Environmental Research Letters

LETTER • OPEN ACCESS

Comparison of multiple $PM_{2.5}$ exposure products for estimating health benefits of emission controls over New York State, USA

Environmental Research Letters

OPEN ACCESS

RECEIVED
8 April 2019

REVISED
27 June 2019

ACCEPTED FOR PUBLICATION
28 June 2019

PUBLISHED
31 July 2019

LETTER

Comparison of multiple $PM_{2.5}$ exposure products for estimating health benefits of emission controls over New York State, USA

Xiaomeng Jin[1,2], Arlene M Fiore[1,2], Kevin Civerolo[3], Jianzhao Bi[4], Yang Liu[4], Aaron van Donkelaar[5], Randall V Martin[5,6,7], Mohammad Al-Hamdan[8], Yuqiang Zhang[9], Tabassum Z Insaf[10,11], Marianthi-Anna Kioumourtzoglou[12], Mike Z He[12] and Patrick L Kinney[13]

1 Department of Earth and Environmental Sciences, Columbia University, New York, NY, United States of America
2 Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, United States of America
3 New York State Department of Environmental Conservation, Albany, NY, United States of America
4 Department of Environmental Health, Emory University, Rollins School of Public Health, Atlanta, GA, United States of America
5 Department of Physics and Atmospheric Science, Dalhousie University, Halifax, NS, Canada
6 Smithsonian Astrophysical Observatory, Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts, United States of America
7 Department of Energy, Environmental & Chemical Engineering, Washington University in St. Louis, St. Louis, Missouri, United States of America
8 Universities Space Research Association, NASA Marshall Space Flight Center, Huntsville, AL, United States of America
9 Nicholas School of the Environment, Duke University, Durham, NC, United States of America
10 New York State Department of Health, Albany, NY, United States of America
11 School of Public Health, University at Albany, Rensselaer, NY, United States of America
12 Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY, United States of America
13 Department of Environmental Health, Boston University School of Public Health, Boston, MA, United States of America

E-mail: xjin@ldeo.columbia.edu

Keywords: PM2.5, mortality burden, exposure assessment, remote sensing

Supplementary material for this article is available online

Abstract

Ambient exposure to fine particulate matter ($PM_{2.5}$) is one of the top global health concerns. We estimate the $PM_{2.5}$-related health benefits of emission reduction over New York State (NYS) from 2002 to 2012 using seven publicly available $PM_{2.5}$ products that include information from ground-based observations, remote sensing and chemical transport models. While these $PM_{2.5}$ products differ in spatial patterns, they show consistent decreases in $PM_{2.5}$ by 28%–37% from 2002 to 2012. We evaluate these products using two sets of independent ground-based observations from the New York City Community Air Quality Survey (NYCCAS) Program for an urban area, and the Saint Regis Mohawk Tribe Air Quality Program for a remote area. Inclusion of satellite remote sensing improves the representativeness of surface $PM_{2.5}$ in the remote area. Of the satellite-based products, only the statistical land use regression approach captures some of the spatial variability across New York City measured by NYCCAS. We estimate the $PM_{2.5}$-related mortality burden by applying an integrated exposure-response function to the different $PM_{2.5}$ products. The multi-product mean $PM_{2.5}$-related mortality burden over NYS decreased by 5660 deaths (67%) from 8410 (95% confidence interval (CI): 4570–12 400) deaths in 2002 to 2750 (CI: 700–5790) deaths in 2012. We estimate a 28% uncertainty in the state-level $PM_{2.5}$ mortality burden due to the choice of $PM_{2.5}$ products, but such uncertainty is much smaller than the uncertainty (130%) associated with the exposure-response function.

1. Introduction

Ambient exposure to fine particulate matter (defined as particles with less than 2.5 μm in aerodynamic diameter) is associated with mortality (Dockery *et al* 1993, Di *et al* 2017), cardiovascular (Gauderman *et al* 2004, Pope *et al* 2002, 2004, 2014), respiratory (Peng *et al* 2009), and other diseases (Pope and Dockery 2012). In the past several decades, efforts have been made to reduce the emissions from stationary and mobile sources in the United States (US) under federal and state regulations (US EPA 2018a). Between 2000 and 2017, the total anthropogenic emissions over the US have declined by 83%, 52%, 47%, 27%, and 7%

for SO_2, NO_x, CO, primary $PM_{2.5}$ and non-methane volatile organic compounds respectively (US EPA 2018a), which led to a 42% decrease in the national annual average $PM_{2.5}$ (US EPA 2018b). The reduction in $PM_{2.5}$ is associated with longer life expectancy (Correia *et al* 2013, Fann *et al* 2017), and decrease in mortality burden over recent decades (Butt *et al* 2017, Wang *et al* 2017, Zhang *et al* 2018).

To quantify the health benefits of emission reduction, an important step is to determine the ambient concentration of ground-level $PM_{2.5}$. In general, ambient $PM_{2.5}$ is estimated using information from at least one of the following three categories: ground-based observations, atmospheric chemical transport model (CTM) simulations, and remote sensing observations. Early studies (e.g. Pope *et al* 2004, Jerrett *et al* 2005) relied on ground-based monitors to estimate $PM_{2.5}$ exposure. For regions without monitors, $PM_{2.5}$ distributions can be filled spatially using geostatistical interpolation techniques such as kriging (Jerrett *et al* 2005, Fann *et al* 2017) and inverse distance weighting (IDW, Lipsett *et al* 2011). Another approach is to build relationships between *in situ* observed $PM_{2.5}$ and land use, meteorological, and geospatial information using statistical methods (Henderson *et al* 2007, Paciorek and Liu 2009, Beckerman *et al* 2013, Wang *et al* 2014, Yanosky *et al* 2014), which can resolve the fine-scale $PM_{2.5}$ spatial gradient, but their skill depends on the availability of ground-based monitors (Lee *et al* 2012). CTMs simulate $PM_{2.5}$ concentrations by solving the mass continuity equations for each PM component given emissions, meteorology, and topography. CTMs have been used to estimate $PM_{2.5}$ exposure and its historical or future trends nationwide (Wang *et al* 2017, Zhang *et al* 2018) and globally (Anenberg *et al* 2010, Silva *et al* 2013, Butt *et al* 2017), and are especially valuable for regions where long-term ground-based measurements are sparse. However, CTMs generally have coarse spatial resolution ($>$ 12 km), limiting their ability to characterize air pollution at local scales (Wang *et al* 2016), and are subject to uncertain emissions, meteorology and chemical processes.

Space-based remote sensing products offer global coverage and more than two decades of continuous observations (Kaufman *et al* 1997, King *et al* 1999, Kaufman *et al* 2002). Satellite retrieved aerosol optical depth (AOD), which is a measure of total light extinction by aerosol, is correlated with the column mass of aerosols (Wang and Christopher 2003, Koelemeijer *et al* 2006). Satellite-derived AOD is generally incorporated into estimates of $PM_{2.5}$ in surface air in two ways: (1) forward geophysical approaches that rely on CTMs to simulate the relationship between $PM_{2.5}$ and AOD (e.g. Liu *et al* 2004, van Donkelaar *et al* 2006, 2014, 2016); (2) statistical approaches that either directly build a relationship between AOD and $PM_{2.5}$ (e.g. Gupta *et al* 2006, Al-Hamdan *et al* 2009, 2014), or add AOD as a predictor along with other land use, meteorological variables in regression models (e.g. Kloog *et al* 2014, Ma *et al* 2014, Just *et al* 2015). Satellite-derived $PM_{2.5}$ is valuable for filling the spatial gaps over regions with sparse monitors (van Donkelaar *et al* 2014, 2016), providing observational constraints to models (Anenberg *et al* 2017, Lacey *et al* 2017), and improving the predictive power of statistical models (Beckerman *et al* 2013). However, using satellite AOD to predict $PM_{2.5}$, especially at shorter time scales, is challenging due to retrieval uncertainties (Martin 2008, van Donkelaar *et al* 2012, Jin *et al* 2019), missing data due to the inability to retrieve over cloud and snow (Gupta and Christopher 2008, Levy *et al* 2009), and the dependence of $PM_{2.5}$-AOD relationship on aerosol speciation, vertical distributions, and aerosol optical properties (Chin *et al* 2002, Gupta *et al* 2006, Jin *et al* 2019).

Over the US, several $PM_{2.5}$ products have become publicly available, owing to the increasing availability of observations, both *in situ* and space-based, and ever-growing computing capacity. However, most epidemiological studies, for practical purposes, rely on a single exposure estimate (e.g. Correia *et al* 2013, Girguis *et al* 2017, Al-Hamdan *et al* 2018, Zhang *et al* 2018). Jerrett *et al* (2017) find a robust association of $PM_{2.5}$ with cardiovascular diseases using multiple $PM_{2.5}$ products, but the derived relative risk factor varies. A comparative study by McGuinn *et al* (2017) over North Carolina finds the urban-rural difference in the relative risk varies with exposure assessment methods. However, objective assessment of the exposure models has long been challenging, mostly due to the lack of externally valid observations (Jerrett *et al* 2017). To address this gap, we use independent ground-based observations to evaluate seven publicly accessible $PM_{2.5}$ products for both urban and rural environments over New York State (NYS). These products include information from ground-based observations, atmospheric models and satellite remote sensing, which cover the most commonly used and up-to-date exposure assessment methods. We then estimate decadal changes in the NYS mortality burden attributable to $PM_{2.5}$ exposure using these $PM_{2.5}$ products, and assess the extent to which health impact analyses are sensitive to the choice of exposure datasets for NYS.

2. Data and methods

2.1. $PM_{2.5}$ products

We collected seven publicly accessible $PM_{2.5}$ exposure products for NYS. These products cover the commonly used approaches to estimate $PM_{2.5}$ exposure, and most of them have been applied to health studies (table 1). Table 1 provides short names for each $PM_{2.5}$ product, along with their spatial and temporal coverage, resolution, and the data sources used to derive $PM_{2.5}$. All products span multiple years from 2002 to 2012, except the CDC WONDER product, which is only available between 2003 and 2011. We compare differences in $PM_{2.5}$ by calculating spatial, temporal and population weighted spatial root mean squared

Table 1. Summary of $PM_{2.5}$ products and ground-based observations used in this study. The spatial and temporal coverage is based on the coverage of the original dataset.

Dataset	Short name	Spatial coverage	Temporal coverage	Spatial resolution	Temporal Resolution	Reference	Data source			Example applications in health studies
							In situ	Remote Sensing	Model	
Global Geophysical Satellite-Based $PM_{2.5}$	Dalhousie_GL[a] ($PM_{2.5\_Dal\_GL}$)	Global	1998–2016	$0.01^\circ \times 0.01^\circ$	Annual	van Donkelaar *et al* (2016)	US EPA AQS[b]	MODIS[c], MISR[d] and SeaWIFS[e] AOD	GEOS-Chem (v9-01-03)	Crouse *et al* (2012), Cohen *et al* (2017)
North America GeophysicalSatellite-Based $PM_{2.5}$	Dalhousie_NA[a] ($PM_{2.5\_Dal\_NA}$)	North America	2000–2016	$0.01^\circ \times 0.01^\circ$	Monthly	van Donkelaar *et al* (2019)	US EPA AQS	MODIS[c], MISR[d] and SeaWIFS[d] AOD	GEOS-Chem (v9-01-03)	None
Statistical Satellite-Based $PM_{2.5}$	Emory[a] ($PM_{2.5\_Emory}$)	NYS	2002–2012	$1 \times 1\ km^2$	Daily	Bi *et al* (2019)	US EPA AQS	MODIS (MAIAC)[e] AOD	None	Girguis *et al* (2017)
CMAQ Simulation	CMAQ ($PM_{2.5\_CMAQ}$)	USA	2002–2012	$12 \times 12\ km^2$	Daily or Hourly	Byun and Schere (2006)	None	None	CMAQ (v4.7)	Zhang *et al* (2018)
Fused Air Quality Surface using Downscaling	FAQSD ($PM_{2.5\_FAQSD}$)	USA	2002–2012	$12 \times 12\ km^2$	Daily	Berrocal *et al* (2010, 2011)	US EPA AQS	None	CMAQ (v4.7)	Breitner *et al* (2016), Hao *et al* (2016), Bravo *et al* (2017)
AQS and Remote Sensing Merged $PM_{2.5}$	CDC WONDER[gf] ($PM_{2.5\_CDC}$)	USA	2003–2011	$10 \times 10\ km^2$	Daily	Al-Hamdan *et al* (2014)	US EPA AQS	MODIS AOD	None	McClure *et al* (2017), Al-Hamdan *et al* (2017, 2018), Loop *et al* (2018)
Inverse distance weighed AQS $PM_{2.5}$	IDW ($PM_{2.5\_IDW}$)	NYS	1999-present	$0.1^\circ \times 0.1^\circ$	Daily	US EPA (2018c)	US EPA AQS	None	None	Lipsett *et al* (2011)
US EPA Air Quality System	AQS ($PM_{2.5\_AQS}$)	USA	1999-present	Point observation	Daily (24 h average)	US EPA (2018c)				
St. Regis Mohawk Tribe Air Quality Program	SRMT ($PM_{2.5\_SRMT}$)	Northern NYS	2002–2012 (with gaps)	Point observation	Daily	Benedict (2011)				
NYC Community Air Quality Survey	NYCCAS ($PM_{2.5\_CAS}$)	New York City	2009–2016	Point observation	2-week average	Matte *et al* (2013)				

[a] The short names are mostly given as the institution of the data developers.
[b] The annual ground-based $PM_{2.5}$ from the global burden disease (GBD) database is used for the development of global $PM_{2.5}$. Over the US, the GBD ground-based $PM_{2.5}$ data are from the US EPA AQS network.
[c] MODIS: MODerate resolution imaging spectroradiometer.
[d] MISR: Multi-angle imaging spectroradiometer.
[e] SeaWiFS: sea-viewing wide field-of-view sensor.
[f] MAIAC: MODIS multi-angle implementation of atmospheric correction.
[g] The official dataset included in the Center for Disease Control and Prevention Wide-ranging ONline Data for Epidemiologic Research (CDCWONDER) database.

differences (RMSD, equations (S1)–(S3) are available online at stacks.iop.org/ERL/14/084023/mmedia), and the spatial and temporal correlation coefficients (R_s and R_T, equations (S4) and (S5)). We define two metrics to characterize the variations in $PM_{2.5}$ across multiple products: the normalized range (equation (S6)) and the uncertainty (δ_{PM}, calculated from the 95% confidence interval (CI) assuming a *t* statistical distribution; equation (S9)). Detailed methods are described in the supplementary material.

Satellite retrieved AOD products are used in four datasets, including the two Dalhousie products (Dalhousie_GL; V4.GL.02 and Dalhousie_NA; V4. NA.03), Emory and CDC WONDER, but the methods used to build the $PM_{2.5}$-AOD relationship differ. The Dalhousie products use a global CTM (GEOS-Chem) to explicitly simulate the $PM_{2.5}$-AOD relationship (van Donkelaar *et al* 2016). Although the Dalhousie products are designed for regional domains or larger, we evaluate their performance at the smaller spatial scale of a single state. The Emory product incorporates satellite AOD as a predictor along with other land use and meteorological variables to a machine learning model (random forest) (Bi *et al* 2019). The CDC WONDER product builds a linear regression model between satellite AOD and ground-based $PM_{2.5}$, and then merges satellite-derived $PM_{2.5}$ with spatially interpolated ground-based $PM_{2.5}$ (Al-Hamdan *et al* 2014). Each of these approaches uses different AOD products (table 1). Four products include simulated $PM_{2.5}$ from global or regional atmospheric chemistry models. The Dalhousie products use GEOS-Chem (v9-01-03) to simulate global distributions of $PM_{2.5}$ and AOD (van Donkelaar *et al* 2012, Boys *et al* 2014, Philip *et al* 2014). The CMAQ simulation of $PM_{2.5}$ was accessed from the US EPA Remote Sensing Information Gateway (RSIG) (US EPA, RSIG 2016). The FAQSD product fuses this CMAQ $PM_{2.5}$ with AQS observations using a space-time downscaling model (Berrocal *et al* 2010, 2011). All products except the CMAQ simulation have been calibrated or merged with ground-based observations of 24 h average $PM_{2.5}$ from the EPA Air Quality System (AQS). To assess the added value of satellite remote sensing and model, we construct another dataset that spatially interpolates the daily AQS observations within NYS using IDW.

2.2. Independent ground-based $PM_{2.5}$ observations

We use ground-based observations from the NYC Community Air Quality Survey (NYCCAS) Program to evaluate these $PM_{2.5}$ products over urban NYC. NYCCAS collected integrated samples for every 2-week period in each season from 2009 to 2016 at 150 distributed sites (figure S1) over NYC, which are chosen to represent a range of land use, traffic intensity and other characteristics (Matte *et al* 2013). While NYCCAS and filter-based AQS data are sampled with different instruments, Matte *et al* (2013) found that the two-week integrated $PM_{2.5\_CAS}$ mirrors $PM_{2.5\_AQS}$ ($R^2 = 0.96$, slope = 1.0).

Over a remote area of upstate NY, we use ground-based measurements collected by the Saint Regis Mohawk Tribe (SRMT) Air Quality Program (Benedict *et al* 2011). SRMT is located in northern NYS, situated in the northwest corner of Franklin County, bordered by St. Lawrence County (figure S1). There are two SRMT sites that collect hourly $PM_{2.5}$ samples continuously with a tapered element oscillating microbalance monitor during our study period of 2002–2012: one located in Saint Lawrence County (hereafter St. Lawrence Site, Latitude: 44.93 °N Longitude: 74.85 °W, AQS code: 360897001), providing data before August 2004; the other located in Franklin County (hereafter Franklin Site, Latitude: 44.98 °N Longitude: 74.69 °W, AQS code: 360337003), providing data since March 2009. Observations from these two sites are not included in the 24 h $PM_{2.5}$ AQS data. The St. Lawrence Site is 37 km away from the nearest 24 h AQS monitor (code: 360893001), but this AQS monitor was discontinued in 2009. Thus, there is no operational AQS site near Franklin Site after 2010, and the evaluation at the Franklin Site represents areas far from monitors (figure S1).

2.3. Calculation of the mortality burden due to $PM_{2.5}$ exposure

We estimate the mortality burden for $PM_{2.5}$ products by resampling them to a common grid of $0.01^\circ \times 0.01^\circ$. We acquire the administrative boundary shapefiles from the Database of Global Administrative Areas (GADM), extract the shapefiles for NYS, and rasterize them to the 0.01° grid, so that each grid cell belongs to one county. The excess mortality attributable to ambient exposure to $PM_{2.5}$ (ΔMort) is estimated using the health impact function (Zhang *et al* 2018):

$$\Delta\mathrm{Mort} = y_0 \times \mathrm{AF} \times \mathrm{Pop}, \tag{1}$$

where y_0 is the baseline mortality rate for specific diseases; Pop is exposed population age 25 years and older; AF is the attributable fraction, which is a function of the relative risk (RR):

$$\mathrm{AF} = 1 - 1/\mathrm{RR}. \tag{2}$$

We use the RR factors from the GBD Study 2010, based on an integrated exposure-response model of Burnett *et al* (2014) developed from a meta-analysis:

$$\text{For } C > C_0\text{: RR}(C) = 1 + \alpha(1 - \exp(-\gamma(C - C_0)^{\delta})), \tag{3}$$

$$\text{For } C < C_0\text{: RR}(C) = 1, \tag{4}$$

where C is the annual average ambient concentration of $PM_{2.5}$; C_0 is the counter-factual level below which no additional risk is assumed; α, γ, and δ are fitting parameters. We acquired the RRs along with their 95% CIs for four causes of diseases, including chronic obstructive pulmonary disease (COPD), ischemic heart disease (IHD), lung cancer (LC), and cerebrovascular

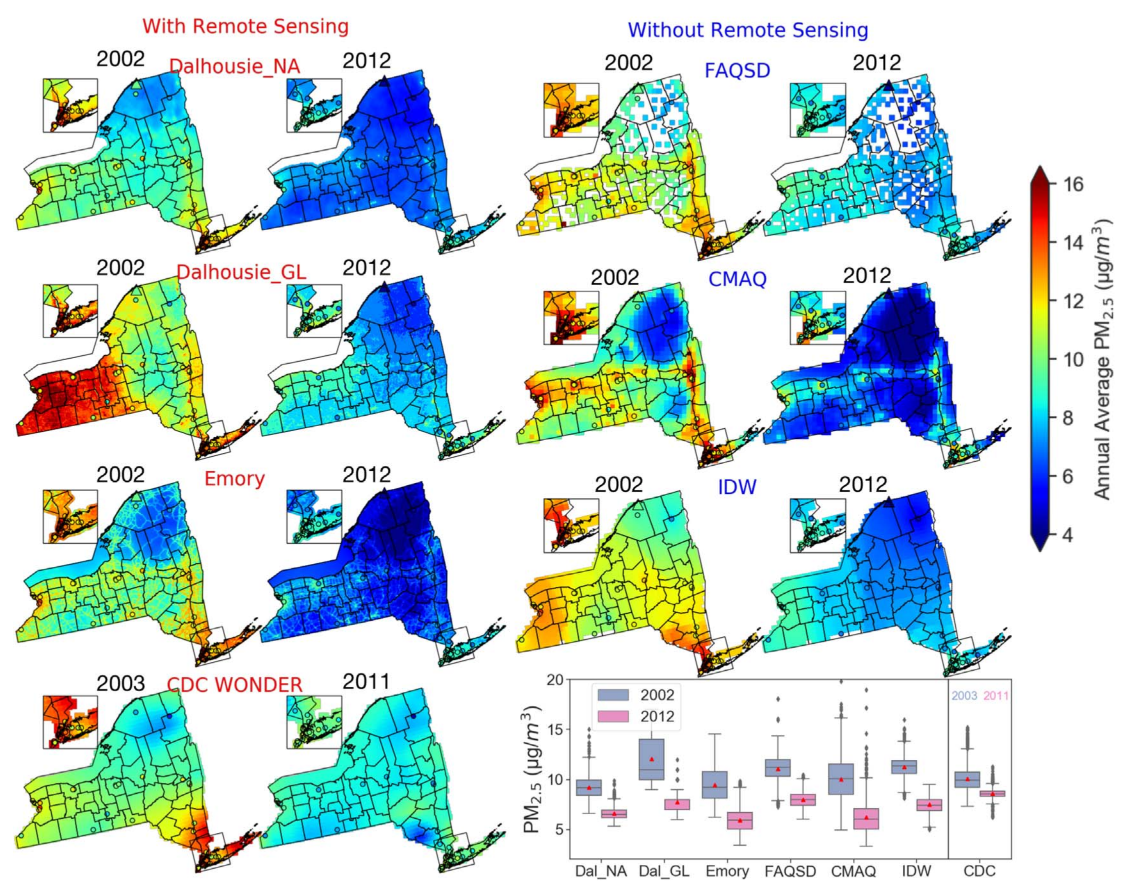

Figure 1. Annual average $PM_{2.5}$ estimated from seven $PM_{2.5}$ products in 2002 and 2012 (2003 and 2011 for CDC WONDER) over New York State, with zoom-in maps for the New York City and surrounding area (upper-left of each panel). $PM_{2.5}$ products labeled in red (left half) include information from satellite remote sensing, and those without remote sensing (right half) are labeled in blue. Annual average $PM_{2.5}$ from the AQS and SRMT sites are shown as circles and triangles respectively. The boxplot shows the range of variation of annual average $PM_{2.5}$ in 2002 and 2012 (2003 and 2011 for CDC WONDER) over New York State. The box shows the inter-quartile (IQR), and the whiskers extend to show the rest of the distribution. Outliers (defined as values either 1.5 × IQR or more above the third quartile or below the first quartile) are shown as single points. The red triangles show the spatial average $PM_{2.5}$.

and ischemic stroke (STROKE) from the Global Burden of Disease Collaborative Network (2013). We use the county-level baseline mortality rate from the National Center for Health Statistics (CDC 2017) from 2002 to 2012 for each specific disease, following the definition of the GBD study (Lim *et al* 2012, Zhang *et al* 2018). We assign the annual county-level baseline-mortality to grid cells falling in the county. County-level population data for age ≧ 25 years are acquired from the CDC WONDER database. Since the population density varies spatially within a county, we distribute the county-level population data for each county by applying the spatial patterns acquired from the Gridded Population of the World (GPW, version 4) data from the Socioeconomic Data and Applications Center (SEDAC). We acquire GPW data for 2000, 2005, and 2010, and linearly interpolate them for each year from 2002 to 2012.

3. Results

3.1. Comparison across $PM_{2.5}$ products at multiple scales

Figure 1 compares the spatial distribution of annual average $PM_{2.5}$ from multiple products in 2002 and 2012 (2003 and 2011 for $PM_{2.5\_CDC}$). The state average $PM_{2.5}$ ranges from 9.2 $\mu g\ m^{-3}$ ($PM_{2.5\_Dal\_NA}$) to 12.1 $\mu g\ m^{-3}$ ($PM_{2.5\_Dal\_GL}$) in 2002, and 5.9 $\mu g\ m^{-3}$ ($PM_{2.5\_Emory}$) to 7.9 $\mu g\ m^{-3}$ ($PM_{2.5\_FAQSD}$) in 2012 (figure 2(a)). All products show similar overall patterns with spatial correlation coefficients (R_S) ranging from 0.65 to 0.90 (table 2). The Emory product shows sharp gradients of $PM_{2.5}$ along the highways, while other products show more spatially homogeneous patterns. $PM_{2.5\_CMAQ}$ shows the largest spread in $PM_{2.5}$ across NYS, overestimating $PM_{2.5}$ over populous urban NYC and underestimating over upstate NY (compared to AQS observations, circles on figure 1), leading to a positive bias of population weighted average (PWA) $PM_{2.5}$ (figure 2(b)), and larger population weighted RMSD with other products (figure S2(b)). $PM_{2.5\_IDW}$, which only relies on the ground-based monitors, tends to smear urban-rural gradients, thus PWA $PM_{2.5\_IDW}$ is lower than other products (figure 2(b)). Excluding the IDW and CMAQ data, the other products show consistent PWA $PM_{2.5}$ with lower than 10% differences (table S1).

While the burden-of-disease studies are typically based on annual average $PM_{2.5}$, building exposure-response functions for acute effects require the $PM_{2.5}$

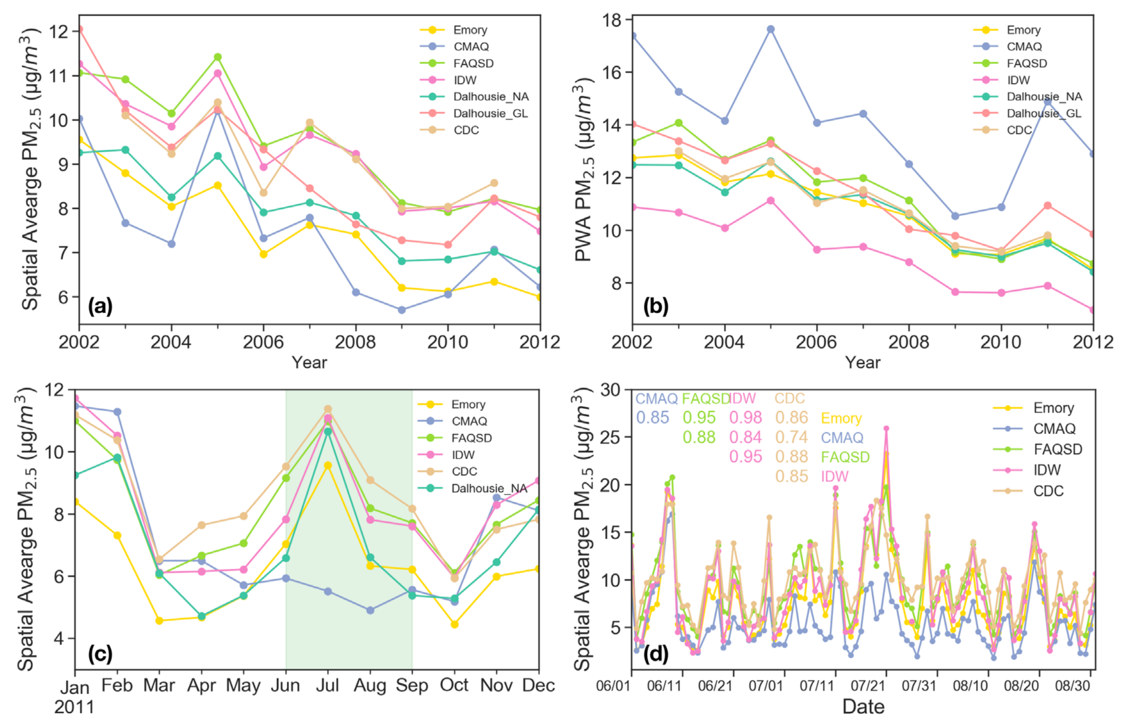

Figure 2. Time series of $PM_{2.5}$ in surface air over NYS from multiple $PM_{2.5}$ products: (a) annual spatial average, (b) annual population weighted average (PWA), (c) monthly spatial average in 2011, (d) daily spatial average in June to August 2011 (green area in (b)). The numbers in (d) are R_T during this period.

Table 2. Spatial/temporal correlation coefficients (R_S/R_T) for different pairs of $PM_{2.5}$ data. R_S is calculated from the multi-year average $PM_{2.5}$ gridded to a common grid of $0.1^\circ \times 0.1^\circ$ resolution (equation (S4)). R_T is calculated from monthly $PM_{2.5}$ averaged across NYS (equation (S5)). The dataset best correlated with independent ground-based observations is highlighted in bold. All products are sampled at each site for comparison with ground-based observations (i.e. AQS, NYCCAS, SRMT).

Dataset name		Dalhousie_GL[a]	Dalhousie_NA	Emory	CMAQ	FAQSD	IDW	CDC WONDER
Dalhousie_NA		0.90						
Emory		0.79	0.86/0.82					
CMAQ		0.82	0.86/0.55	0.85/0.32				
Fused		0.80	0.82/0.88	0.81/0.96	0.88/0.50			
AQS_IDW		0.78	0.83/0.91	0.79/0.92	0.66/0.53	0.65/0.95		
CDC		0.76	0.87/0.77	0.82/0.96	0.69/0.32	0.65/0.93	0.87/0.89	
AQS		0.72	0.88/0.97	0.91/0.99	0.76/0.40	0.87/0.98	0.94/1.0	0.81/0.98
Evaluation with independent ground-based observations								
NYCCAS		0.1	0.33/0.83	**0.62/0.94**	0.41/0.42	0.53/0.93	0.58/0.92	0.31/0.82
SRMT[b]	St. Lawrence	N/A	0.81	**0.89**	0.22	0.74	0.87	0.86
	Franklin	N/A	**0.79**	0.77	0.16	0.58	0.60	0.75

data to accurately capture the temporal variability on shorter time scales. At the monthly scale, the temporal variabilities of statewide average $PM_{2.5\_Emory}$, $PM_{2.5\_IDW}$, and $PM_{2.5\_FAQSD}$ are almost identical ($R_T > 0.9$, table 2), all closely matching the variability of $PM_{2.5\_AQS}$ ($R_T > 0.97$). $PM_{2.5\_Dal\_NA}$ and $PM_{2.5\_CDC}$ show weaker correlations with $PM_{2.5\_Emory}$, $PM_{2.5\_IDW}$, and $PM_{2.5\_FAQSD}$. $PM_{2.5\_CMAQ}$, however, shows weak to no correlation with all of the other products ($R_T < 0.55$). We attribute this difference to the seasonal cycle of $PM_{2.5\_CMAQ}$, which differs from other products (figure 2(c)). At daily scales, $PM_{2.5\_Emory}$, $PM_{2.5\_IDW}$, $PM_{2.5\_FAQSD}$ and $PM_{2.5\_CDC}$ closely match ($R_T > 0.8$, figure 2(d)). Over NYC, where ground-based monitors are densely distributed, we find consistency across all products except for $PM_{2.5\_CMAQ}$ at all scales, with $\delta_{PM} = 10\%$ for annual average $PM_{2.5}$ after excluding $PM_{2.5\_CMAQ}$ (table S1).

3.2. Evaluation with independent ground-based observations

The intensive NYCCAS measurements are ideal for evaluating whether the $PM_{2.5}$ products capture the spatial patterns of $PM_{2.5}$ at the intra-urban scale. Only

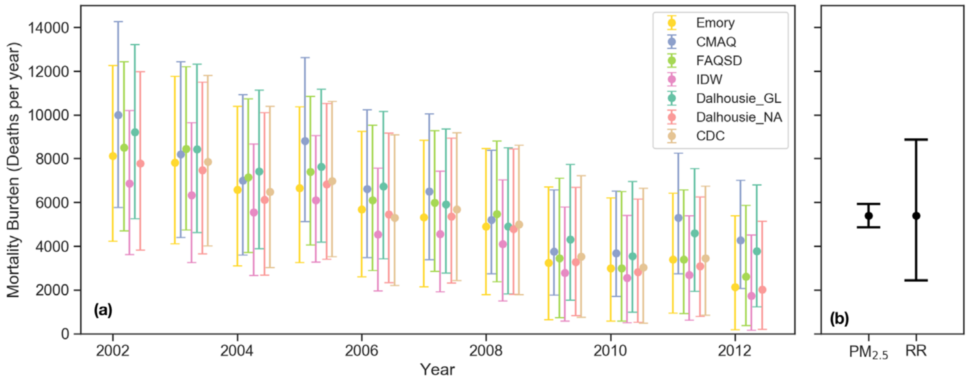

Figure 3. (a) Excess mortality burden attributable to ambient exposure of $PM_{2.5}$ over NYS from 2002 to 2012 using multiple $PM_{2.5}$ products (different colors). The error bars represent uncertainty in the integrated exposure-response function (estimated from the 95% confidence intervals of the relative risk factors). (b) The 95% confidence interval on the 11 year ensemble mean mortality burden due to uncertainty in $PM_{2.5}$ estimates (equation (S7)) versus uncertainty in the relative risk factor.

six pixels cover NYC with the ~10 km resolution of $PM_{2.5\_CMAQ}$, $PM_{2.5\_FAQSD}$, $PM_{2.5\_IDW}$ and $PM_{2.5\_CDC}$ data, but they show moderate spatial correlation with NYCCAS data with R_S ranging from 0.31 to 0.58 (table 2). The Emory product has a finer spatial resolution at 1 km, but it only shows slightly better spatial correlation with $PM_{2.5\_CAS}$ ($R_S = 0.62$). The Dalhousie products show weak ($PM_{2.5\_Dal\_NA}$: $R_S = 0.33$) to no spatial correlation ($PM_{2.5\_Dal\_GL}$: $R_S = 0.1$) with $PM_{2.5\_CAS}$, suggesting limited capability to capture the detailed spatial variability within cities, as expected by the coarser resolution inputs to those datasets. Averaging across all monitors, all products except $PM_{2.5\_CMAQ}$ show strong monthly temporal correlation with $PM_{2.5\_CAS}$ ($R_T > 0.8$, table 2). $PM_{2.5\_CMAQ}$ is overall biased high, and shows an opposite seasonal cycle to $PM_{2.5\_CAS}$ (figure S4).

To evaluate the performance of these $PM_{2.5}$ products over upstate NY, where the ground-based monitors are sparse, we use the $PM_{2.5}$ measurements from two SRMT sites (hereafter $PM_{2.5\_SRMT}$). All products correlate more strongly with $PM_{2.5\_SRMT}$ at the St. Lawrence site than the Franklin site. At the St. Lawrence site, $PM_{2.5\_Emory}$ correlates best with the observed $PM_{2.5\_SRMT}$ ($R_T = 0.89$, table 2), while $PM_{2.5\_CDC}$ has the smallest $RMSD_T$ (1.52 $\mu g\ m^{-3}$, figure S2(c)). At the monthly scale, $PM_{2.5\_IDW}$ and $PM_{2.5\_Emory}$ are more consistent with $PM_{2.5\_SRMT}$ in the cold season (November to March), and $PM_{2.5\_FAQSD}$ is more consistent with $PM_{2.5\_SRMT}$ from May to September, but overestimates $PM_{2.5}$ in winter by 33%. $PM_{2.5\_Dal\_NA}$ overestimates $PM_{2.5}$ in winter, and underestimates in the warm season (figure S4), though it captures the seasonal cycle and the temporal variability ($R_T = 0.81$). At the Franklin site, which is far from the AQS monitors, we find $PM_{2.5\_Dal\_NA}$ best captures the observed temporal variability ($R_T = 0.72$), though it is overall biased high by 40%. $PM_{2.5\_Emory}$ agrees well with $PM_{2.5\_SRMT}$ in summer, but is biased high in winter. $PM_{2.5\_CMAQ}$ shows an opposite seasonal cycle that peaks in January, leading to the lowest R_T value and highest $RMSD_T$ with $PM_{2.5\_SRMT}$ among all products (figure S4).

3.3. Decadal changes in $PM_{2.5}$ and the associated mortality burden

Despite the differences in spatial resolution and $PM_{2.5}$ derivation methods, all products (excluding the $PM_{2.5\_CDC}$) show significant decreases in statewide average $PM_{2.5}$ by 28% ($PM_{2.5\_FAQSD}$) to 37% ($PM_{2.5\_CMAQ}$) from 2002 to 2012 (figure 1). The ensemble average $PM_{2.5}$ over NYS decreased by 33% from 10.5 in 2002 to 7.0 $\mu g\ m^{-3}$ in 2012. The decreasing trend is widespread across all counties with 28%–40% decreases in the ensemble mean of county-level $PM_{2.5}$ (figure S5). The decrease in $PM_{2.5}$ is largely driven by the decrease in secondary inorganic aerosols (Boys *et al* 2014) attributed to anthropogenic emission reductions (US EPA, 2018a, 2018b). The annual average $PM_{2.5}$ shows larger decreases before 2009, and then levels off (figure 2(a)). The stabilization is partly due to the inter-annual variability in meteorology: the near-surface air temperature, which correlates with $PM_{2.5}$ over NYS (Porter *et al* 2015), is overall warmer in 2010 to 2012 than other years over NYS. Squizzato *et al* (2018) suggest $PM_{2.5}$ started to decline again over NYS since 2013.

The consistent decreasing trend provides evidence that $PM_{2.5}$-related air quality has improved significantly over NYS, which should decrease the $PM_{2.5}$-related mortality burden. We apply the integrated exposure-response function of Burnett *et al* (2014) to seven long-term $PM_{2.5}$ products. We estimate a 67% decline in the ensemble mean $PM_{2.5}$-related mortality burden (all causes combined) from 8410 (rounded to three significant figures; 95% CI due to uncertainty in relative risk factor, 4570–12 400) deaths in 2002 to 2753 (CI: 700–5790) deaths in 2012. Depending on the choice of $PM_{2.5}$ products, the estimated annual mortality burden varies from 6860 ($PM_{2.5\_IDW}$, CI: 3630–10 200) to 9990 ($PM_{2.5\_CMAQ}$, CI: 5780–14 300) deaths in 2002, and

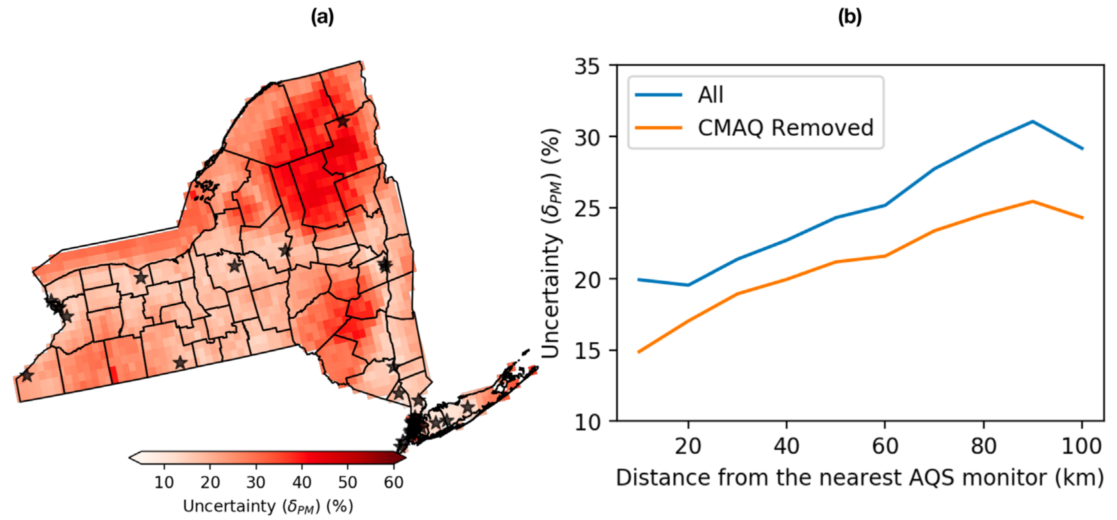

Figure 4. (a) Map of multi-product uncertainty in annual average $PM_{2.5}$ (δ_{PM}) overlaid with locations of AQS monitors in 2012. (b) δ_{PM} as a function of distance to the nearest AQS monitor with all products included (blue) and the outlier product $PM_{2.5\_CMAQ}$ excluded (orange). The calculation is performed on the re-gridded $PM_{2.5}$ products with 0.1° × 0.1° resolution.

1740 ($PM_{2.5\_IDW}$, CI: 162–4520) to 4270 ($PM_{2.5\_CMAQ}$, CI: 2080–7010) deaths in 2012. All products show consistent decreases in the mortality burden (figure 3). Using $PM_{2.5\_Emory}$ yields the largest absolute decrease in mortality burden, by 5990 (CI: 4050–6860) deaths from 2002 to 2012, while using $PM_{2.5\_IDW}$ yields the smallest decrease, by 5130 (CI: 3460–5685) deaths. In terms of relative change, using $PM_{2.5\_Emory}$, $PM_{2.5\_IDW}$, or $PM_{2.5\_Dal\_NA}$ yields the largest decrease in mortality burden (all three at 74%), while using $PM_{2.5\_CMAQ}$ gives the smallest decrease (57%). The decrease in mortality burden combines decreases in $PM_{2.5}$ with decreases in baseline mortality rates: the ensemble mean $PM_{2.5}$-related mortality burden decreases by 46% if the baseline mortality rate is kept constant at 2002 levels, and by 36% if $PM_{2.5}$ concentration is kept constant (figure S6). Among all causes, IHD is the leading cause of $PM_{2.5}$-related mortality in NYS, which contributes 87% of the total mortality (figure S7). The IHD related ensemble mean mortality decreases from 6230 (CI: 3680–8830) deaths in 2002–2030 (CI: 564–4080) deaths in 2012. NYC, the most populated and polluted region in NYS, contributes about half of the total $PM_{2.5}$-related mortality, where the ensemble mean $PM_{2.5}$-related mortality burden decreases by 62% from 4090 (CI: 2480–5690) deaths in 2002 to 1560 (CI: 525–2730) deaths in 2012 (figure S8).

4. Discussion

4.1. Which is the 'best' $PM_{2.5}$ product?

Determining which $PM_{2.5}$ product is the 'best' should take into account at least three criteria—resolution, availability and accuracy (table S2). The statistical satellite-based $PM_{2.5}$ product ($PM_{2.5\_Emory}$) has the finest spatial and temporal resolution, which captures some of the fine-scale patterns of $PM_{2.5}$ by incorporating land use and traffic-related information. Our evaluation with independent observations shows $PM_{2.5\_Emory}$ best agrees with ground-based observations for the urban area ($PM_{2.5\_CAS}$) and the rural external SRMT site that is closer to an AQS monitor. Jerrett *et al* (2017) compare the $PM_{2.5}$ mortality risk estimated using multiple exposure assessment methods, and they also find the best fit with statistical land use regression model. However, $PM_{2.5\_Emory}$ is a localized product designed for a small region (e.g. NYS in this study). The expansion of this product to wider regions is limited by the availability of ground-based monitors and consistent ancillary data. $PM_{2.5\_FAQSD}$ and $PM_{2.5\_CDC}$ are available for the entire US with daily resolution but at coarser spatial resolution (~10 km); we find $PM_{2.5\_FAQSD}$ performs better over urban areas, while $PM_{2.5\_CDC}$ performs better over remote areas (table 2). The global Dalhousie product ($PM_{2.5\_Dal\_GL}$), while limited in temporal resolution, has the widest coverage, which is valuable for assessing the $PM_{2.5}$-related global burden of disease (Cohen *et al* 2017). The regional Dalhousie product ($PM_{2.5\_Dal\_NA}$) is available monthly for North America, and it best correlates with the rural SRMT site farther from any AQS monitor (table 2). Lee *et al* (2012) compare the predictive capabilities of the Dalhousie product versus spatially interpolated $PM_{2.5}$, and they similarly find the Dalhousie product is more accurate than spatially interpolated data for areas 100 km or further away from monitors. In summary, there is no single product that stands out in all three criteria. Depending on the study design, the choice of $PM_{2.5}$ product for epidemiological studies should reflect a trade-off among these criteria.

4.2. How do $PM_{2.5}$ exposure estimates depend on ground-based measurements?

All of the $PM_{2.5}$ products in table 1 (except $PM_{2.5\_CMAQ}$) either merge AQS observations or use AQS observations to train the model, and their temporal variability is thus almost identical to $PM_{2.5\_AQS}$ at AQS sites ($R > 0.97$, table 2), indicating the important role of AQS in driving the temporal variability of these products. Areas surrounding AQS monitors typically have smaller exposure uncertainties than areas where monitors are sparse (figure 4(a)). The largest uncertainty is found over northern NYS, where only one AQS monitor is available. We find all products show better correlation and smaller $RMSD_T$ with $PM_{2.5\_SRMT}$ at the St. Lawrence site than the Franklin site, also suggesting higher confidence of these products over areas closer to AQS monitors. Figure 4(b) shows δ_{PM} as a function of distance to the nearest AQS monitor, and it increases from 20% for areas close to AQS monitors (< 20 km) to 31% for areas far from monitors (> 80 km). The global geophysical satellite $PM_{2.5}$ product ($PM_{2.5\_Dal\_GL}$) is regarded to have the least reliance on ground-based monitors (van Donkelaar *et al* 2016). The regional geophysical satellite-based product ($PM_{2.5\_Dal\_NA}$), mainly differs from $PM_{2.5\_Dal\_GL}$ in how biases are adjusted with ground-based observations. We find a large difference in spatial patterns between $PM_{2.5\_Dal\_NA}$ and $PM_{2.5\_Dal\_GL}$, especially in 2002 (figure 1), suggesting calibration with ground-based monitors is important even in the product with the least reliance on ground-based monitors. Much of NYS has sufficient monitors: more than 90% of the state area contains at least one monitor within 100 km. $PM_{2.5}$ products derived with similar approaches are likely to have larger discrepancies over regions where ground-based monitors are sparse.

4.3. What is the value of satellite remote sensing and model simulations?

Our evaluation with independent observations from SRMT suggests the inclusion of satellite remote sensing improves the representativeness of $PM_{2.5}$ in remote areas (table 2). Of the four satellite-based products, only the statistical approach ($PM_{2.5\_Emory}$) captures some of the urban spatial variability measured by NYCCAS. For the geophysical approach ($PM_{2.5\_Dal\_NA}$ and $PM_{2.5\_Dal\_GL}$), satellite AOD provides observational constraints over the globe with fine spatial resolution, which outperforms unconstrained model simulations (i.e. $PM_{2.5\_CMAQ}$), though the model simulated relationship between AOD-$PM_{2.5}$ often introduces large uncertainties (Jin *et al* 2019). For the AQS-Remote Sensing merged approach ($PM_{2.5\_CDC}$), incorporating satellite-AOD better resolves urban-rural gradients of $PM_{2.5}$ than the product spatially interpolated from AQS observations (i.e. $PM_{2.5\_IDW}$). For the statistical approach, the contribution from satellite AOD is small, less important than land use and meteorological variables (Bi *et al* 2019). Bi *et al* (2019) suggest larger enhancement of $PM_{2.5}$ over roads after incorporating satellite AOD, but the difference is generally small (<0.2 $\mu g\ m^{-3}$). Other studies that use statistical models to predict $PM_{2.5}$ find that models with satellite-based AOD better predict $PM_{2.5}$ than without (Beckerman *et al* 2013, Ma *et al* 2014) .

Among all products, $PM_{2.5\_CMAQ}$ has the least accuracy, whose monthly temporal variability is almost uncorrelated with the others, suggesting that the direct use of this CTM without observational constraints in epidemiological studies will introduce larger uncertainties in exposure estimate, consistent with Jerrett *et al* (2017). $PM_{2.5\_FAQSD}$, which fuses CMAQ with AQS data, shows a stronger correlation with other products. It should be noted that we only evaluate one single model version (CMAQ v4.7) in this study. A newer version of CMAQ (v5.2) improves the organic carbon scheme (Appel *et al* 2017, Murphy *et al* 2017), which is expected to improve the simulation of the seasonal cycle of $PM_{2.5}$. Despite the uncertainties, CTMs have the unique advantage of providing information on aerosol speciation (Di *et al* 2016, Li *et al* 2017, van Donkelaar *et al* 2019), source attribution (Lelieveld *et al* 2015, Silva *et al* 2016a, Hu *et al* 2017), and historical and future trends beyond the period of observations (Silva *et al* 2016b).

4.4. Does the choice of $PM_{2.5}$ products matter for health impact analysis?

Depending on the choice of $PM_{2.5}$ products, we show the estimated mortality burden varies by 43% (equation (S6)). On average, uncertainty in exposure-response function causes 130% uncertainty (equation (S10)) in the estimated mortality burden, which is more than a factor of 4 larger than the uncertainty due to the choice of $PM_{2.5}$ products ($\delta_{PM} = 28\%$). Previous studies similarly suggest uncertainties in exposure-response functions have larger impacts than uncertainty in exposure estimates (Silva *et al* 2013, Ford and Heald 2016). The increasing availability of observations (both *in situ* and space-based) is expected to better constrain the exposure estimate, thus to further reduce uncertainty in $PM_{2.5}$ estimates. All products show consistent decreasing trends in $PM_{2.5}$, and thus decrease in the $PM_{2.5}$-related mortality burden that varies by 26% across the different products. At low $PM_{2.5}$ levels, the relationship between $PM_{2.5}$ and relative risk is approximately linear (Burnett *et al* 2014, Di *et al* 2017), and thus the uncertainty in the exposure-response function should not strongly influence the long-term trend in the mortality burden. However, it should be noted that the integrated model of Burnett *et al* (2014) relies on pooling exposure-response functions from studies using different exposure assessment methods, and uncertainty in exposure could cause errors in building the exposure-response functions (Kioumourtzoglou *et al* 2014, Hart *et al* 2015). Besides, we only consider the uncertainties in the ambient concentration of $PM_{2.5}$, but the measured ambient concentration differs from the true personal exposure, and such difference is expected to introduce larger biases in the estimates of relative risks (Zeger *et al* 2000).

5. Conclusions

We examined seven long-term (2002–2012) publicly available $PM_{2.5}$ products over NYS, which cover the most common exposure assessment methods used in health studies. We use independent ground-based observations to evaluate these products over both urban and rural environments. Among the seven products, the localized statistical satellite-based $PM_{2.5}$ data have the finest spatial and temporal resolution, and best accuracy over areas with dense monitors, while the geophysical satellite-based product correlates best with ground-based $PM_{2.5}$ at the remote site. Inclusion of satellite remote sensing improves the representativeness of $PM_{2.5}$ estimates in a remote area. All products, however, have limited capability to resolve the spatial patterns of $PM_{2.5}$ at the intra-urban scale captured by NYCCAS. While the uncertainty in the state-level PWA $PM_{2.5}$ is small (δ_{PM} <5% after excluding outlier products), we find larger uncertainties over upstate NY where ground-based monitors are sparse. We highlight the importance of ground-based observations to reduce the uncertainties in $PM_{2.5}$ exposure estimate, as well as the independent (i.e. not used to develop the product) observations for objective assessment.

Despite these uncertainties summarized above, all products show a significant decrease of $PM_{2.5}$ by 28%–37% from 2002 to 2012, which we attribute to the implementation of emission controls. We conclude that emission controls have improved public health across NYS: the multi-product ensemble mean $PM_{2.5}$-related mortality burden decreased by 5660 deaths (67%) from 8410 (CI: 4570–12 400) deaths in 2002 to 2750 (CI: 700–5790) deaths in 2012. We estimate a 28% uncertainty in the state total mortality burden due to the choice of exposure assessment method, much less than the uncertainty in the integrated exposure-response function (130%). Overall, we conclude that exposure estimates for $PM_{2.5}$ using combinations of ground-based measurements, remotely sensed and modeled data hold substantial promise, and are rapidly becoming the state of the art for exposure assessment in epidemiological and health impact studies.

Acknowledgments

Support for this project was provided by New York State Energy Research and Development Authority (Grant number: 91268, PI: Fiore), NASA Health and Air Quality Applied Sciences Team (HAQAST, Grant NNX16AQ20G, PI: Fiore), and NASA Applied Sciences Program (Grant NNX16AQ28G, PI: Liu). We acknowledge the National Institutes of Health Institutional Research T32 Training Grant (T32 ES023770), the National Institute of Environmental Health Sciences (NIEHS) Individual Fellowship Grant (F31 ES029372) and Center Core Grant (P30 ES009089). We also acknowledge the free use of the CMAQ model and FAQSD data from EPA RSIG. We thank the New York City Department of Health and Mental Hygiene, Queens College Center for the Biology of Natural Systems, and Zev Ross Spatial Analysis for providing the NYCCAS data. Although this paper was reviewed internally, it does not necessarily reflect the views or policies of the New York State Department of Environmental Conservation.

ORCID iDs

Xiaomeng Jin https://orcid.org/0000-0002-6895-8464
Arlene M Fiore https://orcid.org/0000-0003-0221-2122
Kevin Civerolo https://orcid.org/0000-0003-1536-2664
Jianzhao Bi https://orcid.org/0000-0003-3807-6927
Yang Liu https://orcid.org/0000-0001-5477-2186
Randall V Martin https://orcid.org/0000-0003-2632-8402
Yuqiang Zhang https://orcid.org/0000-0002-9161-7086
Tabassum Z Insaf https://orcid.org/0000-0003-4725-2515
Mike Z He https://orcid.org/0000-0003-2357-3883
Patrick L Kinney https://orcid.org/0000-0003-2801-1003

References

Al-Hamdan A Z, Albashaireh R N, Al-Hamdan M Z and Crosson W L 2017 The association of remotely sensed outdoor fine particulate matter with cancer incidence of respiratory system in the USA *J. Environ. Sci. Health* A **52** 547–54

Al-Hamdan A Z, Preetha P P, Al-Hamdan M Z, Crosson W L and Albashaireh R N 2018 Reconnoitering the linkage between cardiovascular disease mortality and long-term exposures to outdoor environmental factors in the USA using remotely-sensed data *J. Environ. Sci. Health* A **53** 809–18

Al-Hamdan M Z *et al* 2009 Methods for characterizing fine particulate matter using ground observations and remotely sensed data: potential use for environmental public health surveillance *J. Air Waste Manage. Assoc.* **59** 865–81

Al-Hamdan M Z *et al* 2014 Environmental public health applications using remotely sensed data *Geocarto Int.* **29** 85–98

Anenberg S C *et al* 2017 Impacts and mitigation of excess diesel-related NO_x emissions in 11 major vehicle markets *Nature* **545** 1–17

Anenberg S C, Horowitz L W, Tong D Q and West J J 2010 An estimate of the global burden of anthropogenic ozone and fine particulate matter on premature human mortality using

atmospheric modeling *Environm. Health Perspect.* **118** 1189–95
Appel K W *et al* 2017 Description and evaluation of the Community Multiscale Air Quality (CMAQ) modeling system version 5.1 *Geosci. Model Dev.* **10** 1703–32
Beckerman B S, Jerrett M, Serre M, Martin R V, Lee S-J, van Donkelaar A, Ross Z, Su J and Burnett R T 2013 A hybrid approach to estimating national scale spatiotemporal variability of $PM_{2.5}$ in the Contiguous United States *Environ. Sci. Technol.* **47** 7233–41
Benedict A 2011 10 Years Worth of Data, Saint Regis Mohawk Tribe Air Quality Program (https://srmt-nsn.gov/_uploads/environment/aqp-10-years-worth-of-data-2011-V2.pdf) (accessed: 27 January 2019)
Berrocal V J, Gelfand A E and Holland D M 2010 A spatio-temporal downscaler for output from numerical models *J. Agric. Biol. Environ. Stat.* **15** 176–97
Berrocal V J, Gelfand A E and Holland D M 2011 Space–time data fusion under error in computer model output: an application to modeling air quality *Biometrics* **68** 837–48
Bi J, Belle J H, Wang Y, Lyapustin A I, Wildani A and Liu Y 2019 Impacts of snow and cloud covers on satellite-derived $PM_{2.5}$ levels *Remote Sens. Environ.* **221** 665–74
Boys B L *et al* 2014 Fifteen-year global time series of satellite-derived fine particulate matter *Environ. Sci. Technol.* **48** 11109–18
Burnett R T *et al* 2014 An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure *Enviorn. Health Perspect.* **122** 397–403
Butt E W, Turnock S T, Rigby R, Reddington C L, Yoshioka M, Johnson J S, Regayre L A, Pringle K J, Mann G W and Spracklen D V 2017 Global and regional trends in particulate air pollution and attributable health burden over the past 50 years *Environ. Res. Lett.* **12** 104017
Bravo M A, Ebisu K, Dominici F, Wang Y, Peng R D and Bell M L 2017 Airborne fine particles and risk of hospital admissions for understudied populations: effects by urbanicity and short-term cumulative exposures in 708 US counties *Enviorn. Health Perspect.* **125** 594–601
Breitner S *et al* 2016 Associations among plasma metabolite levels and short-term exposure to $PM_{2.5}$ and ozone in a cardiac catheterization cohort *Environ. Int.* **97** 76–84
Byun D and Schere K L 2006 Review of the governing equations, computational algorithms, and other components of the models-3 community multiscale air quality (CMAQ) modeling system *Appl. Mech. Rev.* **59** 51–77
Chin M, Ginoux P, Kinne S, Torres O, Holben B N, Duncan B N, Martin R V, Logan J A, Higurashi A and Nakajima T 2002 Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and Sun photometer measurements *J. Atmos. Sci.* **59** 461–83
Cohen A J *et al* 2017 Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the global burden of diseases study 2015 *Lancet* **389** 1907–18
Correia A W, Pope C A III, Dockery D W, Wang Y, Ezzati M and Dominici F 2013 Effect of air pollution control on life expectancy in the United States *Epidemiology* **24** 23–31
Crouse D L *et al* 2012 Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: a canadian national-level cohort study *Enviorn. Health Perspect.* **120** 708–14
Di Q, Koutrakis P and Schwartz J 2016 A hybrid prediction model for $PM_{2.5}$ mass and components using a chemical transport model and land use regression *Atmos. Environ.* **131** 390–9
Di Q, Wang Y, Zanobetti A, Wang Y, Koutrakis P, Choirat C, Dominici F and Schwartz J D 2017 Air pollution and mortality in the medicare population *New Engl. J. Med.* **376** 2513–22
Dockery D W, Pope C A, Xu X P, Spengler J D, Ware J H, Fay M E, Ferris B G and Speizer F E 1993 An association between air-pollution and mortality in 6 United-States Cities *New Engl. J. Med.* **329** 1753–9
Fann N, Kim S-Y, Olives C and Sheppard L 2017 Estimated changes in life expectancy and adult mortality resulting from declining PM2.5 exposures in the contiguous United States: 1980–2010 *Enviorn. Health Perspect.* **125** 097003
Ford B and Heald C L 2016 Exploring the uncertainty associated with satellite-based estimates of premature mortality due to exposure to fine particulate matter *Atmos. Chem. Phys.* **16** 3499–523
Gauderman W J *et al* 2004 The effect of air pollution on lung development from 10 to 18 years of age *New Engl. J. Med.* **351** 1057–67
Girguis M S, Strickland M J, Hu X, Liu Y, Chang H H, Kloog I, Belanoff C, Bartell S M and Vieira V M 2017 Exposure to acute air pollution and risk of bronchiolitis and otitis media for preterm and term infants *J. Expos. Sci. Environ. Epidemiol.* **28** 348–57
Global Burden of Disease Collaborative Network 2013 *Global Burden of Disease Study 2010 (GBD 2010)—Ambient Air Pollution Risk Model 1990–2010* (Seattle, United States: Institute for Health Metrics and Evaluation (IHME)) (http://ghdx.healthdata.org/record/global-burden-disease-study-2010-gbd-2010-ambient-air-pollution-risk-model-1990-2010) (accessed: 21 February 2019)
Gupta P and Christopher S A 2008 An evaluation of Terra-MODIS sampling for monthly and annual particulate matter air quality assessment over the Southeastern United States *Atmos. Environ.* **42** 6465–71
Gupta P, Christopher S A, Wang J, Gehrig R, Lee Y and Kumar N 2006 Satellite remote sensing of particulate matter and air quality assessment over global cities *Atmos. Environ.* **40** 5880–92
Hao Y, Strosnider H, Balluz L and Qualters J R 2016 Geographic variation in the association between ambient fine particulate matter ($PM_{2.5}$) and term low birth weight in the United States *Enviorn. Health Perspect.* **124** 250–5
Hart J E, Liao X, Hong B, Puett R C, Yanosky J D, Suh H, Kioumourtzoglou M-A, Spiegelman D and Laden F 2015 The association of long-term exposure to $PM_{2.5}$ on all-cause mortality in the Nurses' health study and the impact of measurement-error correction *Environ. Health* **14** 38
Henderson S B, Beckerman B, Jerrett M and Brauer M 2007 Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter *Environ. Sci. Technol.* **41** 2422–8
Hu J, Huang L, Chen M, Liao H, Zhang H, Wang S, Zhang Q and Ying Q 2017 Premature mortality attributable to particulate matter in china: source contributions and responses to reductions *Environ. Sci. Technol.* **51** 9950–9
Jerrett M *et al* 2005 Spatial analysis of air pollution and mortality in Los Angeles *Epidemiology* **16** 727–36
Jerrett M *et al* 2017 Comparing the health effects of ambient particulate matter estimated using ground-based versus remote sensing exposure estimates *Enviorn. Health Perspect.* **125** 1–8
Jin X, Fiore A M, Curci G, Lyapustin A, Civerolo K, Ku M, van Donkelaar A and Martin R V 2019 Assessing uncertainties of a geophysical approach to estimate surface fine particulate matter distributions from satellite-observed aerosol optical depth *Atmos. Chem. Phys.* **19** 295–313
Just A C, Wright R O, Schwartz J, Coull B A, Baccarelli A A, Tellez-Rojo M M, Moody E, Wang Y, Lyapustin A and Kloog I 2015 Using high-resolution satellite aerosol optical depth to estimate daily $PM_{2.5}$ geographical distribution in Mexico city *Environ. Sci. Technol.* **49** 8576–84
Kaufman Y J, Tanré D and Boucher O 2002 A satellite view of aerosols in the climate system *Nature* **419** 215–23
Kaufman Y J, Tanre D, Remer L A, Vermote E F, Chu A and Holben B N 1997 Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer *J. Geophys. Res.* **102** 17051–67
King M D, Kaufman Y J, Tanré D and Nakajima T 1999 Remote Sensing of Tropospheric Aerosols from Space: Past, Present, and Future *Bull. Amer. Meteor. Soc.* **80** 2229–59

Kioumourtzoglou M-A, Spiegelman D, Szpiro A A, Sheppard L, Kaufman J D, Yanosky J D, Williams R, Laden F, Hong B and Suh H 2014 Exposure measurement error in $PM_{2.5}$ health effects studies: a pooled analysis of eight personal exposure validation studies *Environ. Health* **13** 2

Kloog I, Chudnovsky A A, Just A C, Nordio F, Koutrakis P, Coull B A, Lyapustin A, Wang Y and Schwartz J 2014 A new hybrid spatio-temporal model for estimating daily multi-year $PM_{2.5}$ concentrations across northeastern USA using high resolution aerosol optical depth data *Atmos. Environ.* **95** 581–90

Koelemeijer R B A, Homan C D and Matthijsen J 2006 Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe *Atmos. Environ.* **40** 5304–15

Lacey F G, Henze D K, Lee C J, van Donkelaar A and Martin R V 2017 Transient climate and ambient health impacts due to national solid fuel cookstove emissions *Proc. Natl Acad. Sci. USA* **114** 1269

Lee S-J, Serre M, van Donkelaar A, Martin R V, Burnett R T and Jerrett M 2012 Comparison of geostatistical interpolation and remote sensing techniques for estimating long-term exposure to ambient $PM_{2.5}$ concentrations across the continental United States *Enviorn. Health Perspect.* **120** 1727–32

Lelieveld J, Evans J S, Fnais M, Giannadaki D and Pozzer A 2015 The contribution of outdoor air pollution sources to premature mortality on a global scale *Nature* **525** 367–71

Levy R C, Leptoukh G G, Kahn R, Zubko V, Gopalan A and Remer L A 2009 A critical look at deriving monthly aerosol optical depth from satellite data *IEEE Trans. Geosci. Remote Sens.* **47** 2942–56

Li C *et al* 2017 Trends in chemical composition of global and regional population-weighted fine particulate matter estimated for 25 years *Environ. Sci. Technol.* **51** 11185–95

Lim S S *et al* 2012 A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the global burden of disease study 2010 *Lancet* **380** 2224–60

Lipsett M J *et al* 2011 Long-term exposure to air pollution and cardiorespiratory disease in the California teachers study cohort *Am. J. Respir. Crit. Care Med.* **184** 828–35

Liu Y, Park R J, Jacob D J, Li Q, Kilaru V and Sarnat J A 2004 Mapping annual mean ground-level PM2.5 concentrations using multiangle imaging spectroradiometer aerosol optical thickness over the contiguous United States *J. Geophys. Res.: Atmos.* **109** doi: 10.1029/2004JD005025

Loop M S, McClure L A, Levitan E B, Al-Hamdan M Z, Crosson W L and Safford M M 2018 Fine particulate matter and incident coronary heart disease in the REGARDS cohort *Am. Heart J.* **197** 94–102

Ma Z, Hu X, Huang L, Bi J and Liu Y 2014 Estimating ground-level $PM_{2.5}$ in China using satellite remote sensing *Environ. Sci. Technol.* **48** 7436–44

Martin R V 2008 Satellite remote sensing of surface air quality *Atmos. Environ.* **42** 7823–43

McGuinn L A *et al* 2017 Fine particulate matter and cardiovascular disease_ Comparison of assessment methods for long-term exposure *Environ. Res.* **159** 16–23

Matte T D, Ross Z, Kheirbek I, Eisl H, Johnson S, Gorczynski J E, Kass D, Markowitz S, Pezeshki G and Clougherty J E 2013 Monitoring intraurban spatial patterns of multiple combustion air pollutants in New York City: Design and implementation *J. Expos. Sci. Environ. Epidemiol.* **23** 223–31

McClure L A, Loop M S, Crosson W, Kleindorfer D, Kissela B and Al-Hamdan M 2017 Fine particulate matter ($PM_{2.5}$) and the risk of stroke in the REGARDS cohort *J. Stroke Cerebrovascular Dis.* **26** 1739–44

Murphy B N *et al* 2017 Semivolatile POA and parameterized total combustion SOA in CMAQv5.2: impacts on source strength and partitioning *Atmos. Chem. Phys.* **17** 11107–33

Paciorek C J and Liu Y 2009 Limitations of remotely sensed aerosol as a spatial proxy for fine particulate matter *Enviorn. Health Perspect.* **117** 904–9

Peng R D, Bell M L, Geyh A S, McDermott A, Zeger S L, Samet J M and Dominici F 2009 Emergency admissions for cardiovascular and respiratory diseases and the chemical composition of fine particle air pollution *Enviorn. Health Perspect.* **117** 957–63

Philip S *et al* 2014 Global chemical composition of ambient fine particulate matter for exposure assessment *Environ. Sci. Technol.* **48** 13060–8

Pope C A, Burnett R T, Thun M J, Calle E E, Krewski D, Ito K and Thurston G D 2002 Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution *J. Am. Med. Assoc.* **287** 1132–41

Pope C A, Burnett R T, Thurston G D, Thun M J, Calle E E, Krewski D and Godleski J J 2004 Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease *Circulation* **109** 71–7

Pope C A III and Dockery D W 2012 Health effects of fine particulate air pollution: lines that connect *J. Air Waste Manage. Assoc.* **56** 709–42

Pope C A, Turner M C, Burnett R T, Jerrett M, Gapstur S M, Diver W R, Krewski D and Brook R D 2014 Relationships between fine particulate air pollution, cardiometabolic disorders, and cardiovascular mortality *Circ. Res.* **116** 108–15

Porter W C, Heald C L, Cooley D and Russell B 2015 Investigating the observed sensitivities of air-quality extremes to meteorological drivers via quantile regression *Atmos. Chem. Phys.* **15** 10349–66

Silva R A *et al* 2013 Global premature mortality due to anthropogenic outdoor air pollution and the contribution of past climate change *Environ. Res. Lett.* **8** 034005–12

Silva R A *et al* 2016b The effect of future ambient air pollution on human premature mortality to 2100 using output from the ACCMIP model ensemble *Atmos. Chem. Phys.* **16** 9847–62

Silva R A, Adelman Z, Fry M M and West J J 2016a The impact of individual anthropogenic emissions sectors on the global burden of human mortality due to ambient air pollution *Environ. Health Perspect.* **124** 1776–84

Squizzato S, Masiol M, Rich D Q and Hopke P K 2018 $PM_{2.5}$ and gaseous pollutants in New York State during 2005–2016_ Spatial variability, temporal trends, and economic influences *Atmos. Environ.* **183** 209–24

US EPA 2018a State Average Annual Emission Trend, Criteria pollutants State Tier 1 for 1990–2017 (https://epa.gov/sites/production/files/2018-07/state_tier1_caps.xlsx) (accessed: 21 January 2019)

US EPA 2018b Our Nation's Air, Status and Trend Report (https://gispub.epa.gov/air/trendsreport/2017/) (accessed: 24 January 2019)

US EPA 2018c Technical Note-$PM_{2.5}$ Continuous Monitor Comparability Assessment (https://epa.gov/sites/production/files/2016-09/documents/comparabilityassessmenttool.pdf) (accessed: 20 October 2018)

US EPA, RSIG 2016 Technical Information about Fused Air Quality Surface Using Downscaling Tool: Metadata Description (https://epa.gov/sites/production/files/2016-07/documents/data_fusion_meta_file_july_2016.pdf) (accessed: 29 January 2019)

van Donkelaar A, Martin R V, Brauer M and Boys B L 2014 Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter *Enviorn. Health Perspect* **123** 135–43

van Donkelaar A, Martin R V, Brauer M, Hsu N C, Kahn R A, Levy R C, Lyapustin A, Sayer A M and Winker D M 2016 Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors *Environ. Sci. Technol.* **50** 3762–72

van Donkelaar A, Martin R V, Li C and Burnett R T 2019 Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites models, and monitors *Environ. Sci. Technol.* **53** 2595–611

van Donkelaar A, Martin R V and Park R J 2006 Estimating ground-level $PM_{2.5}$ using aerosol optical depth determined from satellite remote sensing *J. Geophys. Res.* **111** D21201–10

van Donkelaar A, Martin R V, Pasch A N, Szykman J J, Zhang L, Wang Y X and Chen D 2012 Improving the accuracy of daily satellite-derived ground-level fine aerosol concentration estimates for North America *Environ. Sci. Technol.* **46** 11971–8

Wang J and Christopher S 2003 Intercomparison between satellite-derived aerosol optical thickness and PM2.5 mass: Implications for air quality studies *Geophys. Res. Lett.* **30** 6807

Wang M *et al* 2014 Performance of multi-city land use regression models for nitrogen dioxide and fine particles *Enviorn. Health Perspect* **112** 843–9

Wang M, Sampson P D, Hu J, Kleeman M, Keller J P, Olives C, Szpiro A A, Vedal S and Kaufman J D 2016 Combining land-use regression and chemical transport modeling in a spatiotemporal geostatistical model for ozone and $PM_{2.5}$ *Environ. Sci. Technol.* **50** 5111–8

Wang J, Xing J, Mathur R, Pleim J E, Wang S, Hogrefe C, Gan C-M, Wong D C and Hao J 2017 Historical trends in $PM_{2.5}$-related premature mortality during 1990–2010 across the Northern Hemisphere *Enviorn. Health Perspect.* **125** 400–8

Yanosky J D, Paciorek C J, Laden F, Hart J E, Puett R C, Liao D and Suh H H 2014 Spatio-temporal modeling of particulate air pollution in the conterminous United States using geographic and meteorological predictors *Environ. Health* **13** 63

Zeger S L, Thomas D, Dominici F, Samet J M, Schwartz J, Dockery D and Cohen A 2000 Exposure measurement error in time-series studies of air pollution: concepts and consequences *Enviorn. Health Perspect.* **108** 419–26

Zhang Y, West J J, Mathur R, Xing J, Hogrefe C, Roselle S J, Bash J O, Pleim J E, Gan C-M and Wong D C 2018 Long-term trends in the ambient $PM_{2.5}$ and O_3 related mortality burdens in the United States under emission reductions from 1990 to 2010 *Atmos. Chem. Phys.* **18** 15003–16