mean, with the largest increases occurring over the Northeast and Upper Midwest, by up to 1-2 and 3-4 μg m^{-3} by mid- (2041-2060; Figure S6) and end-of-century (2081-2100; Figure 4), respectively. These changes are deemed significant if the ensemble mean change exceeds the inter-ensemble range of simulated changes. Grid cells labeled with an 'x' in Figures 4 and S6 do not meet this significance criterion. CESM1 projects smaller ensemble mean $PM_{2.5}$ increases (<1.5 μg m^{-3}) across the Northeast by end-of-century, and decreases over Louisiana, southern Mississippi and Alabama of more than 0.5 μg m^{-3} and 2 μg m^{-3} by mid- and end-of-century, respectively. By 2081-2100, CESM1 also simulates decreases exceeding 0.5 μg m^{-3} over the central Plains, in some of the Northwest and along the southern Atlantic seaboard, though many of these decreases are not significant (Figure 4). In both models, sulfate and organic carbon drive $PM_{2.5}$ increases in the Northeast, with organic carbon contributing more to simulated changes in the Southeast. Significant sulfate increases are projected by both models in parts of the West. GFDL-CM3 also simulates organic carbon increases in the Northwest.

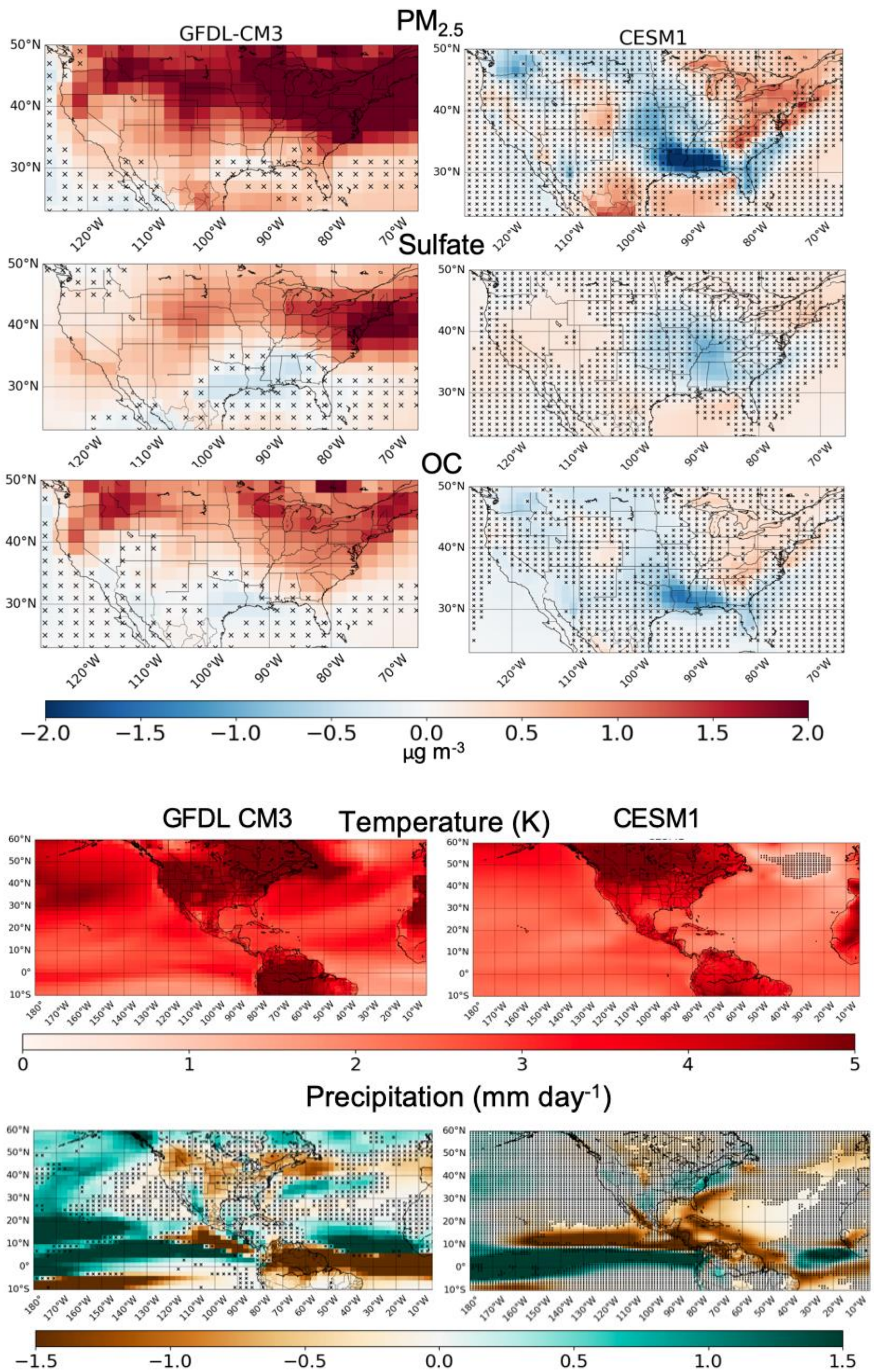

Figure 4. Change in summertime (June-July-August) $PM_{2.5}$, sulfate, organic carbon (OC), daily 2m air temperature (max for GFDL-CM3; mean for CESM1), and precipitation from 2006-2025 to 2081-2100, simulated with GFDL-CM3 (left; 3 ensemble member mean) and CESM1 (right; 12

ensemble member mean) for 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.

Simulated changes in average temperature and precipitation are also shown in Figures 4 and S6. Summertime daily maximum near-surface air temperatures warm in both models, by over 2 K and 4 K by mid- and end-of-century respectively. While GFDL-CM3 simulates a warmer, drier summer over the Northeast, CESM1 warms but wettens (though insignificantly). We do not find evidence that a warmer and drier climate always accompanies higher $PM_{2.5}$, or that more rainfall lowers $PM_{2.5}$. For example, CESM1 simulates declining $PM_{2.5}$ along the Gulf coast without increasing precipitation, which instead increases northeast of this region. Earlier work also demonstrated complex relationships between $PM_{2.5}$ and meteorology that do not simply scale with temperature or precipitation (Dawson et al., 2013; Tai et al., 2010).

For each region and ensemble member, we construct probability density functions by averaging $PM_{2.5}$ over each region on every summer day for the first, middle, and last decade in the RCP8.5_WMGG simulations. The degree to which the ensemble members in different time periods separate from each other offers a measure of significance. GFDL-CM3 simulates decreases in the number of days with low $PM_{2.5}$ concentrations in favor of higher values over the Northeast, Midwest, and Mid-Atlantic regions (light to darker to darkest curves in Figure 5). Little detectable change occurs over the East Texas and Southeast regions as the ensemble members overlap in all three time periods (Figure 5). CESM1 does not project significant changes although a shift towards higher values is perceptible over the Northeast (Figure S7). Overall, this analysis indicates that the uncertainty arising from differences in the model responses to climate change (structural uncertainty) exceeds that from internal variability.

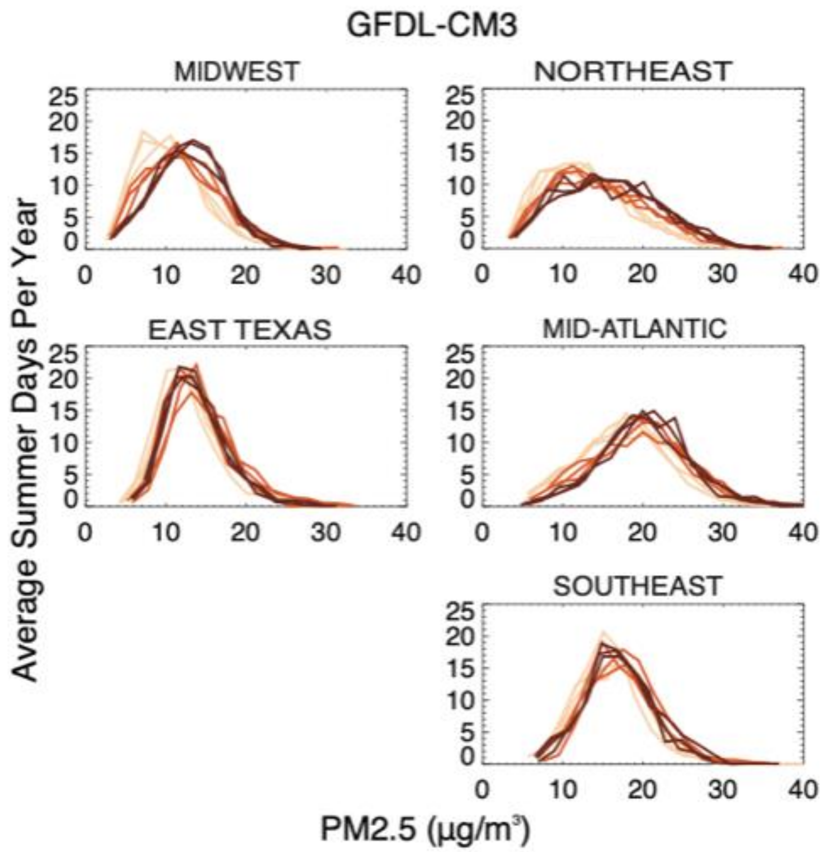

Figure 5. Increasing frequency of high $PM_{2.5}$ events under the RCP8.5_WMGG climate scenario in the GFDL-CM3 model over much of the EUS. Average number of summer days with daily $PM_{2.5}$ falling within 2 µg m^{-3} concentration bins, regionally averaged (where EOF loading > 0.5 in Figure 1) in each GFDL CM3 ensemble member for the years 2006-2015 (light), 2051-2060 (darker) and 2091-2100 (darkest).

4.2 High-$PM_{2.5}$ events: Frequency, duration and intensity

We illustrate our approach with the GFDL-CM3 Northeast EOF for $PM_{2.5}$. We select the upper quartile defined by the full 2006-2100 time series (all values above the red line in Figure S8) and count, separately for each ensemble member, the number of summer days when $PM_{2.5}$ falls in the upper quartile. Over the 21st century, all three GFDL-CM3 ensemble members simulate an increase in this statistic (Figure 6). An ordinary least squares regression suggests an increase in the number of summer days with $PM_{2.5}$ concentrations falling in the upper quartile of 16-20 days (r^2 = 0.3-0.4; range is across ensemble members) by end-of-century. While the changes are not linear with time, this simple metric enables a comparison of changing event frequency over time across ensemble members and variables. Table S4 reports the GFDL-CM3 ensemble mean of these regression statistics for high-PM events, as well as ozone and temperature, in all five regions.

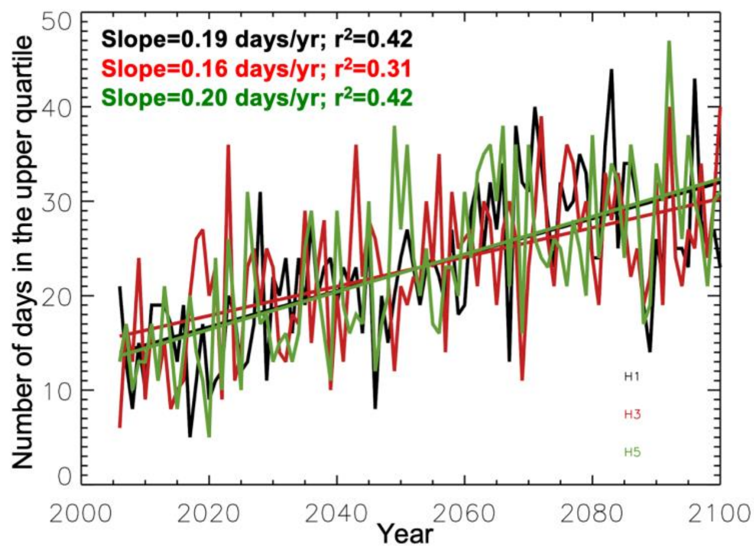

Figure 6. Number of summer days with daily $PM_{2.5}$ falling within the upper quartile defined with respect to the full 2006-2100 period, separately for each GFDL-CM3 ensemble member (colors). Slopes and coefficients of determination (r^2) from ordinary least squares regression are shown in the panel.

To assess changes in the duration of high $PM_{2.5}$ events, we define short (1-2 day), medium (3-4 day) and long (5+ days) durations of top quartile summertime $PM_{2.5}$ events by tracking the number of successive days the PC stays in the upper quartile. For each decade, we sum over all short, medium, and long events. We then average across all ensemble members and report the ensemble mean number of events per decade (colored bars in Figures 7 and S9). Anthropogenic climate change increases the number of 5+ day events over the 21st century, in all regions (green bars in Figures 7 and S9) in GFDL-CM3, although not all changes are significant relative to internal variability (Section 4.3). CESM1 also shows a 21st century increase in 5+ day Northeast $PM_{2.5}$ events but simulates little change or an ensemble average decrease in the longest duration events over other regions (Figures 7 and S9). The differences between the two models in the simulated number of long-lasting events are generally of the same sign as the differences between the simulated summertime mean $PM_{2.5}$ changes in Figure 4.

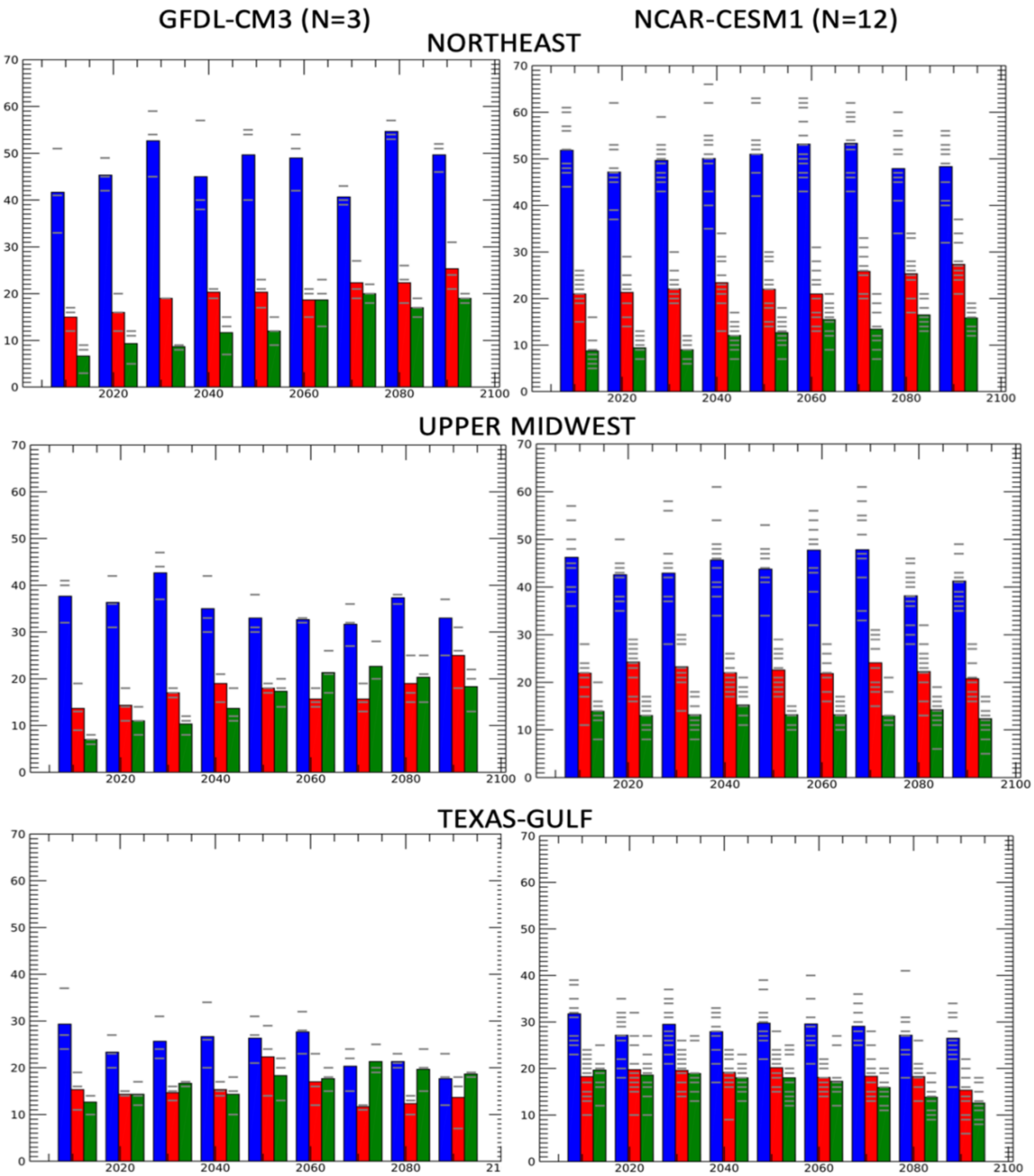

Figure 7. Longer duration upper quartile regional-scale $PM_{2.5}$ events occur under the RCP8.5_WMGG scenario in some regions in the GFDL-CM3 model, but only over the Northeast, and to a lesser extent in CESM1. Shown are the number of times the PC derived from daily mean $PM_{2.5}$ exceeds the upper quartile value, calculated from the full 2006-2100 time period, and stays above that value for 1-2 (blue), 3-4 (red), or 5+ (green) days, summed over each decade within each ensemble member prior to averaging over all GFDL-CM3 (left) and NCAR-CESM1 (right) ensemble members (N) over the Northeast (top), Upper Midwest (middle) and East Texas (bottom) under the RCP8.5_WMGG scenario. The decadal sums for each individual ensemble member are

shown as gray horizontal lines. The range across the gray lines for a given decade is a measure of internal variability. A forced response to rising greenhouse gases is 'detected' when all of the gray lines in a later decade emerge from the range in the early decades.

Our approach thus far defines the upper quartile across the whole time series, which could diagnose a change in duration solely because the frequency changed. Such a change in duration is still relevant from a health impact perspective, especially if extended duration events trigger non-linear health responses. We also investigate the extent to which duration has changed independently from frequency, such as may occur from changing atmospheric circulation. We sample the 10 days each summer with the highest intensity pollution events in GFDL-CM3. We then calculate an average length of episode over the first three vs last three decades of the 21st century. Figure 8 implies that much of the change occurring in Figure 7 is due to changes in frequency. A lengthening of over 0.5 days in the Midwest and almost a full day over the Mid-Atlantic may suggest some underlying fundamental change in ventilation, such as a northward shift of the summertime mid-latitude jet (Barnes & Fiore, 2013; Kerr et al., 2020).

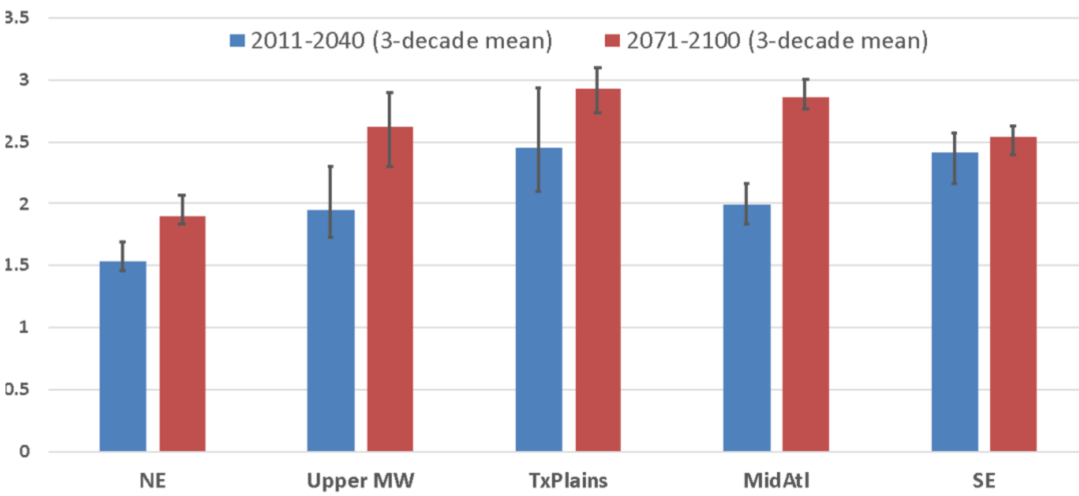

Figure 8. Average length of regional-scale (EOF regions from Figure 1) summertime $PM_{2.5}$ events in the beginning (blue; 2011-2040) versus end (red; 2071-2100) of the 21st century in the GFDL-CM3 model, sampled from 10 days each summer with the highest $PM_{2.5}$ concentrations. The vertical bars indicate the range across the three ensemble members.

As a means of gauging changes in the 'intensity' of events, we construct regional averages of daily $PM_{2.5}$ over the five regions in Figure 1 (where EOFl loadings > 0.5) and report the ensemble mean changes in both models during the 21st century in Table S5. Ensemble mean increases occur in this statistic across all time periods and regions within the GFDL-CM3 model. In CESM1, the ensemble mean increases only over the Northeast, with a slight increase in the upper Midwest by mid-century. We explore the range across individual ensemble members in the next section.

4.3 Changing regional high-$PM_{2.5}$ events in the context of internal climate variability

A novel aspect of our analysis is the use of multiple ensemble members to gauge the significance of changes in high pollution events in light of the variability that arises naturally (internally) in the climate system. The gray lines in Figure 7 denote individual ensemble members (3 GFDL-CM3; 12 NCAR-CESM1). We first consider changes to be significant from one period to another if all ensemble members in the later period fall outside the range of values from the earlier period. GFDL-CM3 simulates significant changes in the longest duration (5+ day) events between the first three and last three decades of the 21st century over the Northeast and mid-Atlantic, and between the first and last two decades of the 21st century in the Southeast (Figures 7 and S9). While the ensemble mean suggests increases in the longest duration events between the early and late 21st century over East Texas and the Upper Midwest (Figure 7), the ensemble member ranges in early versus late decades overlap, implying that these changes are not fully emerging from those that might arise solely due to internal climate variability. While CESM1 indicates a tendency towards increases in the number of 5+ day events, unlike the 3-member GFDL-CM3 ensemble, the ranges across ensemble members in the last few decades do not fully separate from the ranges in the first few decades, even over the Northeast. With a sufficiently large ensemble, such as the multi-model100-member ensembles now being generated for physical climate models, one could better quantify the probability that these changes could arise solely from climate variability, and more cleanly separate inter-model differences from climate variability.

To explore the range of changes one might have diagnosed with a 3-member ensemble as compared to a 12-member ensemble, we consider two end-member cases by comparing the simulated changes in $PM_{2.5}$ event duration from the first two to the last two decades of the 21st century diagnosed by sampling only six of the NCAR-CESM1 ensemble members: three with the smallest (or largest decreases) or largest increases in $PM_{2.5}$. We aim to demonstrate the range that might have occurred if we only had 3 members available, as a way to gauge the potential variability we might have sampled with a larger GFDL-CM3 ensemble. For the longest duration events over the Northeast, increases from the beginning to end-of-century range from 4 to 9.5 days across individual NCAR-CESM1 ensemble members, with the three smallest averaging an increase of 4.2 events per decade that last 5+ days, and the three largest ensemble members averaging an increase of 9.5. For 3-4 day events in the Northeast, the full range spans a decrease of 0.5 to an increase of 11 events, while the averages of the three smallest versus largest ensemble members are 0.83 and 9.5, respectively. We conclude that our limited sampling of three ensemble members in GFDL-CM3 is likely under-representing internal variability, leading to over-confident detection of significant changes in Figure 7.

An analysis of maximum and minimum changes in the 75th percentile daily mean summertime $PM_{2.5}$ values reveals that structural (model response) uncertainty outweighs the role of climate variability (Table S5). The range of changes simulated by the 3-member GFDL ensemble lies completely outside that of the 12-member NCAR ensemble for the Northeast, Midwest, and Mid-Atlantic regions. All three GFDL ensemble members simulate increasing 75th percentile values across all regions except for the East Texas region at mid-Century. In contrast, the sign of the change simulated by CESM1 is only consistent across all 12 ensemble members for the Northeast (increase) and Mid-Atlantic (decrease) by end of century (recall that the mid-Atlantic EOF is displaced inland in CESM1 with respect to observations, Figure 1).

We also select the three NCAR-CESM1 ensemble members with either the smallest or largest changes in 75th percentile daily mean summertime $PM_{2.5}$ concentrations (Table S5). Nearly a factor of 3 range occurs if one considers the average of the 3 NCAR-CESM1 ensemble members with the smallest versus the largest simulated changes over the Northeast. We conclude that inter-model discrepancies reported in the published literature regarding the sign and magnitude of the $PM_{2.5}$ response to climate change reflect not only model structural differences but also internally arising climate variability. This 'climate noise' could be quantified with sufficiently large ensembles that isolate the anthropogenic climate change "signal" (ensemble mean) from the "noise" (ensemble range). Multi-model large ensembles can further distinguish inter-model differences (structural or model response uncertainty) from internal variability (Deser et al., 2020).

5 $PM_{2.5}$-O_3-Temperature Linkages Within and Across Regions

The observational analysis of Schnell and Prather (2017) indicates that extreme events in temperature, MDA8 O_3 and daily mean $PM_{2.5}$ often occur within about a day of each other across the EUS, but the specific relationships vary by region. Climate change induced by rising long-lived greenhouse gases does not change the regional-scale modes of variability in $PM_{2.5}$, O_3, or daily T_{max} as the patterns remain similar throughout the 21st century (Figure S2). Table 1 shows relationships between the 2006-2100 PCs derived from GFDL-CM3 MDA8 O_3, daily mean $PM_{2.5}$ and daily T_{max} within each region. We also examine changes in these relationships over the 21st century by separately analyzing correlations for two decades in the beginning (2006-2025) versus end (2081-2100) of the simulations. The timing of the strongest correlations in Table 1, derived from the daily summertime PCs, are broadly consistent with those emerging from analysis of the 95th percentile of observed warm season pollution and temperature events by Schnell and Prather (2017; see their Figure 4DEF), despite our use of a different metric.

Over all regions and time periods, the strongest correlations in GFDL-CM3 emerge for $PM_{2.5}$ lagging MDA8 O_3 by a day. Future work is needed to determine if these relationships are solely governed by meteorological processes or if, for instance, enhanced O_3 (and OH) production on one day contributes to secondary aerosol formation that accumulates to high $PM_{2.5}$ levels the following day. While secondary inorganic aerosol formation is represented in our GFDL-CM3 configuration, the treatment of secondary organic aerosol is highly simplified and biogenic emissions do not respond to meteorology. Along with the increase in upper quartile $PM_{2.5}$ events discussed in Section 4.2, GFDL-CM3 also projects more frequent O_3 events in both the Northeast and the Mid-Atlantic as well as heat events (Table S4). All three GFDL-CM3 ensemble members show that $PM_{2.5}$-O_3 correlations strengthen or remain similar from 2006-2025 to 2081-2100, with the largest ensemble mean increases occurring over the Southeast (r increases by 0.09) and East Texas (r increases by 0.13) regions.

	T_{max} and O_3			T_{max} and PM			O_3 and PM		
REGION	Lag -1	Lag 0	Lag +1	Lag -1	Lag 0	Lag +1	Lag -1	Lag 0	Lag +1
Over all summers of 2006-2100									
Northeast	0.47	**0.50**	0.38	0.48	**0.57**	**0.57**	0.25	0.58	**0.71**
Mid-Atlantic	**0.67**	**0.67**	0.58	0.33	0.37	0.37	0.40	0.57	**0.65**
Upper Midwest	**0.59**	0.54	0.39	0.51	**0.56**	0.51	0.36	0.61	**0.71**
East Texas	0.03	-0.01	-0.02	0.05	0.04	0.02	0.34	0.51	**0.61**
Southeast	0.03	-0.01	-0.01	0.16	0.15	0.14	0.27	0.46	**0.49**
Only summers of 2006-2025									
Northeast	0.56	**0.59**	0.42	0.52	**0.65**	0.63	0.26	0.58	**0.72**
Mid-Atlantic	**0.71**	**0.71**	0.59	0.26	0.34	0.33	0.33	0.51	**0.60**
Upper Midwest	**0.70**	0.62	0.42	0.50	**0.57**	0.49	0.31	0.58	**0.71**
East Texas	0.12	0.06	0.04	-0.02	-0.05	-0.08	0.26	0.44	**0.55**
Southeast	0.00	-0.05	-0.04	0.07	0.05	0.04	0.23	0.42	**0.45**
Only summers of 2081-2100									
Northeast	0.39	0.42	0.29	0.45	**0.55**	**0.55**	0.19	0.55	**0.70**
Mid-Atlantic	**0.60**	**0.60**	0.51	0.26	0.31	0.30	0.39	0.57	**0.64**
Upper Midwest	**0.53**	0.47	0.31	0.50	**0.56**	0.49	0.31	0.58	**0.68**
East Texas	0.00	-0.02	0.00	-0.01	-0.01	-0.03	0.45	0.61	**0.68**
Southeast	0.06	0.03	0.03	0.17	0.17	0.16	0.32	0.51	**0.54**

Table 1. Ensemble mean correlation coefficients (r) between principal components for pairs of variables simulated by the GFDL-CM3 model (T_{max} is daily maximum temperature at a 2m reference height; O_3 is MDA8 O_3; PM is daily mean $PM_{2.5}$). Correlations are reported for each region on the same day (Lag 0) or with the first variable lagging (Lag -1) or leading (Lag +1) by one 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 $r \geq 0.45$.

The correlation of temperature with O_3 is strongest for zero lag (same day) in the Northeast, and for the same day or O_3 preceding temperature by a day over the Mid-Atlantic, and when ozone precedes temperature by a day over the Upper Midwest. GFDL-CM3 projects a weakening of this temperature-O_3 correlation over the 21st century, with ensemble mean decreases of r=0.17, 0.11, and 0.17 for the Northeast, Mid-Atlantic, and Upper Midwest, respectively. The degraded correlation between O_3 and temperature under climate change was previously shown to occur in this model, and attributed to the summertime mid-latitude jet shifting northward (Barnes & Fiore, 2013). We find no correlation in either of the southern regions (East Texas and Southeast) between temperature and O_3. The absence of an O_3-temperature relationship in GFDL-CM3 agrees with earlier observation-based work showing that humidity offers more explanatory power for O_3 in these regions (Camalier et al., 2007), possibly reflecting a key role for land-atmosphere couplings (Kavassalis & Murphy, 2017; Tawfik & Steiner, 2013).

For temperature and $PM_{2.5}$, we additionally draw on the NCAR-CESM1 simulations. Both models simulate the strongest correlations with zero lag (Tables 1 and S6). All NCAR-CESM1 ensemble members simulate the strongest PM-temperature correlations over the Northeast (ensemble mean r = 0.58, with a range of r=0.55 to 0.60; Table S6), but unlike GFDL-CM3, $PM_{2.5}$ and temperature are not correlated over the displaced Mid-Atlantic region in CESM1 even though an EOF analysis of the CESM1 daily summertime temperature fields reveals a similarly shifted pattern as for $PM_{2.5}$ (Figure S1b). While GFDL-CM3 simulates no relationship between temperature and $PM_{2.5}$ in either southern region, CESM1 indicates a weak temperature-$PM_{2.5}$ anticorrelation for East Texas (Table S6). Prior observation-based work has demonstrated more complex relationships between $PM_{2.5}$ and meteorology (Dawson et al., 2013), in part because individual $PM_{2.5}$ components display different relationships with meteorological variables (Tai et al., 2010; X. Wu et al., 2019). For the highest observed EUS summertime $PM_{2.5}$ events, however, strong relationships with temperature have been found (Porter et al., 2015). The Northeast is the only region for which we identify a consistent change in the $PM_{2.5}$-T relationship across the three GFDL ensemble members from 2006-2025 to 2081-2100, where the correlation declines by an ensemble average of r=0.10 (Table 1).

At present, the EUS climatological summertime near-surface winds are associated with the large-scale circulation around the North Atlantic Subtropical High system, with southerly flow across the southern portion of the domain becoming southwesterly or westerly to the north. A simple inter-regional correlation analysis for MDA8 O_3 and daily mean $PM_{2.5}$ (Table S7) implies a continuation of this circulation pattern over the course of the century. The Northeast and Mid-Atlantic PCs for both air pollutants correlate most strongly with zero lag (r ~ 0.5) whereas both regions tend to lag the Upper Midwest PC by a day (r ~ 0.7). The East Texas and Southeast PCs for O_3 and PM correlate most strongly on the same day (r ~ 0.5, 0.6, respectively), and the Upper Midwest PM PCs correlate most strongly with the East Texas region when lagged by a day (r ~ 0.5). In contrast, the inter-regional correlations for the temperature PCs are almost always strongest for zero lag (not shown). We also conduct this inter-regional correlation analysis separately for simulation years 2006-2025 versus 2081-2100 in each of the 3 GFDL-CM3 ensemble members, but do not detect any robust changes over the course of the century.

6 Discussion and Conclusions

Prior work has shown that some regions experiencing high pollution levels at present will suffer from additional degradation of air quality as the planet continues to warm, if additional controls on air pollutant emissions are not implemented. These studies, however, often conflict (Fiore et al., 2015; Jacob & Winner, 2009; Weaver et al., 2009) and have typically neglected the role of naturally arising internal climate variability by simulating only a small number of years (Deser et al., 2012ab; Garcia-Menendez et al., 2017; Hawkins & Sutton, 2009). With initial condition ensembles in the GFDL-CM3 and CESM1 climate models under a 21st century RCP8.5 scenario with air pollutant emissions frozen in 2005 (denoted RCP8.5_WMGG), we estimate uncertainty due to internal climate variability as the range across the ensemble members available from each model. Relative to this internal variability, we evaluate long-term trends in mean and high air pollution events driven by rising greenhouse gases, as well as model response differences. Differences between the two models serve as a measure of model response (structural) uncertainty.

We demonstrate how Empirical Orthogonal Function (EOF) analysis can be applied to quantify changes in both the frequency and duration of summertime regional-scale pollution episodes over the Eastern United States (EUS). By revealing underlying spatiotemporal patterns of variability, this statistical approach avoids the challenge of bias-correcting individual models, which would be necessary if we were to define high pollution events using an absolute concentration threshold. We find that the models agree best over the Northeast region, where summertime mean surface temperatures increase by over 5 °C during this century, accompanied by a rise in summertime mean $PM_{2.5}$ (up to 1-4 μg m^{-3}). Our analysis of principal components (PCs), the time series accompanying each EOF that indicates how strongly expressed each spatial pattern is on each summer day, reveals an increase in the decadal incidence of upper quartile $PM_{2.5}$ events lasting at least five days over the Northeast that is significant relative to climate variability in GFDL-CM3, and bordering on significant in CESM1 (Figure 7).

The GFDL-CM3 simulations capture, at least qualitatively, observed temporal relationships between EUS MDA8 O_3, daily average $PM_{2.5}$, and daily T_{max}, including those identified by Schnell and Prather (2017). The close temporal occurrence of O_3 and $PM_{2.5}$, and in some cases temperature, events could be relevant to public health, particularly if non-linear responses occur from consecutive or simultaneous exposure. Same-day and consecutive-day exposure to O_3 and $PM_{2.5}$ occurs across the EUS, with GDFL-CM3 projecting a strengthening of this correlation in the southern EUS during the 21st century. Correlated extremes of air pollution and temperature may become more relevant for public health in future decades, particularly in the northern part of our domain where both O_3 and $PM_{2.5}$ remain correlated with temperature (Table 1) and where the frequency and duration of events may increase (Figures 7 and 8). Mascioli et al. (2016) showed that GFDL-CM3 simulates daily T_{max} in excess of the 90th percentile defined relative to 1961-1990 for nearly the entire summer by the 2090s in the RCP8.5 scenario. This standard RCP8.5 scenario warms even more than RCP8.5_WMGG because global aerosols decline, removing the net cooling influence from aerosols, while air quality improves.

The changes we diagnose from GFDL-CM3 imply a trend towards longer-lasting exposures to high pollution events, which may have implications for human and plant health, particularly when accompanied by more intense heat events. By holding anthropogenic emissions fixed in our scenario, we do not consider the potential for human activities to exacerbate or mitigate air pollution levels. This major source of uncertainty has been emphasized in prior studies assessed in Intergovernmental Panel on Climate Change reports (*e.g.*, Kirtman et al., 2013). While we focused on summertime, climate change may extend what is currently 'summer' weather and the accompanying pollutant levels over the EUS into spring and fall, as occurred in October 2010 over the Southeast, triggering high fire and biogenic emissions (Y. Zhang & Wang, 2016). Our study focused on the response of air pollution to changes in meteorology under rising greenhouse gases. Weather-sensitive emission feedbacks such as from wildfires and biogenic emissions were not included in our simulations, and would most likely further amplify pollutant exposure of vulnerable populations and vegetation.

The 12-member NCAR-CESM1 ensemble provides a broader sampling of possible climate states than the 3-member GFDL-CM3 ensemble. Outside of the Northeast, CESM1 simulates different changes in summertime mean $PM_{2.5}$ and upper quartile events, and we find that in some regions, the models do not overlap in their simulated 21st century changes. While three ensemble members is a poor sampling of climate variability, the discrepancies between the two models are sufficiently large as to imply fundamental model differences in their climate responses to rising greenhouse gases. As emphasized by Hawkins & Sutton (2009), uncertain model responses have the potential to be reduced by advancing process-level understanding and improving its representation in models. Air quality projections produced with multi-model chemistry-climate ensembles could transform the capacity to develop probabilistic assessments of changes in regional-scale pollution event frequency and duration, and their co-occurrence with heat, as well as any other metrics of interest for public health or ecosystem welfare. Such ensembles can be parsed separately for uncertainty arising from climate variability versus different model responses.

Our EOF-based approach can be readily applied to any future single or multi-model initial condition chemistry-climate model ensembles. For example, future simulations could sample a wide range of scenarios and incorporate potentially important feedbacks that were neglected in our simulations. A more immediate direction could link EOF patterns to specific meteorological conditions, in which case one could probe existing multi-model initial-condition physical climate model ensembles, with as many as 100 members per model already available (C. Deser et al., 2020), for insights into projected changes in daily MDA8 O_3 and $PM_{2.5}$ events. Understanding and preparing for the range of changes in pollution events that could arise from climate variability may be as important as quantifying the signal from climate change, particularly if climate mitigation leads to less extreme warming scenarios for the 21st century than simulated here.

Acknowledgments, Samples, and Data

We are grateful for helpful discussions with Drs. Dan Bishop, Ron Cohen, Louisa Emmons, Larry Horowitz, Lee Murray, Vaishali Naik, Michael Previdi and Dan Westervelt. We acknowledge Johnny Lin for his publicly available IDL code for varimax rotation varimax_k58.pro. This article was developed under Assistance Agreements 83520601 and 83587801 awarded by the U.S. Environmental Protection Agency to A.M. Fiore. It has not been formally reviewed by EPA. The views expressed in this document are solely those of the authors and do not necessarily reflect those of the Agency. EPA does not endorse any products or commercial services mentioned in this publication. The CESM project is supported primarily by the National Science Foundation (NSF). This material is based upon work supported by the National Center for Atmospheric Research (NCAR), which is a major facility sponsored by the NSF under Cooperative Agreement 1852977. Data and code used to create figures will be provided in Columbia University Academic Commons at ….

References

Abatzoglou, J. T., & Williams, A. P. (2016). Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences*, *113*(42), 11770-11775. https://www.pnas.org/content/pnas/113/42/11770.full.pdf

Andersson, C., & Engardt, M. (2010). European ozone in a future climate: Importance of changes in dry deposition and isoprene emissions. *J. Geophys. Res.*, *115*(D2), D02303. http://dx.doi.org/10.1029/2008JD011690

Austin, J., Horowitz, L. W., Schwarzkopf, M. D., Wilson, R. J., & Levy, H. (2013). Stratospheric Ozone and Temperature Simulated from the Preindustrial Era to the Present Day. *Journal of Climate, 26*(11), 3528-3543. http://dx.doi.org/10.1175/JCLI-D-12-00162.1

Barnes, E. A., & Fiore, A. M. (2013). Surface ozone variability and the jet position: Implications for projecting future air quality. *Geophysical Research Letters, 40*(11), 2839-2844. http://dx.doi.org/10.1002/grl.50411

Bowden, J. H., Nolte, C. G., & Otte, T. L. (2013). Simulating the impact of the large-scale circulation on the 2-m temperature and precipitation climatology. *Climate Dynamics, 40*(7), 1903-1920. https://doi.org/10.1007/s00382-012-1440-y

Boys, B. L., Martin, R. V., van Donkelaar, A., MacDonell, R. J., Hsu, N. C., Cooper, M. J., et al. (2014). Fifteen-Year Global Time Series of Satellite-Derived Fine Particulate Matter. *Environmental Science & Technology, 48*(19), 11109-11118. https://doi.org/10.1021/es502113p

Camalier, L., Cox, W., & Dolwick, P. (2007). The effects of meteorology on ozone in urban areas and their use in assessing ozone trends. *Atmospheric Environment, 41*(33), 7127-7137. http://www.sciencedirect.com/science/article/B6VH3-4NR18M7-2/2/96f87a4d98435ecbf7fd411e2ee1fa4e

Clifton, O. E., Fiore, A. M., Correa, G., Horowitz, L. W., & Naik, V. (2014). Twenty-first century reversal of the surface ozone seasonal cycle over the northeastern United States. *Geophysical Research Letters, 41*(20), 7343-7350. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014GL061378

Coffel, E. D., Horton, R. M., & de Sherbinin, A. (2017). Temperature and humidity based projections of a rapid rise in global heat stress exposure during the 21st century. *Environmental Research Letters, 13*(1), 014001. http://dx.doi.org/10.1088/1748-9326/aaa00e

Cooper, O. R., Gao, R.-S., Tarasick, D., Leblanc, T., & Sweeney, C. (2012). Long-term ozone trends at rural ozone monitoring sites across the United States, 1990–2010. *Journal of Geophysical Research: Atmospheres, 117*(D22), D22307. http://dx.doi.org/10.1029/2012JD018261

Dawson, J. P., Bloomer, B. J., Winner, D. A., & Weaver, C. P. (2013). Understanding the meteorological drivers of U.S. particulate matter concentrations in a changing climate. *Bulletin of the American Meteorological Society*. http://dx.doi.org/10.1175/BAMS-D-12-00181.1

Deser, C., Knutti, R., Solomon, S., & Phillips, A. S. (2012a). Communication of the role of natural variability in future North American climate. *Nature Clim. Change, 2*(12), 888-888. 10.1038/nclimate1779. http://dx.doi.org/10.1038/nclimate1779

Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N., et al. (2020). Insights from Earth system model initial-condition large ensembles and future prospects. *Nature Climate Change, 10*(4), 277-286. https://doi.org/10.1038/s41558-020-0731-2

Deser, C., Phillips, A. S., Alexander, M. A., & Smoliak, B. V. (2013). Projecting North American Climate Over the Next 50 Years: Uncertainty due to Internal Variability. *Journal of Climate*. http://dx.doi.org/10.1175/JCLI-D-13-00451.1

Deser, C., Phillips, A., Bourdette, V., & Teng, H. (2012b). Uncertainty in climate change projections: the role of internal variability. *Climate Dynamics, 38*(3-4), 527-546. http://dx.doi.org/10.1007/s00382-010-0977-x

Donner, L. J., Wyman, B. L., Hemler, R. S., Horowitz, L. W., Ming, Y., Zhao, M., et al. (2011). The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3. *Journal of Climate, 24*(13), 3484-3519. http://dx.doi.org/10.1175/2011JCLI3955.1

East, J., & Garcia-Menendez, F. (2020). Internal climate variability and initial condition ensembles in air quality projections. *U.S. CLIVAR Variations* (Summer 2020, 18, No. 2). https://opensky.ucar.edu/islandora/object/usclivar:125

Eder, B. K., Davis, J. M., & Bloomfield, P. (1993). A characterization of the spatiotemporal variability of non-urban ozone concentrations over the eastern United States. *Atmospheric Environment. Part A. General Topics, 27*(16), 2645-2668. http://www.sciencedirect.com/science/article/pii/096016869390035W

Filleul, L., Cassadou, S., Médina, S., Fabres, P., Lefranc, A., Eilstein, D., et al. (2006). The relation between temperature, ozone, and mortality in nine French cities during the heat wave of 2003. *Environmental Health Perspectives, 114*(9), 1344-1347. https://pubmed.ncbi.nlm.nih.gov/16966086

Fiore, A. M., Jacob, D. J., Mathur, R., & Martin, R. V. (2003). Application of empirical orthogonal functions to evaluate ozone simulations with regional and global models. *J. Geophys. Res., 108*(D14), 4431. http://dx.doi.org/10.1029/2002JD003151

Fiore, A. M., Naik, V., & Leibensperger, E. M. (2015). Air Quality and Climate Connections. *Journal of the Air & Waste Management Association, 65*(6), 645-685. https://doi.org/10.1080/10962247.2015.1040526

Fiore, A. M., Naik, V., Spracklen, D. V., Steiner, A., Unger, N., Prather, M., et al. (2012). Global air quality and climate. *Chemical Society Reviews, 41*(19), 6663-6683. http://dx.doi.org/10.1039/C2CS35095E

Frost, G. J., McKeen, S. A., Trainer, M., Ryerson, T. B., Neuman, J. A., Roberts, J. M., et al. (2006). Effects of changing power plant NOx emissions on ozone in the eastern United States: Proof of concept. *Journal of Geophysical Research: Atmospheres, 111*(D12), D12306. http://dx.doi.org/10.1029/2005JD006354

Fu, T.-M., & Tian, H. (2019). Climate Change Penalty to Ozone Air Quality: Review of Current Understandings and Knowledge Gaps. *Current Pollution Reports, 5*(3), 159-171. https://doi.org/10.1007/s40726-019-00115-6

García-Herrera, R., Díaz, J., Trigo, R. M., Luterbacher, J., & Fischer, E. M. (2010). A Review of the European Summer Heat Wave of 2003. *Critical Reviews in Environmental Science and Technology, 40*(4), 267-306. https://doi.org/10.1080/10643380802238137

Garcia-Menendez, F., Monier, E., & Selin, N. E. (2017). The role of natural variability in projections of climate change impacts on U.S. ozone pollution. *Geophysical Research Letters, 44*(6), 2911-2921. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL071565

Ghan, S. J., Liu, X., Easter, R. C., Zaveri, R., Rasch, P. J., Yoon, J.-H., & Eaton, B. (2012). Toward a Minimal Representation of Aerosols in Climate Models: Comparative Decomposition of Aerosol Direct, Semidirect, and Indirect Radiative Forcing. *Journal of Climate, 25*(19), 6461-6476. https://doi.org/10.1175/JCLI-D-11-00650.1

Hawkins, E., & Sutton, R. (2009). The Potential to Narrow Uncertainty in Regional Climate Predictions. *Bulletin of the American Meteorological Society*, 1095-1108.

Hong, C., Zhang, Q., Zhang, Y., Davis, S. J., Tong, D., Zheng, Y., et al. (2019). Impacts of climate change on future air quality and human health in China. *Proceedings of the National Academy of Sciences, 116*(35), 17193-17200. https://www.pnas.org/content/pnas/116/35/17193.full.pdf

Horton, D. E., Harshvardhan, & Diffenbaugh, N. S. (2012). Response of air stagnation frequency to anthropogenically enhanced radiative forcing. *Environmental Research Letters*, *7*(4), 044034. http://dx.doi.org/10.1088/1748-9326/7/4/044034

Horton, D. E., Skinner, C. B., Singh, D., & Diffenbaugh, N. S. (2014). Occurrence and persistence of future atmospheric stagnation events. *Nature Climate Change, 4*(8), 698-703. https://doi.org/10.1038/nclimate2272

Hou, P., & Wu, S. (2016). Long-term Changes in Extreme Air Pollution Meteorology and the Implications for Air Quality. *Scientific Reports, 6*(1), 23792. https://doi.org/10.1038/srep23792

Jacob, D. J., & Winner, D. A. (2009). Effect of climate change on air quality. *Atmospheric Environment, 43*(1), 51-63. http://www.sciencedirect.com/science/article/B6VH3-4TNWH49-4/2/30b98e804a7d8cab841c22038bc0c264

Kavassalis, S. C., & Murphy, J. G. (2017). Understanding ozone-meteorology correlations: A role for dry deposition. *Geophysical Research Letters*, *44*(6), 2922-2931. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL071791

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., et al. (2015). The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability. *Bulletin of the American Meteorological Society, 96*(8), 1333-1349. https://doi.org/10.1175/BAMS-D-13-00255.1

Kerr, G. H., Waugh, D. W., Steenrod, S. D., Strode, S. A., & Strahan, S. E. (2020). Surface Ozone-Meteorology Relationships: Spatial Variations and the Role of the Jet Stream. *Journal of Geophysical Research: Atmospheres, 125*(21), e2020JD032735. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JD032735

Kerr, G. H., Waugh, D. W., Strode, S. A., Steenrod, S. D., Oman, L. D., & Strahan, S. E. (2019). Disentangling the Drivers of the Summertime Ozone-Temperature Relationship Over the United States. *Journal of Geophysical Research: Atmospheres*, *124*(19), 10503-10524. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD030572

Kirtman, B., S.B. Power, J.A. Adedoyin, G.J. Boer, R. Bojariu, I. Camilloni, et al. (2013). Near-term Climate Change: Projections and Predictability. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. M. Midgley (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Konovalov, I. B., Beekmann, M., Kuznetsova, I. N., Yurova, A., & Zvyagintsev, A. M. (2011). Atmospheric impacts of the 2010 Russian wildfires: integrating modelling and measurements of an extreme air pollution episode in the Moscow region. *Atmos. Chem. Phys.*, *11*(19), 10031-10056. https://acp.copernicus.org/articles/11/10031/2011/

Lamarque, J.-F., Kyle, G. P., Meinshausen, M., Riahi, K., Smith, S., van Vuuren, D., et al. (2011). Global and regional evolution of short-lived radiatively-active gases and aerosols in the Representative Concentration Pathways. *Climatic Change*, *109*(1-2), 191-212. http://dx.doi.org/10.1007/s10584-011-0155-0

Lehman, J., Swinton, K., Bortnick, S., Hamilton, C., Baldridge, E., Eder, B., & Cox, B. (2004). Spatio-temporal characterization of tropospheric ozone across the eastern United States. *Atmospheric Environment*, *38*(26), 4357-4369. https://www.sciencedirect.com/science/article/pii/S1352231004003838

Leibensperger, E. M., Mickley, L. J., & Jacob, D. J. (2008). Sensitivity of US air quality to mid-latitude cyclone frequency and implications of 1980–2006 climate change. *Atmos. Chem. Phys.*, *8*(23), 7075-7086. http://www.atmos-chem-phys.net/8/7075/2008/

http://www.atmos-chem-phys.net/8/7075/2008/acp-8-7075-2008.pdf

Li, L. F., Li, W. H., & Deng, Y. (2013). Summer rainfall variability over the Southeastern United States and its intensification in the 21st century as assessed by CMIP5 models. *Journal of Geophysical Research-Atmospheres*, *118*(2), 340-354. <Go to ISI>://000317838100008

Lin, C. Y. C., Jacob, D. J., & Fiore, A. M. (2001). Trends in exceedances of the ozone air quality standard in the continental United States, 1980-1998. *Atmospheric Environment*, *35*(19), 3217-3228. http://www.sciencedirect.com/science/article/B6VH3-433W771-1/2/4ee0c5d9076c1f8aad96415ff5cb353f

Liu, X., Easter, R. C., Ghan, S. J., Zaveri, R., Rasch, P., Shi, X., et al. (2012). Toward a minimal representation of aerosols in climate models: description and evaluation in the Community Atmosphere Model CAM5. *Geosci. Model Dev.*, *5*(3), 709-739. https://gmd.copernicus.org/articles/5/709/2012/

Logan, J. A. (1989). Ozone in Rural Areas of the United States. *J. Geophys. Res.*, *94*(D6), 8511-8532.

Mascioli, N. R., Fiore, A. M., Previdi, M., & Correa, G. (2016). Temperature and Precipitation Extremes in the United States: Quantifying the Responses to Anthropogenic Aerosols and Greenhouse Gases, *Journal of Climate*, *29*(7), 2689-2701. Retrieved Jun 30, 2021, from https://journals.ametsoc.org/view/journals/clim/29/7/jcli-d-15-0478.1.xml

Mickley, L. J., Jacob, D. J., Field, B. D., & Rind, D. (2004). Effects of future climate change on regional air pollution episodes in the United States. *Geophys. Res. Lett.*, *31*(24), L24103. http://dx.doi.org/10.1029/2004GL021216

Murphy, D. M., Chow, J. C., Leibensperger, E. M., Malm, W. C., Pitchford, M., Schichtel, B. A., et al. (2011). Decreases in elemental carbon and fine particle mass in the United States. *Atmos. Chem. Phys.*, *11*(10), 4679-4686. https://acp.copernicus.org/articles/11/4679/2011/

Naik, V., Horowitz, L. W., Fiore, A. M., Ginoux, P., Mao, J., Aghedo, A. M., & Levy, H. (2013). Impact of preindustrial to present-day changes in short-lived pollutant emissions on atmospheric composition and climate forcing. *Journal of Geophysical Research: Atmospheres, 118*(14), 8086-8110. http://dx.doi.org/10.1002/jgrd.50608

Nolte, C. G., Spero, T. L., Bowden, J. H., Mallard, M. S., & Dolwick, P. D. (2018). The potential effects of climate change on air quality across the conterminous US at 2030 under three Representative Concentration Pathways. *Atmos. Chem. Phys.*, *18*(20), 15471-15489. https://acp.copernicus.org/articles/18/15471/2018/

Oswald, E. M., Dupigny-Giroux, L.-A., Leibensperger, E. M., Poirot, R., & Merrell, J. (2015). Climate controls on air quality in the Northeastern U.S.: An examination of summertime ozone statistics during 1993–2012. *Atmospheric Environment*, *112*, 278-288. http://www.sciencedirect.com/science/article/pii/S1352231015300200

Phalitnonkiat, P., Hess, P. G. M., Grigoriu, M. D., Samorodnitsky, G., Sun, W., Beaudry, E., et al. (2018). Extremal dependence between temperature and ozone over the continental US. *Atmos. Chem. Phys.*, *18*(16), 11927-11948. https://acp.copernicus.org/articles/18/11927/2018/

Porter, W. C., & Heald, C. L. (2019). The mechanisms and meteorological drivers of the summertime ozone–temperature relationship. *Atmos. Chem. Phys.*, *19*(21), 13367-13381. https://www.atmos-chem-phys.net/19/13367/2019/

Porter, W. C., Heald, C. L., Cooley, D., & Russell, B. (2015). Investigating the observed sensitivities of air-quality extremes to meteorological drivers via quantile regression. *Atmos. Chem. Phys.*, *15*(18), 10349-10366. https://www.atmos-chem-phys.net/15/10349/2015/

Previdi, M., & Fiore, A. M. (2019). The Importance of Sampling Variability in Assessments of ENSO-PM2.5 Relationships: A Case Study for the South Central United States. *Geophysical Research Letters*, *46*(12), 6878-6884. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL082250

Rao, S. T., Zalewsky, E., & Zurbenko, I. G. (1995). Determining Temporal and Spatial Variations in Ozone Air Quality. *Journal of the Air & Waste Management Association*, *45*(1), 57-61. https://doi.org/10.1080/10473289.1995.10467342

Rieder, H. E., Fiore, A. M., Horowitz, L. W., & Naik, V. (2015). Projecting policy-relevant metrics for high summertime ozone pollution events over the Eastern United States due to climate and emission changes during the 21st century. *Journal of Geophysical Research: Atmospheres*, 2014JD022303. http://dx.doi.org/10.1002/2014JD022303

Schmidt, D. F., & Grise, K. M. (2019). Impacts of Subtropical Highs on Summertime Precipitation in North America. *Journal of Geophysical Research: Atmospheres*, *124*(21), 11188-11204. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD031282

Schnell, J. L., Holmes, C. D., Jangam, A., & Prather, M. J. (2014). Skill in forecasting extreme ozone pollution episodes with a global atmospheric chemistry model. *Atmos. Chem. Phys.*, *14*(15), 7721-7739. http://www.atmos-chem-phys.net/14/7721/2014/

Schnell, J. L., & Prather, M. J. (2017). Co-occurrence of extremes in surface ozone, particulate matter, and temperature over eastern North America. *Proceedings of the National Academy of Sciences*, 201614453. https://www.pnas.org/content/pnas/early/2017/02/21/1614453114.full.pdf

Schnell, J. L., Prather, M. J., Josse, B., Naik, V., Horowitz, L. W., Cameron-Smith, P., et al. (2015). Use of North American and European air quality networks to evaluate global chemistry–

climate modeling of surface ozone. *Atmos. Chem. Phys., 15*(18), 10581-10596. https://acp.copernicus.org/articles/15/10581/2015/

Schnell, J. L., Prather, M. J., Josse, B., Naik, V., Horowitz, L. W., Zeng, G., et al. (2016). Effect of climate change on surface ozone over North America, Europe, and East Asia. *Geophysical Research Letters, 43*(7), 3509-3518. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL068060

Shaposhnikov, D., Revich, B., Bellander, T., Bedada, G. B., Bottai, M., Kharkova, T., et al. (2014). Mortality related to air pollution with the moscow heat wave and wildfire of 2010. *Epidemiology (Cambridge, Mass.), 25*(3), 359-364. https://pubmed.ncbi.nlm.nih.gov/24598414

Shen, L., Mickley, L. J., & Gilleland, E. (2016). Impact of increasing heat waves on U.S. ozone episodes in the 2050s: Results from a multimodel analysis using extreme value theory. *Geophysical Research Letters, 43*(8), 4017-4025. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL068432

Simon, H., Reff, A., Wells, B., Xing, J., & Frank, N. (2015). Ozone Trends Across the United States over a Period of Decreasing NOx and VOC Emissions. *Environmental Science & Technology, 49*(1), 186-195. https://doi.org/10.1021/es504514z

Skyllakou, K., Garcia Rivera, P., Dinkelacker, B., Karnezi, E., Kioutsioukis, I., Hernandez, C., et al. (2021). Changes in PM2.5 concentrations and their sources in the US from 1990 to 2010. *Atmos. Chem. Phys. Discuss., 2021*, 1-34. https://acp.copernicus.org/preprints/acp-2021-495/

Solomon, P. A., Crumpler, D., Flanagan, J. B., Jayanty, R. K. M., Rickman, E. E., & McDade, C. E. (2014). U.S. National PM2.5 Chemical Speciation Monitoring Networks—CSN and IMPROVE: Description of networks. *Journal of the Air & Waste Management Association, 64*(12), 1410-1438. https://doi.org/10.1080/10962247.2014.956904

Spracklen, D. V., Mickley, L. J., Logan, J. A., Hudman, R. C., Yevich, R., Flannigan, M. D., & Westerling, A. L. (2009). Impacts of climate change from 2000 to 2050 on wildfire activity and carbonaceous aerosol concentrations in the western United States. *J. Geophys. Res., 114*(D20), D20301. http://dx.doi.org/10.1029/2008JD010966

Sun, W., Hess, P., Chen, G., & Tilmes, S. (2019). How waviness in the circulation changes surface ozone: a viewpoint using local finite-amplitude wave activity. *Atmos. Chem. Phys., 19*(20), 12917-12933. https://acp.copernicus.org/articles/19/12917/2019/

Sun, W., Hess, P., & Liu, C. (2017). The impact of meteorological persistence on the distribution and extremes of ozone. *Geophysical Research Letters, 44*(3), 1545-1553. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL071731

Tai, A. P. K., L.J. Mickley, & Jacob, D. J. (2010). Correlations between fine particulate matter (PM2.5) and meteorological variables in the United States: Implications for the sensitivity of PM2.5 to climate change. *Atmospheric Environment, 44*(32), 3976-3984. http://www.sciencedirect.com/science/article/B6VH3-50GWNDS-3/2/df35554565c3e1b79bac2a652ca96aa4

Tai, A. P. K., Mickley, L. J., Jacob, D. J., Leibensperger, E. M., Zhang, L., Fisher, J. A., & Pye, H. O. T. (2012). Meteorological modes of variability for fine particulate matter (PM2.5) air quality in the United States: implications for PM2.5 sensitivity to climate change. *Atmos. Chem. Phys., 12*(6), 3131-3145. http://www.atmos-chem-phys.net/12/3131/2012/

http://www.atmos-chem-phys.net/12/3131/2012/acp-12-3131-2012.pdf

Tang, Y., Winkler, J., Zhong, S., Bian, X., Doubler, D., Yu, L., & Walters, C. (2017). Future changes in the climatology of the Great Plains low-level jet derived from fine resolution multi-model simulations. *Scientific Reports, 7*(1), 5029. https://doi.org/10.1038/s41598-017-05135-0

Tawfik, A. B., & Steiner, A. L. (2013). A proposed physical mechanism for ozone-meteorology correlations using land–atmosphere coupling regimes. *Atmospheric Environment, 72*, 50-59. https://www.sciencedirect.com/science/article/pii/S1352231013001672

Turnock, S. T., Allen, R. J., Andrews, M., Bauer, S. E., Emmons, L., Good, P., et al. (2020). Historical and future changes in air pollutants from CMIP6 models. *Atmos. Chem. Phys. Discuss., 2020*, 1-40. https://www.atmos-chem-phys-discuss.net/acp-2019-1211/

Vukovich, F. M. (1995). Regional-scale boundary layer ozone variations in the eastern United States and their association with meteorological variations. *Atmos. Environ., 29*, 2259-2273.

Wang, J. X. L., & Angell, J. K. (1999). Air stagnation climatology for the United States (1948-1998). *NOAA/Air Resources Laboratory ATLAS No. 1*.

Weaver, C. P., Cooter, E., Gilliam, R., Gilliland, A., Grambsch, A., Grano, D., et al. (2009). A Preliminary Synthesis of Modeled Climate Change Impacts on U.S. Regional Ozone Concentrations. *Bulletin of the American Meteorological Society, 90*(12), 1843-1863. http://journals.ametsoc.org/doi/abs/10.1175/2009BAMS2568.1

Wilks, D. S. (1995). *Statistical Methods in the Atmospheric Sciences*. San Diego, California: Academic.

Wu, S. L., Mickley, L. J., Leibensperger, E. M., Jacob, D. J., Rind, D., & Streets, D. G. (2008). Effects of 2000-2050 global change on ozone air quality in the United States. *Journal of Geophysical Research - Atmospheres, 113*(D6), D06302.

Wu, X., Xu, Y., Kumar, R., & Barth, M. (2019). Separating Emission and Meteorological Drivers ofMid-21st-Century Air Quality Changes in IndiaBased on Multiyear Global-RegionalChemistry-Climate Simulations. *Journal of Geophysical Research: Atmospheres, 124*(23), 13420-13438. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD030988

Xu, Y., & Lamarque, J.-F. (2018). Isolating the Meteorological Impact of 21st Century GHG Warming on the Removal and Atmospheric Loading of Anthropogenic Fine Particulate Matter Pollution at Global Scale. *Earth's Future, 6*(3), 428-440. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017EF000684

Xu, Y., Wu, X., Kumar, R., Barth, M., Diao, C., Gao, M., et al. (2020). Substantial Increase in the Joint Occurrence and Human Exposure of Heatwave and High-PM Hazards Over South Asia in the Mid-21st Century. *AGU Advances, 1*(2), e2019AV000103. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019AV000103

Zhang, J., Gao, Y., Luo, K., Leung, L. R., Zhang, Y., Wang, K., & Fan, J. (2018). Impacts of compound extreme weather events on ozone in the present and future. *Atmos. Chem. Phys., 18*(13), 9861-9877. https://acp.copernicus.org/articles/18/9861/2018/

Zhang, Y., & Wang, Y. (2016). Climate-driven ground-level ozone extreme in the fall over the Southeast United States. *Proceedings of the National Academy of Sciences, 113*(36), 10025-10030. https://www.pnas.org/content/pnas/113/36/10025.full.pdf

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

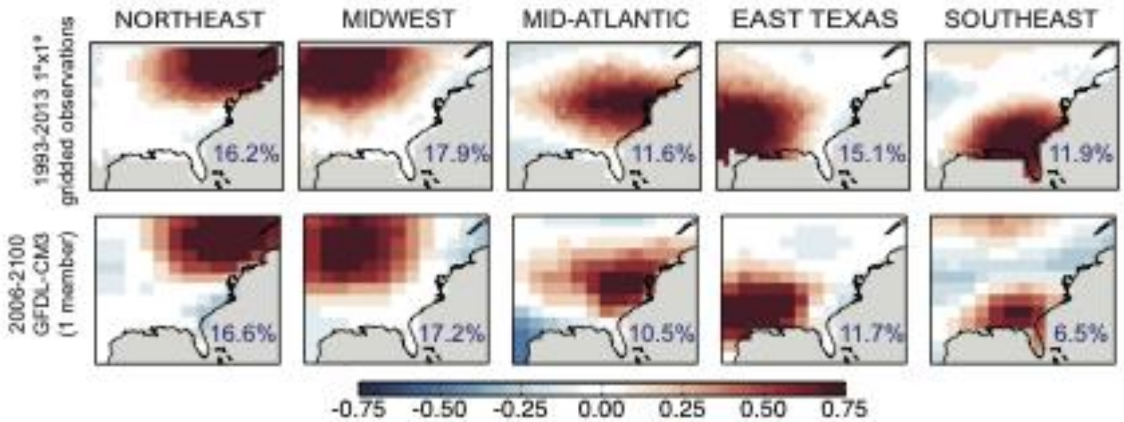

Figure S1a. Regions emerging from an EOF analysis on standardized anomalies of summertime daily maximum 8-hour average (MDA8) O_3 over the EUS. Shown are the EOF pattern loadings derived from (top) gridded observations, (bottom) one of three ensemble members in the GFDL-CM3 chemistry-climate model. Blue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

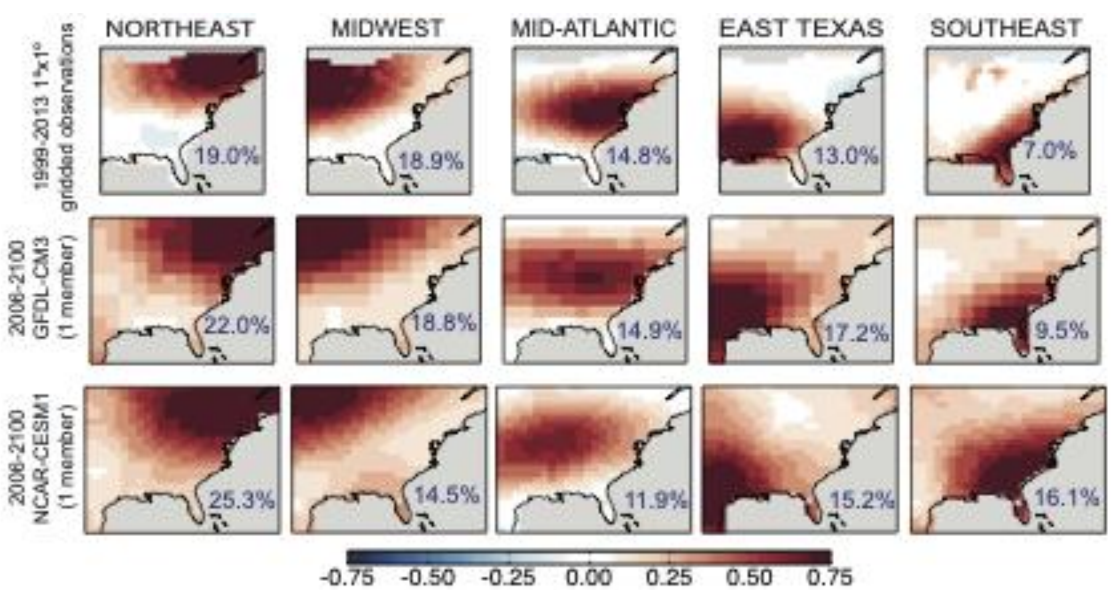

Figure S1b. Regions emerging from an EOF analysis on standardized anomalies of daily temperature over the EUS during summer (June-July-August). Shown are the EOF pattern loadings derived from (top) gridded observations of daily T_{max}, (middle row) daily T_{max} simulated by one of three ensemble members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

.

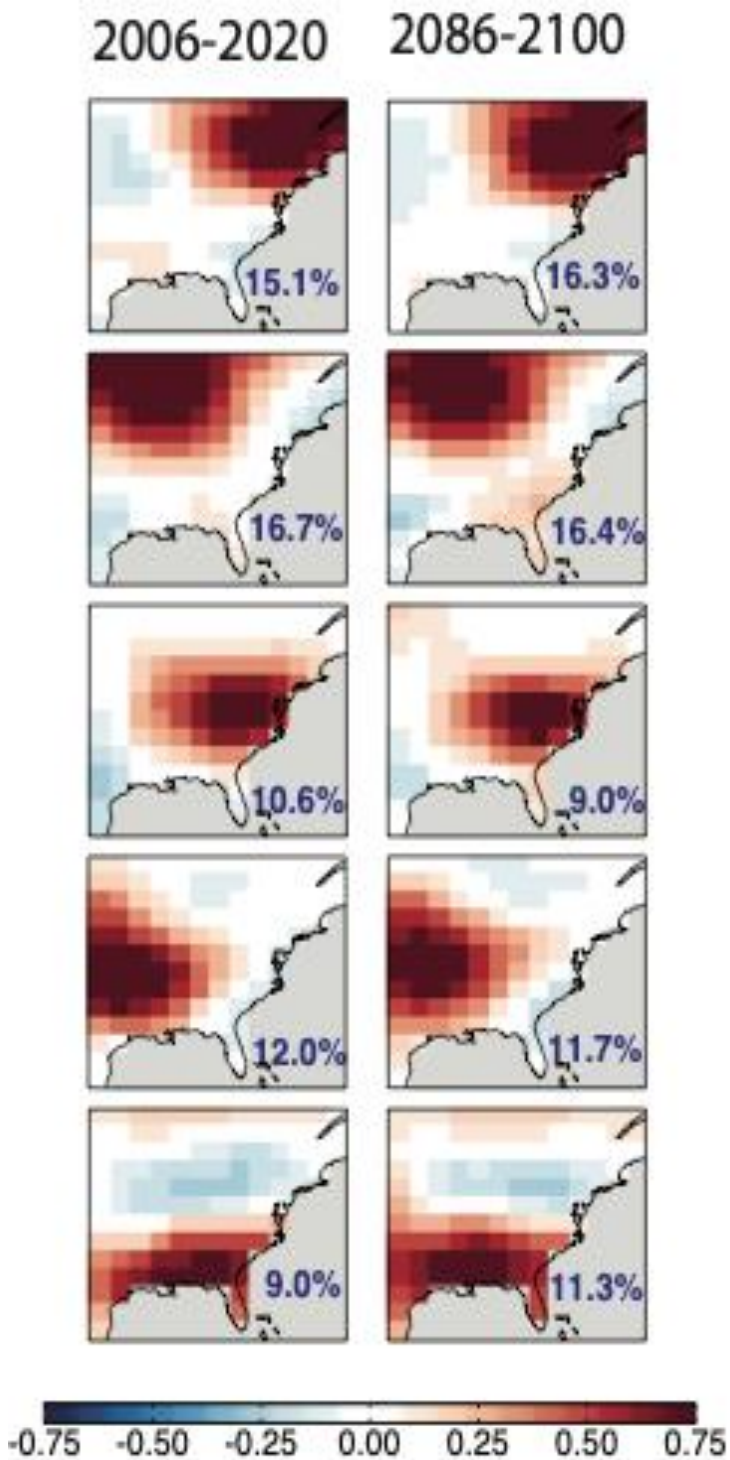

Figure S2. EOFs derived from daily mean $PM_{2.5}$ at the end of the century versus the beginning of the century in the GFDL-CM3 simulation (ensemble member 1).

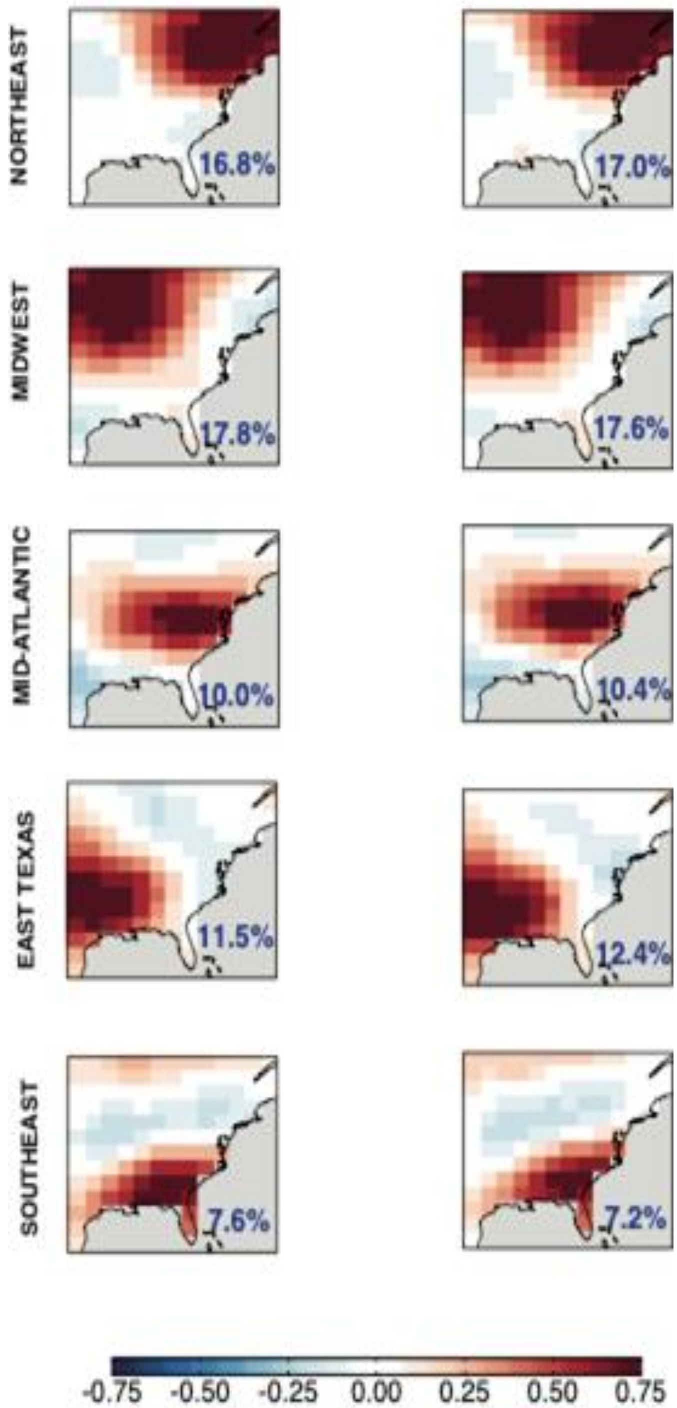

Figure S3a. EOFs derived from daily mean $PM_{2.5}$ in the two GFDL-CM3 ensemble members not shown in Figure 1. Blue text shows the percentage variance explained by each EOF.

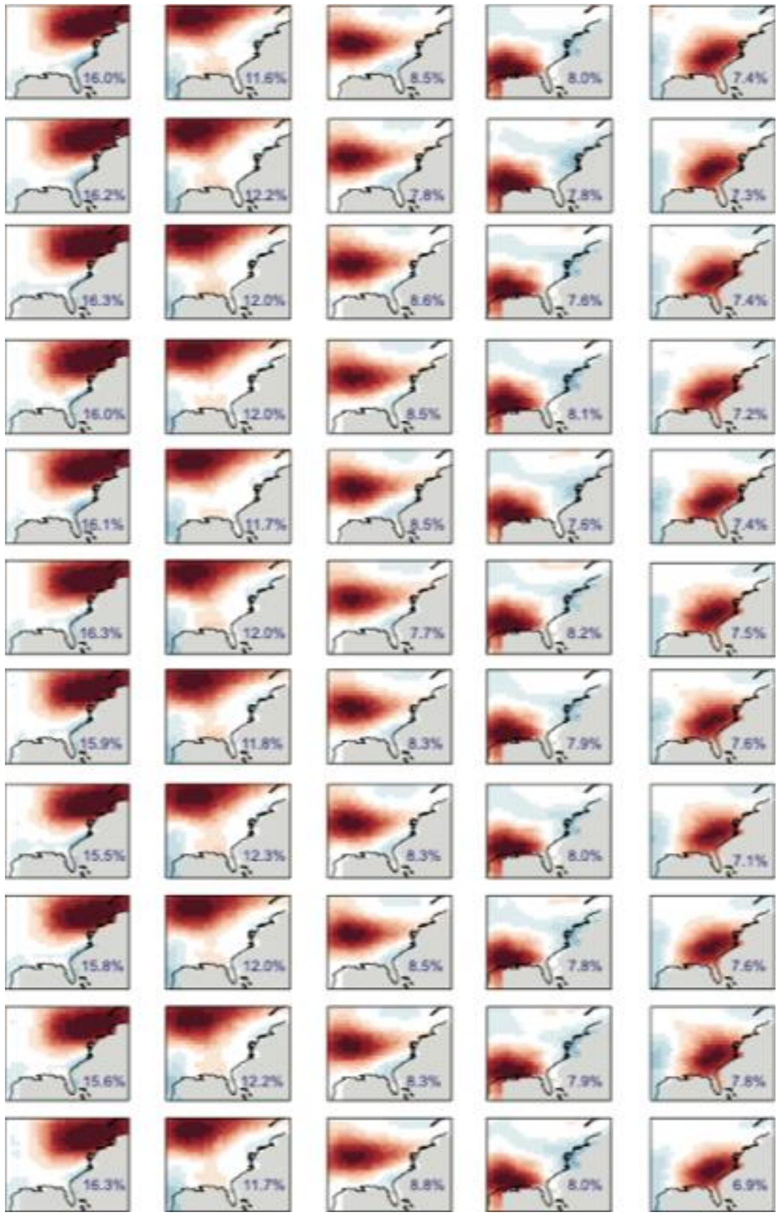

Figure S3b. EOFs derived from daily mean $PM_{2.5}$ in the 11 NCAR-CESM ensemble members not shown in Figure 1. Blue text shows the percentage variance explained by each EOF.

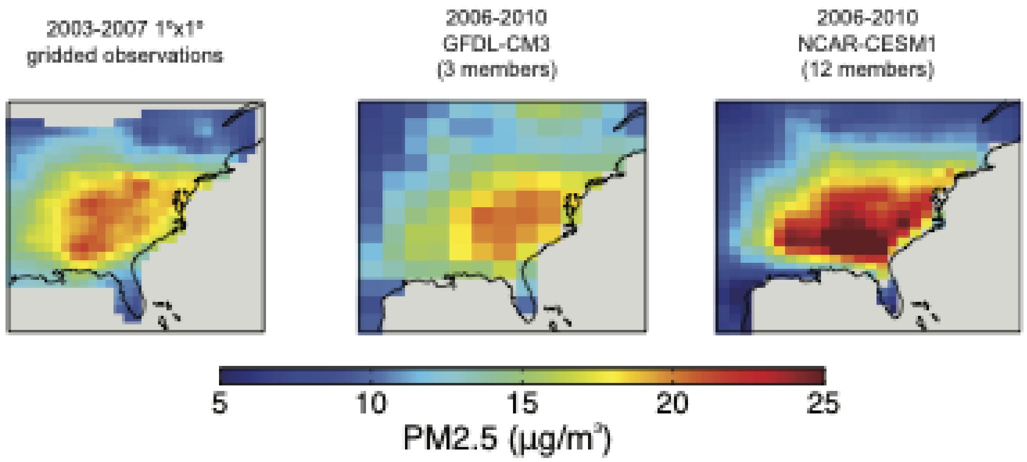

Figure S4a. Model evaluation of summertime ensemble mean $PM_{2.5.}$ with the gridded observational dataset used to derive the EOFs from daily data in Figure 1. Observations are averaged from 2003-2007, centered around 2005, the year for which emissions are perpetually repeated in the model, to avoid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

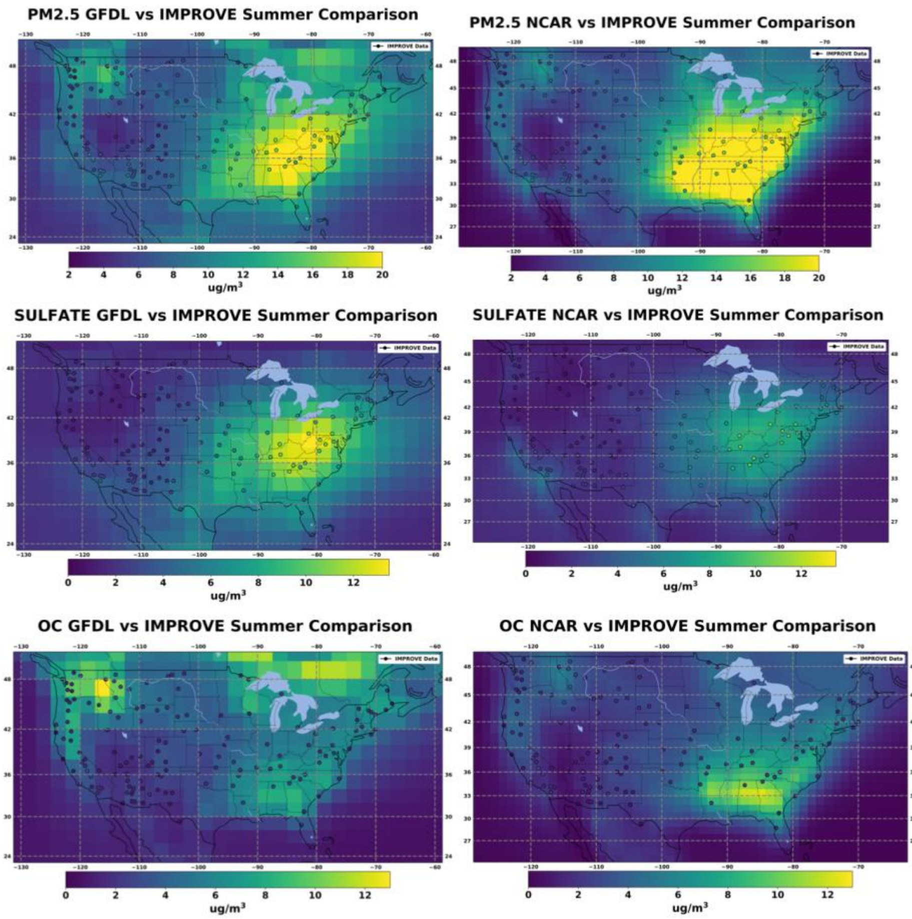

Figure S4b. Model evaluation of summertime ensemble mean $PM_{2.5.}$ with measurements from the IMPROVE network. Both models overestimate observed EUS $PM_{2.5}$ in summer (June-July-August) but differ in their simulation of individual components. Shown are summertime, ensemble mean surface PM2.5 (top), 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.

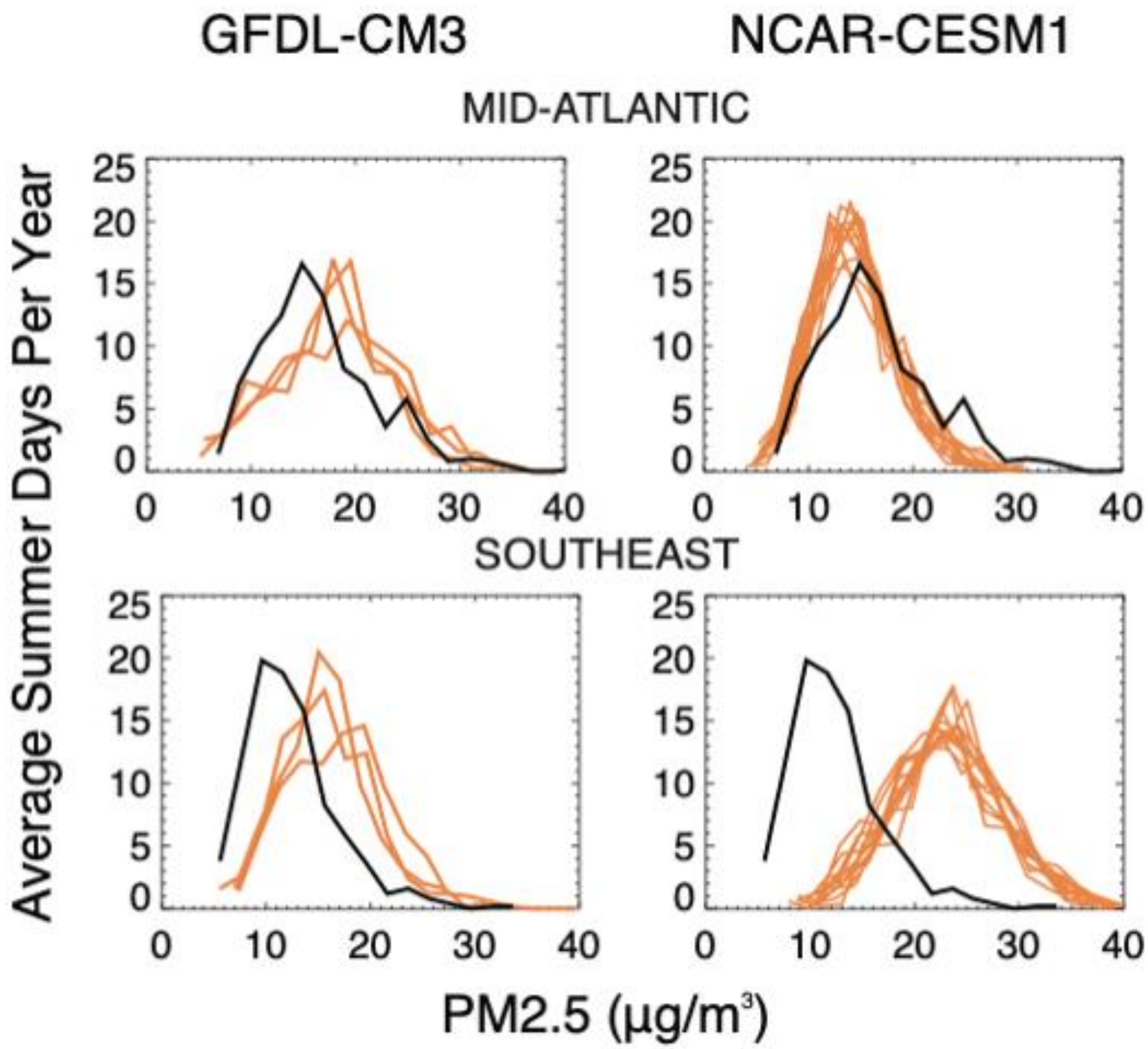

Figure S5. Average number of summer days with daily $PM_{2.5}$ falling within 2 μg m^{-3} concentration bins, regionally averaged (where EOF loading in Figure 1 exceeds > 0.5) in the observations (black) for the years 2003-2007 and in the individual ensemble members (orange) for GFDL-CM3 (left) and NCAR-CESM1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

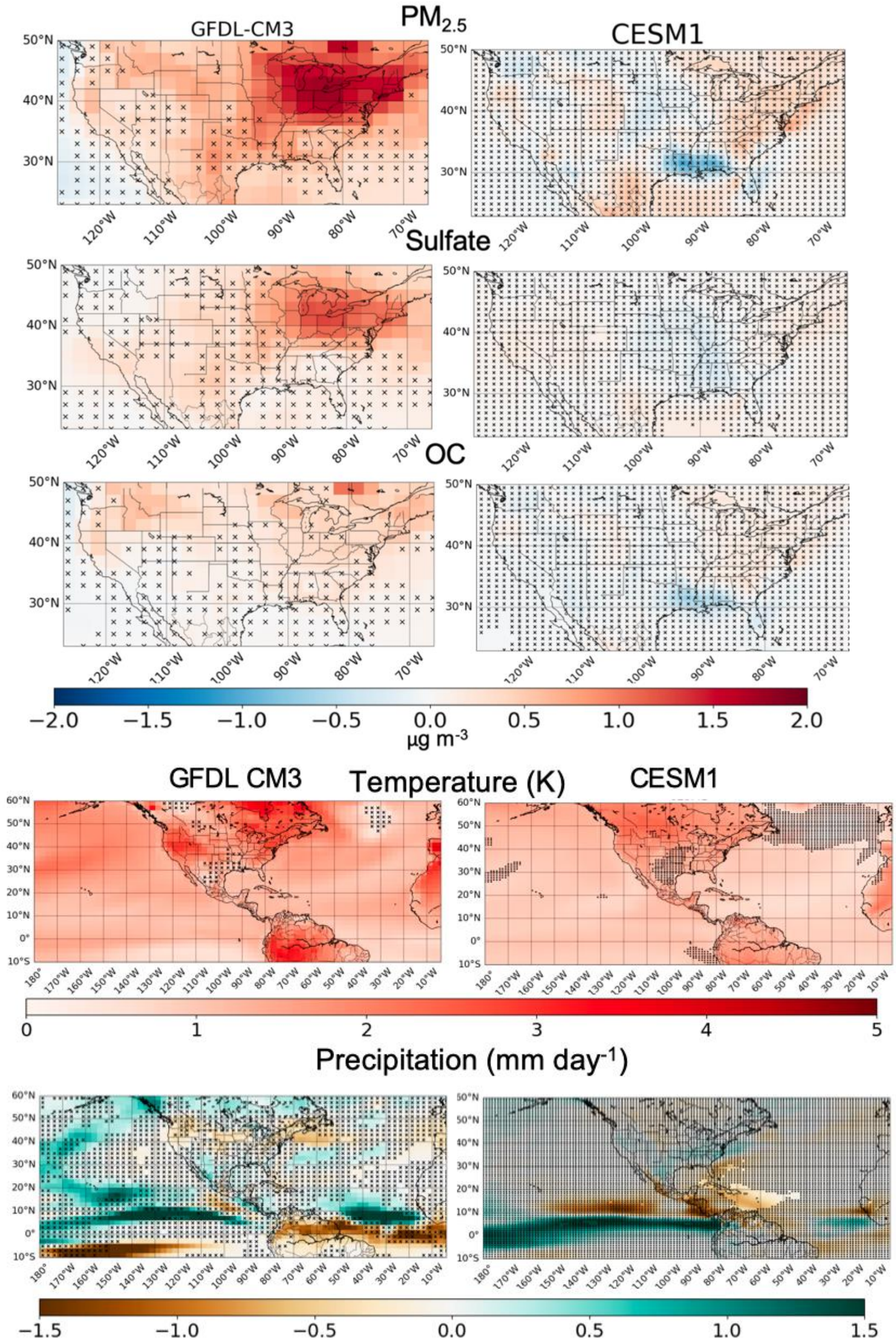

Figure S6. Change in summertime (June-July-August) $PM_{2.5}$, sulfate, organic carbon (OC), daily 2m air temperature (max for GFDL-CM3; mean for CESM1), and precipitation from

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.

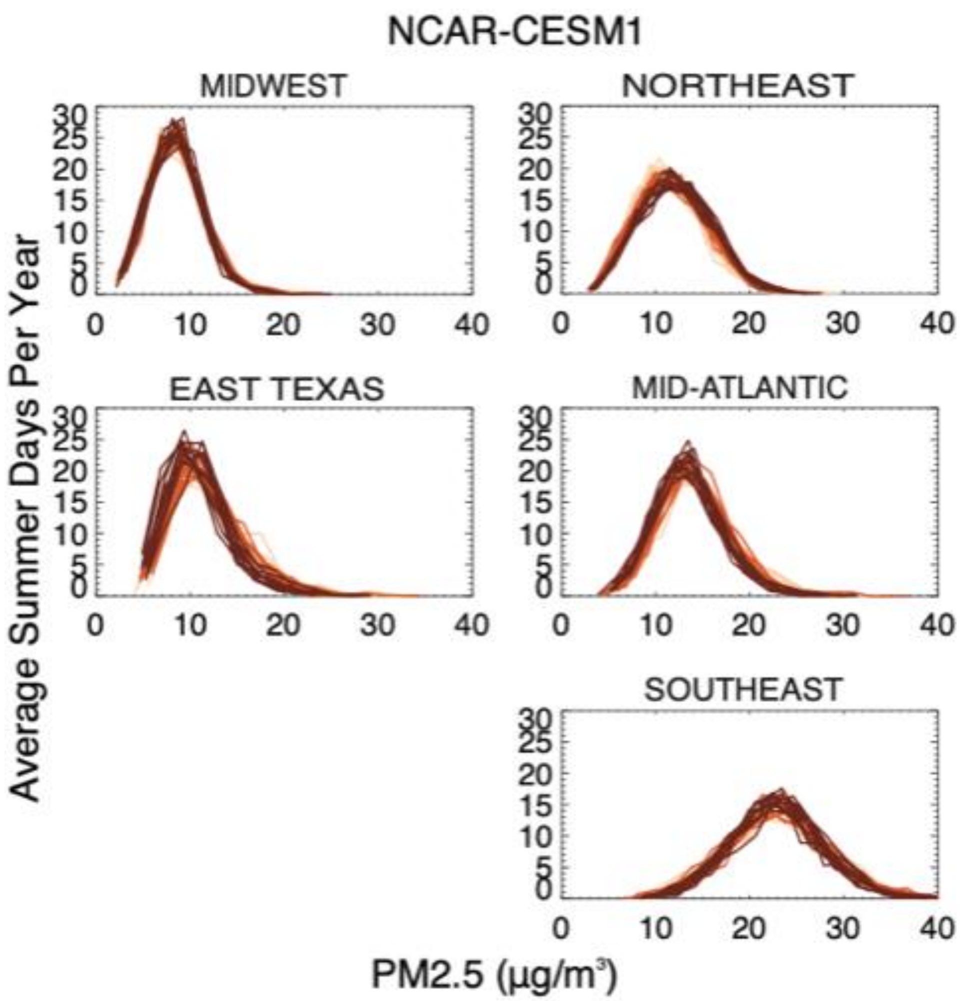

Figure S7. Little detectable change in the surface $PM_{2.5}$ distributions under the RCP8.5_WMGG scenario in the 12-member NCAR-CESM1 ensemble. Average number of summer days with daily $PM_{2.5}$ falling within 2 μg m^{-3} concentration bins, regionally averaged (where EOF loading > 0.5 in Figure 1)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

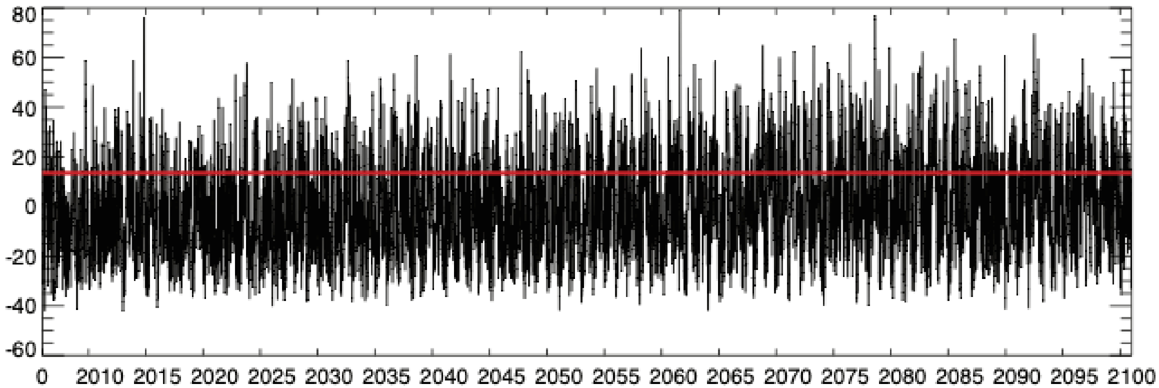

Figure S8. Principal component accompanying the Northeast EOF derived from the GFDL-CM3 model for the first of three ensemble members under the RCP8.5_WMGG scenario from 2006-2100. The red line indicates the 75th percentile.

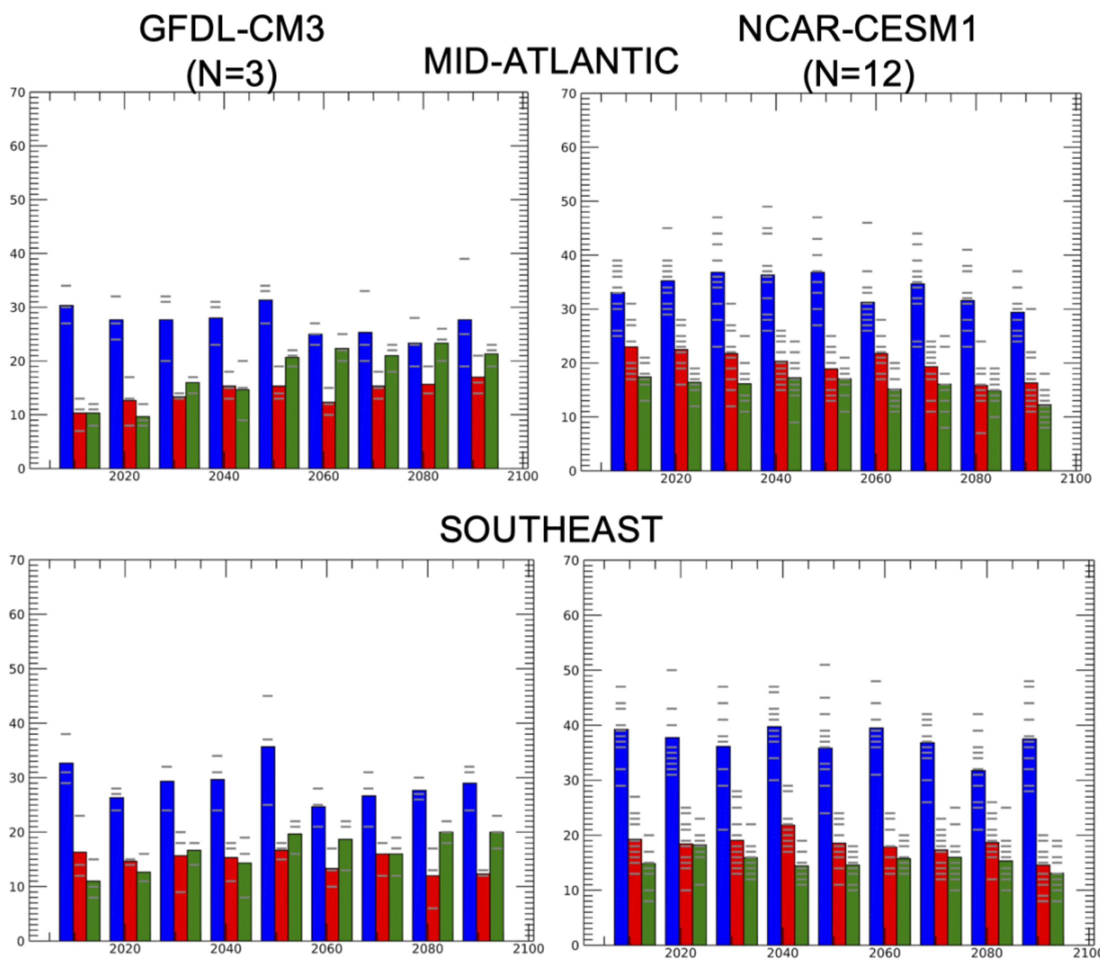

Figure S9. As in Figure 7 in the main text, but for the Mid-Atlantic and Southeast.

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 $r \geq 0.4$.