Taylor & Francis
Taylor & Francis Group

Journal of the Air & Waste Management Association

ISSN: 1096-2247 (Print) 2162-2906 (Online) Journal homepage: https://www.tandfonline.com/loi/uawm20

A hybrid model for spatially and temporally resolved ozone exposures in the continental United States

Qian Di, Sebastian Rowland, Petros Koutrakis & Joel Schwartz

To cite this article: Qian Di, Sebastian Rowland, Petros Koutrakis & Joel Schwartz (2017) A hybrid model for spatially and temporally resolved ozone exposures in the continental United States, Journal of the Air & Waste Management Association, 67:1, 39-52, DOI: 10.1080/10962247.2016.1200159

To link to this article: https://doi.org/10.1080/10962247.2016.1200159

View supplementary material

Published online: 22 Jun 2016.

Submit your article to this journal

Article views: 1636

View related articles

View Crossmark data

Citing articles: 35 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=uawm20

JOURNAL OF THE AIR & WASTE MANAGEMENT ASSOCIATION
2017, VOL. 67, NO. 1, 39–52
http://dx.doi.org/10.1080/10962247.2016.1200159

Taylor & Francis
Taylor & Francis Group

TECHNICAL PAPER

A hybrid model for spatially and temporally resolved ozone exposures in the continental United States

Qian Di, Sebastian Rowland, Petros Koutrakis, and Joel Schwartz

Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA

ABSTRACT

Ground-level ozone is an important atmospheric oxidant, which exhibits considerable spatial and temporal variability in its concentration level. Existing modeling approaches for ground-level ozone include chemical transport models, land-use regression, Kriging, and data fusion of chemical transport models with monitoring data. Each of these methods has both strengths and weaknesses. Combining those complementary approaches could improve model performance. Meanwhile, satellite-based total column ozone, combined with ozone vertical profile, is another potential input. The authors propose a hybrid model that integrates the above variables to achieve spatially and temporally resolved exposure assessments for ground-level ozone. The authors used a neural network for its capacity to model interactions and nonlinearity. Convolutional layers, which use convolution kernels to aggregate nearby information, were added to the neural network to account for spatial and temporal autocorrelation. The authors trained the model with the Air Quality System (AQS) 8-hr daily maximum ozone in the continental United States from 2000 to 2012 and tested it with left out monitoring sites. Cross-validated R^2 on the left out monitoring sites ranged from 0.74 to 0.80 (mean 0.76) for predictions on 1 km × 1 km grid cells, which indicates good model performance. Model performance remains good even at low ozone concentrations. The prediction results facilitate epidemiological studies to assess the health effect of ozone in the long term and the short term.

Implications: Ozone monitors do not provide full data coverage over the United States, which is an obstacle to assess the health effect of ozone when monitoring data are not available. This paper used a hybrid approach to combine satellite-based ozone measurements, chemical transport model simulations, land-use terms, and other auxiliary variables to obtain spatially and temporally resolved ground-level ozone estimation.

PAPER HISTORY
Received February 15, 2016
Revised April 28, 2016
Accepted April 28, 2016

Introduction

Ground-level ozone is a serious public health concern. The adverse effects of ozone are well documented including respiratory symptoms (Schwartz et al., 1994; Hao et al., 2015; Gent et al., 2003), the development of asthma (McConnell et al., 2002; Sousa et al., 2013), airway inflammation (Koren et al., 1989; Tank et al., 2011), and mortality (Franklin and Schwartz, 2008; Turner et al., 2015; Atkinson et al., 2016; Bell and Dominici, 2008). These health effects have been reported for both long- and short-term exposures (Jerrett et al., 2009; Bell et al., 2004). Ozone is one of criteria pollutants regulated by the U.S. Environmental Protection Agency (EPA) based on maximum of 8-hr average. Ground-level ozone is a product of photochemical reactions involving nitric oxide (NO), nitrogen dioxide (NO_2), hydrocarbons, nitrogen oxides (NO_x), and volatile organic compounds (VOCs). Ground-level ozone concentration is typically characterized by a diurnal variability with peak concentrations occurring at daytime. Many parameters, including local combustion sources, land-surface characteristics, and atmospheric conditions, influence ozone formation and removal, resulting in high spatial and temporal variability of ozone concentration. Therefore, predicting ozone concentrations is challenging, especially at fine resolutions.

Fine spatial and temporal resolutions are critical to assessing human exposures for health studies. Many early epidemiological studies used ozone measurements from the nearest monitoring sites to assign exposure (Jerrett et al., 2009). This approach introduces nondifferential measurement error, because it fails to capture ozone scavenging by NO and other sources of local variability.

Other approaches of accessing ozone concentrations involve spatial interpolation, land-use regression, satel-

CONTACT Qian Di qid335@mail.harvard.edu Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA 02215, USA.

Supplemental data for this article can be accessed on the publisher's website.

© 2017 A&WMA

lite-based data modeling, and chemical transport model. *Spatial interpolation*, such as inverse distance weighting (Breton et al., 2012) and Kriging (Tranchant and Vincent, 2000), was used to estimate ozone exposures for epidemiology studies. Often, a radius threshold is chosen in interpolation (Bell, 2006). Spatial interpolation has the advantage of low computation cost and reduces measurement error, but often generates oversmoothed distributions, which inadequately represents local variability (Abraham and Comrie, 2004). Due to complex transport and chemistry, terrain variability can cause ozone concentration to vary remarkably within a short distance, which imposes an even greater challenge for spatial interpolation (Loibi et al., 1994). *Land-use regression* (LUR) assumes that land-use terms are predictors for ozone level and uses covariates such as traffic, population density, and elevation to model ozone (Malmqvist et al., 2014). LUR is relatively easy to implement and has satisfying model performance at small scales, but has limited capacity to capture temporal variations and can miss some short-term and regional patterns (Hoek et al., 2008). *Satellite* observations measure ozone over larger spatial and temporal scales than most LURs. Most satellite ozone measurements are column-based, such as total ozone mapping spectrometer (TOMS), global ozone monitoring experiment (GOME) (Burrows et al., 1999), and ozone monitoring instrument (OMI) (Levelt et al., 2006). Some satellite measurements also provide vertical distribution of ozone, including solar backscatter ultra violet (SBUV), GOME, and later OMI. Two OMI ozone data products, produced by the OMI-TOMS and the OMI Differential Optical Absorption Spectroscopy (OMI-DOAS) retrieval algorithms, demonstrate high agreement with total column ozone observation at a global scale, with about 1% disagreement (Balis et al., 2007; McPeters et al., 2008). At ground level, OMI ozone observations are close to ground monitor-based mean concentrations, but at higher elevations these observations deviate from the monitors (Wang et al., 2011). The discrepancy can be as large as 20% (Liu et al., 2010a).

A *chemical transport model* (CTM) is a more advanced tool of estimating ozone, which simulates the formation, dispersion, and deposition of ozone. CTMs, such as GEOS-Chem (Bey et al., 2001), MOZART (Model of Ozone and Related Tracers; Brasseur et al., 1998), and CMAQ (Community Multiscale Air Quality Model; Byun and Schere, 2006) have been applied to estimate ground-level ozone at city level (Lei et al., 2007; Sokhi et al., 2006), country level (Liu et al., 2010b; Tong and Mauzerall, 2006), continent level (Fusco and Logan, 2003; Pfister et al., 2008), or beyond. Due to limitations of both computational capacity and the spatial resolution of emission inventories, ozone estimation from CTM is usually not spatially resolved enough to assess exposure at local scale. Typical scales are 4° × 5°, 2.0° × 2.5°, 0.500° × 0.667°, or 0.2500° × 0.3125°, although CMAQ-Urban can produce very-fine-scale predictions in selected urban locations with good emission inventories. However, CTMs deviate substantially from real-world measurements due to imperfect data and chemistry, and these errors tend to increase at finer time or spatial scales. One limitation of many ozone models is that their performance is only tested against the monitoring sites used to train the models, which does not test the validity of the model in areas without monitoring data. Cross-validation can test model validity at unmonitored areas by leaving out monitors during model training, and subsequently testing the correlation between the model and the left out monitors.

With both strengths and weaknesses, the aforementioned approaches are complementary to each other. This study proposes a hybrid approach, which integrates informative variables and existing ground-level ozone modeling approaches into a neural network–based framework. Ten-fold cross-validation was used to test model performance and avoid overfitting. After model training, we predicted ground-level ozone at nationwide 1 km × 1 km grid cells and produced spatially and temporally resolved ozone exposure assessments, which can be used by epidemiologists to assess the acute and chronic health effects of ozone.

A similar hybrid approach has been applied to assess human exposures to $PM_{2.5}$ (particulate matter with an aerodynamic diameter <2.5 μm) mass and chemical components (Di et al., 2015, 2016; Kloog et al., 2011, 2014). This study applies a hybrid approach similar to the previous model of $PM_{2.5}$ but incorporates additional variables due to ozone's distinct gaseous nature and chemical characteristics. We present a new model for ground-level ozone that relies on multiple data sources and the application of neural network with convolutional layers.

Data and methods

Study domain

The spatial area is the continental United States, which includes the 48 contiguous states and Washington, D.C. The study period is 2000–2012, covering 4749 days.

Monitoring data

Monitoring data for ozone concentrations across the study area were collected by the EPA Air Quality System (AQS). There were 1877 monitoring sites available within the study area during the study period, but some of them reported data for a subset of the study

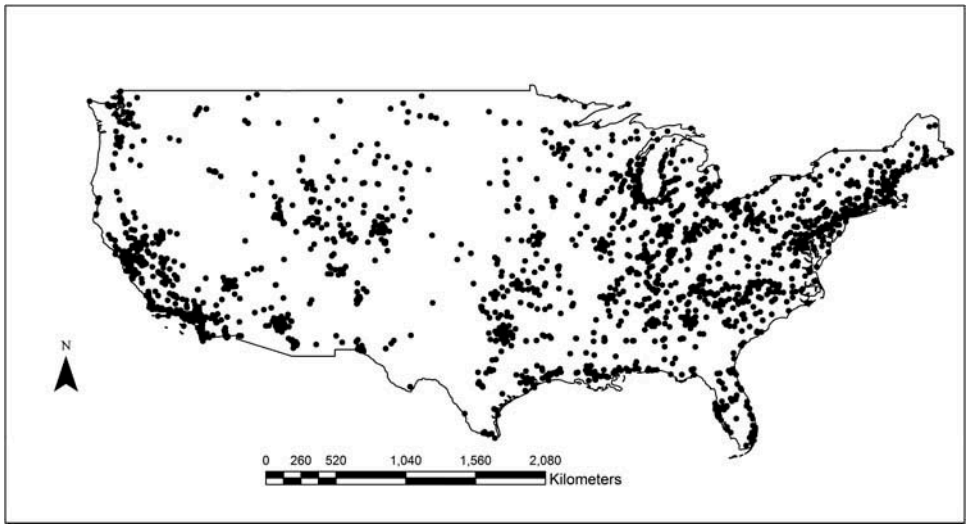

Figure 1. Ozone monitoring sites in the United States.

period or reported data intermittently. Monitoring sites were densely located in the eastern United States and the western coast, whereas the Mountain region and other remote areas had fewer monitoring sites (Figure 1). We calibrated the model to the 8-hr daily maximum ozone (daily 8hr-max ozone). In this paper, unless specified otherwise, the term "ozone" refers to daily 8hr-max ozone at ground level.

Chemical transport model output

We used GEOS-Chem version 9.0.2 to simulate ozone formation, dispersion, and deposition. GEOS-Chem incorporates meteorological inputs, emission inventories, and atmospheric chemical reactions. Its methodology has been described in previous literature (Bey et al., 2001). We first performed a global 2.0° × 2.5° simulation and exported boundary conditions. We then performed a nested grid simulation at 0.500° × 0.667° for the North America. For years from 2000 to 2004, 2.0° × 2.5° outputs were used instead because meteorological inputs at 0.500° × 0.667° were not available.

Satellite-based ozone measurements

The OMI instrument is on board the Earth Observing System (EOS)-Aura satellite, which was launched in July 2004 (Levelt et al., 2006). OMI's raw data were processed by two distinct algorithms, which yielded two different data products. Data product OMTO3e (version 003) was produced from the TOMS version 8.5 algorithm, which is based on the TOMS version 8 algorithm (Bhartia and Wellemeyer, 2002). The other data product, OMDOAO3e (version 003), was produced from the OMI-DOAS algorithms (Veefkind et al., 2006). The two algorithms generally agree with each other, with a mean difference in the total column ozone below 3%, although larger differences occur at high latitude areas and over clouds (Kroon et al., 2008). Both data products have a spatial resolution of 0.25° × 0.25° and are available since July 2004.

Ozone vertical profile

Satellite instruments measure total column ozone; however, the vertical distribution profile is needed to obtain ground-level ozone concentration. We adopted an approach similar to the approach used in modeling $PM_{2.5}$, where aerosol optical depth (AOD) is a column measurement of aerosol and researchers used the vertical profile from a chemical transport model to calibrate AOD to ground-based $PM_{2.5}$ (Liu et al., 2004; Van Donkelaar et al., 2010). GEOS-Chem simulates ozone concentrations at different layers. We defined a scaling factor as the fraction of ground-level ozone in the total column ozone and used this factor to calibrate satellite-based column ozone to ground-level ozone. One advantage of GEOS-Chem ozone vertical profile is the absence of missing values. GEOS-Chem tropospheric ozone predictions agree with monitor observations in terms of the overall characteristics, but significant differences exist by region and by season (Liu et al., 2006). OMI also provides ozone vertical profile (data product OMO3PR version 003) (Ahmad et al., 2003), in which an optimal estimation algorithm adjusts ozone in each atmospheric layer based on a priori information and minimizes the difference between modeled and measured ozone (Rodgers, 2000). Although some missing values occur

occasionally, comparison of retrieved and measured ozone indicates good agreement (Veefkind et al., 2009). The OMI ozone profile has a spatial resolution of 13 km × 48 km. We linearly interpolated the data at all missing values.

NO_x, SO_2, and VOC data

Ozone precursors include nitrogen oxides (NO_x), carbon monoxide (CO), methane (CH_4), and volatile organic compounds (VOCs). Ozone precursors react with the presence of sunlight and form ozone. NO, in contrast, decreases ozone concentration by inducing ozone scavenging (Graedel et al., 1977). Although emission inventories of these compounds are used in the GEOS-Chem model, they lack the temporal resolution of the monitoring data. To account for those relevant atmospheric reactions, we included AQS daily measurements of sulfur dioxide (SO_2), nitrogen dioxide (NO_2), NO_x, and VOCs into our ozone model. AQS measurements are point measurements and sparsely located. We applied distance decay functions to aggregate point data from monitors into convolutional layers (section Convolutional Layer, Supplemental Material).

In order to obtain higher spatial and temporal coverage, we also used satellite-based total column SO_2 and total column NO_2 from OMI data products (OMSO2e version 003 and OMNO2d version 003) (Krotkov et al., 2011).

Meteorological data

Our model used meteorological fields from the National Centers for Environmental Prediction (NCEP) North American Regional Reanalysis data. This data set assimilates multiple measurements from land-surface, ship, radiosonde, pibal, aircraft, satellite, and other sources, with a resolution of 0.3° (about 32 km) at the daily level (Kalnay et al., 1996). The reanalysis data set was chosen because it has both relatively high spatiotemporal resolution and no missing values. We used 16 meteorological variables in order to fully capture meteorological conditions and account for complex atmospheric processes. The variables included air temperature, accumulated total precipitation, downward shortwave radiation flux, accumulated total evaporation, planetary boundary layer height, low cloud area fraction, precipitation rate, precipitable water for the entire atmosphere, pressure, specific humidity at 2 m, visibility, wind speed, medium cloud area fraction, high cloud area fraction, and albedo. Wind speed was computed as the vector sum of u-wind (east-west component of the wind) at 10 m and v-wind (north-south component) at 10 m.

Land-use terms

Land-use terms are proxies for ozone formation or removal and capture spatial variations at local scale, which may not be measured by satellite or modeled by GEOS-Chem. The detailed procedure of processing elevation, road density, NEI (National Emissions Inventory), population density, percentage of urban, and NDVI (normalized difference vegetation index) has been specified somewhere else (Kloog et al., 2012). We used two variables to approximate vegetation: the percentage of vegetation from NCEP North American Regional Reanalysis data and 16-day 1-km MODIS (Moderate Resolution Imaging Spectroradiometer) NDVI data product MOD13A2 (Didan, 2015). For days without NDVI values, we linearly interpolated values from neighboring days.

Regional and monthly dummy

Regional and monthly dummy variables were used to capture different associations between the above variables and monitored ozone by season and climate type. The major climate types were used to define the regional dummy variable (Kottek et al., 2006).

Neural network

We used a neural network for its capacity to model nonlinearity and interactions among variables (Bishop, 1995; Haykin, 2004). The target variable was monitored ozone from the AQS network, and the predictor variables included the aforementioned variables. The input variables were available for the entire study area. Some variables had a small proportion of missing values, and we estimated the missing data using linear interpolation (Table S2, Supplemental Material). Not all variables were available during the entire study period. For each year, we fitted a neural network with available variables in that year. Most existing studies fitted models with in situ information, the values of each variable at the monitoring sites; however, information about neighboring areas can be also informative. For instance, nearby traffic volume influences in situ ozone levels by either providing ozone precursors or scavenging ozone. To incorporate the nearby information into the neural network, we used convolutional layers (LeCun and Bengio, 1995). A convolutional layer is computed by applying a convolution kernel (e.g., mean, inverse distance–weighted mean) to the inputs in order to compute a scalar summary of the neighboring cells, which is then used as an additional predictor. By choosing kernels, we obtained different aggregations of neighboring information, which gave the neural network more flexibility to capture spatial

autocorrelation and improved model fit. We computed convolutional layers for each land-use variable, predicted ozone of nearby areas, and predicted ozone of proceeding and subsequent days. To create the convolutional layers for predicted ozone, we first fitted the neural network and obtained intermediate ozone predictions. Then we computed spatial and temporal convolutional layers for predicted ozone and fitted the neural network again with those convolutional layers (Figure S2). The details of convolutional layers and fitting a neural network are presented in Supplemental Material.

We used 10-fold cross-validation to validate neural network results, in which all monitors were randomly divided into 10 splits. We then trained a neural network with 9 splits of the monitors and made ozone predictions for the remaining 1 split. The process was repeated nine times and made ozone predictions for the other 9 splits. Combining the predicted ozone from the 10 splits together yielded ozone predictions for all monitors. We calculated total R^2, spatial R^2, and temporal R^2 for all monitors as well as by region and season to evaluate model performance. Calculations of R^2 and other metrics of model performance (bias and slope) are specified in Supplemental Material.

To make ozone predictions, we trained a neural network with all monitors. The trained neural network was used to predict ozone at 1 km × 1 km grid cells for the whole study area during the entire study period. We prepared input variables at 1 km × 1 km grid cells and made ozone predictions with the trained neural network. We linearly interpolated the data if missing values were present. All programming work was implemented in MATLAB (version 2014a; The MathWorks, Inc., Natick, MA).

Results

After conducting 10-fold cross-validation, total R^2 ranged from 0.74 to 0.80, with mean R^2 = 0.76 (Table 1). Slope was near 1; bias was about 1.20 ppb for the whole concentration range and 2.82 ppb below 75 ppb (Tables 1, S3). Model performance did not vary much by year, nor was there any temporal trend in model fit. In contrast, model performance varied by season, with the highest R^2 observed in autumn, followed by summer, spring, and winter (Table 2). By region, model performance in the Middle Atlantic, South Atlantic, East North Central, West South Central, and Pacific states was near or above the national average, whereas the New England, Mountain, and West North Central states were below the national average (Table 3). The above regional division is from the U.S. Census Bureau (Table S1, Figure S7). Figure 2 visualizes model fits for the study area. Wyoming, Montana, western Colorado, eastern Washington State, eastern Tennessee, and Marine had lower fits than other states.

Figure 3 visualizes the spatial pattern of ozone in the study area. The Mountain states had the highest ozone levels for all seasons. Areas around the Appalachian Mountains also witnessed high ozone levels, although less so. The eastern United States, with much lower

Table 1. Cross-validated total R^2, spatial R^2, temporal R^2, and corresponding MSE between monitored and predicted ozone in each year for the study area.

Year	Total R^2	RMSE	Spatial R^2	RMSE	Temporal R^2	RMSE	Bias	Slope
2000	0.79	8.09	0.82	3.30	0.78	7.42	1.38	0.99
2001	0.78	8.12	0.80	3.51	0.77	7.34	1.21	0.99
2002	0.80	8.09	0.82	3.30	0.80	7.39	1.12	0.99
2003	0.77	7.98	0.78	3.21	0.77	7.33	1.33	0.99
2004	0.76	7.44	0.76	2.98	0.76	6.83	1.34	0.99
2005	0.78	7.84	0.78	3.11	0.77	7.22	1.37	0.99
2006	0.77	7.52	0.78	2.95	0.76	6.94	1.10	0.99
2007	0.77	7.32	0.85	2.88	0.76	6.76	1.13	0.99
2008	0.76	7.12	0.82	2.75	0.75	6.58	1.09	0.99
2009	0.76	6.61	0.79	2.57	0.75	6.13	1.18	0.98
2010	0.74	6.90	0.82	2.58	0.73	6.42	1.31	0.98
2011	0.75	7.01	0.80	2.72	0.74	6.50	1.08	0.99
2012	0.78	6.76	0.83	2.59	0.77	6.29	0.96	0.99
Mean	0.76	7.36	0.80	2.91	0.75	6.79	1.20	0.99

Note. Unit for RMSE (root mean square error) is ppb; the unit for bias is ppb.

Table 2. Cross-validated total R^2, spatial R^2, temporal R^2, and corresponding MSE between monitored and predicted ozone in each year divided by season.

Year	Spring R^2	RMSE	Summer R^2	RMSE	Autumn R^2	RMSE	Winter R^2	RMSE	Annual R^2	RMSE
2000	0.70	7.78	0.73	9.23	0.77	7.72	0.69	6.30	0.79	8.09
2001	0.71	7.54	0.71	9.59	0.74	7.61	0.69	6.27	0.78	8.12
2002	0.67	6.97	0.75	9.70	0.80	7.87	0.66	6.34	0.80	8.09
2003	0.71	7.55	0.71	9.19	0.75	7.70	0.67	6.10	0.77	7.98
2004	0.68	7.03	0.70	8.42	0.77	7.22	0.65	6.36	0.76	7.44
2005	0.67	7.16	0.70	9.18	0.77	7.52	0.70	6.05	0.78	7.84
2006	0.65	7.08	0.70	8.93	0.73	6.89	0.68	5.66	0.77	7.52
2007	0.71	6.76	0.71	8.59	0.75	7.06	0.66	5.91	0.77	7.32
2008	0.68	6.48	0.70	8.34	0.74	6.80	0.69	5.47	0.76	7.12
2009	0.68	6.34	0.70	7.43	0.75	6.41	0.66	5.83	0.76	6.61
2010	0.64	6.53	0.69	7.77	0.75	6.73	0.65	6.16	0.74	6.90
2011	0.63	6.51	0.69	8.08	0.75	6.68	0.65	5.36	0.75	7.01
2012	0.69	6.28	0.72	8.09	0.73	6.39	0.65	5.95	0.78	6.76
Mean	0.68	6.92	0.71	8.66	0.75	7.12	0.67	5.98	0.76	7.36

Note. The definition of seasons was specified in Figure 2. Unit for RMSE (root mean square error) is ppb; the unit for bias is ppb.

Table 3. Cross-validated total R^2, spatial R^2, temporal R^2, and corresponding MSE between monitored and predicted ozone in each year divided by U.S. Census Divisions.

Year	New England	Middle Atlantic	East North Central	West North Central	South Atlantic	East South Central	West South Central	Mountain	Pacific	National
2000	0.70	0.81	0.79	0.74	0.78	0.74	0.79	0.77	0.80	0.79
2001	0.70	0.83	0.81	0.73	0.79	0.74	0.76	0.73	0.78	0.78
2002	0.74	0.85	0.85	0.76	0.82	0.79	0.76	0.74	0.78	0.80
2003	0.68	0.81	0.81	0.70	0.77	0.74	0.76	0.76	0.78	0.77
2004	0.63	0.81	0.77	0.71	0.76	0.72	0.76	0.74	0.78	0.76
2005	0.69	0.82	0.81	0.76	0.79	0.74	0.76	0.73	0.76	0.78
2006	0.71	0.82	0.79	0.74	0.76	0.72	0.77	0.73	0.76	0.77
2007	0.69	0.82	0.82	0.73	0.78	0.75	0.77	0.73	0.75	0.77
2008	0.71	0.82	0.77	0.69	0.80	0.73	0.74	0.72	0.76	0.76
2009	0.72	0.78	0.77	0.72	0.75	0.75	0.76	0.69	0.77	0.76
2010	0.71	0.80	0.72	0.69	0.76	0.68	0.74	0.68	0.75	0.74
2011	0.66	0.78	0.75	0.67	0.75	0.72	0.77	0.68	0.75	0.75
2012	0.69	0.82	0.81	0.78	0.75	0.76	0.75	0.74	0.77	0.78
Mean	0.60	0.73	0.69	0.62	0.68	0.62	0.69	0.67	0.73	0.70

Note. Unit for RMSE (root mean square error) is ppb; the unit for bias is ppb.

ozone year round, experienced higher ozone levels in summer. Figure 3 also presents low concentrations in cities and along highways. In terms of temporal trend, Figure 4 presents a general decreasing trend of ozone, although less obvious in some regions.

Discussion

This study proposed a hybrid model framework, which integrated satellite-based data, CTM outputs, ozone vertical profiles, meteorological variables, land-use terms, and other atmospheric compounds that were related to ozone formation or deposition. Convolutional layers aggregated nearby information and improved model fit. The average cross-validated R^2 between predicted and monitored daily 8hr-max ozone was 0.76 (0.74~0.80 by year). Few existing studies have ever modeled 8-hr maximum ground-level ozone at daily basis or attempted to make predictions at nationwide 1 km × 1 km grid cells. We believe that this level of temporal/spatial coverage and model performance is an improvement over previous ozone prediction approaches. Epidemiological studies investigating the acute and chronic effects of ozone will benefit from more accurate and granular exposure assessments.

Our hybrid approach has several advantages and innovations. First, model performance surpasses existing studies. Some previous studies adopted land-use regression, Kriging or other methods and achieved root mean square error (RMSE) > 10 ppb in Belgium (Hooyberghs et al., 2006); RMSE > 10 ppb in Italy (Carnevale et al., 2008); and daily R^2 = 0.653 in Quebec (Adam-Poupart et al., 2014). Our hybrid model outperformed land-use regression results, with averaged cross-validated annual R^2 = 0.76 and RMSE = 7.36 ppb. Another improvement is that LUR is usually constrained to specific locations, whereas our hybrid model covers the entire continental United States. In terms of CTMs, some CMAQ simulations achieved normalized mean error (NME) less than 35% over the continental United States in summer (Tong and Mauzerall, 2006), improved to NME 17.9% but focused on the eastern United States (Appel et al., 2007), and continued to obtain NME between 17.7% and 21.7% (Zhang et al., 2009). CMAQ simulation was becoming better over time, but our hybrid model still outperformed it with a cross-validated NME = 13.13%. Combing multiple CTM simulations and comparing with monitored ozone, some researchers obtained mean R^2 = 0.57 for the continental United States for the whole year (Reidmiller et al., 2009), compared with mean R^2 = 0.76 in our study. This indicates that our hybrid model surpasses CTM simulations as a whole. Besides, convolutional layers take neighboring information into account, which is also applicable to other studies. Other methods, such as Kriging, have been widely used to aggregate nearby information in ozone modeling. For a convolutional layer, the specific aggregation depends on the kernel function, which is more versatile than Kriging. More importantly, being an input layer of a neural network, a convolutional layer can have complex interaction with other variables, which can better capture much more complex nonlinear atmospheric processes. By introducing convolutional layers, this study introduces a new way of incorporating neighboring information to improve model performance.

We integrated multiple data sources into a single ozone modeling framework and improved model fit. Not all of the variables contributed equally to model performance. Satellite-based ozone measurement, GEOS-Chem simulations, and land-use terms were critical to model performance. Hence, previous studies also combined land-use regression with chemical transport model (Akita et al., 2014), or land-use regression with Kriging (Wang et al., 2015), at the

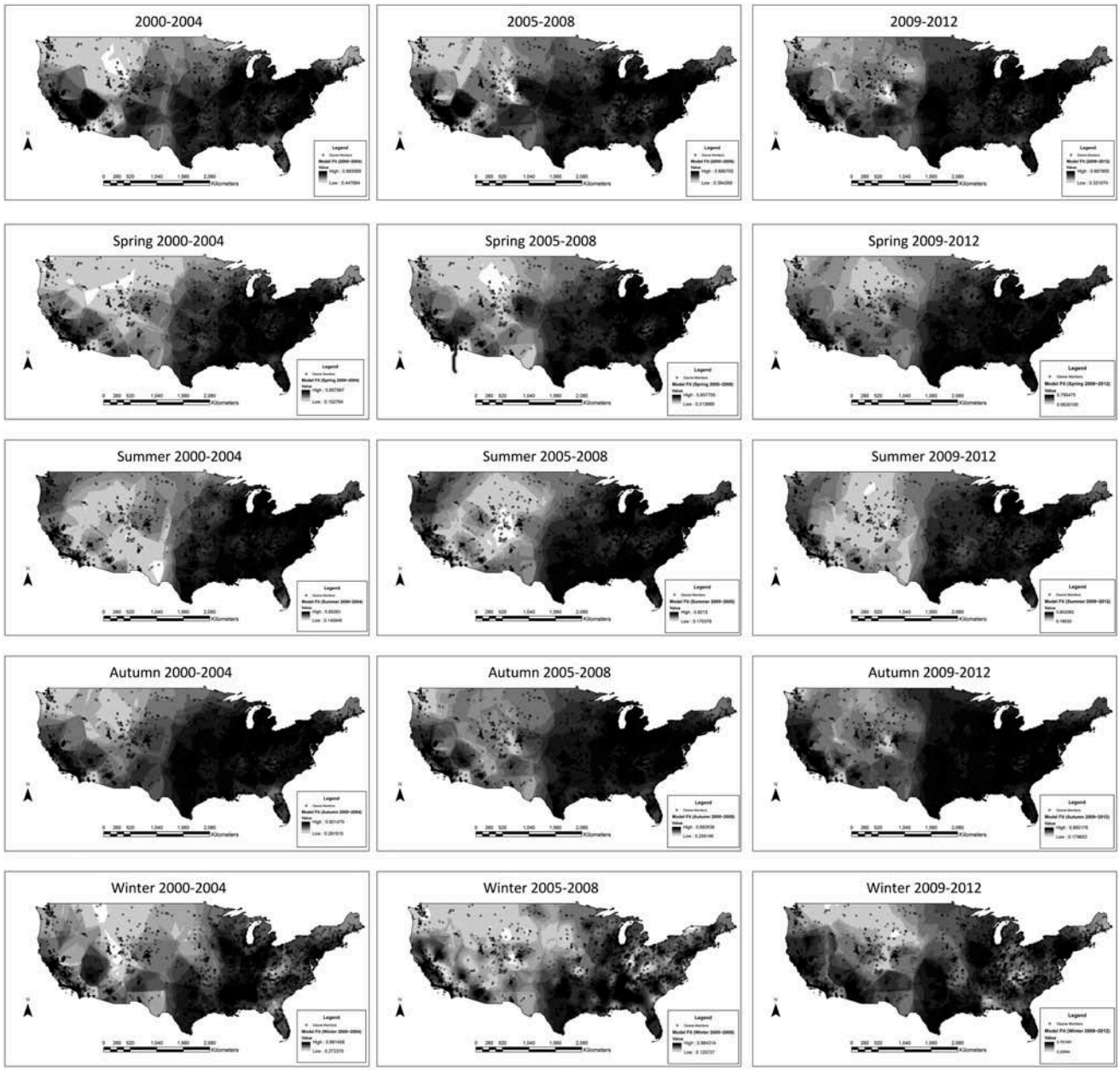

Figure 2. Model performance in the continental United States. This figure visualizes the total R^2 between monitored and predicted ozone. We interpolated R^2 to areas without monitors using Kriging interpolation. Spring was defined as March to May; summer was defined as June to August; autumn was defined as September to November; winter was from December to February of the next year.

regional or municipal scale. Other variables, including regional dummy variables, and certain meteorological variables played an auxiliary role. Some variables are complementary to each other. For example, satellite-based instruments, such as OMI, have daily measurements with a large spatial coverage, but their values are averaged column measurements of ozone for a large volume of air. AQS monitors measure ground-level ozone at specific locations. Thus, satellite-based measurements cannot capture variability at small scales like monitors do (Wang et al., 2011). On the other hand, land-use terms are proxies for local emission, which gives rise to local variability, but they usually do not provide much information on the temporal variability. Land-use terms and satellite observations are complementary to each other because land-use terms are at small local scales and satellite observations have wide time and space coverage. Combining both data sets overcomes weaknesses and improves the model. The use of neural network rather than a regression did not singularly drive model performance; a study also used neural network with only

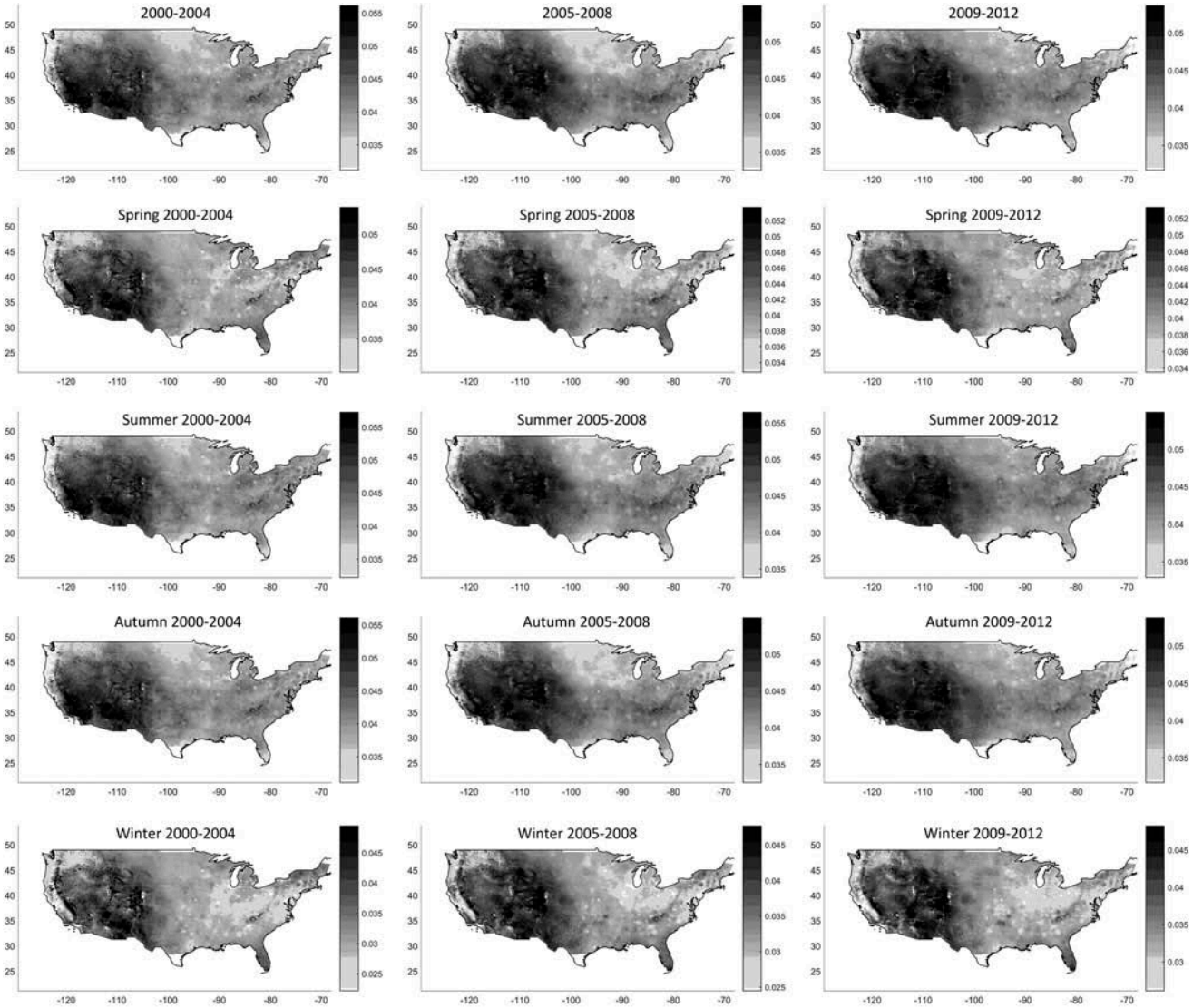

Figure 3. Spatial distribution of predicted ozone. The trained neural network predicted ozone at 1 km × 1 km grid cells. Those figures visualize annual averages and seasonal averages of predicted ozone for 2000~2004, 2005~2008, and 2009~2012. (Seasons were defined as in Figure 2.)

land-use terms and achieved model performance inferior to ours (RMSE > 10 ppb) (Carnevale et al., 2008).

We found an east-west gradient of ozone concentration (Figure 3). High concentrations in the western United States and Mountain states are attributable to factors including high elevation, deep boundary layer, large-scale subsidence, slow ozone deposition to the arid terrain, and slow ozone loss caused by dry conditions (Fiore et al., 2002). The high ground-level ozone in the Mountain states reflects stratospheric intrusion, which can produce some transient peak ozone concentrations at ground level (Davies and Schuepbach, 1994). Compared with the high concentrations in the Mountain states, urban areas had lower ozone. Other air pollutants (e.g., NO) react with ozone and cause ozone scavenging in urban areas, such as San Francisco, Los Angeles, New York City, Houston, and Chicago as well as areas along highways (Figure 3). For the same reason, we observed higher ozone concentrations in rural areas than urban areas in general (Figure 5). We found a general trend of decreasing concentrations over time that agrees with trends observed in monitoring data alone (Camalier et al., 2007; Cox and Chu, 1996), but the trend is less evident at the national level and in several regions (Figure 4). Figure 5 presents temporal trend by season. In spring and autumn, there is an increasing trend over time,

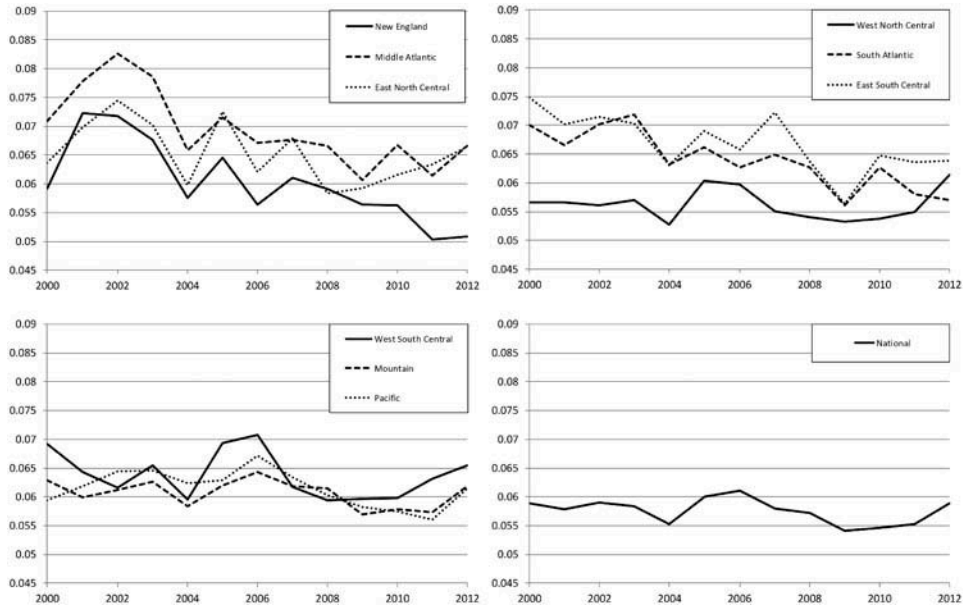

Figure 4. Regional trend in ozone levels at national and regional levels. Concentration is defined by annual fourth maximums of daily 8hr-max.

because NO emission controls in recent years have reduced ozone scavenging and raised background ozone levels. The decreasing summer averages reflect the implemented emission control policies for ozone precursors, but this trend was reversed after the recession. The temporal trend in each region may deviate from the national trend (Figure S4). The regional discrepancies and different effects of emission control in spring, summer, and autumn have been described in previous literature (Cooper et al., 2012). The increasing trend in winter is almost consistent in all regions, which is related to suppressed ozone scavenging due to decreasing NO_x concentration via NO_x titration (Austin et al., 2014; Jhun et al.,

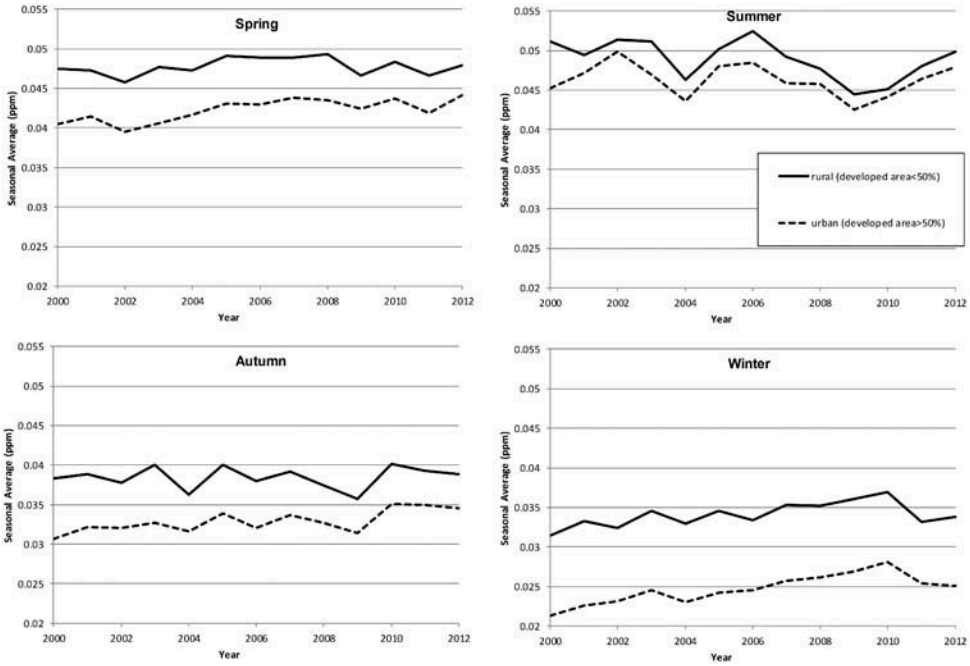

Figure 5. Seasonal averages by urban and rural area. Seasonal averages were computed by averaging predicted ozone at all 1 km × 1 km grid cells in urban or rural areas. Urban areas are defined by developed areas above 50% based on National Land Cover Database (NLCD) 2001, 2006, and 2011 (Fry et al., 2011; Xian et al., 2011).

2014). This suggests a side effect of controlling air pollutant: pollution emission control (e.g., NO_x) may ironically lead to ozone increase under certain conditions (Li et al., 2013).

Model performance was good at typical concentrations. Figure 6 presents that the linearity between predicted and monitored ozone held below 110 ppb. Furthermore, model performance was still good with mean R^2 almost unchanged below 75 ppb, the EPA 8hr-max ozone standard (Table S3). This performance will enable epidemiologists to assess the adverse effect of ozone even at low concentrations. Conversely, the model's linearity had much uncertainty above 120 ppb due to insufficient data (Figure 6); meanwhile, model performance dropped at high concentrations (Figure S5). The inability to accurately predict extreme values is a limitation of our model, which may limit its usage in epidemiological studies that focus on peak concentrations. In terms of model performance over time, there was a slight decreasing trend in temporal R^2, which may result from out-of-date land-use variables. Population density was retrieved for year 2000 and assumed to be constant over time. Population density data for year 2000 do not reflect population density in recent years. Updating population density to be time-varying may improve model performance. This hybrid approach used daily 8hr-max ozone as ozone metric, which avoided noisy ozone fluctuations at night and improved model fit. Although our model performed less well in some remote and sparsely populated areas at daily basis, the annual average demonstrated less discrepancy (Figure S6).

Some limitations remain in our hybrid approach. First, this hybrid approach combines multiple data sets into a single framework and thus requires many variables that may not be available to countries where public available data sets are sparse. Second, the prediction interval is not available in the prediction results. A formal assessment of uncertainty level is critical in epidemiological studies to determine statistical power. Both issues are worthy of further investigations.

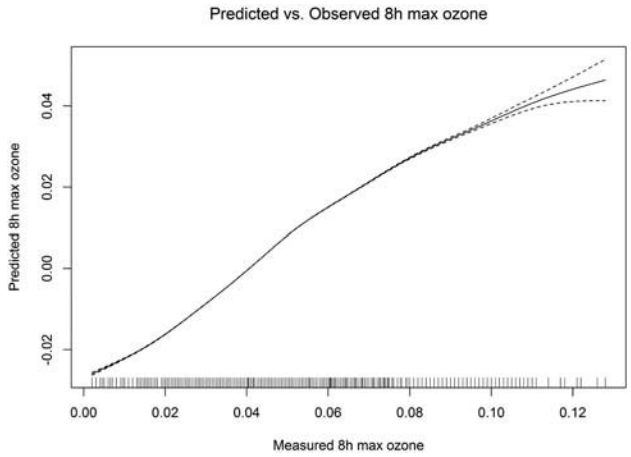

Figure 6. Relationship between measured and predicted ozone. We fitted a regression of predicted ozone on monitored ozone with penalized spline. To assess the linearity between predicted and monitored ozone, we did not specify the degrees of freedom. This figure is for year 2009.

Conclusion

In this paper, we introduced a hybrid model that predicts daily 8hr-max ozone across the continental United States. The main feature of this model is its ability to integrate information from multiple data sources. Specifically, we integrated data from satellite-based ozone measurements, ozone vertical profile, CTM outputs, land-use terms, meteorological variables, concentrations of ozone precursors and other air pollutants, NDVI, and regional/monthly dummy variables. The hybrid model used neural network with convolutional layers, which aggregated information from neighborhood to improve model fit. We calibrated the model using AQS daily 8hr-max ozone measurements. Mean cross-validated R^2 was 0.76, ranging from 0.74 to 0.80 for the entire United States. The model performed better in the eastern United States. The trained neural network predicted daily 8hr-max ozone at nationwide 1 km × 1 km grid cells from 2000 to 2012. These ozone assessments can help scientists investigate the health effect of ozone.

Acknowledgment

The authors thank the China Section of the Air & Waste Management Association for the generous scholarship we received to cover the cost of page charges, which made the publication of this paper possible.

Funding

This publication was made possible by EPA grants RD83479801 and 83587201-0, National Institute of Environmental Health Sciences (NIEHS) grant R01 ES024332-01A1, Health Effects Institute (HEI) grant 4953-RFA14-3/16-4, and National Institutes of Health (NIH grant R21 ES024012 and R01 ES024332. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the EPA. Further, EPA do not endorse the purchase of any commercial products or services mentioned in the publication.

About the authors

Qian Di is a graduate student at the Department of Environmental Health, Harvard T.H. Chan School of Public Health.

Sebastian Rowland is a graduate student at the Department of Environmental Health, Harvard T.H. Chan School of Public Health.

Petros Koutrakis is a professor at the Department of Environmental Health, Harvard T.H. Chan School of Public Health.

Joel Schwartz is a professor at the Department of Environmental Health, Harvard T.H. Chan School of Public Health.

References

Abraham, J.S., and A.C. Comrie. 2004. Real-time ozone mapping using a regression-interpolation hybrid approach, applied to Tucson, Arizona. *J. Air Waste Manage. Assoc.* 54:914–25. doi: 10.1080/10473289.2004.10470960

Adam-Poupart, A., A. Brand, M. Fournier, M. Jerrett, and A. Smargiassi. 2014. Spatiotemporal modeling of ozone levels in Quebec (Canada): A comparison of kriging, land-use regression (LUR), and combined Bayesian maximum entropy–LUR approaches. *Environ. Health Perspect.* 122:970–6. doi: 10.1289/ehp.1306566

Ahmad, S.P., P.F. Levelt, P.K. Bhartia, E. Hilsenrath, G.W. Leppelmeier, and J.E. Johnson. 2003. Atmospheric products from the ozone monitoring instrument (OMI), optical science and technology. In *Proceedings of SPIE conference on Earth Observing Systems VIII, San Diego, California*, 3–8 August *2003*, 619–30. Bellingham WA, USA: International Society for Optics and Photonics.

Akita, Y., J.M. Baldasano, R. Beelen, M. Cirach, K. De Hoogh, G. Hoek, M. Nieuwenhuijsen, M.L. Serre, and A. De Nazelle. 2014. Large scale air pollution estimation method combining land use regression and chemical transport modeling in a geostatistical framework. *Environ. Sci. Technol.* 48:4452–9. doi: 10.1021/es405390e

Appel, K.W., A.B. Gilliland, G. Sarwar, and R.C. Gilliam. 2007. Evaluation of the Community Multiscale Air Quality (CMAQ) model version 4.5: Sensitivities impacting model performance: Part I—Ozone. *Atmos. Environ.* 41:9603–15. doi: 10.1016/j.atmosenv.2007.08.044

Atkinson, R.W., B.K. Butland, C. Dimitroulopoulou, M.R. Heal, J.R. Stedman, N. Carslaw, D. Jarvis, C. Heaviside, S. Vardoulakis, and H. Walton. 2016. Long-term exposure to ambient ozone and mortality: A quantitative systematic review and meta-analysis of evidence from cohort studies. *BMJ Open* 6:e009493. doi: 10.1136/bmjopen-2015-009493

Austin, E., A. Zanobetti, B. Coull, J. Schwartz, D.R. Gold, and P. Koutrakis. 2014. Ozone trends and their relationship to characteristic weather patterns. *J. Expo. Sci. Environ. Epidemiol.* 25:532–42. doi: 10.1038/jes.2014.45

Balis, D., M. Kroon, M. Koukouli, E. Brinksma, G. Labow, J. Veefkind, and R. McPeters. 2007. Validation of Ozone Monitoring Instrument total ozone column measurements using Brewer and Dobson spectrophotometer ground-based observations. *J. Geophys. Res. Atmos.* 112:D24S46. doi: 10.1029/2007JD008796

Bell, M.L. 2006. The use of ambient air quality modeling to estimate individual and population exposure for human health research: A case study of ozone in the Northern Georgia Region of the United States. *Environ. Int.* 32:586–93. doi: 10.1016/j.envint.2006.01.005

Bell, M.L., and F. Dominici. 2008. Effect modification by community characteristics on the short-term effects of ozone exposure and mortality in 98 US communities. *Am. J. Epidemiol.* 167:986–97. doi: 10.1093/aje/kwm396

Bell, M.L., A. McDermott, S.L. Zeger, J.M. Samet, and F. Dominici. 2004. Ozone and short-term mortality in 95 US urban communities, 1987–2000. *JAMA* 292:2372–8. doi: 10.1001/jama.292.19.2372

Bey, I., D.J. Jacob, R.M. Yantosca, J.A. Logan, B.D. Field, A.M. Fiore, Q. Li, H.Y. Liu, L.J. Mickley, and M.G. Schultz. 2001. Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *J. Geophys. Res.* 106:23073–95. doi: 10.1029/2001JD000807

Bhartia, P.K., and C. Wellemeyer. 2002. TOMS-V8 total O_3 algorithm. OMI Algorithm Theoretical Basis Document. Greenbelt, MD: NASA Goddard Space Flight Center. http://eospso.gsfc.nasa.gov/sites/default/files/atbd/ATBD-OMI-02.pdf (accessed December 31, 2015).

Bishop, C.M. 1995. *Neural Networks for Pattern Recognition*. Oxford, UK: Oxford University Press.

Brasseur, G., D. Hauglustaine, S. Walters, P. Rasch, J.F. Müller, C. Granier, and X. Tie. 1998. MOZART, a global chemical transport model for ozone and related chemical tracers: 1. Model description. *J. Geophys. Res. Atmos.* 103:28265–89. doi: 10.1029/98JD02397

Breton, C.V., X. Wang, W.J. Mack, K. Berhane, M. Lopez, T.S. Islam, M. Feng, F. Lurmann, R. McConnell, and H. N. Hodis. 2012. Childhood air pollutant exposure and carotid artery intima-media thickness in young adults. *Circulation* 126:1614–20. doi: 10.1161/CIRCULATIONAHA.112.096164

Burrows, J.P., M. Weber, M. Buchwitz, V. Rozanov, A. Ladstätter-Weißenmayer, A. Richter, R. DeBeek, R. Hoogen, K. Bramstedt, and K.-U. Eichmann. 1999. The global ozone monitoring experiment (GOME): Mission concept and first scientific results. *J. Atmos. Sci.* 56:151–75. doi: 10.1175/1520-0469(1999)056<0151:TGOMEG>2.0.CO;2

Byun, D., and K.L. Schere. 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. doi: 10.1115/1.2128636

Camalier, L., W. Cox, and P. Dolwick. 2007. The effects of meteorology on ozone in urban areas and their use in assessing ozone trends. *Atmos. Environ.* 41:7127–37. doi: 10.1016/j.atmosenv.2007.04.061

Carnevale, C., G. Finzi, E. Pisoni, V. Singh, and M. Volta. 2008. Neural networks and co-kriging techniques to forecast ozone concentrations in urban areas. In *Proceedings of the iEMSs Fourth Biennial Meeting, Barcelona, Spain*, 21 September 2008, 1125–32. Barcelona, Spain: International Environmental Modelling and Software Society.

Cooper, O.R., R.S. Gao, D. Tarasick, T. Leblanc, and C. Sweeney. 2012. Long-term ozone trends at rural ozone monitoring sites across the United States, 1990–2010. *J. Geophys. Res. Atmos.* 117:D22307. doi: 10.1029/2012JD018261

Cox, W.M., and S.-H. Chu. 1996. Assessment of interannual ozone variation in urban areas from a climatological

perspective. *Atmos. Environ*. 30:2615–25. doi: 10.1016/1352-2310(95)00346-0

Davies, T., and E. Schuepbach. 1994. Episodes of high ozone concentrations at the earth's surface resulting from transport down from the upper troposphere/lower stratosphere: A review and case studies. *Atmos. Environ*. 28:53–68. doi: 10.1016/1352-2310(94)90022-1

Di, Q., I. Kloog, P. Koutrakis, A. Lyapustin, Y. Wang, and J. Schwartz. 2015. Assessing $PM_{2.5}$ exposures with high spatio-temporal resolution across the continental United States. *Environ. Sci. Technol*. 50:4712–4721. doi: 10.1021/acs.est.5b06121.

Di, Q., J. Schwartz, and P. Koutrakis. 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. doi: 10.1016/j.atmosenv.2016.02.002

Didan, K. 2015. *MOD13A2 MODIS/Terra Vegetation Indices 16-Day L3 Global 1km SIN Grid V006*. NASA EOSDIS Land Processes DAAC. Sioux Falls, SD: USGS Earth Resources Observation and Science (EROS) Center. doi: 10.5067/MODIS/MOD13A2.006

Fiore, A.M., D.J. Jacob, I. Bey, R.M. Yantosca, B.D. Field, A.C. Fusco, and J.G. Wilkinson. 2002. Background ozone over the United States in summer: Origin, trend, and contribution to pollution episodes. *J. Geophys. Res. Atmos*. 107:D15. doi: 10.1029/2001JD000982

Franklin, M., and J. Schwartz. 2008. The impact of secondary particles on the association between ambient ozone and mortality. *Environ. Health Perspect*. 116:453–8. doi: 10.1289/ehp.10777

Fry, J.A., G. Xian, S. Jin, J.A. Dewitz, C.G. Homer, Y. LIMIN, C.A. Barnes, N.D. Herold, and J. D Wickham. 2011. Completion of the 2006 National Land Cover Database for the conterminous United States. *Photogramm. Eng. Rem. S*. 77:858–64.

Fusco, A.C., and J.A. Logan. 2003. Analysis of 1970–1995 trends in tropospheric ozone at Northern Hemisphere midlatitudes with the GEOS-CHEM model. *J. Geophys. Res. Atmos*. 108:D15. doi: 10.1029/2002JD002742

Gent, J.F., E.W. Triche, T.R. Holford, K. Belanger, M.B. Bracken, W.S. Beckett, and B.P. Leaderer. 2003. Association of low-level ozone and fine particles with respiratory symptoms in children with asthma. *JAMA* 290:1859–67. doi: 10.1001/jama.290.14.1859

Graedel, T.E., L.A. Farrow, and T.A. Weber. 1977. Photochemistry of the "Sunday effect". *Environ. Sci. Technol*. 11:690–4. doi: 10.1021/es60130a005

Hao, Y., L. Balluz, H. Strosnider, X. Wen, C. Li, and J.R. Qualters. 2015. Ozone, fine particulate matter, and chronic lower respiratory disease mortality in the United States. *Am. J. Respir. Crit. Care* 192:337–41. doi: 10.1164/rccm.201410-1852OC

Haykin, S. 2004. *Neural networks: A comprehensive foundation*, 2nd ed. Hamilton, Ontario, Canada: McMaster University.

Hoek, G., R. Beelen, K. de Hoogh, D. Vienneau, J. Gulliver, P. Fischer, and D. Briggs. 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos. Environ*. 42:7561–78. doi: 10.1016/j.atmosenv.2008.05.057

Hooyberghs, J., C. Mensink, G. Dumont, and F. Fierens. 2006. Spatial interpolation of ambient ozone concentrations from sparse monitoring points in Belgium. *J. Environ. Monit*. 8:1129–35. doi: 10.1039/B612607N

Jerrett, M., R.T. Burnett, C.A. Pope III, K. Ito, G. Thurston, D. Krewski, Y. Shi, E. Calle, and M. Thun. 2009. Long-term ozone exposure and mortality. *N. Engl. J. Med*. 360:1085–95. doi: 10.1056/NEJMoa0803894

Jhun, I., B.A. Coull, A. Zanobetti, and P. Koutrakis. 2014. The impact of nitrogen oxides concentration decreases on ozone trends in the USA. *Air Qual. Atmos. Health*. 8:283–92. doi: 10.1007/s11869-014-0279-2

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, R. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K.C. Mo, C. Ropelewski, J. Wang, R. Jenne, and D. Joseph. 1996. The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Am. Meteorol. Soc*. 77:437–71. doi: 10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Kloog, I., F. Nordio, B.A. Coull, and J. Schwartz. 2012. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal $PM_{2.5}$ exposures in the Mid-Atlantic states. *Environ. Sci. Technol*. 46:11913–21. doi: 10.1021/es302673e

Koren, H.S., R.B. Devlin, D.E. Graham, R. Mann, M.P. McGee, D.H. Horstman, W. J. Kozumbo, S. Becker, D.E. House, and W.F. McDonnell. 1989. Ozone-induced inflammation in the lower airways of human subjects. *Am. Rev. Respir. Dis*. 139:407–15. doi: 10.1164/ajrccm/139.2.407

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel. 2006. World map of the Köppen-Geiger climate classification updated. *Meteorol. Z*. 15:259–63. doi: 10.1127/0941-2948/2006/0130

Kroon, M., J.P. Veefkind, M. Sneep, R. McPeters, P. Bhartia, and P. Levelt. 2008. Comparing OMI-TOMS and OMI-DOAS total ozone column data. *J. Geophys. Res. Atmos*. 113: D16. doi: 10.1029/2007JD008798

Krotkov, N., P. Leonard, M. Walters, P. Leonard, and P. Lead. 2012. OMI/Aura Sulfur Dioxide (SO_2) Total Column Daily L3 Best Pixel Global 0.25deg Lat/Lon Grid, version 003. Greenbelt, MD: Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC) (accessed August 30, 2015). doi: 10.5067/Aura/OMI/DATA3008

LeCun, Y., and Y. Bengio. 1995. Convolutional networks for images, speech, and time series. In *The Handbook of Brain Theory and Neural Networks*, ed. M.A. Arbib, 255–8. Cambridge, MA: MIT Press.

Lei, W., B.d. Foy, M. Zavala, R. Volkamer, and L. Molina. 2007. Characterizing ozone production in the Mexico City Metropolitan Area: A case study using a chemical transport model. *Atmos. Chem. Phys*. 7: 1347–66. doi: 10.5194/acp-7-1347-2007

Levelt, P.F., G.H. Van den Oord, M.R. Dobber, A. Mälkki, H. Visser, J. De Vries, P. Stammes, J. O. Lundell, and H. Saari. 2006. The ozone monitoring instrument. *IEEE Trans. Geosci. Remote Sens*. 44:1093–101. doi: 10.1109/TGRS.2006.872333

Li, Y., A.K. Lau, J.C. Fung, J. Zheng, and S. Liu. 2013. Importance of NO_x control for peak ozone reduction in the Pearl River Delta region. *J. Geophys. Res. Atmos.* 118:9428–43. doi: 10.1002/jgrd.50659

Liu, X., P. Bhartia, K. Chance, L. Froidevaux, R. Spurr, and T. Kurosu. 2010a. Validation of Ozone Monitoring Instrument (OMI) ozone profiles and stratospheric ozone columns with Microwave Limb Sounder (MLS) measurements. *Atmos. Chem. Phys.* 10:2539–49. doi: 10.5194/acp-10-2539-2010

Liu, X., K. Chance, C.E. Sioris, T.P. Kurosu, R.J. Spurr, R.V. Martin, T.M. Fu, J.A. Logan, D.J. Jacob, and P.I. Palmer. 2006. First directly retrieved global distribution of tropospheric column ozone from GOME: Comparison with the GEOS-Chem model. *J. Geophys. Res. Atmos.* 111:D2. doi: 10.1029/2005JD006564

Liu, X.-H., Y. Zhang, J. Xing, Q. Zhang, K. Wang, D.G. Streets, C. Jang, W.-X. Wang, and J.-M. Hao. 2010b. Understanding of regional air pollution over China using CMAQ, part II. Process analysis and sensitivity of ozone and particulate matter to precursor emissions. *Atmos. Environ.* 44:3719–27. doi: 10.1016/j.atmosenv.2010.03.036

Liu, Y., R.J. Park, D.J. Jacob, Q. Li, V. Kilaru, and J.A. Sarnat. 2004. Mapping annual mean ground-level $PM_{2.5}$ concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States. *J. Geophys. Res.* 109:D22. doi: 10.1029/2004JD005025

Loibi, W., W. Winiwarter, A. Kopsca, J. Zufger, and R. Baumann. 1994. Estimating the spatial distribution of ozone concentrations in complex terrain. *Atmos. Environ.* 28:2557–66. doi: 10.1016/1352-2310(94)90430-8

Malmqvist, E., D. Olsson, A. Hagenbjörk-Gustafsson, B. Forsberg, K. Mattisson, E. Stroh, M. Strömgren, E. Swietlicki, L. Rylander, and G. Hoek. 2014. Assessing ozone exposure for epidemiological studies in Malmö and Umeå, Sweden. *Atmos. Environ.* 94:241–8. doi: 10.1016/j.atmosenv.2014.05.038

McConnell, R., K. Berhane, F. Gilliland, S.J. London, T. Islam, W.J. Gauderman, E. Avol, H.G. Margolis, and J.M. Peters. 2002. Asthma in exercising children exposed to ozone: A cohort study. *Lancet* 359:386–91. doi: 10.1016/S0140-6736(02)07597-9

McPeters, R., M. Kroon, G. Labow, E. Brinksma, D. Balis, I. Petropavlovskikh, J.P. Veefkind, P. Bhartia, and P. Levelt. 2008. Validation of the AURA Ozone Monitoring Instrument total column ozone product. *J. Geophys. Res. Atmos.* 113:D15. doi: 10.1029/2007JD008802

Pfister, G., L. Emmons, P. Hess, J. F. Lamarque, A. Thompson, and J. Yorks. 2008. Analysis of the Summer 2004 ozone budget over the United States using Intercontinental Transport Experiment Ozonesonde Network Study (IONS) observations and Model of Ozone and Related Tracers (MOZART-4) simulations. *J. Geophys. Res. Atmos.* 113:D23. doi: 10.1029/2008JD010190

Reidmiller, D., A.M. Fiore, D. Jaffe, D. Bergmann, C. Cuvelier, F. Dentener, B.N. Duncan, G. Folberth, M. Gauss, and S. Gong. 2009. The influence of foreign vs. North American emissions on surface ozone in the US. *Atmos. Chem. Phys.* 9:5027–42. doi: 10.5194/acp-9-5027-2009

Rodgers, C.D. 2000. *Inverse Methods for Atmospheric Sounding: Theory and Practice*. Singapore: World Scientific

Schwartz, J., D.W. Dockery, L.M. Neas, D. Wypij, J. H. Ware, J.D. Spengler, P. Koutrakis, F.E. Speizer, and B.G. Ferris Jr. 1994. Acute effects of summer air pollution on respiratory symptom reporting in children. *Am. J. Respir. Crit. Care* 150:1234–42. doi: 10.1164/ajrccm.150.5.7952546

Sokhi, R.S., R. San José, N. Kitwiroon, E. Fragkou, J.L. Pérez, and D. Middleton. 2006. Prediction of ozone levels in London using the MM5–CMAQ modelling system. *Environ. Modell. Softw.* 21:566–76. doi: 10.1016/j.envsoft.2004.07.016

Sousa, S., M. Alvim-Ferraz, and F. Martins. 2013. Health effects of ozone focusing on childhood asthma: What is now known—A review from an epidemiological point of view. *Chemosphere* 90:2051–8. doi: 10.1016/j.chemosphere.2012.10.063

Tank, J., H. Biller, K. Heusser, O. Holz, A. Diedrich, T. Framke, A. Koch, A. Grosshennig, W. Koch, and N. Krug. 2011. Effect of acute ozone induced airway inflammation on human sympathetic nerve traffic: A randomized, placebo controlled, crossover study. *PLoS ONE* 6: e18737. doi: 10.1371/journal.pone.0018737

Tong, D.Q., and D.L. Mauzerall. 2006. Spatial variability of summertime tropospheric ozone over the continental United States: Implications of an evaluation of the CMAQ model. *Atmos. Environ.* 40:3041–56. doi: 10.1016/j.atmosenv.2005.11.058.

Tranchant, B., and A. Vincent. 2000. Statistical interpolation of ozone measurements from satellite data (TOMS, SBUV and SAGE II) using the kriging method. *Ann. Geophys.* 18:666–78. doi: 10.1007/s00585-000-0666-x

Turner, M.C., M., C. Jerrett, A. Pope, D. Krewski, S.M. Gapstur, W.R. Diver, B.S. Beckerman, J.D. Marshall, J. Su, and D.L. Crouse. 2015. Long-term ozone exposure and mortality in a large prospective study. *Am. J. Respir. Crit. Care* doi: 10.1164/rccm.201508-1633OC

Van Donkelaar, A., R. V. Martin, M. Brauer, R. Kahn, R. Levy, C. Verduzco, and P.J. Villeneuve. 2010. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application. *Environ. Health Perspect.* 118:847–55. doi: 10.1289/ehp.0901623

Veefkind, J.P., J.F. De Haan, E.J. Brinksma, M. Kroon, and P.F. Levelt. 2006. Total ozone from the Ozone Monitoring Instrument (OMI) using the DOAS technique. *IEEE Trans. Geosci. Remote Sens.* 44:1239–44. doi: 10.1109/TGRS.2006.871204

Wang, L., M. Newchurch, A. Biazar, X. Liu, S. Kuang, M. Khan, and K. Chance. 2011. Evaluating AURA/OMI ozone profiles using ozonesonde data and EPA surface measurements for August 2006, *Atmos. Environ.* 45:5523–30. doi: 10.1016/j.atmosenv.2011.06.012

Wang, M., J.P. Keller, S.D. Adar, S.Y. Kim, T.V. Larson, C. Olives, P.D. Sampson, L. Sheppard, A.A. Szpiro, and S. Vedal. 2015. Development of long-term spatiotemporal models for ambient ozone in six metropolitan regions of

the United States: The MESA Air study. *Atmos. Environ*. 123:79–87. doi: 10.1016/j.atmosenv.2015.10.042

Xian, G., C. Homer, J. Dewitz, J. Fry, N. Hossain, and J. Wickham. 2011. Change of impervious surface area between 2001 and 2006 in the conterminous United States. *Photogramm. Eng. Rem. S.* 77:758–62

Zhang, Y., K. Vijayaraghavan, X.Y. Wen, H.E. Snell, and M.Z. Jacobson. 2009. Probing into regional ozone and particulate matter pollution in the United States: 1. A 1 year CMAQ simulation and evaluation using surface and satellite data. J. *Geophys. Res. Atmos*. 114:D22304. doi: 10.1029/2009JD011898