JGR Atmospheres

RESEARCH ARTICLE
10.1029/2021JD035985

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

Arlene M. Fiore[1,2,3], George P. Milly[2], Sarah E. Hancock[4,5], Laurel Quiñones[6,7], Jared H. Bowden[8], Erik Helstrom[2,9], Jean-François Lamarque[10], Jordan Schnell[11,12], J. Jason West[13], and Yangyang Xu[14]

[1]Department of Earth and Environmental Science, Columbia University, Palisades, NY, USA, [2]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA, [3]Now at Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA, [4]Department of Computer Science, Columbia University, New York, NY, USA, [5]Now at Department of Environmental Science and Engineering, Harvard University, Cambridge, MA, USA, [6]Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, USA, [7]Now at Department of Mechanical Engineering, Columbia University, New York, NY, USA, [8]Department of Applied Ecology, North Carolina State University, Raleigh, NC, USA, [9]Now at Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA, [10]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA, [11]Department of Earth and Planetary Sciences, Institute for Sustainability and Energy at Northwestern University, Evanston, IL, USA, [12]Now at Cooperative Institute for Research in Environmental Sciences, NOAA Global Systems Laboratory, University of Colorado Boulder, Boulder, CO, USA, [13]Department of Environmental Sciences & Engineering, University of North Carolina, Chapel Hill, NC, USA, [14]Department of Atmospheric Sciences, Texas A&M University, College Station, TX, USA

Key Points:

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

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. M. Fiore,
amfiore@mit.edu

Citation:

Fiore, A. M., Milly, G. P., Hancock, S. E., Quiñones, L., Bowden, J. H., Helstrom, E., et al. (2022). Characterizing changes in eastern U.S. pollution events in a warming world. *Journal of Geophysical Research: Atmospheres*, *127*, e2021JD035985. https://doi.org/10.1029/2021JD035985

Received 6 OCT 2021
Accepted 7 APR 2022

Abstract Risk assessments of air pollution impacts on human health and ecosystems would ideally consider a broad set of climate and emission scenarios, as well as natural internal climate variability. We analyze initial condition chemistry-climate ensembles to gauge the significance of greenhouse-gas-induced air pollution changes relative to internal climate variability, and consider 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 4°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 three days doubles in GFDL-CM3 and increases by ~50% in CESM1. In other EUS regions, inter-model differences in $PM_{2.5}$ responses to climate change cannot be explained solely 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 regional-scale pollution and heat event frequency and duration, while obviating the need to bias-correct concentration-based thresholds separately in individual models.

Plain Language Summary Prior studies concluded 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. Differentiating the relative contributions of these uncertainties 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 air pollution over five eastern U.S. regions 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, both models simulate longer-lasting and more frequent Northeast $PM_{2.5}$ episodes, which could exacerbate public health burdens, especially given correlations with temperature and ozone. Projecting changes in other Eastern United States regions is limited by inter-model differences that exceed the uncertainty attributable to climate variability. While our ensembles are small relative to those generated now 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.

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 USA (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). This correlation between air pollution and heat events reflects common meteorological drivers including local stagnation under anticyclonic conditions moreso than temperature-driven responses of precursor emissions and chemistry (e.g., Vukovich, 1995; Porter & Heald, 2019). 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; 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 (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, 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., 2019, 2020; 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 (Fiore et al., 2015; Fu & Tian, 2019; Jacob & Winner, 2009; Kirtman et al., 2013; Nolte et al., 2018; Nolte et al., 2021; 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 ("signal" due to rising greenhouse gases plus aerosols) from internal variability (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., 2012, 2015; Jacob & Winner, 2009; Weaver et al., 2009). Advances in computational power now permit large ensemble simulations with physical climate models (Deser, Knutti et al., 2012, Deser, Phillips et al., 2012, 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. 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 future regional meteorological trends (Deser, Knutti, et al., 2012, Deser, Phillips et al., 2012) to which air pollution will respond. Each ensemble member describes one possible 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. Below, we demonstrate a novel approach with chemistry-climate models to characterize the role of internal climate variability on air pollution trends, aiming to discern more clearly robust responses of air quality to anthropogenic climate change.

Schnell et al. (2014, 2015) 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 chemistry-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., Horton et al., 2012), but also each region of interest (Schnell et al., 2015; Turnock et al., 2020).

To overcome these challenges, we adopt a statistical approach that avoids individual model bias correction. 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. Specifically, we examine changes in the frequency and duration of high pollution events over five distinct EUS regions. 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 $PM_{2.5}$, ozone, 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 the impacts of climate change on air pollution during the 21st century. The GFDL-CM3 model includes fully coupled ocean-atmosphere-sea ice-dynamic vegetation land models, and stratospheric and tropospheric gas-phase chemistry and a bulk aerosol scheme (Austin et al., 2013; Donner et al., 2011; Naik et al., 2013), enabling us to examine consistent changes in $PM_{2.5}$, ozone, and temperature. Particulate matter (PM) and ozone precursor emissions are held fixed at 2005 levels as described by Clifton et al. (2014) while Well-Mixed Greenhouse Gases (WMGG) follow the RCP8.5 scenario. We refer to this scenario as "RCP8.5_WMGG". The native model resolution is a c48 cubed sphere which is post-processed to a $2° \times 2°$ 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.

We also draw on an available 12-member "RCP8.5_WMGG" ensemble with the CESM1 coupled atmosphere-ocean-sea ice-land model at $1° \times 1°$ horizontal resolution generated at NCAR (Xu & Lamarque, 2018). CESM1 includes an interactive aerosol scheme with three internally mixed modes (MAM3; Ghan et al., 2012; Liu et al., 2012). The NCAR-CESM1 RCP8.5_WMGG simulations do not include fully interactive tropospheric chemistry needed to simulate changes in oxidants, and thus we can only use them to provide additional context for the changes in high-$PM_{2.5}$ and temperature events diagnosed with the GFDL-CM3 ensemble. Aerosol and precursor emissions, as well as the monthly varying oxidant fields (ozone, hydrogen peroxide, and the hydroxyl radical from Tilmes et al., 2015), are held fixed at 2005 levels. 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).

Both models apply monthly varying dry deposition, biogenic, and biomass burning emissions that repeat every year (Lamarque et al., 2011; Naik et al., 2013). Diurnal cycles are imposed for isoprene emissions and ozone dry deposition in GFDL-CM3. Wet deposition, lightning NO_x, dust and sea salt are interactive with meteorology in both models. Dimethyl sulfide and marine organic aerosol are coupled to the simulated meteorology and thus respond to changes in climate in GFDL-CM3 (Naik et al., 2013) but are monthly varying and repeat annually in CESM1 (Lamarque et al., 2011). A simple representation of secondary organic aerosol (SOA) is included in GFDL-CM3 with monthly varying, annually repeating sources from biogenic terpenes and anthropogenic

butane calculated offline; while nitrate aerosol is included, it is not coupled to radiation (Naik et al., 2013). In CESM1, SOA precursor is emitted as a gas and converted to aerosol using a fixed yield, but nitrate is not included. The simulations in both models 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, brought about mainly by changing the meteorology.

Hourly surface ozone, daily maximum temperature at 2 m 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 SOA), the smallest size bin (of five) for dust, and the smallest two size bins (of five) for sea salt. We calculate maximum daily 8-hr average ozone (MDA8 O_3) from the hourly ozone fields. $PM_{2.5}$ is defined in CESM1 as the sum of daily mean sulfate, dust, black carbon, and primary and SOA in the Aitken and accumulation modes in the lowest atmospheric layer, which we convert from the native model mass mixing ratio (kg/kg) to mass density ($\mu g/m^3$). While neglecting nitrate aerosol could lead to underestimates in $PM_{2.5}$, we find that $PM_{2.5}$ is generally high compared to observations over the EUS in summer (Section 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 O_3 measured at U.S. and Canadian ground-based networks that were optimally interpolated to a 1° x 1° 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 1° x 1° temperature fields that Schnell and Prather (2017) regridded from the 0.5° × 0.5° European Centre for Medium-Range Weather Forecasting 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 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; polluted/clean air and hot/cold temperatures occur across the region indicated by the EOF at the same time.

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 domains shown in Figure 1. The EOFs derived from the observed MDA8 O_3 (Figure S1a) and daily maximum temperature (Figure S1b) datasets each capture 73% of the total variance. Table S1 in Supporting Information S1 lists the variance explained by the first 10 EOFs. The order in which the EOFs emerge is not identical, but the same EOFs always emerge in the top five. We apply Varimax rotation to the first five EOFs across all variables, which we selected by considering a change point in the amount of variance explained by each successive EOF derived from observations (Wilks, 1995). Varimax rotation re-distributes the total variance explained by the five retained EOFs, which can align the rotated EOFs more closely with underlying physical processes. Our analysis uses these rotated EOFs and accompanying PCs. The EOF analysis reduces the data set size for temporal analysis from the number of individual grid cells (>400) to five EUS regions. Below we refer to the EOFs by the region names shown in Figure 1. For some of our analysis we define a regional mask where the EOF loading exceeds 0.5.

Prior analysis of summertime daily ground-level ozone over the EUS revealed similar EOFs to those in Figure 1 (Eder et al., 1993; Fiore et al., 2003; Lehman et al., 2004). Our approach, and our choice to retain the first five EOFs for Varimax rotation closely follows Lehman et al. (2004). Even though the time periods in these studies differ, similar EOF patterns emerge. The EOFs derived from summertime ground-level MDA8 O_3 observations (Figure S1a) spatially correlate with those for $PM_{2.5}$ (r = 0.93–0.99, highest in the Northeast). EOFs for daily temperature (Figure S1b) also correlate with those for $PM_{2.5}$ (r = 0.85–0.95, highest in the Northeast and Upper Midwest), implying a common underlying dependence on weather-driven daily variability.

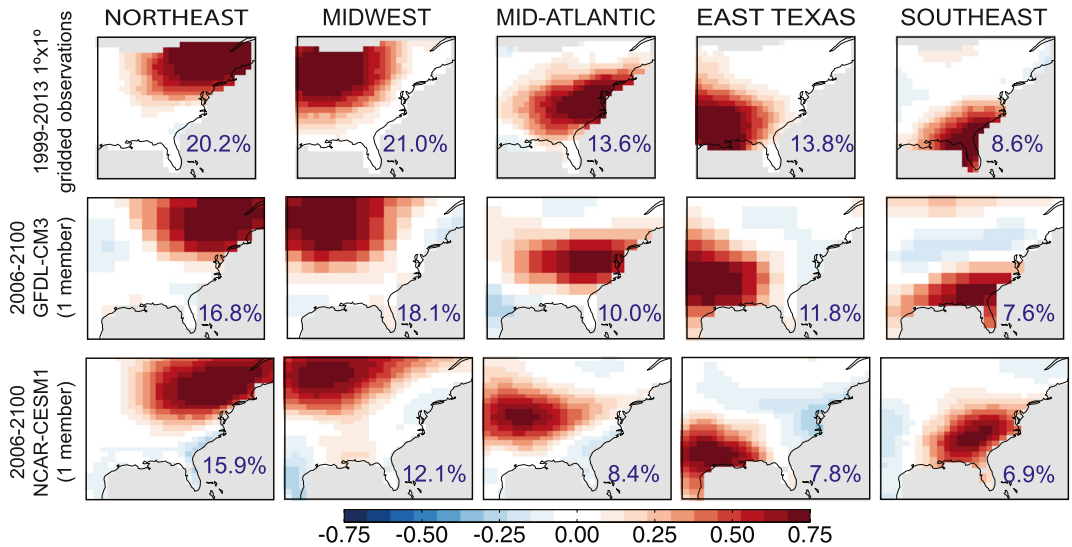

Figure 1. Regions emerging from an Empirical Orthogonal Function (EOF) analysis on standardized anomalies of summertime daily surface $PM_{2.5}$ over the Eastern United States. Shown are the Varimax-rotated EOF pattern loadings derived from (top) gridded observations (middle) one of three ensemble members 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.

We apply a parallel analysis to the model data (Figures 1, S1a, S1b); see Tables S2-S3ab in Supporting Information S1 for variance explained by the first 10 raw EOFs in GFDL-CM3 ($PM_{2.5}$, ozone and temperature) and CESM1 ($PM_{2.5}$ and temperature only). The spatial dimension decreases from 113 grid cells in GFDL-CM3 and 465 in CESM1 to five regions. De-trending prior to the EOF analysis does not change the spatial patterns. The correspondence of the five EOFs in the three variables we consider with those emerging from the ozone analysis of Lehman et al. (2004; see their Figure 2) suggests that the EOFs physically represent regions experiencing similar daily weather-driven variability. 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 data set. 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 (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 to 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 ozone events, we count the number of days on which the PC exceeds this threshold. From 1999–2005 to 2007–2013 (time periods indicated by the blue dashed vertical lines in Figure 2), the number of observed days with high pollution over the Northeast drops: from 239 to 80 days for $PM_{2.5}$ and from 221 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.

Our analysis does not focus on the magnitude of the pollution concentrations during these events. Rather, our primary interest is to define changes in event frequency and duration, and co-occurrence of high $PM_{2.5}$, ozone,

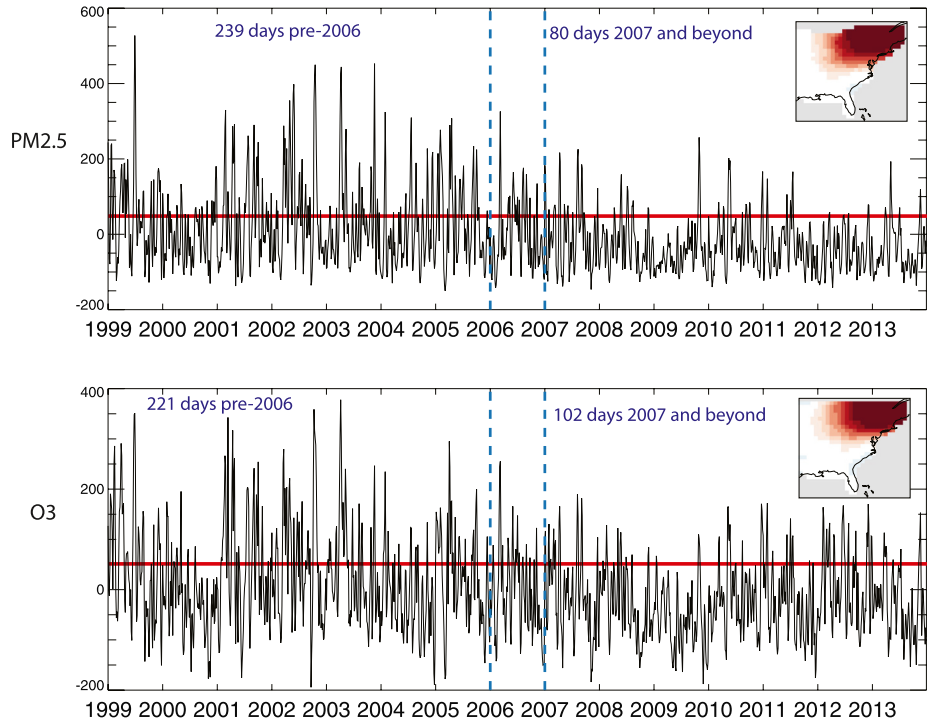

Figure 2. Proof-of-concept application of Empirical Orthogonal Function (EOF) analysis to track observed decreases in high pollution events as emission controls were phased in during the 2000s. Northeast principal components derived from observed summertime (top) daily mean $PM_{2.5}$ accompanying the EOF shown in the top left of Figure 1 and (bottom) MDA8 O_3 from 1999 to 2013 accompanying the EOF shown in the top left of Figure S1a. Shown are the 75th percentile thresholds (red lines) used to define and count the number of high regional-scale pollution events in two separate time periods (blue dashed lines).

and temperature under the RCP8.5_WMGG climate change scenario for the 21st century. In any case, the largest, longest-lasting pollution episodes – especially those with coincident high heat, high ozone, and high $PM_{2.5}$ – are typically associated with peak pollutant concentrations (Schnell & 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 stochastic climate variability present in the real atmosphere, for example year-to-year variations, and imprinted on the air pollutant measurements. Second, the simulations cannot capture observed trends due to changing anthropogenic emissions because they hold air pollutant emissions constant at 2005 levels. In light of these challenges, we adopt a statistical approach to evaluate three aspects of the GFDL-CM3 and NCAR-CESM1 $PM_{2.5}$ simulations: (a) multi-year summertime average $PM_{2.5}$ and the dominant chemical components (sulfate and organic carbon), (b) the EOFs derived from daily $PM_{2.5}$, and (c) probability distributions of regionally averaged daily $PM_{2.5}$. We also summarize prior evaluation of ground-level MDA8 O_3 and temperature in GFDL-CM3 to support the cross-correlative analysis of these variables and $PM_{2.5}$ in Section 5. 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.

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 (Figure S4a). The observed spatial mean, median, and maximum values are 14.5, 14.2, and 21.6 μg m^{-3}, respectively. The simulated statistics fall closer to observed in GFDL-CM3 (13.9, 13.9, and 20.6 μg m^{-3}) than in CESM1

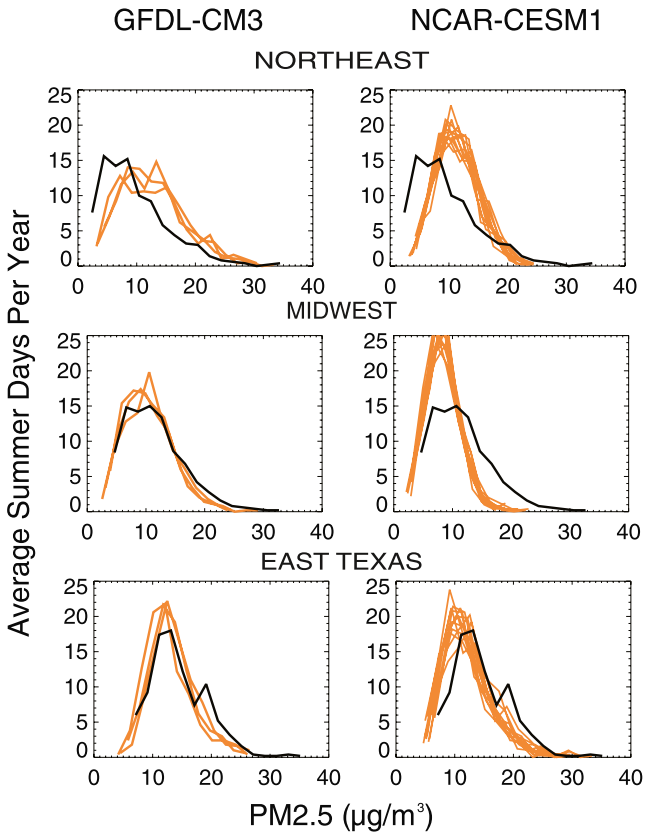

Figure 3. Distributions of the average number of summer days with regionally averaged daily $PM_{2.5}$ falling within 2 μg m^{-3} concentration bins. Averages are taken over the regions where the Empirical Orthogonal Function loading in Figure 1 exceeds >0.5 in the observations (black) for the years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

(13.7, 12.0, and 30.4 μg m^{-3}, highest over the Southeast). Comparison with the IMPROVE network (Solomon et al., 2014) suggests both models overestimate $PM_{2.5}$ by more than 2 μg m^{-3} at some EUS rural sites (Figure S4b), although the comparison in Figure S4a with gridded observations at spatial scales closer to the model horizontal resolutions is most pertinent. Evaluating chemical composition at the IMPROVE sites reveals excessive organic carbon (over 5 μg m^{-3}) in CESM1 in the Southeast, although sulfate is too low (by at least 2 μg m^{-3}; Figure S4b). GFDL-CM3 overestimates both species by at least 2 μg m^{-3} at several EUS IMPROVE sites.

EOFs derived from summertime daily $PM_{2.5}$. The EOF analysis applied to observed daily surface $PM_{2.5}$ from 1999 to 2013 identifies the same underlying modes of variability as 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 ~50% of the overall variance in the modeled data set. Four of the EOFs correspond to those derived from observations (Figure 1), but 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 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 select grid cells where the EOF loading exceeds 0.5 to define a regional mask separately for each model and the observations. We apply this mask to calculate daily regional mean $PM_{2.5}$ for 2006–2010 and 2003–2007 for the models and observations, respectively, and then sort into 2 μg m^{-3} concentration bins. The mismatch of time periods reflects a compromise to align the model with constant year 2005 emissions and the observations. The 5-year period is 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 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 $PM_{2.5}$ concentrations >26 μg m^{-3}, as in the observations. The GFDL-CM3 distributions over the Northeast, mid-Atlantic and Southeast reflect the mean positive bias evident from Figure S4ab. In CESM1, the positive bias is even higher over the Northeast (and Southeast), with little similarity to the observed distribution shape. We place more emphasis below on the GFDL-CM3 simulations from which we have ozone, $PM_{2.5}$, and temperature available. The NCAR-CESM1 simulations provide a broader context on inter-model differences and on climate variability as measured by the range across ensemble members.

Summertime MDA8 O_3 and Temperature. The GFDL-CM3 simulations include interactive gas-phase chemistry, enabling us to examine connections with $PM_{2.5}$ and temperature in Section 5. Prior evaluation indicates that the model represents changes in the mean and high tail of ground-level ozone observed at eastern U.S. Clean

Air Status and Trends Network sites during summer, despite a systematic high bias (Naik et al., 2013; Clifton et al., 2014; Rieder et al., 2015, 2018). The model has previously been shown to capture observed relationships between ozone and temperature, and between relative ozone variability and the mid-latitude jet location, over the Northeast region (Barnes & Fiore, 2013; Rasmussen et al., 2012). GFDL-CM3 also captures the spatial distribution of observed 1976–2005 summertime monthly mean and extreme temperatures (Mascioli et al., 2016). Figures S1a and S1b show matching EOF patterns derived from the RCP8.5_WMGG simulations and from observations of MDA8 O_3 and daily temperature, with the exception of the NCAR-CESM1 Mid-Atlantic EOF for temperature (recall NCAR-CESM1 does not simulate ozone interactively).

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 USA 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), respectively (Figure 4; note the colorbar saturates to allow comparison with CESM1 and the individual components, with a maximum increase of 4.3 μg m^{-3}). We consider changes as significant if the ensemble mean change exceeds the 95th percentile of a sample constructed from differences in 2006–2025 means across pairs of ensemble members (Text S1 in Supporting Information S1). 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 by over 0.5 μg m^{-3} and 2 μg m^{-3} (maximum decrease is 2.5 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 (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 some Western regions. GFDL-CM3 also simulates organic carbon increases across the Northwest, in contrast to CESM1. While it is possible that the lack of interactive tropospheric chemistry and therefore reliance on off-line oxidant fields in CESM1 could produce some of the inter-model differences in $PM_{2.5}$, the inter-model differences in organic carbon in Figure 4 suggest larger roles for changing meteorology.

Simulated changes in average temperature and precipitation are also shown in Figures 4 and S6. Summertime daily near-surface air temperatures warm in both models, by over 2 and 4 K by mid- and end-of-century respectively. While GFDL-CM3 simulates a warmer, drier summer over the EUS, CESM1 warms but wettens although significant changes only occur over a limited area. 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 a significant increase in precipitation. Earlier work 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 ensemble member, we construct distributions by averaging $PM_{2.5}$ over each EOF region (Figure 1) on every summer day of the first, middle, and last decades of the 21st century. These distributions are displayed as boxplots in Figures 5 and S7. We consider forced changes as detected relative to internal climate variability in cases where a given statistic for all ensemble members does not overlap in different time periods. GFDL-CM3 simulates increases across the interquartile range (25th, 50th, and 75th percentiles) over the Northeast, Midwest, and Mid-Atlantic regions, whereas changes are not detectable over East Texas or the Southeast regions as these statistics overlap across ensemble members in all three time periods. Over the Northeast, the interquartile statistics also increase in CESM1, but no shift is detectable over the Midwest, and 75th percentile values decline over East Texas. These changes in the two models are consistent with the long-term trends in the PCs associated with the EOFs in Figure 1: increasing 10-summer smoothed PCs from 2006 to 2100 in the Northeast, Midwest, and Mid-Atlantic derived from the GFDL-CM3 simulations (Figure S8a), and increasing Northeast but decreasing East Texas 10-summer smoothed PCs in the NCAR-CESM1 (Figure S8b). Taken together with the mean changes in Figure 4, these findings suggest that the uncertainty in the model responses to rising greenhouse gases is larger than the uncertainty arising from internal variability, except in the Northeast where both models project $PM_{2.5}$ increases.

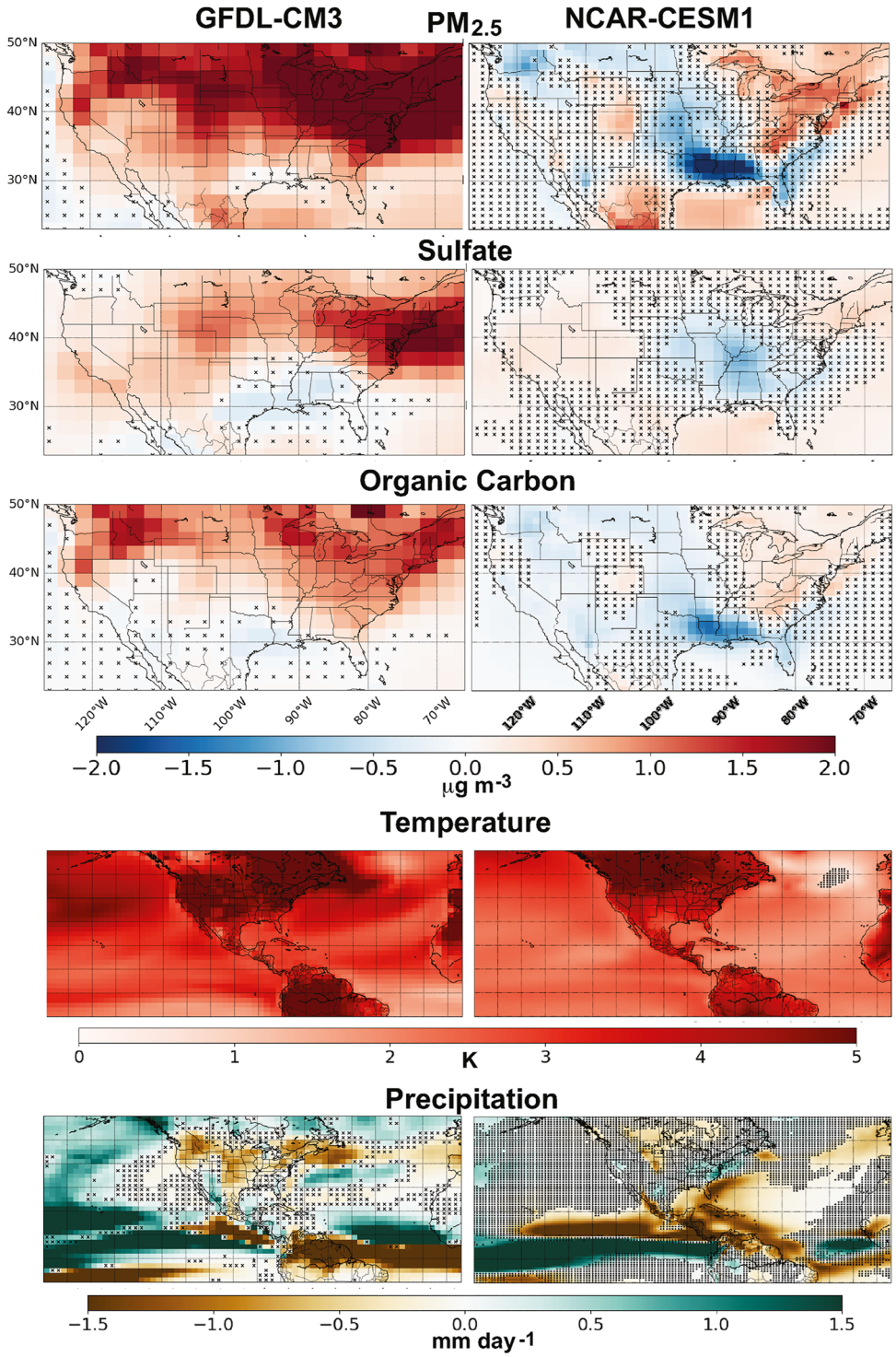

Figure 4. Change in summertime (June-July-August) $PM_{2.5}$, sulfate, organic carbon, daily 2m air temperature (max for GFDL-CM3; mean for CESM1), and precipitation from 2006 to 2025 to 2081–2100, simulated with GFDL-CM3 (left; 3 ensemble member mean) and CESM1 (right; 12 ensemble member mean) for the RCP8.5_WMGG scenario. Grid cells marked with an "x" indicate that the ensemble mean change is not significant relative to changes arising solely from internal variability (see Text S1 in Supporting Information S1 for details).

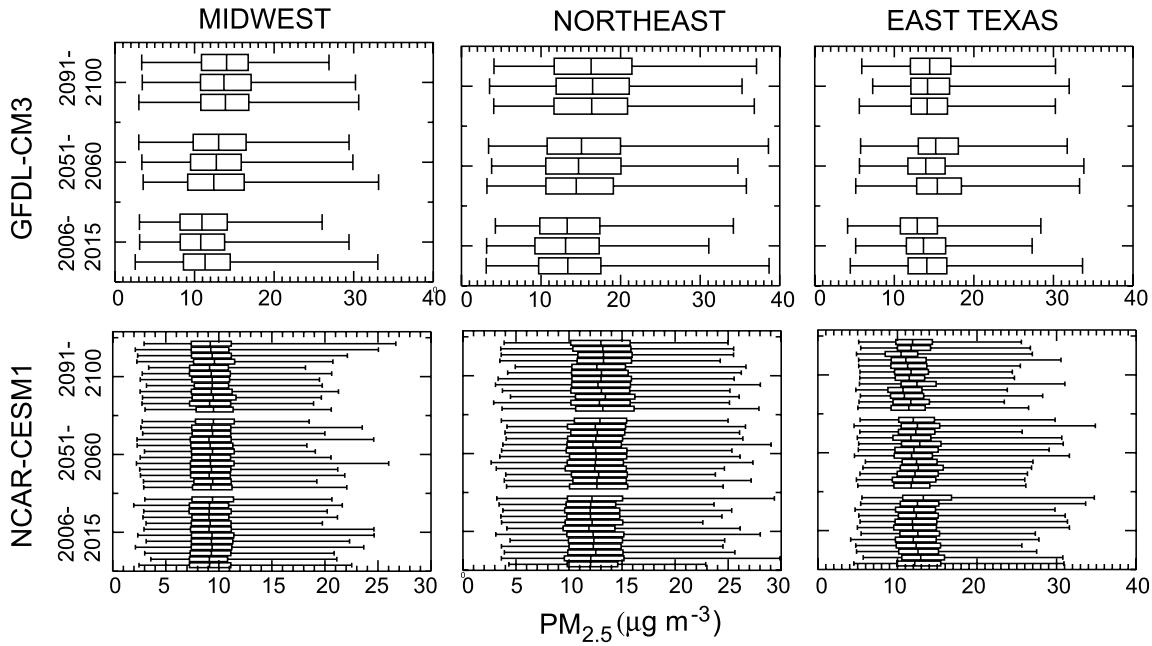

Figure 5. Distributions of regional average daily $PM_{2.5}$ (μg m^{-3}) in each ensemble member of the GFDL-CM3 model (top) and NCAR-CESM1 (bottom) RCP8.5_WMGG simulations over the Midwest (left), Northeast (middle) and East Texas (right) regions at the beginning, middle, and end of the 21st century; see Figure S7 for the Mid-Atlantic and Southeast.

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 (*e.g.,* all values above the red line in Figures S8a and S8b) 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 frequency over time across ensemble members and variables. Table S4 in Supporting Information S1 reports the GFDL-CM3 ensemble mean of these regression statistics for high-PM events, as well as ozone and temperature, in all five regions.

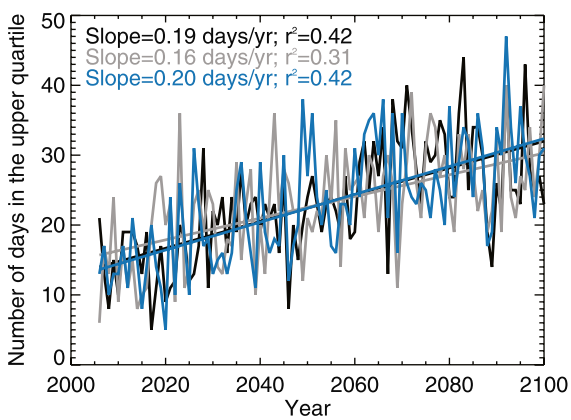

Figure 6. Increasing frequency of summertime high-$PM_{2.5}$ days over the Northeast. ''High'' days are defined as the Northeast principal component derived from summertime daily $PM_{2.5}$ falling within the upper quartile defined with respect to the full 2006–2100 period, separately for each GFDL-CM3 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 days) versus long (3+ day) 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 1–2 versus 3+ day events. We then average across all ensemble members and report the ensemble mean number of events per decade (filled circles in Figures 7 and S9). Anthropogenic climate change increases the number of 3+ day events (red symbols in Figures 7 and S9) in the GFDL-CM3 ensemble mean over the 21st century in all regions except for the Southeast, although not all increases are significant relative to internal variability or to events of different duration lengths (Section 4.3). While ensemble mean 3+ day upper quartile $PM_{2.5}$ events also increase in CESM1 over the Northeast, CESM1 simulates little change in this statistic over other regions (Figures 7 and S9).

An increase in longer duration events could occur simply because the frequency of upper quartile events, where the upper quartile is defined relative to the full 2006–2100 period, is greater toward the end of the simulation. Even a small long-term upward trend in a PC, as is evident from the smoothed

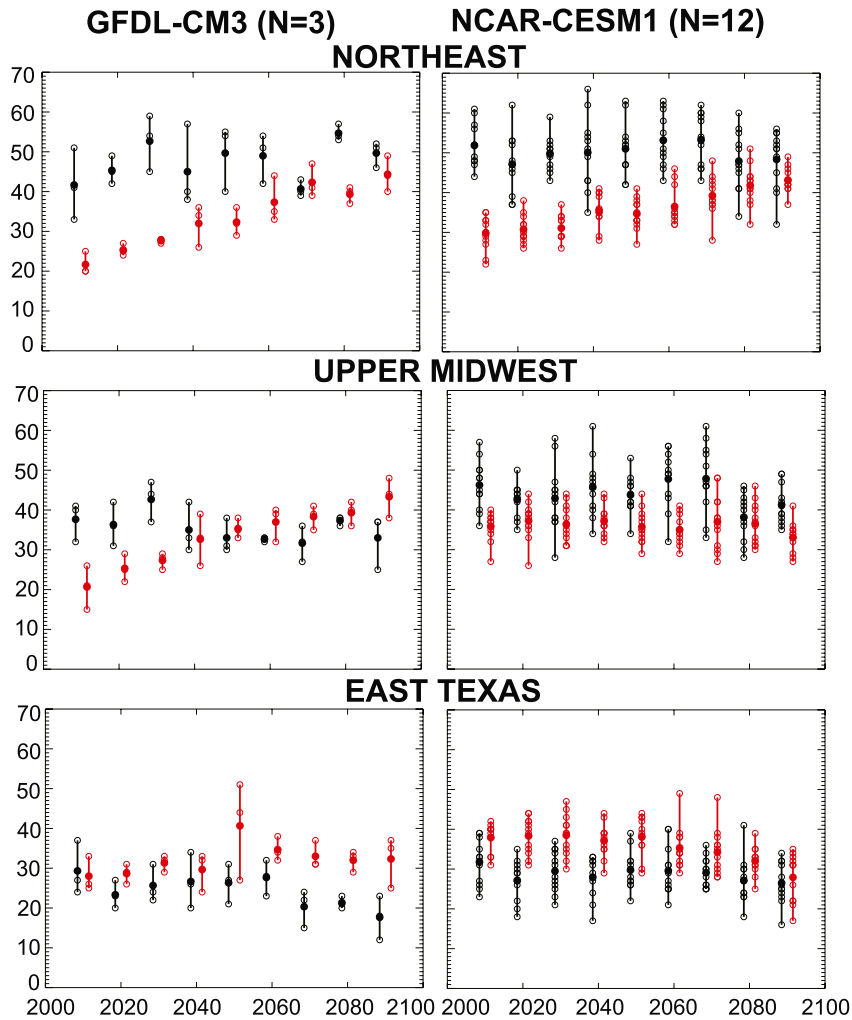

Figure 7. Longer duration upper quartile regional-scale $PM_{2.5}$ events occur more frequently under the RCP8.5_WMGG scenario in some regions in GFDL-CM3, but only over the Northeast in CESM1, while the frequency of short events changes little. Shown are the number of times the Principal Component derived from daily mean summertime $PM_{2.5}$ exceeds the upper quartile value (calculated from the full 2006–2100 time period) and stays above that value for 1–2 (black) or 3+ (red) days, summed over each decade within each ensemble member (open circles) prior to averaging over all GFDL-CM3 (left) and NCAR-CESM1 (right) ensemble members (N; filled circles) over the Northeast (top), Upper Midwest (middle) and East Texas (bottom) under the RCP8.5_WMGG scenario. The vertical range for a given decade is a measure of internal variability. A forced response to rising greenhouse gases is ''detected'' when all of the open circles in a later decade are outside the range simulated in the early decades.

black curve in Figure S8a for the GFDL-CM3 Northeast, Mid-Atlantic, and Southeast EOFs, could be responsible for increasing both event frequency and duration. Changes in duration may be relevant for understanding changes in health burdens, especially if extended duration events trigger non-linear health responses. We also aim to determine if duration has changed independently from frequency, such as might occur if regional atmospheric circulation becomes more sluggish in response to rising greenhouse gases. We examine changes in duration alone by sampling the highest decile of the PCs derived from summertime daily mean $PM_{2.5}$ in GFDL-CM3 in the first versus last three decades of the 21st century. We then calculate an average length of episode for each time period. Figure 8 shows the ensemble mean duration (colored bars) and the range over the three ensemble members (black vertical bars). The ensemble member range in the last three decades exceeds the range in the first three decades over the Northeast, Midwest and the Mid-Atlantic, with ensemble mean increases of <0.2, <0.4, and <0.6 days, respectively, suggestive of an underlying change in ventilation affecting these regions, such as a northward shift of the summertime mid-latitude jet (Barnes & Fiore, 2013; Kerr et al., 2020). These relatively small increases

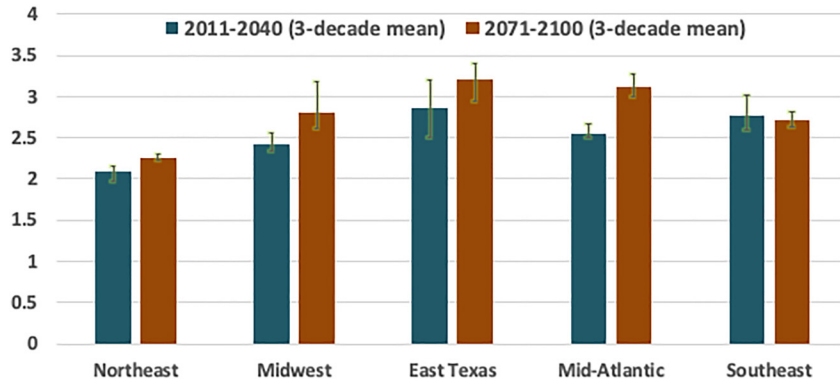

Figure 8. Increasing duration of summertime high-$PM_{2.5}$ events over the Northeast, Midwest, and Mid-Atlantic. Average length (days) of regional-scale (EOF regions in Figure 1) summertime $PM_{2.5}$ events in the beginning (green; 2011–2040) versus end (brown; 2071–2100) of the 21st century in the GFDL-CM3 RCP8.5_WMGG simulations. For each period, we sample the days when the Principal Component values derived from daily $PM_{2.5}$ fall in the upper decile, defined separately for early versus late 21st century to ensure the same number of days in each period. Vertical bars indicate the range across the three ensemble members.

in duration, however, imply that much of the change occurring in Figure 7 is due to changes in frequency. We conclude that increases in duration independent from frequency occur in the GFDL-CM3 simulations, but that the changes in event length in Figure 7 are strongly influenced by the long-term upward trend in the PC (Figure S8a) resulting in a higher frequency of upper quartile events by the end of the century (Figure 6).

As a means of gauging changes in the ''intensity'' of events, we construct regional averages of daily $PM_{2.5}$ over the five regions (where EOF loadings >0.5 in Figure 1) and report the ensemble mean changes in both models by the 2050s and 2090s in Table S5 in Supporting Information S1. Ensemble mean increases occur in this statistic across all regions within the GFDL-CM3 model (+0.8–3.7 μg m^{-3} by the 2090s; largest over the Northeast). In CESM1, the ensemble mean increases by the 2090s only over the Northeast (by 1.0 μg m^{-3}). 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 variability arising naturally (internally) in the climate system. The open circles in Figures 7 and S9 denote individual ensemble members. We first consider changes as significant if all ensemble members in the later period fall outside the range of values from the earlier period. GFDL-CM3 simulates significant changes in 3+ day events between the first three and last three decades of the 21st century over the Northeast, Upper Midwest, and mid-Atlantic (Figures 7 and S9). An increase in longer duration events over the Northeast is robustly detected for different event duration definitions (4+, 4–6, 5+, and 6+ days, Figure S10a). Increases are also detected as significant against internal variability for all event durations except 4–6 days over the Mid-Atlantic, and except for 5+ or 6+ day durations over the Upper Midwest. While the GFDL-CM3 ensemble mean number of 3+ day events also increases between the early and late 21st century over East Texas and the Southeast (Figures 7 and S9), the ensemble member ranges in early versus late decades overlap, indicating that these changes are not fully emerging from those solely due to internal climate variability. CESM1 only indicates a significant increase in the number of 3+ day events between the first and last decades over the Northeast (Figures 7 and S9), and does not simulate significant change over any other region. Furthermore, detection of Northeast changes significant against internal climate variability with CESM1 is not robust to duration length choice (Figure S10b). Our analysis below aims to differentiate the role of internal variability versus model response uncertainty in leading to these inter-model differences.

An analysis of maximum and minimum changes in the 75th percentile daily mean summertime $PM_{2.5}$ values reveals that differences in the model responses (such as arise from different meteorological responses to climate change and different representations of chemical responses) outweigh the role of climate variability (Table S5 in Supporting Information S1). The range of changes simulated by the 3-member GFDL ensemble lies completely

Table 1
Regional $PM_{2.5}$-Ozone-Temperature Relationships in Summer (June-July-August)

	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
Only Summers of 2006–2015									
Northeast	**0.55–0.57**	**0.57–0.63**	0.38–0.46	0.47–0.55	**0.61–0.67**	**0.60–0.65**	0.17–0.34	0.55–0.62	**0.70–0.72**
Mid-Atlantic	**0.69–0.72**	**0.68–0.74**	0.54–0.61	0.23–0.32	0.30–0.41	0.30–0.40	0.27–0.36	0.48–0.55	**0.57–0.62**
Upper Midwest	**0.68–0.71**	0.60–0.65	0.41–0.43	**0.44–0.54**	**0.50–0.63**	**0.42–0.54**	0.27–0.35	0.56–0.61	**0.68–0.71**
East Texas	NA	NA	NA	NA	NA	NA	0.16–0.30	0.37–0.46	**0.50–0.56**
Southeast	NA	NA	NA	NA	NA	NA	0.12–0.31	**0.31–0.49**	**0.31–0.51**
Only Summers of 2091–2100									
Northeast	0.35–0.40	0.37–0.45	0.24–0.31	0.39–0.47	**0.49–0.58**	**0.49–0.58**	0.11–0.19	0.46–0.54	**0.65–0.70**
Mid-Atlantic	**0.53–0.60**	**0.53–0.60**	0.43–0.52	0.22–0.29	0.27–0.34	0.26–0.33	0.34–0.44	0.55–0.62	**0.63–0.67**
Upper Midwest	**0.49–0.53**	0.40–0.47	0.22–0.33	**0.45–0.54**	**0.50–0.61**	**0.43–0.53**	0.27–0.34	0.53–0.58	**0.67–0.67**
East Texas	NA	NA	NA	NA	NA	NA	0.37–0.50	0.54–0.64	**0.61–0.69**
Southeast	NA	NA	NA	NA	NA	NA	0.22–0.35	**0.42–0.53**	**0.48–0.59**

Note. Minimum - maximum correlation coefficients (r) between principal components for pairs of variables simulated by the three GFDL-CM3 model ensemble members (T_{max} is daily maximum temperature at a 2m reference height; O_3 is MDA8 O_3; PM is daily mean $PM_{2.5}$) in the first versus the last decade of the simulations. 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 1 day. NA denotes $r < 0.20$; the strongest correlations with $r > 0.4$ within each region are shown in bold. Multiple bold entries for a given relationship and region indicate overlapping ranges.

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 the mid-Atlantic EOF is displaced inland in CESM1; 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 in Supporting Information S1). 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 by the 2090s 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 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, such as the multi-model 100-member ensembles now being generated for physical climate models, can further distinguish inter-model differences (model response uncertainty) from internal variability (Deser et al., 2020).

5. $PM_{2.5}$-Ozone-Temperature Linkages

We turn next to the GFDL-CM3 simulations to examine relationships between ozone, $PM_{2.5}$ and heat, and how climate change may alter these relationships. Along with the increase in upper quartile $PM_{2.5}$ events discussed in Section 4.2, GFDL-CM3 also projects more frequent ozone events in both the Northeast and the Mid-Atlantic as well as heat events (Table S4 in Supporting Information S1). Increasing exposure to multiple pollutants and heat could imply an increase in the public health burden, particularly if non-linear responses are triggered by coincident or consecutive exposure to high events. The observational analysis of Schnell and Prather (2017) indicates that EUS extreme events in temperature, MDA8 O_3 and daily mean $PM_{2.5}$ often occur within a day of each other but the specific temporal relationships vary by region. Since we find little change in the regional-scale modes of variability (EOFs) in $PM_{2.5}$, ozone, or daily T_{max} induced by rising long-lived greenhouse gases (Figure S2), we examine here any changes in the relationships between the PCs. Table 1 shows the correlations between the PCs derived from GFDL-CM3 MDA8 O_3, daily mean $PM_{2.5}$ and daily T_{max} during the first (2006–2015) versus last

(2091–2100) decade of the RCP8.5_WMGG simulations. The lags associated with the strongest correlations in Table 1 are broadly consistent with those identified by analysis of the 95th percentile of observed warm season pollution and temperature events (Schnell & Prather, 2017; see their Figure 4def), despite our use of a different metric. Modeled relationships in Table 1 are consistent with observational analysis indicating weak temperature-ozone relationships in the Southeast (e.g., Camalier et al., 2007).

Strong correlations emerge in all regions when $PM_{2.5}$ lags MDA8 O_3 by a day. $PM_{2.5}$-MDA8 O_3 correlations strengthen or remain similar from 2006–2015 to 2091–2100 in the GFDL-CM3 ensemble. Future work is needed to determine if these relationships are governed by meteorology or chemistry. For instance, enhanced ozone (and hydroxyl radical (OH)) production on one day may contribute to secondary aerosol formation that accumulates to high $PM_{2.5}$ concentrations the following day. While secondary inorganic aerosol formation is represented in GFDL-CM3, the treatment of SOA is simplified and may miss feedbacks. Since biogenic emissions and wildfires do not respond to meteorology, these simulations likely underestimate the response of air pollution to meteorology, and climate change.

The correlation of temperature with ozone is strongest for for the same day or ozone preceding temperature by a day over the Northeast and Mid-Atlantic, and when ozone precedes temperature by a day over the Upper Midwest. GFDL-CM3 projects a significantly weaker temperature-ozone correlation in all of these regions over the 21st century (Table 1). The degraded correlation between ozone 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 ozone. The absence of an ozone-temperature relationship in GFDL-CM3 agrees with earlier observation-based work showing that humidity offers more explanatory power for ozone 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. All ensemble members in both models simulate the strongest correlations with zero lag (Tables 1 and S6 in Supporting Information S1). All NCAR-CESM1 ensemble members simulate the strongest PM-temperature correlations over the Northeast (ensemble mean $r = 0.58$; Table S6 in Supporting Information S1), 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 in Supporting Information S1). 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; 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). Over the Northeast, we identify a change (weakening correlation) in the $PM_{2.5}$-T relationship across the three GFDL ensemble members from 2006–2015 to 2091–2100 (Table 1).

6. Discussion and Conclusions

Prior work has shown that some regions experiencing high pollution at present will suffer from poorer 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, Knutti et al., 2012, Deser, Phillips et al., 2012; 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 changes in mean and high air pollution events driven by rising greenhouse gases, as well as differences in model responses to rising gases. Model response differences can reflect choices regarding process representation (e.g., which emissions interact with meteorology, and fixed oxidant fields in the NCAR-CESM1 simulations).

We demonstrate how 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. This approach will work best in cases where biases are systematic, where models capture EOFs derived from observations, and where these underlying modes of spatiotemporal variability are not changing in response to forcings applied in the simulations. We find that the models agree best over the Northeast region, where summertime mean surface temperatures increase by over 4°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 3+ days over the Northeast that is significant relative to climate variability in GFDL-CM3 for the first three versus last three decades, as indicated by no overlap of ensemble members in the two periods, and significant in CESM1 from the first to the last decade (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 ozone 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 ozone 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 ozone 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. We find that T_{max} falls in the upper quartile defined relative to 2006–2100 on most summer days by the end of the century (Table S4 in Supporting Information S1).

The changes we diagnose from GFDL-CM3 imply a trend toward 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. This major source of uncertainty has been emphasized in prior studies assessed in Intergovernmental Panel on Climate Change reports (Kirtman et al., 2013). While we focused on summertime, climate change may extend what is currently "summer" weather and the accompanying pollutant concentrations over the EUS into spring and fall, as occurred in October 2010 over the Southeast, triggering high fire and biogenic emissions (Zhang & Wang, 2016). 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 than GFDL-CM3, 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 and 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 (Deser et al., 2020), for insights into projected changes in ozone 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.

Data Availability Statement

Data and code underlying our analysis and displayed in figures is available in Dryad at https://doi.org/10.5061/dryad.d2547d83s.

Acknowledgments
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, and for constructive comments from three anonymous reviewers. 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, which is a major facility sponsored by the NSF under Cooperative Agreement 1852977.

References

Abatzoglou, J. T., & Williams, A. P. (2016). Impact of anthropogenic climate change on wildfire across Western U.S. forests. *Proceedings of the National Academy of Sciences*, *113*(42), 11770–11775. Retrieved from 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. *Journal of Geophysical Research*, *115*(D2), D02303. https://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. https://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. https://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 and 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. Retrieved from 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://doi.org/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. https://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*, *117*(D22), D22307. https://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

Deser, C., Knutti, R., Solomon, S., & Phillips, A. S. (2012). Communication of the role of natural variability in future North American climate. *Nature Climate Change*, *2*(11), 775–779. https://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., Bourdette, V., & Teng, H. (2012). Uncertainty in climate change projections: The role of internal variability. *Climate Dynamics*, *38*(3–4), 527–546. https://doi.org/10.1007/s00382-010-0977-x

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*. https://doi.org/10.1175/JCLI-D-13-00451.1

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. https://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. Retrieved from 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. Retrieved from 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. Retrieved from 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. *Journal of Geophysical Research*, *108*(D14), 4431. https://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. https://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*, *111*(D12), D12306. https://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://doi.org/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* (pp. 1095–1108). Bulletin of the American Meteorological Society.

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. https://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. Retrieved from 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://doi.org/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://doi.org/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://doi.org/10.1029/2019JD030572

Kirtman, B., Power, S. B., Adedoyin, J. A., Boer, G. J., Bojariu, R., Camilloni, I., 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 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. *Atmospheric Chemistry and Physics*, *11*(19), 10031–10056. Retrieved from 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. https://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. Retrieved from 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. *Atmospheric Chemistry and Physics*, *8*(23), 7075–7086. Retrieved from http://www.atmos-chem-phys.net/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.

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. Retrieved from 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. *Geoscientific Model Development*, *5*(3), 709–739. Retrieved from https://gmd.copernicus.org/articles/5/709/2012/

Logan, J. A. (1989). Ozone in rural areas of the United States. *Journal of Geophysical Research*, *94*(D6), 8511–8532. https://doi.org/10.1029/jd094id06p08511

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 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. *Geophysical Research Letters*, *31*(24), L24103. https://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. *Atmospheric Chemistry and Physics*, *11*(10), 4679–4686. Retrieved from 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. https://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. *Atmospheric Chemistry and Physics*, *18*(20), 15471–15489. https://acp.copernicus.org/articles/18/15471/2018/

Nolte, C. G., Spero, T. L., Bowden, J. H., Sarofim, M. C., Martinich, J., & Mallard, M. S. (2021). Regional temperature-ozone relationships across the US under multiple climate and emissions scenarios. *Journal of the Air & Waste Management Association*, *71*(10), 1251–1264. https://doi.org/10.1080/10962247.2021.1970048

Oswald, E. M., Dupigny-Giroux, L.-A., Leibensperger, E. M., Poirot, R., & Merrell, J. (2015). Climate controls on air quality in the Northeastern US: An examination of summertime ozone statistics during 1993–2012. *Atmospheric Environment*, *112*, 278–288. Retrieved from 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. *Atmospheric Chemistry and Physics*, *18*(16), 11927–11948. Retrieved from 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. *Atmospheric Chemistry and Physics*, *19*(21), 13367–13381. Retrieved from 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. *Atmospheric Chemistry and Physics*, *15*(18), 10349–10366. Retrieved from 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://doi.org/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

Rasmussen, D. J., Fiore, A. M., Naik, V., Horowitz, L. W., McGinnis, S. J., & Schultz, M. G. (2012). Surface ozone-temperature relationships in the eastern U.S.: A monthly climatology for evaluating chemistry-climate models. *Atmospheric Environment*, *47*, 142–153. https://doi.org/10.1016/j.atmosenv.2011.11.021

Rieder, H. E., Fiore, A. M., Clifton, O. E., Correa, G., Horowitz, L. W., & Naik, V. (2018). Combining model projections with site-level observations to estimate changes in distributions and seasonality of ozone in surface air over the USA. *Atmospheric Environment*, *193*, 302–315. https://doi.org/10.1016/j.atmosenv.2018.07.042

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*, *120*(2), 784–800. 2014JD022303. https://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://doi.org/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. *Atmospheric Chemistry and Physics*, *14*(15), 7721–7739. Retrieved from 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. Retrieved from 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. *Atmospheric Chemistry and Physics*, *15*(18), 10581–10596. Retrieved from 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://doi.org/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*, *25*(3), 359–364. Retrieved from 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://doi.org/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 and 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 $PM_{2.5}$ concentrations and their sources in the U.S. from 1990 to 2010. *Atmospheric Chemistry and Physics*, *21*, 17115–17134. https://doi.org/10.5194/acp-21-17115-2021

Solomon, P. A., Crumpler, D., Flanagan, J. B., Jayanty, R. K. M., Rickman, E. E., & McDade, C. E. (2014). U.S. National $PM_{2.5}$ 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. *Journal of Geophysical Research*, *114*(D20), D20301. https://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. *Atmospheric Chemistry and Physics*, *19*(20), 12917–12933. Retrieved from 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://doi.org/10.1002/2016GL071731

Tai, A. P. K., Mickley, L. J., & Jacob, D. J. (2010). Correlations between fine particulate matter ($PM_{2.5}$) and meteorological variables in the United States: Implications for the sensitivity of $PM_{2.5}$ to climate change. *Atmospheric Environment*, *44*(32), 3976–3984. Retrieved from 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 ($PM_{2.5}$) air quality in the United States: Implications for $PM_{2.5}$ sensitivity to climate change. *Atmospheric Chemistry and Physics*, *12*(6), 3131–3145. Retrieved from http://www.atmos-chem-phys.net/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://doi.org/10.1016/j.atmosenv.2013.03.002

Tilmes, S., Lamarque, J.-F., Emmons, L. K., Kinnison, D. E., Ma, P.-L., Liu, X., et al. (2015). Description and evaluation of tropospheric chemistry and aerosols in the Community Earth System Model (CESM1.2). *Geoscientific Model Development*, *8*(5), 1395–1426. https://doi.org/10.5194/gmd-8-1395-2015

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. *Atmospheric Chemistry and Physics*, *20*, 14547–14579. https://doi.org/10.5194/acp-20-14547-2020

Vukovich, F. M. (1995). Regional-scale boundary layer ozone variations in the eastern United States and their association with meteorological variations. *Atmospheric Environment*, *29*(17), 2259–2273. https://doi.org/10.1016/1352-2310(95)00146-p

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

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. https://doi.org/10.1175/2009BAMS2568.1

Wilks, D. S. (1995). *Statistical methods in the atmospheric sciences*. 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. https://doi.org/10.1029/2007jd008917

Wu, X., Xu, Y., Kumar, R., & Barth, M. (2019). Separating Emission and Meteorological Drivers of Mid-21st-Century Air Quality Changes in India based on Multiyear Global-Regional Chemistry-Climate Simulations. *Journal of Geophysical Research: Atmospheres*, *124*(23), 13420–13438. https://doi.org/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://doi.org/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://doi.org/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. *Atmospheric Chemistry and Physics*, *18*(13), 9861–9877. Retrieved from 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. Retrieved from https://www.pnas.org/content/pnas/113/36/10025.full.pdf