Atmospheric Environment 276 (2022) 119031

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

Application of an improved gas-constrained source apportionment method using data fused fields: A case study in North Carolina, USA

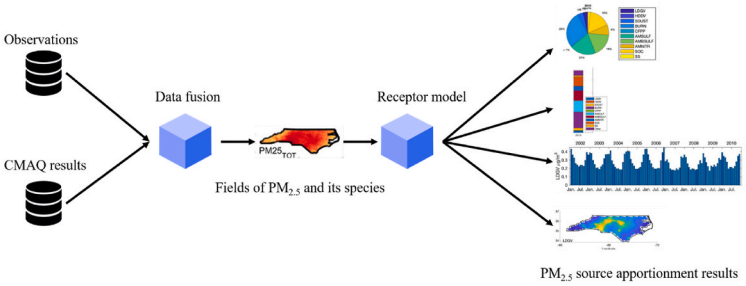
Ran Huang [a], Zongrun Li [a], Cesunica E. Ivey [b], Xinxin Zhai [a], Guoliang Shi [c], James A. Mulholland [a], Robert Devlin [d], Armistead G. Russell [a,*]

[a] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA
[b] Department of Chemical and Environmental Engineering, University of California Riverside, Riverside, CA, USA
[c] State Environmental Protection Key Laboratory of Urban Ambient Air Particulate Matter Pollution Prevention and Control, Center for Urban Transport Emission Research, College of Environmental Science and Engineering, Nankai University, Tianjin, China
[d] Center for Public Health and Environmental Assessment, US Environmental Protection Agency, Research Triangle Park, NC, USA

HIGHLIGHTS

- Fusing observations and simulations to generate pollution fields
- A computationally effective method based on data fusion and receptor model to apportion $PM_{2.5}$
- Spatial and temporal variations of sources to $PM_{2.5}$ in North Carolina

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
CMAQ
CMB
Data fusion
$PM_{2.5}$
Source apportionment

ABSTRACT

A number of studies have found differing associations of disease outcomes with $PM_{2.5}$ components (or species) and sources (e.g., biomass burning, diesel vehicles and gasoline vehicles). Here, a unique method of fusing daily chemical transport model (Community Multiscale Air Quality Modeling) results with observations has been utilized to generate spatiotemporal fields of the concentrations of major gaseous pollutants (CO, NO_2, NO_x, O_3, and SO_2), total $PM_{2.5}$ mass, and speciated $PM_{2.5}$ (including crustal elements) over North Carolina for 2002–2010. The fused results are then used in chemical mass balance source apportionment model, CMBGC-Iteration, which uses both gas constraint and particulate matter concentrations to quantify source impacts. The method, as applied to North Carolina, quantifies the impacts of ten source categories and provides estimates of source contributions to $PM_{2.5}$ concentrations. The ten source categories include both primary sources (diesel vehicles, gasoline vehicles, dust, biomass burning, coal-fired power plants and sea salt) and secondary components (ammonium sulfate, ammonium bisulfate, ammonium nitrate and secondary organic carbon). The results show a steady decrease in anthropogenic source impacts, especially from diesel vehicles and coal-fired power plants. Secondary pollutant components accounted for approximately 70% of $PM_{2.5}$ mass. This study demonstrates an ability to provide spatiotemporal fields of both PM components and source impacts using a chemical transport

* Corresponding author.
E-mail address: ted.russell@ce.gatech.edu (A.G. Russell).

https://doi.org/10.1016/j.atmosenv.2022.119031
Received 18 May 2021; Received in revised form 1 January 2022; Accepted 24 February 2022
Available online 26 February 2022

model fused with observation data, linked to a receptor-based source apportionment method, to develop spatiotemporal fields of multiple pollutants.

1. Introduction

A number of epidemiologic studies have found adverse health effects associated with exposure to particular matter (PM) (Fann et al., 2018; Iskandar et al., 2012; Krall et al., 2017; Liu et al., 2016; Peng et al., 2009; Sarnat et al., 2008, 2015; Silverman and Ito, 2010; Stafoggia et al., 2013; Ye et al., 2018). Particulate matter with an aerodynamic diameter less than 2.5 μm ($PM_{2.5}$) has the most substantial evidence indicating a relationship between short- and long-term exposures with multiple health outcomes. Considering that $PM_{2.5}$ is a mixture of species that come from different sources, it has been hypothesized that the associations of health outcomes with these species and sources may differ, and studies have found this to be the case (Huang et al., 2019; Kim et al., 2012; Krall et al., 2017; Lippmann, 2014; Lippmann et al., 2013; Sarnat et al., 2008, 2015; Vedal et al., 2013).

Studies focusing on associations between total $PM_{2.5}$ exposure (Rhee et al., 2019), its components (Ito et al., 2011) and sources (Bell et al., 2014) historically were based on using ambient air concentrations measured at monitoring sites. Use of ground-based observations, alone, lacks detailed spatial information of air pollution due to the sparse distribution of monitoring sites, which can lead to exposure estimation error and limited ability to assess health impacts in epidemiologic studies (Adams et al., 2015; Goldman et al., 2011; Sarnat et al., 2007; Sheppard et al., 2005; Strand et al., 2006; Zeger et al., 2000). Modeling approaches such as dispersion models (Holmes and Morawska, 2006), chemical transport models (Byun and Schere, 2006) (CTM), and land use regression (Hoek et al., 2008), have frequently been used to generate exposure fields that provide more spatial information. Meanwhile, hybrid approaches that combine observations from monitoring sites, output from CTMs, satellite observations and land use variables using statistical (Hu et al., 2014) and machine learning (Meng et al., 2018) methods have become available. Hybrid approaches decrease the bias in the CTM spatiotemporal fields while using the detailed spatial information not available in ground-based monitoring networks (Ivey et al., 2015). Some of these methods are typically applied to total $PM_{2.5}$ mass with limited application to $PM_{2.5}$ species, especially for element species, which are important tracer species for quantifying the sources of $PM_{2.5}$ components. A hybrid approach developed by Friberg et al. (2016) provides spatiotemporal fields of both total and speciated $PM_{2.5}$ (and, here, is extended to include element species). The method fuses daily results from a CTM with monitoring site observations to develop spatiotemporal fields of pollutant concentrations. The exposure fields of pollutants from data fusion are further used in source apportionment methods to obtain spatiotemporal fields of source impacts.

In this work, we use CMBGC-Iteration (Shi et al., 2018), an improved gas-constrained (GC) source apportionment method based on the Chemical Mass Balance (CMB) model (Watson et al., 2004), a widely used receptor model, to quantify the impacts of ten distinct sources on $PM_{2.5}$ in North Carolina based on the exposure fields generated from the data fusion method. The ten sources include both primary sources (diesel vehicles, gasoline vehicles, dust, biomass burning, coal combustion, and sea salt) and secondary components (ammonium sulfate, ammonium bisulfate, ammonium nitrate, and secondary organic carbon). Total $PM_{2.5}$ mass and its species including elemental carbon (EC), organic carbon (OC), sulfate, ammonium, nitrate and ten element species (Al, Ca, Cu, Fe, K, Mn, Na, Pb, Si, and Zn), five gaseous species (CO, NO_2, NO_x, O_3 and SO_2), and the ten source contributions to total $PM_{2.5}$ mass were estimated for the 2002–2010 period in support of the CATHGEN study of associations between pollutants and health endpoints in a cohort study in North Carolina, USA.

2. Methods

2.1. Ambient monitor data

The air quality data used for data fusion come from the Chemical Speciation Network (CSN) (Watson et al., 2004) and Interagency Monitoring of Protected Visual Environments (IMPROVE) (Watson et al., 2004) network. Concentrations of pollutants were obtained for total $PM_{2.5}$ mass, five major $PM_{2.5}$ species (EC, OC, sulfate, ammonium, nitrate), ten element species (Al, Ca, Cu, Fe, K, Mn, Na, Pb, Si, and Zn), and five gaseous species (carbon monoxide (CO), nitrogen dioxide (NO_2), nitrogen oxides (NO_x), ozone (O_3) and sulfur dioxide (SO_2)) (Fig. 1). Due to the limited number of monitoring sites for some species (e.g., $PM_{2.5}$ element species, CO, NO_2, NO_x, and SO_2) in North Carolina, we also included monitoring sites in neighboring states that are in proximity to the state border. Simulated pollutant fields come from the Community Multiscale Air Quality (CMAQ, version 4.5) model with a horizontal resolution of 12 km (provided by U.S. EPA (Appel et al., 2008)).

2.2. Data fusion approach

We applied a data fusion method developed by Friberg et al. (2016), which blends daily averaged observations and CMAQ results based on spatial correlation analysis between observations and CMAQ simulations. The resulting product is a new field that captures the temporal variations in local observations, as well as spatial variability in CMAQ simulations. The method has previously been used to estimate daily $PM_{2.5}$ and its major species (EC, OC, sulfate, ammonium, and nitrate) and gaseous pollutants (CO, NO_x, and NO_2) for North Carolina from 2006 to 2008 (Huang et al., 2018). Here, we extend the method to include element species and apply the method for a longer time period, which is more conducive for longer-term epidemiologic analyses. Fields for Cu, Fe, K, Mn, Zn were derived by fusing the observations and the corresponding simulated trace species in CMAQ. Due to lack of simulated concentrations for those tracer species for some years, we applied

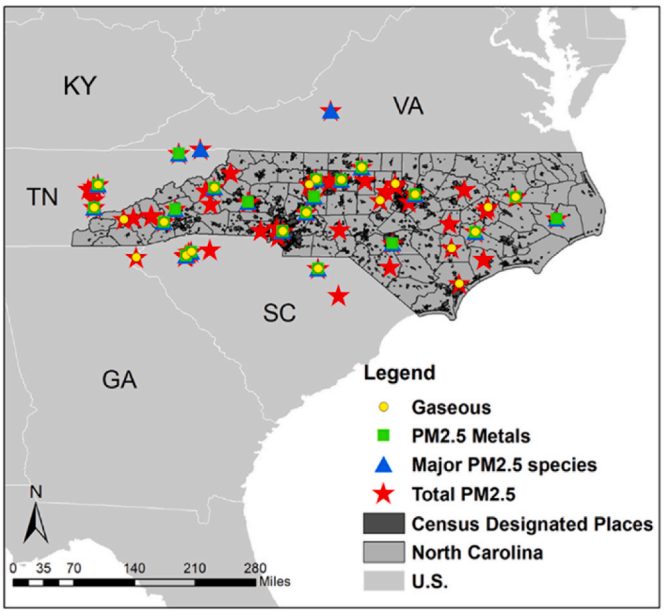


Fig. 1. Locations of monitoring sites used in this study.

the ratios that come from source profile of those tracer species and multiplied them with simulated total $PM_{2.5}$ to obtain the simulated concentrations. Additional details of the data fusion method and application in the CATHGEN study can be found in previous publications (Huang et al., 2018; McGuinn et al., 2017).

2.3. CMBGC-iteration model

The CMBGC-Iteration (Shi et al., 2018) model extends the traditional CMB approach, using the ratio of gases to $PM_{2.5}$ as constraints while considering the uncertainties of source profiles and receptor datasets. Gas-to-particle ratios (e.g., $SO_2/PM_{2.5}$, $CO/PM_{2.5}$, and $NO_x/PM_{2.5}$ ratios) were introduced into CMBGC-Iteration to identify potential sources and their contributions. Ten sources were identified using this model, including six primary sources (diesel vehicles, gasoline vehicles, dust, biomass burning, coal combustion, and sea salt) and four secondary components (ammonium sulfate, ammonium bisulfate, ammonium nitrate and secondary organic carbon). "Coal combustion" only includes the primary components from coal combustion and does not include contributions from the secondary formation of $PM_{2.5}$ due to SO_2 or NO_x emissions, or the associated ammonium. The source profiles used in this CMBGC-Iteration application were developed by Ivey et al. (2017), where nonlinear optimization was used to update static profiles to reflect spatial and seasonal nuances. Detailed descriptions of the CMBGC-Iteration model are found in a previous publication (Shi et al., 2018).

The approach chosen here, combining data-fused species fields and a receptor-based source apportionment model differs from the method developed by Ivey et al. (2015) which uses a CTM to directly estimate *a priori* source impact fields, then integrates observations to modify those fields. That approach is computationally intensive. As applied here, the data fusion approach can utilize prior model results from the CTM without the need for CTM-developed source impact fields, which is attractive as there is a long record of CTM simulations available at a variety of horizontal resolutions. Both approaches reproduce observed species levels very well because they integrate observations, and prior studies have thoroughly evaluated the approaches using data withholding and cross validation (Friberg et al., 2016; Huang et al., 2018). However, it is not currently possible to evaluate how well actual source impacts are estimated, as there are no direct observations for comparison. Both methods capture the temporal and spatial trends in pollutants, however the following approach applied in this work is computationally attractive and can be utilized when both CTM simulations (including PM and gaseous species) and observations are available.

3. Results and discussions

3.1. Observations and CMAQ results

Prior to the application of the data fusion method, monthly and annual averages of CMAQ simulation results were evaluated against observations from monitoring sites from 2002 to 2010. Quantities evaluated included total $PM_{2.5}$ mass, its species including EC, OC, sulfate, ammonium, nitrate, and 10 $PM_{2.5}$ element species (Al, Ca, Cu, Fe, K, Mn, Na, Pb, Si, and Zn), and five gaseous pollutants (Figs. S1 and S2, Table 1). Simulations from CMAQ have good correlations with observations for $PM_{2.5}$ major species with R (Pearson correlation) larger than 0.57.

The annual average of observed total $PM_{2.5}$ mass shows a decreasing trend from 2002 to 2009 with a slight increase from 2009 to 2010. Overall, total $PM_{2.5}$ mass decreases about 24% from 2002 to 2010. The CMAQ simulation captures a similar decreasing and increasing trend but underestimates the concentrations. Concentrations of some of the major $PM_{2.5}$ species, e.g., sulfate, ammonium, and nitrate also show a similar pattern, while EC and OC do not. Those two species have larger observed annual average concentrations during 2006–2008, possibly because

Table 1
Statistical performance of CMAQ, data fusion (DF) and data withholding (WH) for $PM_{2.5}$ and $PM_{2.5}$ species and gaseous species.

SPECIES (#)	PM25_TOT (60)			EC (22)			OC (19)			SO4 (22)			NO3 (22)			NH4 (22)			CO (9)		
OBS VS Simulations	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH
BIAS (NME)	−0.17	−0.02	−0.02	0.14	0.04	0.05	−0.35	0.06	0.05	−0.18	0.01	0.03	0.31	0.01	0.07	−0.03	−0.01	−0.01	−0.27	0.03	0.04
NRMSE	0.50	0.13	0.22	1.08	0.45	0.56	0.77	0.44	0.52	0.50	0.13	0.23	1.69	0.48	0.69	0.55	0.13	0.26	0.55	0.29	0.35
R	0.58	0.97	0.91	0.57	0.86	0.79	0.57	0.86	0.78	0.75	0.98	0.95	0.59	0.91	0.82	0.58	0.98	0.91	0.42	0.79	0.78
SPECIES (#)	NO2 (5)			NOx (8)			O3 (21)			SO2 (11)			Si (23)			Ca (23)			K (18)		
OBS VS Simulations	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH
BIAS (NME)	−0.08	−0.004	−0.005	0.65	−0.11	−0.12	0.53	0.02	0.04	0.29	−0.02	−0.02	3.62	−0.03	−0.04	5.19	0.03	0.05	2.57	0.01	0.07
NRMSE	0.49	0.32	0.34	1.65	0.87	0.94	0.62	0.20	0.23	1.12	0.73	0.80	4.84	0.43	0.68	6.49	0.81	0.96	3.69	0.80	1.3
R	0.58	0.84	0.82	0.54	0.85	0.83	0.61	0.85	0.82	0.52	0.72	0.70	0.04	0.95	0.86	0.23	0.79	0.65	0.20	0.92	0.78
SPECIES (#)	Fe (23)			Cu (23)			Al (23)			Zn (23)			Mn (23)			Pb (23)			Na (18)		
OBS VS Simulations	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH	CMAQ	DF	WH
BIAS (NME)	2.40	0.04	0.07	2.18	0.22	0.32	4.32	−0.03	−0.05	−0.46	−0.76	−0.84	4.47	0.00	0.02	0.58	0.02	0.04	−0.85	−0.01	−0.03
NRMSE	3.16	0.49	0.59	3.78	2.31	2.5	5.85	0.87	1.21	2.16	2.04	2.27	5.62	0.85	1.05	1.81	1.22	1.44	1.50	0.49	1.08
R	0.36	0.86	0.80	0.17	0.56	0.32	0.03	0.91	0.81	0.03	0.84	0.80	0.26	0.79	0.57	0.24	0.74	0.48	0.34	0.93	0.60

these years have larger wildfire burn areas compared to other years (North Carolina Forest Service). Further, the OC, in large part, will be due to biogenic emissions (Offenberg et al., 2011). For EC, the CMAQ simulation shows an increasing trend rather than the decreasing trend seen in the observations from 2007 to 2010, which may be due to the fire emissions used in the simulation. In 2008, there was a large wildfire (Evans Road Wildfire) (North Carolina Climate Office) in Eastern North Carolina. According to the 2008 National Emission Inventory (NEI) (EPA, 2008), 64% of the total $PM_{2.5}$ emissions come from fires in North Carolina. Emissions used in CMAQ for those years (2007–2010) were based on the 2008 NEI and projected to other years. Considering the exceptional event in 2008, simulations using the 2008 NEI fire emissions may lead to the increasing trend for EC since 2007. Correcting such a bias shows the importance of the data fusion step.

Observed trends in most crustal element species are flat from 2002 to 2010 except for aluminum, sodium, lead, and zinc. Annual average of sodium decreases about 70% from 0.22 μg/m^3 to 0.07 μg/m^3 while aluminum and zinc increased by about 145%. CMAQ-derived concentrations of crustal elements are overestimated except sodium and zinc as compared to observations. Only observations above detection limits were considered for this evaluation. Observed O_3 concentrations are relatively low and stable over these years. The annual average of other gaseous pollutants decreases over 40% in 2010 compared to 2002. SO_2 shows the most significant change in the annual average, with decreases of about 70% from 2.99 ppb (2002) to 0.80 ppb (2010). CMAQ simulations have better performance for gaseous pollutants than most PM species, with an averaged correlation of 0.53 for the gases (Table 1). CMAQ does not capture the decrease in observed CO concentrations (about 22%) from 2004 to 2005 and underestimates the CO concentrations throughout the entire period. However, CO performance improved later (e.g., smaller difference compared to observations) despite missing the higher CO in earlier years, which may be due to the larger wildfire emissions inputs. CMAQ is biased high in simulating the ozone concentrations but captures the year-to-year variations, with a correlation of 0.61 compared to observations (Table 1).

Observed total $PM_{2.5}$ mass and five major PM species display a strong seasonal variation. Total $PM_{2.5}$, OC, sulfate, and ammonium are high in summer and low in winter while EC and nitrate are high in winter and low in summer. CMAQ simulations do not capture the same seasonal variation for total $PM_{2.5}$, EC, and OC. Of the element species, only silicon and zinc have obvious seasonal variations, with high observed concentrations in summer and low observed concentrations in winter. Observed potassium has several peaks during summer due to wildfires. For gaseous pollutants, seasonal variations are strong for all species. All the gaseous pollutants, except O_3, have higher concentrations in the winter compared to the summer. The correlations between observed and CMAQ-simulated concentrations of element species are mostly smaller than 0.4 (Table 1).

3.2. Data fusion results

Comparisons between observations with data fusion results have better performance than that with CMAQ with smaller normalized mean errors and root mean square errors and higher correlations due to the method mechanism. To evaluate uncertainty of data fusion method, an exhaustive 10% data withholding of observations is conducted and the results still outperform the raw CMAQ results (Table 1). Local observations are averaged over 24 h, over which time air parcels have typically moved 10's of km's, and thus have been impacted by sources over a relatively broad region. However, they may be highly impacted by a local source, thus skewing the observations such that they no longer represent regional conditions. Using the air quality model and interpolation will help decrease this misalignment, but it still can exist. Total $PM_{2.5}$ mass is high in the Piedmont region (the middle region of the state (North Carolina Department of Public Instruction)) and low in the mountain (western NC) and coastal (Eastern NC) regions (Fig. 2,

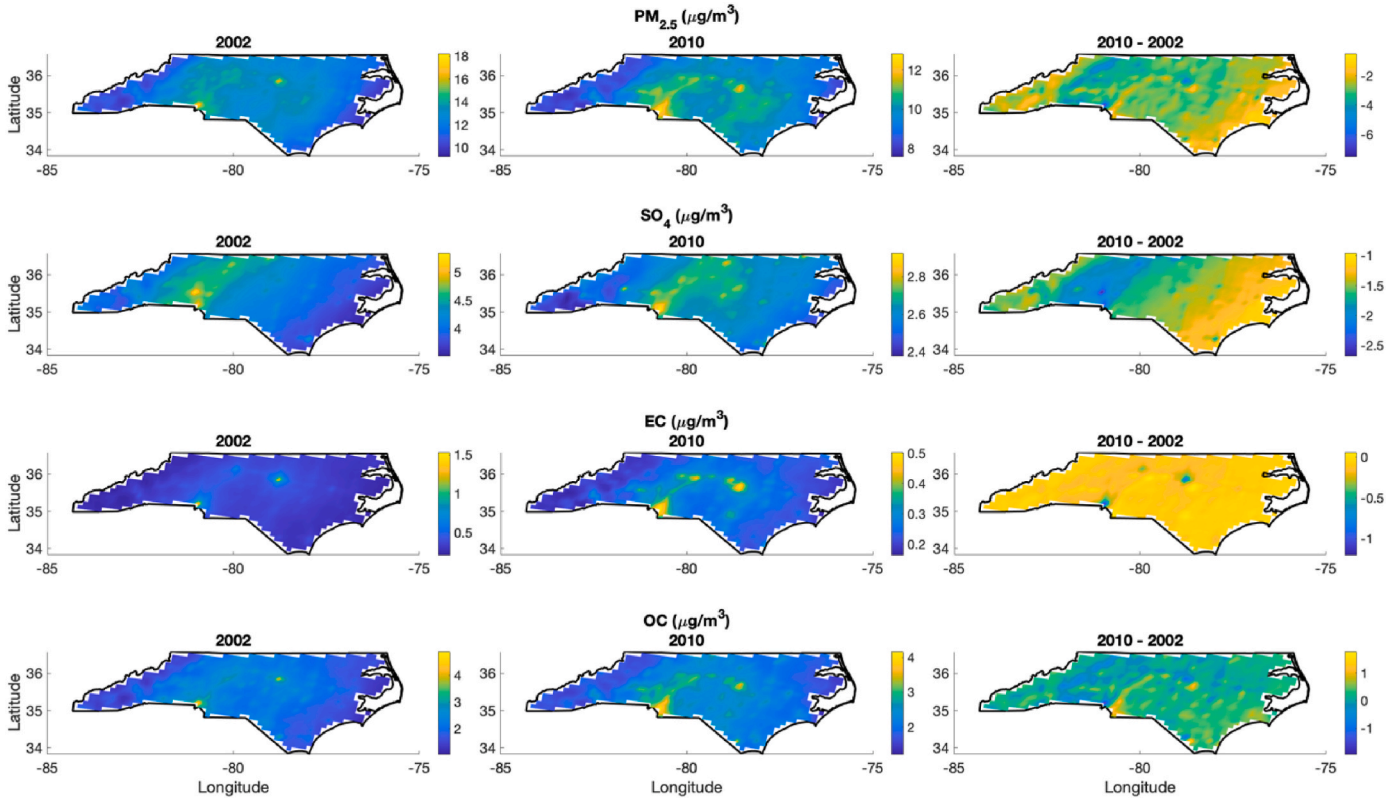


Fig. 2. Annual average spatial distribution fields from data fusion for total $PM_{2.5}$, EC, OC and sulfate in 2002 and 2010 and the concentration differences between these two years, μg/m^3.

Fig. S3). The concentrations of $PM_{2.5}$ were lower across North Carolina in 2010 compare to 2002. Urban areas have higher EC and OC concentrations than rural areas (Fig. 2, Figs. S4 and S5). Compare to 2002, EC is lower in 2010 while OC increases in some regions in North Carolina. The middle region has higher sulfate concentrations (Fig. 2, Fig. S6) due to SO_2 emissions from coal-fired power plants and the concentrations decrease dramatically from 2002 to 2010, especially in the center of the state. Nine of the state's 14 major coal-fired power plants were located in this region (Li and Gibson, 2014). Higher ammonium (Fig. S7) and nitrate (Fig. S8) concentrations in the southeastern part of the state are due largely to the intensive hog operations located in this area (Wing et al., 2000). Spatial distributions of crustal elements (Fig. S9 – Fig. S18) are more homogeneous than for major $PM_{2.5}$ species. For the gaseous pollutants (Fig. S19 – Fig. S23), all the species show higher concentrations in the areas with greater emissions, with the exception of O_3 which is more spatially homogeneous.

OC and sulfate are the dominant species in total $PM_{2.5}$, accounting for about 20% and 30%, respectively (Fig. 3). 2009 has the lowest total $PM_{2.5}$ mass with the lowest sulfate, ammonium, and nitrate concentrations among the reporting years. The large decrease of these secondary pollutant concentrations is mainly due to the reduction of SO_2 and NO_x emissions from coal-fired power plants as a response to the North Carolina Clean Smokestacks Act (North Carolina Department of Environmental Quality), which was aimed at reducing emissions of SO_2 and NO_x from coal-fired power plants by 49% and 77%, respectively, by 2009 compared to 1998.

3.3. Source apportionment (CMBGC-Iteration) results

3.3.1. Source contributions

From 2002 to 2010, total $PM_{2.5}$ concentration decreases by 20.7% while most source contributions decrease during the same period, with the exception of biomass burning and soil dust. Diesel vehicle contribution decreases by about 81.7% and coal-fired power plant contribution decreases by 66.7%, and these are the two largest decreases among the sources. Ammonium bisulfate, ammonium sulfate and ammonium nitrate decrease by about 43.1%, 27.3% and 38.3%, respectively. Secondary organic carbon and gasoline vehicles decrease by 3.5% and 6%, respectively, which are the two lowest rates of decrease. Biomass burning increases by 15.3%, and soil dust increases by 31.4% (Fig. 4).

In North Carolina, biomass burning, ammonium bisulfate and ammonium sulfate are the main sources, comprising about 60% of the total $PM_{2.5}$. Contributions from dust and coal-fired power plants (coal combustion) are about 4% and 0.06% in 2010, which are relatively low compared to 20% and 8% from the 2011 NEI (EPA, 2011) for North Carolina. However, this is not unexpected given the large amount of

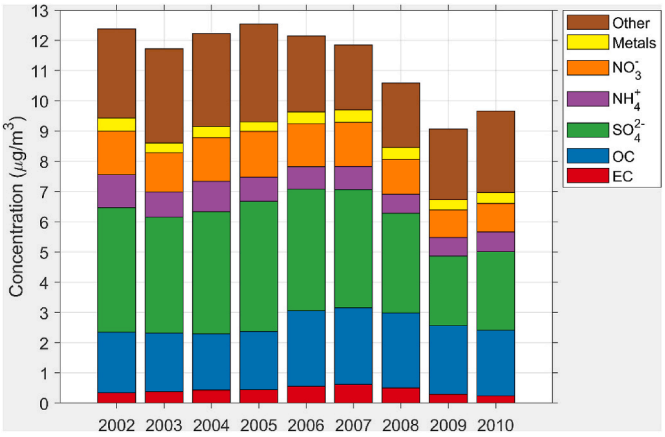


Fig. 3. Total $PM_{2.5}$ mass and its species concentrations from data fusion, averaged statewide: 2002–2010, μg/m^3.

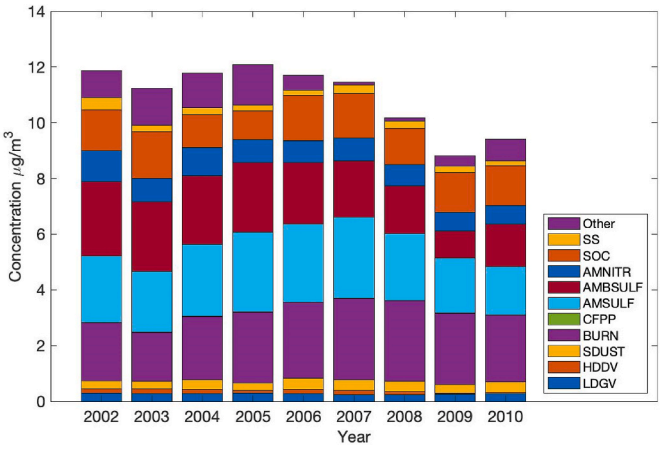


Fig. 4. Annual average sources contributions in total $PM_{2.5}$ mass, μg/m^3. (SS: sea salt, SOC: secondary organic carbon, AMNITR: ammonium nitrate, AMBSULF: ammonium bisulfate, AMSULF: ammonium sulfate, CFPP: coal-fired power plants, BURN: biomass burning, SDUST: soil dust, HDDV: heavy-duty diesel vehicles, LDGV: light-duty gasoline vehicles).

secondary $PM_{2.5}$ formation and that the coal combustion emissions are often elevated and away from cities. The contribution from biomass burning is about 28% in 2010, somewhat larger than the 15% of primary $PM_{2.5}$ from biomass burning reported in the 2011 NEI. It may be due to the collinearity of sources with similar tracer species. For biomass burning, the largest tracer species is organic carbon, which also highly depends on vehicles and secondary organic carbon (Fig. 5).

3.3.2. Spatial distributions

Annual averages of the CMBGC-Iteration model daily values show that light-duty gasoline vehicle and heavy-duty diesel vehicle impacts have higher contributions along highways and in urban areas, while soil dust has higher impacts along the coast and in the Appalachian Mountains of western North Carolina than urban areas (Fig. 6, Fig. S24-Fig. S33). Compare with 2002, heavy-duty diesel vehicle impacts are lower especially in urban areas while light-duty gasoline vehicle impacts increase in urban areas in 2010. Biomass burning impacts are relatively homogenous throughout the state. And it increases in most of region in 2010 compare with 2002 (Fig. 6). The increasing pattern is consistent with increased burning area in North Carolina (Fig. S35) (National Interagency Fire Center). Contributions from coal-fired power plants are mainly in the Piedmont region of central North Carolina where many of the larger coal-fired power plants are located. The obvious reduction is near the coal-fired power plants region (Fig. 6). It can be explained by the responses to North Carolina Clean Smokestacks Act (North Carolina Department of Environmental Quality). Ammonium sulfate and ammonium nitrate are elevated in southeastern North Carolina, location of many hog farms that emit ammonia (Wing et al., 2000). The elevated ammonia leads to lower ammonium bisulfate as sulfate forms ammonium sulfate rather than ammonium bisulfate when ammonia concentrations are high (Cheng et al., 2019). Secondary organic carbon is also high in the Piedmont region and is more homogeneous than diesel vehicles and gasoline vehicles contributions. Sodium (Na^+) levels are relatively homogeneous and low, though slightly higher over the mountains than the coast. This unanticipated result is explained by the lack of observations near the coast leading to a negative exponent in regression model of data fusion and it leads to negative correlation between data fusion result and CMAQ result (Table S1). The CMBGC-Iteration model was run both with, and without, the sea salt source, leading to very similar levels of impacts from sources other than sea salt (Fig. S34).

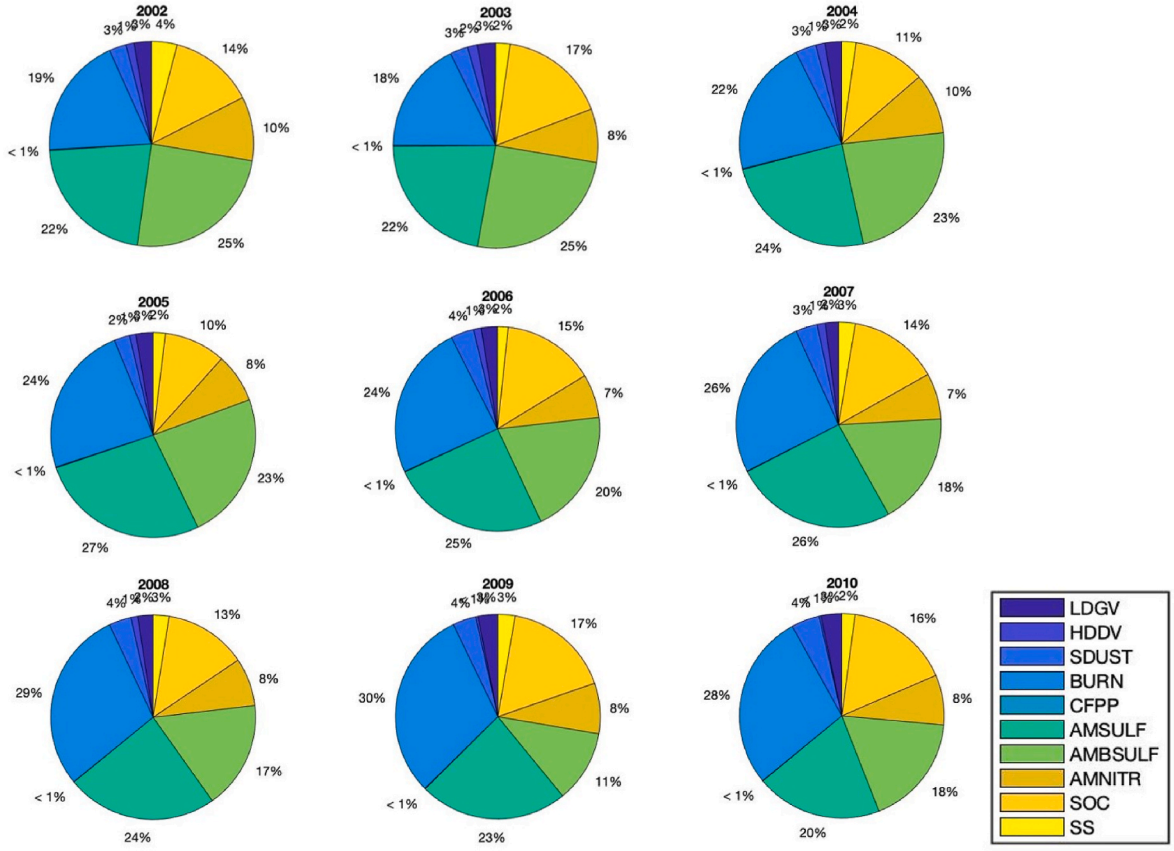


Fig. 5. Annual average contributions of $PM_{2.5}$ from ten sources from 2002 to 2010 in North Carolina.

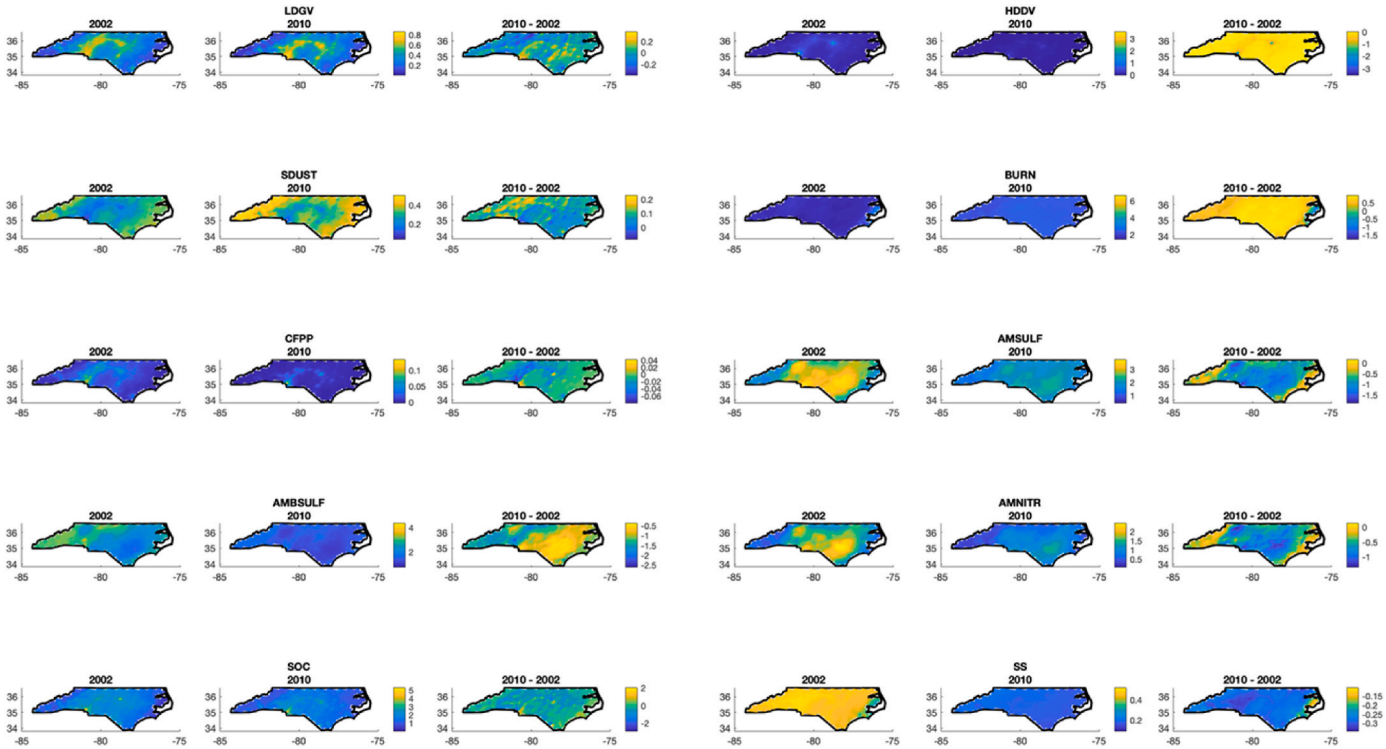


Fig. 6. Annual average spatial distribution of $PM_{2.5}$ from ten sources in 2002 and 2010 in North Carolina and differences between these two years, $\mu g/m^3$.

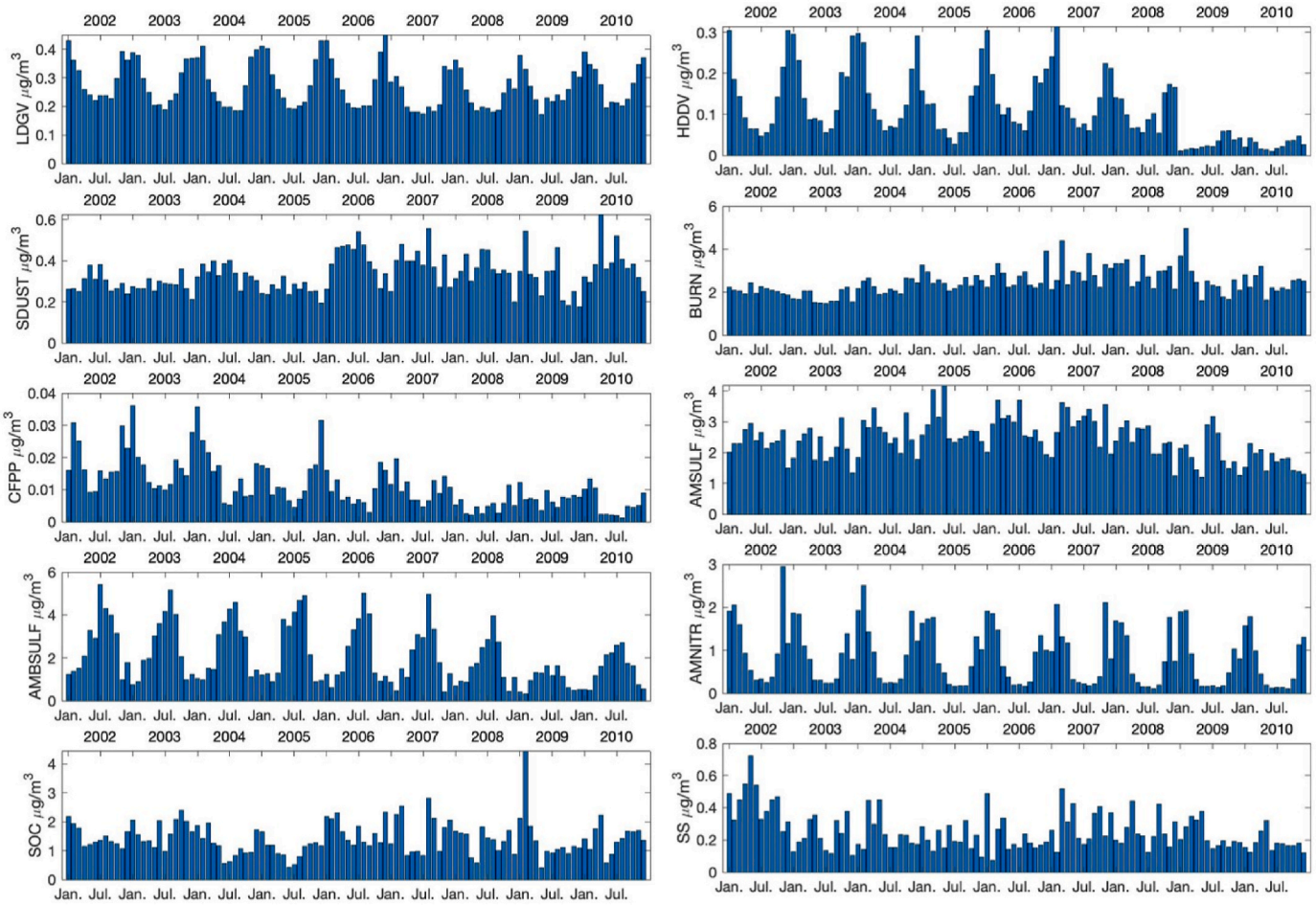


Fig. 7. Monthly average source contributions to total $PM_{2.5}$ mass: 2002–2010, μg/m^3.

3.3.3. *Temporal distributions*

$PM_{2.5}$ sources show strong seasonal variations (Fig. 7). Contributions of diesel vehicles, gasoline vehicles and coal-fired power plants are higher in winter than in summer due to lower mixing heights in winter. Ammonium nitrate is also high as its formation is thermodynamically more favorable at low temperatures (Russell et al., 1967; Stelson and Seinfeld, 1967). Ammonium sulfate and ammonium bisulfate are higher in summer than in winter as SO_2 is rapidly oxidized, and it is lost more slowly to wet deposition. Of interest, ammonium bisulfate increases more in the summer than ammonium sulfate as the amount of sulfate formed becomes greater than the ammonia available to form ammonium sulfate, leading to higher aerosol acidities (Shi et al., 2017; Weber et al., 2016). The concentration of soil dust is also higher in summer.

3.3.4. *Source apportionments evaluation*

Since it is hard to evaluate the apportionment results because direct observations of source impacts are not available, we compare our results with air pollutant emission trend data for 2002 to 2010 in North Carolina and previous apportionment of $PM_{2.5}$ results in southeastern United States (EPA, 2021; Zhai et al., 2017). For biomass burning, annual averaged biomass burning impacts from our source apportionment study are compared with annual burned area due to absence of prescribed burning emissions during 2002–2007 in the trend data (National Interagency Fire Center). Biomass burning impacts and burned area increase from previous year to recent year and are positively correlated with an $R^2 = 0.71$ (Fig. 8). Mobile emissions decrease 37% while the impacts from gasoline and diesel vehicles decrease 30%. The similarity is evidenced by the R^2 of 0.88 (Fig. 8). Coal-fired power plant contributions decrease by 67% while fuel combustion from electricity utility decreases by 66%. The decreases of ammonium sulfate and ammonium bisulfate can be explained by decreases of SO_2 emissions (Fig. 8). Meanwhile, reduction of ammonium nitrate is due to NO_x emission controls, including those due to the North Carolina Clean Smokestacks Act (North Carolina Department of Environmental Quality). Ammonium nitrate impacts are positively correlated with NO_x emission with an $R^2 = 0.75$ (Fig. 8). To evaluate our novel method, we also compare our results with long-term source apportionments of $PM_{2.5}$ study in Georgia based on observations and CMB modeling (Zhai et al., 2017). The study shows the largest contributions are from sulfate, biomass burning, and other OC and it is consistent with our results. Biomass burning is increasing in North Carolina while the Georgia study found decreases. The difference is tied to different burned area trends in North Carolina and Georgia. However, the relative contributions of biomass burning increased due to the reduced $PM_{2.5}$ in both studies. Decreases are detected in the other sources. Both of the studies found similar temporal patterns. In previous study, contributions from primary coal combustion and vehicles are spatially heterogeneous while ammonium sulfate, ammonium bisulfate and ammonium nitrate are homogeneous. Our results show that primary $PM_{2.5}$ from coal combustion and $PM_{2.5}$ from vehicles are concentrated in urban areas, especially near the road and power plants. For ammonium sulfate, ammonium bisulfate and ammonium nitrate, high concentrations are in southeastern North Carolina with more homogeneous pattern. As demonstrated, this method provides additional spatial information beyond applying CMB in measurement locations and interpolating those results.

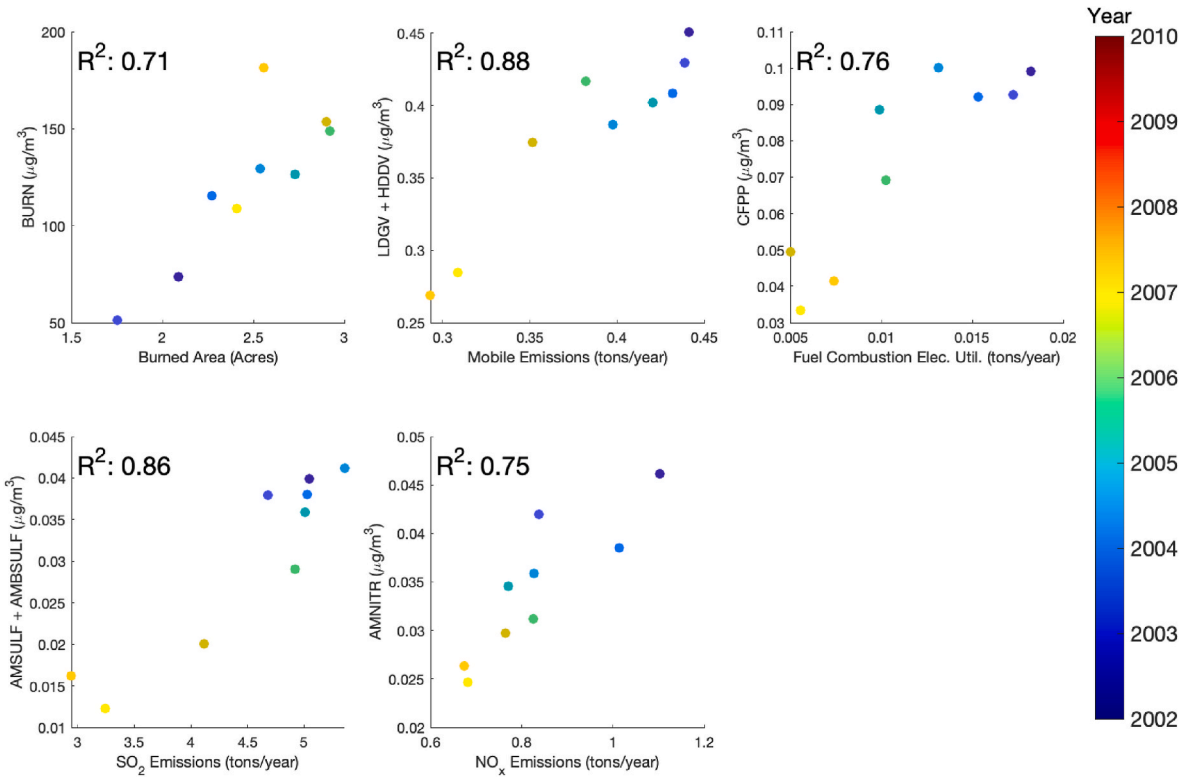


Fig. 8. Biomass burning vs. burned area during 2002–2010 in North Carolina; source apportionments vs. air pollutant emission trend during 2002–2010 in North Carolina.

4. Conclusions

Daily pollutant concentrations of total and speciated $PM_{2.5}$ and five gaseous pollutants are estimated in North Carolina at 12 km horizontal resolution from 2002 to 2010 using a data fusion method. The fused data are used in the CMBGC-Iteration source apportionment method to provide of $PM_{2.5}$ impacts from ten sources across the state. The results show a significant decrease over time in both PM species and source impacts, especially from diesel vehicles and coal combustion. Secondary pollutants are dominant over the state, though sulfate, nitrate, and the related ammonium, are decreasing. The methods presented in this study provide a long period of pollutant concentrations and source impacts that are readily useable for long-term, state-wide epidemiologic studies. The pollutant exposure and source impacts fields are presently being used in a health study to identify the associations between $PM_{2.5}$ components and sources and coronary heart disease (McGuinn et al., 2017; Slawsky et al., 2021). The approach is readily applicable across a variety of spatial scales when observations and CTM results are available.

CRediT author statement

Ran Huang: Conceptualization, Methodology, Formal analysis, Writing-Original draft preparation. **Zongrun Li:** Formal analysis, Writing-Review & Editing. **Cesunica E. Ivey:** Investigation, Resources. **Xinxin Zhai:** Software, Methodology, Resources. **Guoliang Shi:** Software, Methodology, Resources. **James A. Mulholland:** Supervision. **Robert Devlin:** Supervision. **Armistead G. Russell:** Supervision, Conceptualization.

Disclaimer

The research described in this article has been reviewed by the Center for Public Health and Environmental Assessment, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Agency, nor does the mention of trade names of commercial products constitute endorsement or recommendation for use.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We gratefully acknowledge the U.S. EPA, especially Valerie Garcia and K. Wyat Appel, for supplying CMAQ modeling results. This publication was funded, in part, by US EPA Grant Number R834799 and an NIH ENVISION Grant Number R01ES027892. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the US government. Further, the US government does not endorse the purchase of any commercial products or services mentioned in the publication. We also acknowledge the Southern Company and the Electric Power Research Institute (EPRI) for their support. We thank everyone who helped us with the manuscript, in particular Mr. Tommy N. Skipper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosenv.2022.119031.

References

Adams, K., Greenbaum, D.S., Shaikh, R., van Erp, A.M., Russell, A.G., 2015. Particulate matter components, sources, and health: systematic approaches to testing effects. J. Air Waste Manag. Assoc. 65, 544–558.

Appel, K.W., Bhave, P.V., Gilliland, A.B., Sarwar, G., Roselle, S.J., 2008. Evaluation of the community multiscale air quality (CMAQ) model version 4.5: sensitivities impacting model performance; part II—particulate matter. Atmos. Environ. 42, 6057–6066.
Bell, M.L., Ebisu, K., Leaderer, B.P., Gent, J.F., Lee, H.J., Koutrakis, P., et al., 2014. Associations of PM2. 5 constituents and sources with hospital admissions: analysis of four counties in Connecticut and Massachusetts (USA) for persons≥ 65 years of age. Environ. Health Perspect. 122, 138–144.
Byun, D., 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.
Cheng, T., Luo, L., Yang, L., Fan, H., Wu, H., 2019. Formation and emission characteristics of ammonium sulfate aerosols in flue gas downstream of selective catalytic reduction. Energy Fuels 33, 7861–7868.
EPA, 2008. National Emissions Inventory (NEI) Data.
EPA, 2011. National Emissions Inventory (NEI) Data.
EPA, 2021. Air Pollutant Emissions Trends Data.
Fann, N., Alman, B., Broome, R.A., Morgan, G.G., Johnston, F.H., Pouliot, G., et al., 2018. The health impacts and economic value of wildland fire episodes in the US: 2008–2012. Sci. Total Environ. 610, 802–809.
Friberg, M.D., Zhai, X., Holmes, H.A., Chang, H.H., Strickland, M.J., Sarnat, S.E., et al., 2016. Method for fusing observational data and chemical transport model simulations to estimate spatiotemporally resolved ambient air pollution. Environ. Sci. Technol. 50, 3695–3705.
Goldman, G.T., Mulholland, J.A., Russell, A.G., Strickland, M.J., Klein, M., Waller, L.A., et al., 2011. Impact of exposure measurement error in air pollution epidemiology: effect of error type in time-series studies. Environ. Health 10, 61.
Hoek, G., Beelen, R., De Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., et al., 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmos. Environ. 42, 7561–7578.
Holmes, N.S., Morawska, L., 2006. A review of dispersion modelling and its application to the dispersion of particles: an overview of different dispersion models available. Atmos. Environ. 40, 5902–5928.
Hu, X., Waller, L.A., Lyapustin, A., Wang, Y., Al-Hamdan, M.Z., Crosson, W.L., et al., 2014. Estimating ground-level PM2. 5 concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model. Remote Sens. Environ. 140, 220–232.
Huang, M., Ivey, C., Hu, Y., Holmes, H.A., Strickland, M.J., 2019. Source apportionment of primary and secondary PM2. 5: associations with pediatric respiratory disease emergency department visits in the US State of Georgia. Environ. Int. 133, 105167.
Huang, R., Zhai, X., Ivey, C.E., Friberg, M.D., Hu, X., Liu, Y., et al., 2018. Air pollutant exposure field modeling using air quality model-data fusion methods and comparison with satellite AOD-derived fields: application over North Carolina, USA. Air Qual. Atmos. & Health 11, 11–22.
Iskandar, A., Andersen, Z.J., Bønnelykke, K., Ellermann, T., Andersen, K.K., Bisgaard, H., 2012. Coarse and fine particles but not ultrafine particles in urban air trigger hospital admission for asthma in children. Thorax 67, 252–257.
Ito, K., Mathes, R., Ross, Z., Nádas, A., Thurston, G., Matte, T., 2011. Fine particulate matter constituents associated with cardiovascular hospitalizations and mortality in New York City. Environ. Health Perspect. 119, 467–473.
Ivey, C., Holmes, H., Shi, G., Balachandran, S., Hu, Y., Russell, A.G., 2017. Development of PM2. 5 source profiles using a hybrid chemical transport-receptor modeling approach. Environ. Sci. Technol. 51, 13788–13796.
Ivey, C.E., Holmes, H.A., Hu, Y., Mulholland, J.A., Russell, A., 2015. Development of PM 2.5 source impact spatial fields using a hybrid source apportionment air quality model. Geosci. Model Dev. (GMD) 8, 2153–2165.
Kim, S.-Y., Peel, J.L., Hannigan, M.P., Dutton, S.J., Sheppard, L., Clark, M.L., et al., 2012. The temporal lag structure of short-term associations of fine particulate matter chemical constituents and cardiovascular and respiratory hospitalizations. Environ. Health Perspect. 120, 1094–1099.
Krall, J.R., Mulholland, J.A., Russell, A.G., Balachandran, S., Winquist, A., Tolbert, P.E., et al., 2017. Associations between source-specific fine particulate matter and emergency department visits for respiratory disease in four US cities. Environ. Health Perspect. 125, 97–103.
Li, Y.R., Gibson, J.M., 2014. Health and air quality benefits of policies to reduce coal-fired power plant emissions: a case study in North Carolina. Environ. Sci. Technol. 48, 10019–10027.
Lippmann, M., 2014. Toxicological and epidemiological studies of cardiovascular effects of ambient air fine particulate matter (PM2. 5) and its chemical components: coherence and public health implications. Crit. Rev. Toxicol. 44, 299–347.
Lippmann, M., Chen, L.C., Gordon, T., Ito, K., Thurston, G.D., 2013. National Particle Component Toxicity (NPACT) Initiative: Integrated Epidemiologic and Toxicologic Studies of the Health Effects of Particulate Matter Components. Research Report, vol. 5. Health Effects Institute.
Liu, S., Ganduglia, C.M., Li, X., Delclos, G.L., Franzini, L., Zhang, K., 2016. Short-term associations of fine particulate matter components and emergency hospital admissions among a privately insured population in Greater Houston. Atmos. Environ. 147, 369–375.
McGuinn, L.A., Ward-Caviness, C., Neas, L.M., Schneider, A., Di, Q., Chudnovsky, A., et al., 2017. Fine particulate matter and cardiovascular disease: comparison of assessment methods for long-term exposure. Environ. Res. 159, 16–23.
Meng, X., Hand, J.L., Schichtel, B.A., Liu, Y., 2018. Space-time trends of PM2. 5 constituents in the conterminous United States estimated by a machine learning approach, 2005–2015. Environ. Int. 121, 1137–1147.
National Interagency Fire Center. Historical year-end fire statistics by state. Date access. https://www.nifc.gov/fire-information/statistics. (Accessed 13 December 2021).
North Carolina Climate Office, 2019. NC Extremes: NC Pocosins a Hotspot for Large Wildfires.
North Carolina Department of Environmental Quality, 2019. Clean Smokestacks Act (CSA).
North Carolina Department of Public Instruction, 2019. Our State Geography in a Snap: the Piedmont Region. NCpedia.
North Carolina Forest Service, 2019. Number of Fires and Acres Burned in NC, 1928 through 2015.
Offenberg, J.H., Lewandowski, M., Jaoui, M., Kleindienst, T.E., 2011. Contributions of biogenic and anthropogenic hydrocarbons to secondary organic aerosol during 2006 in Research Triangle Park, NC. Aerosol Air Qual. Res. 11, 99–108.
Peng, R.D., Bell, M.L., Geyh, A.S., McDermott, A., Zeger, S.L., Samet, J.M., et al., 2009. Emergency admissions for cardiovascular and respiratory diseases and the chemical composition of fine particle air pollution. Environ. Health Perspect. 117, 957–963.
Rhee, J., Dominici, F., Zanobetti, A., Schwartz, J., Wang, Y., Di, Q., et al., 2019. Impact of long-term exposures to ambient PM2. 5 and ozone on ARDS risk for older adults in the United States. Chest 156, 71–79.
Russell, A.G., McRae, G.J., Cass, G.R., 1967. Mathematical modeling of the formation and transport of ammonium nitrate aerosol. Atmos. Environ. 17, 949–964, 1983.
Sarnat, J.A., Marmur, A., Klein, M., Kim, E., Russell, A.G., Sarnat, S.E., et al., 2008. Fine particle sources and cardiorespiratory morbidity: an application of chemical mass balance and factor analytical source-apportionment methods. Environ. Health Perspect. 116, 459–466.
Sarnat, J.A., Wilson, W.E., Strand, M., Brook, J., Wyzga, R., Lumley, T., 2007. Panel discussion review: session 1—exposure assessment and related errors in air pollution epidemiologic studies. J. Expo. Sci. Environ. Epidemiol. 17, S75–S82.
Sarnat, S.E., Winquist, A., Schauer, J.J., Turner, J.R., Sarnat, J.A., 2015. Fine particulate matter components and emergency department visits for cardiovascular and respiratory diseases in the St. Louis, Missouri–Illinois, metropolitan area. Environ. Health Perspect. 123, 437–444.
Sheppard, L., Slaughter, J.C., Schildcrout, J., Liu, L.S., Lumley, T., 2005. Exposure and measurement contributions to estimates of acute air pollution effects. J. Expo. Sci. Environ. Epidemiol. 15, 366–376.
Shi, G., Liu, J., Wang, H., Tian, Y., Wen, J., Shi, X., et al., 2018. Source apportionment for fine particulate matter in a Chinese city using an improved gas-constrained method and comparison with multiple receptor models. Environ. Pollut. 233, 1058–1067.
Shi, G., Xu, J., Peng, X., Xiao, Z., Chen, K., Tian, Y., et al., 2017. pH of aerosols in a polluted atmosphere: source contributions to highly acidic aerosol. Environ. Sci. Technol. 51, 4289–4296.
Silverman, R.A., Ito, K., 2010. Age-related association of fine particles and ozone with severe acute asthma in New York City. J. Allergy Clin. Immunol. 125, 367–373 e5.
Slawsky, E., Ward-Caviness, C.K., Neas, L., Devlin, R.B., Cascio, W.E., Russell, A.G., et al., 2021. Evaluation of PM2.5 air pollution sources and cardiovascular health. Environ. Epidemiol. 5, e157.
Stafoggia, M., Samoli, E., Alessandrini, E., Cadum, E., Ostro, B., Berti, G., et al., 2013. Short-term associations between fine and coarse particulate matter and hospitalizations in Southern Europe: results from the MED-PARTICLES project. Environ. Health Perspect. 121, 1026–1033.
Stelson, A., Seinfeld, J.H., 1967. Relative humidity and temperature dependence of the ammonium nitrate dissociation constant. Atmos. Environ. 16, 983–992, 1982.
Strand, M., Vedal, S., Rodes, C., Dutton, S.J., Gelfand, E.W., Rabinovitch, N., 2006. Estimating effects of ambient PM 2.5 exposure on health using PM 2.5 component measurements and regression calibration. J. Expo. Sci. Environ. Epidemiol. 16, 30–38.
Vedal, S., Campen, M.J., McDonald, J.D., Larson, T.V., Sampson, P.D., Sheppard, L., et al., 2013. National Particle Component Toxicity (NPACT) Initiative Report on Cardiovascular Effects. Research Report. Health Effects Institute, pp. 5–8.
Watson, J.G., Chow, J., Fujita, E., 2004. Protocol for Applying and Validating the CMB Model for PM2. 5 and VOC. US Environmental Protection Agency, Research Triangle Park, NC.
Weber, R.J., Guo, H., Russell, A.G., Nenes, A., 2016. High aerosol acidity despite declining atmospheric sulfate concentrations over the past 15 years. Nat. Geosci. 9, 282–285.
Wing, S., Cole, D., Grant, G., 2000. Environmental injustice in North Carolina's hog industry. Environ. Health Perspect. 108, 225–231.
Ye, D., Klein, M., Mulholland, J.A., Russell, A.G., Weber, R., Edgerton, E.S., et al., 2018. Estimating acute cardiovascular effects of ambient PM 2.5 metals. Environ. Health Perspect. 126, 027007.
Zeger, S.L., Thomas, D., Dominici, F., Samet, J.M., Schwartz, J., Dockery, D., et al., 2000. Exposure measurement error in time-series studies of air pollution: concepts and consequences. Environ. Health Perspect. 108, 419–426.
Zhai, X., Mulholland, J.A., Russell, A.G., Holmes, H.A., 2017. Spatial and temporal source apportionment of PM2.5 in Georgia, 2002 to 2013. Atmos. Environ. 161, 112–121.