Atmospheric Environment 241 (2020) 117470

Contents lists available at ScienceDirect

Atmospheric Environment

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

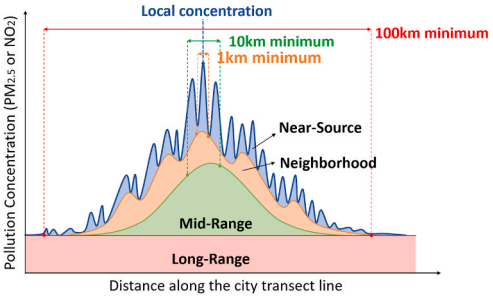
Short communication

Spatial decomposition analysis of NO_2 and $PM_{2.5}$ air pollution in the United States

Yuzhou Wang [a], Matthew J. Bechle [a], Sun-Young Kim [b,c], Peter J. Adams [d,e], Spyros N. Pandis [f], C. Arden Pope III [g], Allen L. Robinson [e,h], Lianne Sheppard [b,i], Adam A. Szpiro [i], Julian D. Marshall [a,*]

[a] Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, 98195, United States
[b] Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, WA, 98195, United States
[c] Department of Cancer Control and Population Health, Graduate School of Cancer Science and Policy, National Cancer Center, Goyang-si, Gyeonggi-do, South Korea
[d] Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA, 15213, United States
[e] Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA, 15213, United States
[f] Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA, 15213, United States
[g] Department of Economics, Brigham Young University, Provo, UT, 84602, United States
[h] Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA, 15213, United States
[i] Department of Biostatistics, University of Washington, Seattle, WA, 98195, United States

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Spatial decomposition
Spatial increment
NO_2
$PM_{2.5}$

ABSTRACT

Length scales for spatial variability of air pollution concentrations depend on the pollutant and the location. In this paper, we develop a readily scalable algorithm based on "spatial-increment", to decompose the air pollution concentration into four spatial components: long-range, mid-range, neighborhood, and near-source. We apply the algorithm to annual-average concentrations of outdoor nitrogen dioxide (NO_2) and fine particulate matter ($PM_{2.5}$) for all census blocks in the contiguous US. For NO_2, "neighborhood" and "mid-range" components dominate both within-city and between-city concentration differences (both components are ~5-fold larger in large urbanized areas than rural areas). For $PM_{2.5}$, the "long-range" component dominates; this component varies by region (e.g., is three times greater in the Midwest [7 μg/m^3] than in the West [2.3 μg/m^3]), whereas variation by urban area size is relatively minor. Our study provides the first nation-level fine-scale decomposed pollution

* Corresponding author. Department of Civil & Environmental Engineering, University of Washington, 201 More Hall, Box 352700, Seattle, WA, 98195-2700, United States.
E-mail address: jdmarsh@uw.edu (J.D. Marshall).

https://doi.org/10.1016/j.atmosenv.2020.117470
Received 9 December 2019; Received in revised form 19 March 2020; Accepted 2 April 2020
Available online 10 April 2020
1352-2310/© 2020 Published by Elsevier Ltd.

surfaces to date; this dataset is publicly available. Results can be used to estimate, at least to a zeroth order, the contribution of sources at different distances from the receptor to the annual average pollution in a location of interest.

1. Introduction

Outdoor concentrations of air pollutants vary on length scales from meters, to 100s or 1000s of km. Concentrations in the environment reflect spatial patterns in emissions, physical and chemical processes governing production and removal of pollutants, and turbulent advection and dispersion that drives transport and dilution. The degree of spatial heterogeneity varies by time, location, and pollutant.

Here, we use spatial decomposition to investigate how annual-average concentrations across the United States vary at length scales ranging from under 1 km to over 100 km. A motivation is that spatial patterns at a location may reveal information on source contributions: spatially homogenous concentrations suggest dominance of regional sources and/or secondary pollutants (e.g., sulfate $PM_{2.5}$ formed from SO_2 emissions from coal power plants), whereas heterogeneities on short length-scales suggest influence by local emission sources (e.g., NO_x emissions from traffic in an urban area).

The literature has quantified spatial components of air pollution concentrations (e.g., local, urban, long-range components) based on source-receptor modeling and concentration increments. The modeling approach (Guo et al., 2016; Sciare et al., 2010; Wu et al., 2013) often involves comparing results from a chemical transport model (CTM) under different scenarios (e.g., in the model, switching on/off local traffic, total urban emissions, or upwind power plants) (Diamantopoulou et al., 2016; Viana et al., 2008). This method is typically used for a single location (a specific city or region); application to many locations is computationally expensive (Thunis, 2018). Recently developed source-apportionment algorithms, such as Particulate Source Apportionment Technology (PSAT) and Integrated Source Apportionment Model (ISAM), running in parallel with CTMs have reduced this cost by a factor of 10–20, but they still can be applied to only a few tens of receptors (Chang et al., 2019; Chen et al., 2017; Wagstrom et al., 2008; Wagstrom and Pandis, 2011). Some integrated assessment tools (e.g., Air Benefit and Cost and Attainment Assessment System (ABaCAS), Greenhouse Gas–Air Pollution Interactions and Synergies (GAINS), the Intervention Model for Air Pollution (InMAP) Source-Receptor Matrix (ISRM), Screening for High Emission Reduction Potentials for Air Quality (SHERPA), FAst Scenario Screening Tool (TM5-FASST)) retrieve source-appointment relationships through calculated emission-concentration sensitivities of full CTMs, and can be efficiently applied to hundreds of receptors in a larger domain (Crippa et al., 2019; Goodkind et al., 2019; Kiesewetter et al., 2014; Tessum et al., 2017; Thunis et al., 2018). Those and other modeling approaches could potentially be applied to questions considered here. Limitations of the modeling approaches include (1) problems of nonlinearity (or non-additivity) for secondary pollutants (Clappier et al., 2017; Thunis et al., 2019; Zhao et al., 2017), (2) limitations of input data (e.g., uncertainty in the emission inventory), and (3) limited spatial resolution for national-scale simulations (Kiesewetter et al., 2014; Paolella et al., 2018; Thunis et al., 2016).

Concentration-based approaches use spatial concentration increments, typically from field measurements, to infer spatial or source contributions (Apte et al., 2017; Both et al., 2011; Beekmann et al., 2015; Cyrys et al., 2008; Diamantopoulou et al., 2016; Gómez-Losada et al., 2016; Lenschow et al., 2001; Querol et al., 2004; Squizzato et al., 2012; Watson and Chow, 2001). Lenschow et al. (2001) utilized such a method to estimate the local, urban, and regional background concentrations of PM10 for Berlin by subtracting the roadside, urban background, and regional background measurement. Approaches based on the analysis of temporally -resolved pollutant concentrations have also been used (Apte et al., 2017; Both et al., 2011; Diamantopoulou et al., 2016; Gómez-Losada et al., 2016; Watson and Chow, 2001). These methods are limited by the availability of the necessary measurements and have previously been applied to a single or a few locations at a time. Only a few studies have considered spatial decomposition for a broader region relying mainly on empirical modeling prediction surfaces and concentration increment approaches. Antonelli et al. (2017) used an image decomposition method, called wavelet decomposition, to decompose a 1 km × 1 km spatial surface of year 2003–2008 daily $PM_{2.5}$ predictions in the New England region of the US into three spatial components, determined through visual inspection of the decomposed surfaces. Beelen et al. (2009) estimated regional background, rural and urban concentration surfaces separately using universal kriging from rural and urban background sites, and covariates representing their respective spatial scales and sources, to produce 1 km × 1 km composite EU-wide maps of year 2001 annual-averaged NO_2, PM_{10} and O_3 concentrations.

Here, we explore a readily scalable algorithm to spatially decompose ambient air pollution concentrations using spatial increments, and apply this method to a national fine-scale dataset of outdoor NO_2 and $PM_{2.5}$ predictions. To our knowledge, this paper provides the first national high-resolution, spatially decomposed air pollution surfaces for the US. The results, which we are making publicly available, provide useful information for the contiguous US regarding the contribution of air pollution sources at spatial scales from local to regional.

2. Material and methods

$PM_{2.5}$ and NO_2 concentrations employed here are outdoor annual-average predicted concentrations at all (n ≈ 6 million) census block centroids with non-zero population in the contiguous US. The concentration predictions are derived from the Center for Air, Climate, and Energy Solutions (CACES) empirical models (Kim et al., 2020). The predictions incorporate satellite-derived estimates of ambient concentrations and land-cover, land-use and other geographic datasets, and ground-level monitoring data in a universal Kriging framework. Publicly-available predictions are for six air pollutants and multiple-years (1979–2015) (https://www.caces.us). We report decomposition analyses for year-2010 in the main paper; results for other years are in the supplementary material (see below).

We decompose annual-average concentrations into four spatial components: "long-range", "mid-range", "neighborhood", and "near-source":

$$C_i^{long-range} = \min\{C_j | d_{i,j} \le 100\ km\},$$

$$C_i^{mid-range} = \min\{C_j | d_{i,j} \le 10\ km\} - \min\{C_j | d_{i,j} \le 100\ km\},$$

$$C_i^{neighborhood} = \min\{C_j | d_{i,j} \le 1\ km\} - \min\{C_j | d_{i,j} \le 10\ km\}, \text{ and}$$

$$C_i^{near-source} = C_i - \min\{C_j | d_{i,j} \le 1\ km\}.$$

Here, C_i is the model-predicted concentration at location i; the four superscripts on C_i ("long-range", "mid-range", "neighborhood", "near-source") are the four spatial components considered here; $d_{i,j}$ is the distance between two locations i and j; and $\min\{C_j | d_{i,j} \le "x"\ km\}$ represents the minimum concentrations from all block centroids within an "x" km circular buffer of location i (where "x" takes the values above: 1, 10, and 100 km).

The 1, 10 and 100 km length scales are commonly used for

neighborhood, city, and regional scales (Kumar et al., 2014; Lin et al., 2014). This approach is straightforward, intuitive, and computationally demanding, but feasible to implement for a national dataset, across multiple pollutants and years. While these four components are inherently defined by the spatial increments of predicted concentrations, they may provide insight into source contribution. "Long-range" (>100 km) likely represents regional background and long-range transport. "Mid-range" (10–100 km) likely represents, e.g., urbanized areas, and agricultural regions. "Neighborhood" (1–10 km) likely represents localized sources such as commercial districts, industrial areas, and intersections of major highways. "Near-source" (<1 km) likely represents hyper-local enhancements (e.g., roadways).

We apply the spatial decompositions to each of the approximately 6 million census block centroids in the contiguous US and analyze results nationally, by state, region, and urbanicity. Urbanicity levels are defined in the 2010 Census (Manson et al., 2019) based on population: urbanized areas have 50,000 or more people (we subdivided them into small/medium/large urbanized areas, by population tertiles); urban clusters have 2500–50,000 people; rural areas are all remaining census blocks. Rural areas contain 59 million people; urban clusters (n = 3087) contain 29 million people; small urban areas (n = 440) contain 75 million people; medium urban areas (n = 47) contain 75 million people; large urban areas (n = 10) contain 73 million people.

To further explore urban-scale patterns, we selected six cities to investigate more closely: New York, NY (year-2010 urban area population: 18.4 million); Los Angeles, CA (12.1 million); Seattle, WA (3.1 million); Minneapolis, MN (3.1 million); Spokane, WA (0.5 million); Tuscaloosa, AL (0.1 million). These cities were selected to represent a range of sizes, pollution sources, and geographies (e.g., region of the US, climate, distance-to-coast, and regional economy) across the US. For each city, we conduct the following analyses to understand urban-scale variability. Following Novotny et al. (2011), we consider concentrations along transect lines across the urban center. Here, we apply the decompositions on the points at each 10-m interval along the transect lines, using empirical-model predictions from the nearest block centroid as an approximation of the total concentrations at each 10-m interval. As a sensitivity analysis to explore the robustness of the approach, we compare results for multiple years and for several transects per city (12 transects passing through the city center, at 15° intervals). Also, we compare results with different buffer radii to provide insight on our choice of buffer lengths.

To understand the relationship between the model structure of the prediction surface and the spatial decomposition components, we calculate how much each empirical regression component (i.e., Kriging and each independent variable) contributes to the year-2010 spatial decomposition results for both pollutants. For each block centroid prediction and for both pollutants, we separate the without-Kriging concentration predictions (i.e., the ultimate prediction results minus the Kriging adjustment values), and then apply the same spatial decomposition algorithm to the without-Kriging predictions. The contributions of the Kriging to the four spatial components are calculated by the population-weighted averages of the concentration differences between initial and without-Kriging decompositions. For other regression components, we run the no-intercept multiple regressions of without-Kriging decomposed concentrations on the model-selected independent variables for total concentration estimation. The contributions of each independent variable are calculated as the population-weighted-average product of the variable values and coefficients; the contributions are then aggregated to nine categories according to the variable type. Following Kim et al. (2020), the categories are traffic, urban land use, rural land use, population, elevation, emission, imperviousness, vegetation, and satellite.

3. Results

3.1. Within-urban decomposition results

Figs. 1 and 2 illustrate, for a case-study city (Seattle), the results of the spatial decomposition approach developed here. Across a ~100-km urban transect (Fig. 1), the "long-range" values are almost constant. That result is expected; those values likely represent the regional background. For Seattle, "long-range" values account for 21% (for NO_2) and 38% ($PM_{2.5}$) of the average predictions. The "mid-range" increments are low at the urban edge and increase gradually approaching the urban center at approximately 50 km along the line, presumably reflecting the effects of aggregated urban emissions in the urban center. The "neighborhood" and "near-source" values vary at much shorter spatial scales, with multiple peaks along the transect, presumably reflecting localized emission sources. Spatial variability is greater for NO_2 than $PM_{2.5}$, and the "near-source" component is larger for NO_2 (15%) than $PM_{2.5}$ (7%).

Fig. 2 shows the block centroid decomposed concentrations for

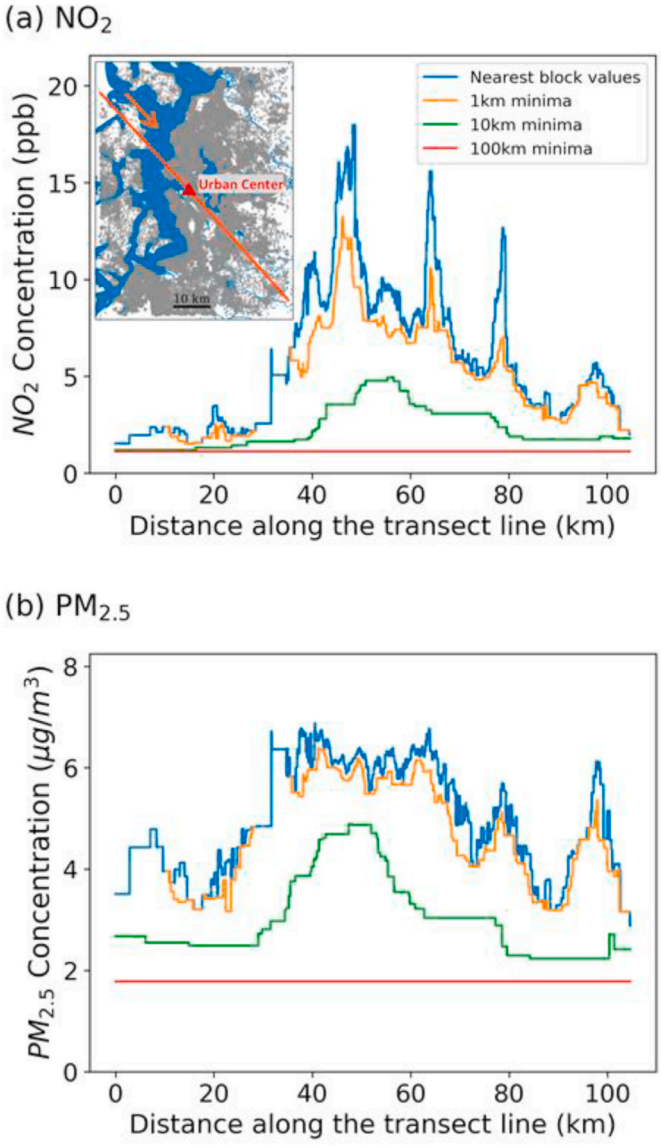


Fig. 1. Decomposition results for (a) NO_2 and (b) $PM_{2.5}$ along a transect line from Northwest to Southeast across the center of Seattle. The inset map in the upper left of (a) shows the Seattle census block centroids (shown in grey) and transect line.

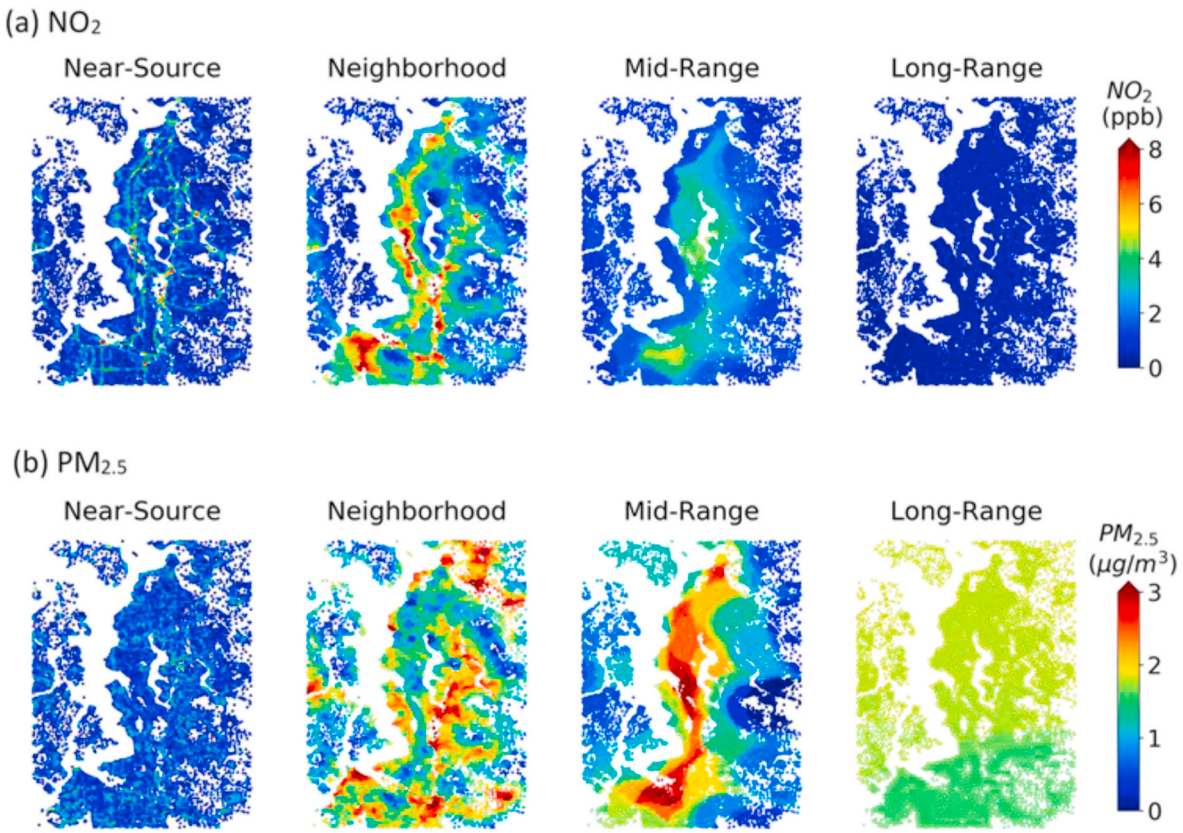


Fig. 2. The block level decomposed concentration maps of (a) NO_2 and (b) $PM_{2.5}$ for Seattle.

Seattle (see Fig. S1 for 1, 10, and 100 km buffer minimum concentrations). "Near-source" concentrations of NO_2 reflect the major road network (see Fig. S2). For $PM_{2.5}$, "near-source" concentrations are relatively low (<1 μg/m^3), with no clear spatial patterns. The "neighborhood" component has several hot spots for both pollutants (e.g., downtowns, intersections of major highways).

The transect-line decompositions for different transect directions (Fig. S3-S9), different years (Fig. S10-S11), and block-level decomposition maps for the additional five cities (Fig. S12-S16) are in the SI. For each case-study city, the decomposition results are generally stable over different years and transects, and spatial patterns for each of the components are in general consistent with expectations. For example, Fig. S11 captures the elevated concentrations and contributions of "long-range" $PM_{2.5}$ component in Spokane (and to a lesser degree Seattle) in 2015 owing to the large wildfire season in northeastern Washington State (Engel et al., 2019). The decompositions for different cities reveal broadly similar spatial patterns of variability, though the partitions of spatial components differ. For example, for both pollutants, the "mid-range" component (likely representing elevated regional urban concentrations) is typically less pronounced for the smaller cities (Spokane, Tuscaloosa), whereas for bigger cities (New York City, Los Angeles) this component often has a larger contribution and remains elevated throughout the urban portion of the transect, only decreasing when the transects extend into unpopulated areas such as mountainous areas and open water. Relative to the other cities, the "long-range" contributions are comparatively smaller for $PM_{2.5}$ in Seattle and for NO_2 in Los Angeles.

3.2. Decomposition results by state, region and urbanicity

After applying the spatial decomposition to all block centroids in the US (Fig. S17), we calculate the corresponding population weighted averages by state (Fig. 3). For NO_2, high levels of the "neighborhood" and "mid-range" components are found in states with large cities (e.g., Los Angeles, New York, Chicago). For $PM_{2.5}$, the overall concentration and the "long-range" component are higher in the East than in the West.

Fig. 4 summarizes the total and decomposed concentrations by urban area size and geographic region. For NO_2, concentrations are ~3-fold higher in large urban areas (average: 13 ppb) than rural areas (4 ppb). "Neighborhood" and "mid-range" components (both about 4 ppb) dominate that concentration difference; "long-range" components are nearly the same in absolute terms (<1 ppb difference) across urban/rural groups. In relative terms, the "long-range" component is 21% overall (16% [large urban areas], 41% [rural]; Fig. S18). The "near-source" component is 13% overall (12% [large urban areas], 7% [rural]). The differences in decomposed NO_2 concentrations among regions are mostly driven by regional differences amongst the ten largest urbanized areas (see Fig. S19). The decomposition patterns suggest that NO_2 is of urban origin and varies locally, and that climate or geographic differences have less effects on the NO_2 concentrations. The mid-range and neighborhood components are greater for the large urbanized areas than for other locations (see Fig. 4), consistent with larger cities having comparatively denser traffic and more urban emissions.

In contrast, geographic regions dominate the differences of decomposed $PM_{2.5}$ concentrations, mainly in the "long-range" component. That results likely reflects that $PM_{2.5}$ is a regional pollutant with a strong secondary component. The "long-range" concentration in the Midwest (7 μg/m^3) is almost three times than in the West (2.3 μg/m^3); it contributes 30% of the total $PM_{2.5}$ in the West, versus >50% in the remaining three regions. We hypothesize that these regional differences mainly reflect differences in emissions and emissions density (e.g., traffic, coal-fired power plants, agriculture), but other aspects (e.g., topography, climate, meteorology) may also play a role. The $PM_{2.5}$ concentration differences by urban area size are relatively minor compared with regional differences. $PM_{2.5}$ concentrations are higher for large urban areas than for small urban areas, and urban concentrations

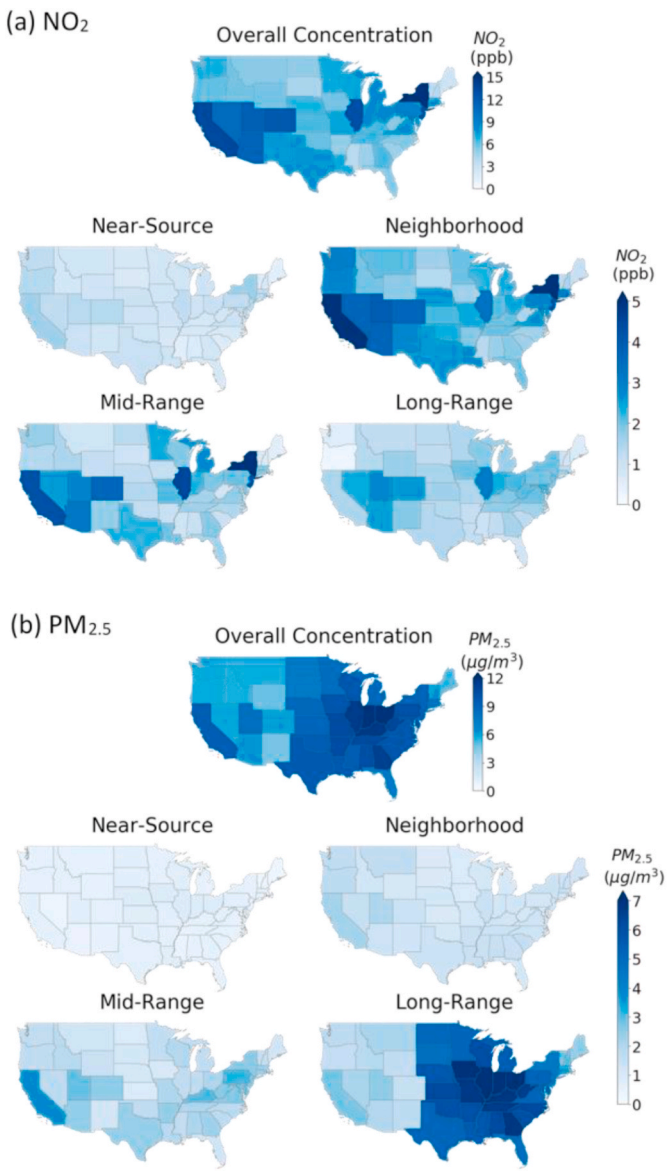


Fig. 3. Population-weighted averages of overall and decomposed concentrations for (a) NO_2 and (b) $PM_{2.5}$ at ~6 million census blocks by state.

are larger than rural concentrations, but the differences are relatively modest (<2 μg/m^3); those differences are mainly driven by the increase of "mid-range" concentrations. "Long-range" components dominate (≥50%) $PM_{2.5}$ concentrations in both rural and urban areas.

3.3. Contributions of empirical regression terms

Fig. 5 summarizes contributions of each empirical regression term (traffic, land use, satellite data, Kriging, etc.) for the overall pollution concentration to the four decomposed concentrations. A version of Fig. 5, but without the satellite contributions, is in the supplementary material (Fig. S20). The goal of this analysis is to find out how the underlying model structure contributes to the spatial decomposition patterns, not to make causal inference about mechanisms. The results are aggregated to all US blocks, urban blocks, and rural blocks, separately. For both pollutants, satellite concentration estimates dominate the "long-range" component. For NO_2 in urban blocks, two variables that strongly contribute, in addition to satellite estimates, are vegetation (a "negative" contribution: more vegetation corresponds to lower concentrations of NO_2) for the "mid-range" component and imperviousness for the "neighborhood" and "near-source" components. For $PM_{2.5}$, satellite estimates dominate; in addition, impervious surfaces also contribute mainly for the "neighborhood" and "mid-range" components. Land use types (both rural and urban land uses) play a most important role for "near-source" $PM_{2.5}$. In comparison, for rural blocks, the contributions of imperviousness are negligible for both pollutants.

4. Discussion and conclusion

Our results reveal within-city and national patterns of the decomposed concentrations. NO_2 is an air pollutant of predominantly urban origin: differences in mean concentrations are mainly dominated by "neighborhood" and "mid-range" components, and concentrations vary by urbanicity level. $PM_{2.5}$ is dominated by "long-range" transport; concentrations vary at a state and regional level. Those findings are consistent with previous research and understandings (Eeftens et al., 2015; Hewitt, 1991; Wang et al., 2015). Our results provide new and useful quantification of the spatial patterns of these two pollutants, on a consistent basis with high spatial precision and throughout the US.

One limitation of our results is the use of model-estimated concentrations; those models, while exhibiting good predictive performance (R^2: 0.84 [NO_2], 0.85 [$PM_{2.5}$]; root-mean-square error: 2.2 ppb [NO_2], 1.2 μg/m^3 [$PM_{2.5}$]), are inherently less variable than the actual concentrations. The models are trained on regulatory monitoring data typically designed to capture ambient concentrations. In addition, the empirical models represent national patterns and average relationships between land use and pollution concentrations; they are unable to predict hot-spots caused by atypical conditions. In some locations, near source contributions will be underestimated. Besides, concentrations from empirical models do not provide quantification of source contribution or transport trajectories, thus we cannot make firm causal statements about mechanisms.

Another aspect of our decompositions is that they are performed on the census block centroid locations. An advantage of the census blocks over equal-distanced grids is that blocks are dense in populated areas where people live and pollution gradients are typically largest. Each block centroid has, on average, 50 additional block centroids within its 1-km circular buffer. However, 578,971 (9.3%) and 938 (0.015%) blocks centroids have no neighbors within 1 and 10 km buffers, respectively. Those blocks represent 5.1% [no 1-km neighboring block] and 0.007% [no 10-km neighboring block] of the contiguous US population. This outcome generally occurs in sparsely populated regions (Fig. S21). Those regions would likely have few local sources but the true average local contribution would be nonzero. However, in our approach, the local component is, by definition, zero for locations with the minimum concentration within 1 km (which includes all locations with no 1-km neighboring Blocks). Since the true value is nonzero but the estimated value is zero, we are underestimating (Fig. S22). Because the local contributions are expected to be small, the amount of underestimation is likely small in absolute terms; however, other techniques would be needed to quantify the "local" component in these locations.

An additional aspect of our study that is important for interpreting the results is the use of fixed radii (1, 10, and 100 km) to calculate the four spatial components. Fixed radii reflect an approach that is easy to understand and straightforward to apply nationally. However, because of differences in size and shape of built-up areas, the impact of the radii may vary across the US. For example, the "mid-range" component employs a 10 km buffer, yet city-size varies. In the 2010 US Census (Manson et al., 2019), the mean area of an urban area is 462 km^2 (equivalent radius for a circular layout: 12 km), with an interquartile range of 112 km^2 (equivalent radius: 6 km) to 413 km^2 (11 km) and an overall range of 25.7 km^2 (3 km) to 8936 km^2 (53 km). Thus, the 10 km buffer can be smaller or larger than an urban area. The values used here (1, 10, 100 km) were selected because they represent orders of magnitude and because in preliminary investigations they seemed to capture the components better than alternative values (Fig. S23-S24). Moreover, the use

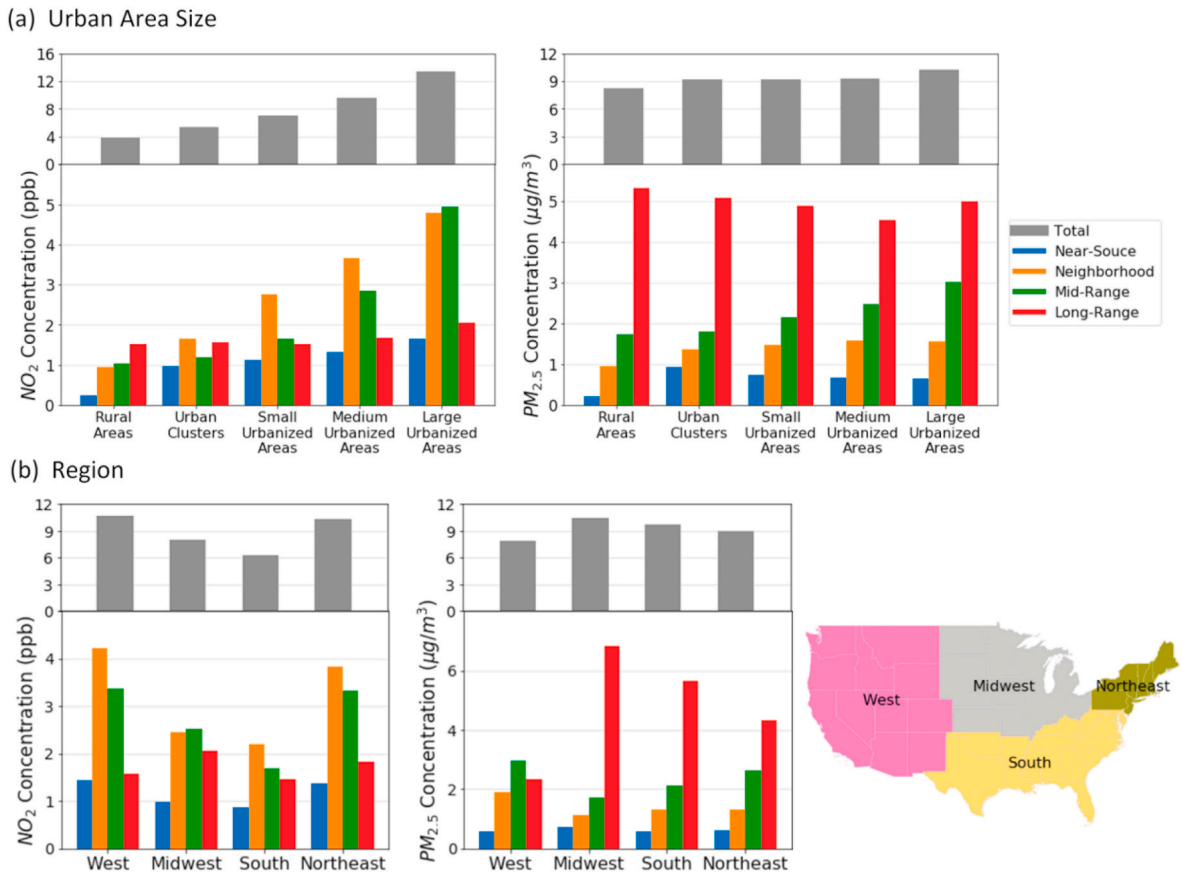


Fig. 4. Population-weighted averages of overall (grey bars) and decomposed concentrations (colored bars) for NO_2 and $PM_{2.5}$ at ~6 million census blocks by (a) urban area size; (b) US region.

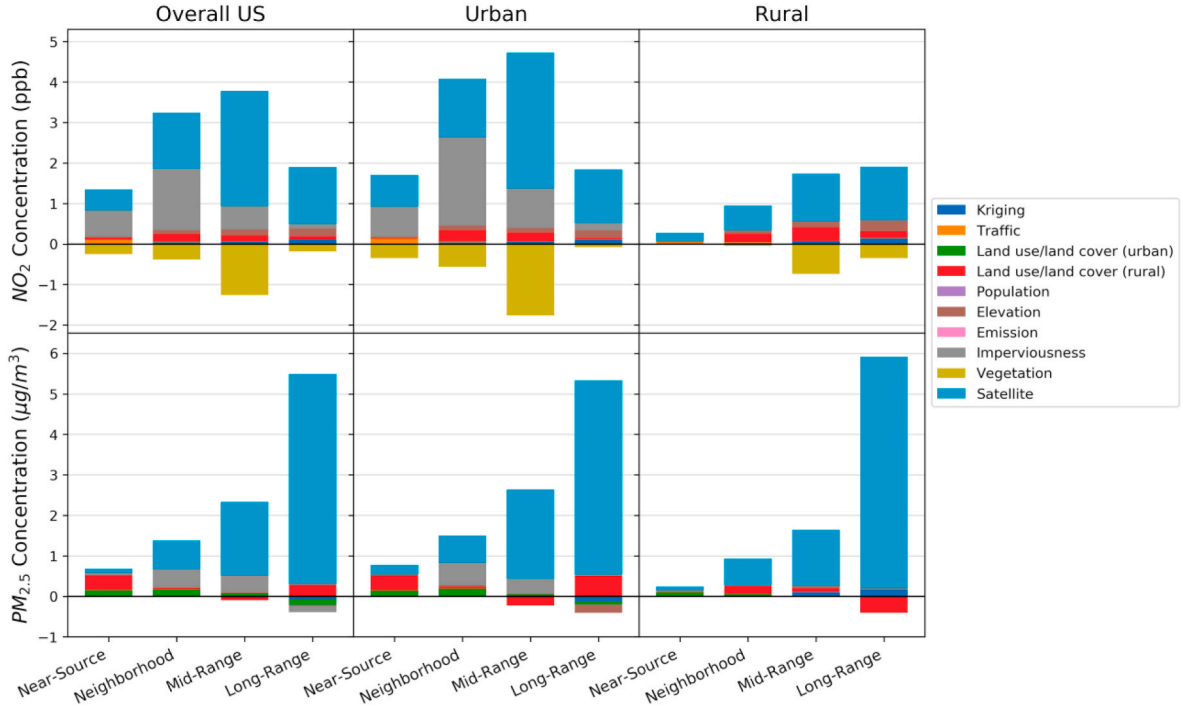


Fig. 5. Contributions of each empirical regression term for the overall pollution concentration to the four decomposed concentrations for NO_2 and $PM_{2.5}$ for the overall US, all urban blocks and all rural blocks.

of a circular buffer assumes equal contribution from all directions and ignores meteorology and topography, which could lead to mischaracterization in some locations. For example, the national block-level "long-range" NO_2 and $PM_{2.5}$ maps (Fig. S17) exhibit blotchy circular features, which are typically driven by several "clean" points that have much lower concentrations than surroundings, or by sharp gradients in these areas and few points (such as mountainous regions). The blotchy features diminish with aggregation (e.g., to state level [Fig. 3]).

Strengths of our study include national application of a straightforward approach for spatial decomposition; shedding new light by applying our method at fine resolution but across a broad domain; and investigating results for the whole US and separately by city, state, region, and level of urbanicity. The spatial increment and moving window approach are easy to understand and computationally feasible for the resolution and domain considered here. Weaknesses include that the empirical model is imperfect (e.g., nonzero error and bias; outlier concentrations may be underestimated), the empirical model is predicted only at centroids, and that the approach does not shed light on mechanisms.

To facilitate future research using our decomposition results, data are freely available online (http://dx.doi.org/10.17632/wdx2ntjznj.1).

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.

CRediT authorship contribution statement

Yuzhou Wang: Data curation, Methodology, Investigation, Formal analysis, Validation, Visualization, Writing - original draft. **Matthew J. Bechle:** Data curation, Methodology, Investigation, Validation, Writing - review & editing. **Sun-Young Kim:** Data curation, Validation, Writing - review & editing. **Peter J. Adams:** Conceptualization, Validation, Writing - review & editing. **Spyros N. Pandis:** Conceptualization, Validation, Writing - review & editing. **C. Arden Pope:** Conceptualization, Validation, Writing - review & editing. **Allen L. Robinson:** Conceptualization, Validation, Writing - review & editing. **Lianne Sheppard:** Conceptualization, Validation, Writing - review & editing. **Adam A. Szpiro:** Conceptualization, Validation, Writing - review & editing. **Julian D. Marshall:** Conceptualization, Methodology, Investigation, Formal analysis, Validation, Writing - review & editing, Supervision.

Acknowledgements

This publication was developed as part of the Center for Air, Climate, and Energy Solutions (CACES), which was supported under Assistance Agreement No. R835873 awarded by the U.S. Environmental Protection Agency. It has not been formally reviewed by EPA. The views expressed in this document are solely those of authors and do not necessarily reflect those of the Agency. EPA does not endorse any products or commercial services mentioned in this publication.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.atmosenv.2020.117470.

References

Apte, J.S., Messier, K.P., Gani, S., Brauer, M., Kirchstetter, T.W., Lunden, M.M., Marshall, J.D., Portier, C.J., Vermeulen, R.C.H., Hamburg, S.P., 2017. High-resolution air pollution mapping with google street view cars: exploiting big data. Environ. Sci. Technol. 51, 6999–7008. https://doi.org/10.1021/acs.est.7b00891.

Antonelli, J., Schwartz, J., Kloog, I., Coull, B.A., 2017. Spatial multiresolution analysis of the effect of PM2.5 on birth weights. Ann. Appl. Stat. 11, 792–807. https://doi.org/10.1214/16-AOAS1018.

Beekmann, M., Prévôt, A.S.H., Drewnick, F., Sciare, J., Pandis, S.N., Gon, D., van der, C. H.A., Crippa, M., Freutel, F., Poulain, L., Ghersi, V., Rodriguez, E., Beirle, S., Zotter, P., Weiden-Reinmüller, S.-L. von der, Bressi, M., Fountoukis, C., Petetin, H., Szidat, S., Schneider, J., Rosso, A., El Haddad, I., Megaritis, A., Zhang, Q.J., Michoud, V., Slowik, J.G., Moukhtar, S., Kolmonen, P., Stohl, A., Eckhardt, S., Borbon, A., Gros, V., Marchand, N., Jaffrezo, J.L., Schwarzenboeck, A., Colomb, A., Wiedensohler, A., Borrmann, S., Lawrence, M., Baklanov, A., Baltensperger, U., 2015. In situ, satellite measurement and model evidence on the dominant regional contribution to fine particulate matter levels in the Paris megacity. Atmos. Chem. Phys. 15, 9577–9591. https://doi.org/10.5194/acp-15-9577-2015.

Beelen, R., Hoek, G., Pebesma, E., Vienneau, D., de Hoogh, K., Briggs, D.J., 2009. Mapping of background air pollution at a fine spatial scale across the European Union. Sci. Total Environ. 407, 1852–1867. https://doi.org/10.1016/j.scitotenv.2008.11.048.

Both, A.F., Balakrishnan, A., Joseph, B., Marshall, J.D., 2011. Spatiotemporal aspects of real-time PM2.5: low- and middle-income neighborhoods in Bangalore, India. Environ. Sci. Technol. 45, 5629–5636. https://doi.org/10.1021/es104331w.

Chang, X., Wang, S., Zhao, B., Xing, J., Liu, X., Wei, L., Song, Y., Wu, W., Cai, S., Zheng, H., Ding, D., Zheng, M., 2019. Contributions of inter-city and regional transport to PM2.5 concentrations in the Beijing-Tianjin-Hebei region and its implications on regional joint air pollution control. Sci. Total Environ. 660, 1191–1200. https://doi.org/10.1016/j.scitotenv.2018.12.474.

Chen, D., Liu, X., Lang, J., Zhou, Y., Wei, L., Wang, X., Guo, X., 2017. Estimating the contribution of regional transport to PM2. 5 air pollution in a rural area on the North China Plain. Sci. Total Environ. 583, 280–291. https://doi.org/10.1016/j.scitotenv.2017.01.066.

Clappier, A., Belis, C.A., Pernigotti, D., Thunis, P., 2017. Source apportionment and sensitivity analysis: two methodologies with two different purposes, Geosci. Model Dev. 10, 4245–4256. https://doi.org/10.5194/gmd-10-4245-2017.

Crippa, M., Janssens-Maenhout, G., Guizzardi, D., Dingenen, R.V., Dentener, F., 2019. Contribution and uncertainty of sectorial and regional emissions to regional and global PM 2.5 health impacts. Atmos. Chem. Phys. 19 (7), 5165–5186. https://doi.org/10.5194/acp-19-5165-2019.

Cyrys, J., Pitz, M., Heinrich, J., Wichmann, H.-E., Peters, A., 2008. Spatial and temporal variation of particle number concentration in Augsburg, Germany. Sci. Total Environ. 401, 168–175. https://doi.org/10.1016/j.scitotenv.2008.03.043.

Diamantopoulou, M., Skyllakou, K., Pandis, S.N., 2016. Estimation of the local and long-range contributions to particulate matter levels using continuous measurements in a single urban background site. Atmos. Environ. 134, 1–9. https://doi.org/10.1016/j.atmosenv.2016.03.015.

Eeftens, M., Phuleria, H.C., Meier, R., Aguilera, I., Corradi, E., Davey, M., Ducret-Stich, R., Fierz, M., Gehrig, R., Ineichen, A., Keidel, D., Probst-Hensch, N., Ragettli, M.S., Schindler, C., Künzli, N., Tsai, M.-Y., 2015. Spatial and temporal variability of ultrafine particles, NO2, PM2.5, PM2.5 absorbance, PM10 and PMcoarse in Swiss study areas. Atmos. Environ. 111, 60–70. https://doi.org/10.1016/j.atmosenv.2015.03.031.

Engel, R.A., Marlier, M.E., Lettenmaier, D.P., 2019. On the causes of the summer 2015 Eastern Washington wildfires. Environ. Res. Commun. 1, 011009 https://doi.org/10.1088/2515-7620/ab082e.

Goodkind, A.L., Tessum, C.W., Coggins, J.S., Hill, J.D., Marshall, J.D., 2019. Fine-scale damage estimates of particulate matter air pollution reveal opportunities for location-specific mitigation of emissions. Proc. Natl. Acad. Sci. Unit. States Am. 116, 8775–8780. https://doi.org/10.1073/pnas.1816102116.

Gómez-Losada, Á., Pires, J.C.M., Pino-Mejías, R., 2016. Characterization of background air pollution exposure in urban environments using a metric based on Hidden Markov Models. Atmos. Environ. 127, 255–261. https://doi.org/10.1016/j.atmosenv.2015.12.046.

Guo, J., He, J., Liu, Hongli, Miao, Y., Liu, Huan, Zhai, P., 2016. Impact of various emission control schemes on air quality using WRF-Chem during APEC China 2014. Atmos. Environ. 140, 311–319. https://doi.org/10.1016/j.atmosenv.2016.05.046.

Hewitt, C.N., 1991. Spatial variations in nitrogen dioxide concentrations in an urban area. Atmos. Environ. Part B - Urban Atmos. 25, 429–434. https://doi.org/10.1016/0957-1272(91)90014-6.

Kim, S.-Y., Bechle, M., Hankey, S., Sheppard, L., Szpiro, A.A., Marshall, J.D., 2020. Concentrations of criteria pollutants in the contiguous U.S., 1979 – 2015: role of prediction model parsimony in integrated empirical geographic regression. PLoS One 15, e0228535. https://doi.org/10.1371/journal.pone.0228535.

Kiesewetter, G., Borken-Kleefeld, J., Schöpp, W., Heyes, C., Thunis, P., Bessagnet, B., Terrenoire, E., Gsella, A., Amann, M., 2014. Modelling NO2 concentrations at the street level in the GAINS integrated assessment model: projections under current legislation. Atmos. Chem. Phys. 14, 813–829. https://doi.org/10.5194/acp-14-813-2014.

Kumar, P., Morawska, L., Birmili, W., Paasonen, P., Hu, M., Kulmala, M., Harrison, R.M., Norford, L., Britter, R., 2014. Ultrafine particles in cities. Environ. Int. 66, 1–10. https://doi.org/10.1016/j.envint.2014.01.013.

Lenschow, P., Abraham, H.-J., Kutzner, K., Lutz, M., Preuß, J.-D., Reichenbächer, W., 2001. Some ideas about the sources of PM10. Atmos. Environ. 35, S23–S33. https://doi.org/10.1016/S1352-2310(01)00122-4. Selected Papers Presented at the Venice Conference.

Lin, M., Hang, J., Li, Y., Luo, Z., Sandberg, M., 2014. Quantitative ventilation assessments of idealized urban canopy layers with various urban layouts and the same building packing density. Build. Environ. 79, 152–167. https://doi.org/10.1016/j.buildenv.2014.05.008.

Manson, S., Schroeder, J., Van Riper, D., Ruggles, S., 2019. IPUMS National Historical Geographic Information System: Version 14.0 [database]. IPUMS, Minneapolis, MN.

Novotny, E.V., Bechle, M.J., Millet, D.B., Marshall, J.D., 2011. National satellite-based land-use regression: NO2 in the United States. Environ. Sci. Technol. 45, 4407–4414. https://doi.org/10.1021/es103578x.

Paolella, D.A., Tessum, C.W., Adams, P.J., Apte, J.S., Chambliss, S., Hill, J., Muller, N.Z., Marshall, J.D., 2018. Effect of model spatial resolution on estimates of fine particulate matter exposure and exposure disparities in the United States. Environ. Sci. Technol. Lett. 5, 436–441. https://doi.org/10.1021/acs.estlett.8b00279.

Querol, X., Alastuey, A., Ruiz, C.R., Artiñano, B., Hansson, H.C., Harrison, R.M., Buringh, E., ten Brink, H.M., Lutz, M., Bruckmann, P., Straehl, P., Schneider, J., 2004. Speciation and origin of PM10 and PM2.5 in selected European cities. Atmos. Environ. 38, 6547–6555. https://doi.org/10.1016/j.atmosenv.2004.08.037.

Sciare, J., d'Argouges, O., Zhang, Q.J., Sarda-Estève, R., Gaimoz, C., Gros, V., Beekmann, M., Sanchez, O., 2010. Comparison between simulated and observed chemical composition of fine aerosols in Paris (France) during springtime: contribution of regional versus continental emissions. Atmos. Chem. Phys. 10, 11987–12004. https://doi.org/10.5194/acp-10-11987-2010.

Squizzato, S., Masiol, M., Innocente, E., Pecorari, E., Rampazzo, G., Pavoni, B., 2012. A procedure to assess local and long-range transport contributions to PM2.5 and secondary inorganic aerosol. J. Aerosol Sci. 46, 64–76. https://doi.org/10.1016/j.jaerosci.2011.12.001.

Tessum, C.W., Hill, J.D., Marshall, J.D., 2017. InMAP: a model for air pollution interventions. PLoS One 12, e0176131. https://doi.org/10.1371/journal.pone.0176131.

Thunis, P., 2018. On the validity of the incremental approach to estimate the impact of cities on air quality. Atmos. Environ. 173, 210–222. https://doi.org/10.1016/j.atmosenv.2017.11.012.

Thunis, P., Degraeuwe, B., Pisoni, E., Trombetti, M., Peduzzi, E., Belis, C.A., Wilson, J., Clappier, A., Vignati, E., 2018. PM2.5 source allocation in European cities: a SHERPA modelling study. Atmos. Environ. 187, 93–106. https://doi.org/10.1016/j.atmosenv.2018.05.062.

Thunis, P., Clappier, A., Tarrason, L., Cuvelier, C., Monteiro, A., Pisoni, E., Wesseling, J., Belis, C.A., Pirovano, G., Janssen, S., Guerreiro, C., Peduzzi, E., 2019. Source apportionment to support air quality planning: strengths and weaknesses of existing approaches. Environ. Int. 130, 104825. https://doi.org/10.1016/j.envint.2019.05.019.

Thunis, P., Miranda, A., Baldasano, J.M., Blond, N., Douros, J., Graff, A., Janssen, S., Juda-Rezler, K., Karvosenoja, N., Maffeis, G., Martilli, A., 2016. Overview of current regional and local scale air quality modelling practices: assessment and planning tools in the EU. Environ. Sci. Pol. 65, 13–21. https://doi.org/10.1016/j.envsci.2016.03.013.

Viana, M., Kuhlbusch, T.A.J., Querol, X., Alastuey, A., Harrison, R.M., Hopke, P.K., Winiwarter, W., Vallius, M., Szidat, S., Prévôt, A.S.H., Hueglin, C., Bloemen, H., Wåhlin, P., Vecchi, R., Miranda, A.I., Kasper-Giebl, A., Maenhaut, W., Hitzenberger, R., 2008. Source apportionment of particulate matter in Europe: a review of methods and results. J. Aerosol Sci. 39, 827–849. https://doi.org/10.1016/j.jaerosci.2008.05.007.

Wagstrom, K.M., Pandis, S.N., 2011. Contribution of long range transport to local fine particulate matter concerns. Atmos. Environ. 45, 2730–2735. https://doi.org/10.1016/j.atmosenv.2011.02.040.

Wagstrom, K.M., Pandis, S.N., Yarwood, G., Wilson, G.M., Morris, R.E., 2008. Development and application of a computationally efficient particulate matter apportionment algorithm in a three-dimensional chemical transport model. Atmos. Environ. 42, 5650–5659. https://doi.org/10.1016/j.atmosenv.2008.03.012.

Wang, L., Liu, Z., Sun, Y., Ji, D., Wang, Y., 2015. Long-range transport and regional sources of PM2.5 in Beijing based on long-term observations from 2005 to 2010. Atmos. Res. 157, 37–48. https://doi.org/10.1016/j.atmosres.2014.12.003.

Watson, J.G., Chow, J.C., 2001. Estimating middle-, neighborhood-, and urban-scale contributions to elemental carbon in Mexico City with a rapid response aethalometer. J. Air Waste Manag. Assoc. 51, 1522–1528. https://doi.org/10.1080/10473289.2001.10464379.

Wu, D., Fung, J.C.H., Yao, T., Lau, A.K.H., 2013. A study of control policy in the Pearl River Delta region by using the particulate matter source apportionment method. Atmos. Environ. 76, 147–161. https://doi.org/10.1016/j.atmosenv.2012.11.069. Improving Regional Air Quality over the Pearl River Delta and Hong Kong: Sci. Pol.

Zhao, B., Wu, W., Wang, S., Xing, J., Chang, X., Liou, K.-N., Jiang, J.H., Gu, Y., Jang, C., Fu, J.S., Zhu, Y., Wang, J., Lin, Y., Hao, J., 2017. A modeling study of the nonlinear response of fine particles to air pollutant emissions in the Beijing–Tianjin–Hebei region. Atmos. Chem. Phys. 17, 12031–12050. https://doi.org/10.5194/acp-17-12031-2017.