

Sources of ambient PM_{2.5} exposure in 96 global cities

Mei W. Tessum^a, Susan C. Anenberg^b, Zoe Chafe^c, Daven K. Henze^d, Gary Kleiman^e, Iyad Kheirbek^c, Julian D. Marshall^f, Christopher W. Tessum^{g,*}

^aDepartment of Agricultural and Biological Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States

^bDepartment of Environmental and Occupational Health, George Washington University, Washington, DC, United States

^cC40 Cities Climate Leadership Group Inc., New York, New York, United States

^dDepartment of Mechanical Engineering, University of Colorado, Boulder, Colorado, United States

^eOrbis Air, LLC. Concord, Massachusetts, United States

^fDepartment of Civil and Environmental Engineering, University of Washington, Seattle, Washington, United States

^gDepartment of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States

* Email for correspondence: ctessum@illinois.edu

Abstract

To improve air quality, knowledge of the sources and locations of air pollutant emissions is critical. However, for many global cities, no previous estimates exist of how much exposure to fine particulate matter (PM_{2.5}), the largest environmental cause of mortality, is caused by emissions within the city vs. outside its boundaries. We use the Intervention Model for Air Pollution (InMAP) global-through-urban reduced complexity air quality model with a high-resolution, global inventory of pollutant emissions to quantify the contribution of emissions by source type and location for 96 global cities. Among these cities, we find that the fraction of PM_{2.5} exposure caused by within-city emissions varies widely ($\mu=51\%$; $\sigma=23\%$) and is not well-explained by surrounding population density. The list of most-important sources also varies by city. Compared to a more mechanistically detailed model, InMAP predicts urban measured concentrations with less bias but more error. Predictive accuracy in urban areas is not particularly high with either model, suggesting an opportunity for improving global urban air emission inventories. We expect the results herein can be useful as a screening tool for policy options and in many cases may be robust enough to inform policy action to improve public health.

Keywords: Environmental policy, fine particulate matter, air quality, pollution, metropolitan, air quality modeling, chemical transport modeling

1. Introduction

Air pollution is the greatest single environmental health risk worldwide. According to the World Health Organization, ambient air pollution prematurely kills 7 million people per year (WHO, 2021), with an estimated economic cost of ~ \$3 trillion USD, or 3.3% of global GDP (Mylllyvirta, 2020). Among air pollutants, fine particulate matter (PM_{2.5}) has the largest health impact in monetized terms: more than half of the global population is exposed to annual-average ambient concentrations exceeding the first interim target from the World Health Organization, 35 µg/m³ (HEI, 2019). Efforts to reduce PM_{2.5} concentrations have not been uniformly successful (UNEP, 2021).

Making effective plans for improving air pollution requires prioritization, which in turn requires knowing the emission sources that contribute most to poor ambient air quality and to the resulting health effects. However, the complexity of the atmospheric system and of the human and natural systems that cause emissions make that task challenging; in many cases it can be difficult for scientists and policymakers to determine which sources to target to maximally reduce the population exposure.

Air quality models and other decision support tools that relate pollution emissions to the resulting ambient concentrations and health impacts can be important for designing effective policies to improve air quality. Mechanistically detailed Eulerian chemical transportation models (CTMs) are commonly used for this purpose, but owing to high requirements for user training, computational resources, and input data, are often unavailable for urban-level policy analysis. Multiple reduced-complexity air quality models (RCMs) have been designed to fill this gap (Levy et al. 2009; Heo et al. 2017; Casey et al. 2018; Mikati et al. 2018). Among these RCMs, the Intervention Model for Air Pollution (InMAP; Tessum et al., 2017) has proven useful in health impact assessment and environmental justice applications (Goodkind et al. 2019; Tessum et al. 2019; Tschofen et al. 2019; Hill et al. 2019; Thind et al. 2019; Liu et al. 2019, Tessum et al. 2021) owing in part to its use of a variable spatial resolution computational grid which focuses computational effort on areas with high population density. The recent creation of a global-through-urban version of the InMAP model (Thakrar et al., 2021) now provides the opportunity to estimate the exposure consequences of primary and secondary PM_{2.5} concentrations at high spatial resolution in nearly all densely populated areas globally.

Metropolitan-level policy can play an important role in urban air quality (Friedman et al. 2001; Streets et al. 2007; Tonne et al. 2008; Slovic and Ribeiro 2018; Izquierdo et al. 2020), but policymakers in many global cities have not previously had access to detailed information regarding where their ambient PM_{2.5} pollution comes from. Common initial questions from policy makers, in considering how to address air pollution, include (1) Source apportionment: which sources (electricity generation, transportation, industry, etc.) contribute substantially to ambient pollution levels? and (2) Local influence and authority: how much pollution is generated within the city versus transported in from outside? Previous globally-scoped studies have provided this information at national or subcontinental resolution (e.g., Lelieveld et al. 2015; Anenberg et al. 2019), which is of limited use for decision-making at the urban level, or have provided information on the sources contributing to pollution in cities based on a relatively low-

resolution emissions inventory but no information regarding whether those sources are located within or external to the city boundaries ([McDuffie et al. 2021](#)).

Here, we use InMAP to provide scoping-level answers to the two questions above for 96 global cities. Specifically, we estimate the contribution of 12 emission source sectors, both within and outside of the city boundaries, for primary (i.e., directly emitted) and secondary (i.e., formed in the atmosphere from primary emissions of gaseous pollutants) PM_{2.5}. We additionally provide an evaluation of model performance in these urban areas and discuss opportunities for future improvements in model accuracy. For many of the cities we study, information provided herein is the only quantitative information that exists regarding the within-city vs. out-of-city contribution of ambient PM_{2.5}. Results reported here provide both information for stakeholders and an analysis of opportunities to improve their accuracy in future work.

2. Methods

2.1 PM_{2.5} Emission Source Estimation Inventory Data

We study 96 cities that are members of C40, a network of mayors of global cities dedicated to delivering action on climate change ([C40, 2021](#)). For each city, we consider anthropogenic, biogenic, mineral dust, soil denitrification, and biomass burning emissions. For anthropogenic emissions, we use data from the Community Emissions Data System (CEDS) for year 2014 with eight sectors (Table 1): non-combustion agriculture (AGR); energy transformation and extraction (ENE); industrial combustion and processes (IND); surface transportation (road, rail, other) (TRA); residential, commercial, and other (RCO); solvents (SLV); waste disposal and handling (WST); and international shipping (SHP) ([Hoesly et al., 2018](#)). The CEDS emission species used here and their mappings to InMAP species are summarized in Table S1. The CEDS emissions dataset is available at 0.5×0.5 degree spatial resolution, meaning that a single emissions grid cell is larger than many of the cities we study. Section 2.2 describes our methods for downscaling these emissions data to produce higher-resolution estimates.

We use biogenic volatile organic chemical (VOC) emissions generated by the MEGAN model ([Guenther et al. 2006](#)), including the individual VOC components that are considered secondary organic aerosol (SOA) precursors by GEOS-Chem. We use mineral dust emissions generated by the “DustDead” GEOS-Chem algorithm ([Zender et al. 2003](#)). The DustDead model simulates emissions of mineral dust that are mobilized by wind (excluding road dust). The “EMIS_DST1” and “EMIS_DST2” variables were used in this study to represent primary PM_{2.5} emissions. We use soil NO_x emissions from the GEOS-Chem soil NO_x extension ([Hudman et al. 2012](#)). These emissions are at 0.25×0.3125 degree spatial resolution and for year 2016, downloaded from the GEOS-Chem FTP website ([GEOS-Chem authors, 2019](#)).

We use biomass burning emissions from the fourth generation Global Fire Emissions Database (GFED4; [Giglio et al. 2013](#)) at 0.25×0.25 degree spatial resolution for year 2016. These emissions are separate from domestic biomass burning for residential energy use, which is included in the CEDS dataset.

2.2 Spatial Surrogates

We downscale anthropogenic emissions from the native 0.5×0.5 degree CEDS spatial resolution to InMAP grid cells—which vary in size with a minimum edge length of 0.039×0.03125 degrees—using surrogate spatial data, which allows us to represent the spatial distribution of emissions within each CEDS grid cell. We do not apply additional spatial processing to non-anthropogenic emissions. To spatially downscale anthropogenic emissions, we employ spatial datasets that are global in scope and freely available, allowing us to scale our approach to a large number of cities. In cases where the datasets provide no information for a given surrogate within a given city—for example, some cities do not have agricultural areas within their boundaries—we assume emissions are evenly distributed throughout the city area. We also assume all emission sources except electricity generation occur at ground-level. Spatial surrogates used for each CEDS emission sector are as follows:

For *energy transformation and extraction*, we use a database of SO_2 emissions from global electricity generating units (EGUs) (Tong et al. 2018). We use the spatial distribution of SO_2 emissions (rather than another pollutant) because SO_2 emissions are responsible for the vast majority of overall health impacts from EGUs (Fann et al. 2012). EGUs typically have tall emissions stacks, and their emissions plumes often continue to rise after release, owing to their upward velocity exiting the stack and their higher temperature relative to surrounding air. To incorporate stack height and plume rise, we assume that EGUs have stack parameters equal to mean values for EGUs in the year-2014 US EPA National Emissions Inventory (US EPA, 2016) as processed by Tessum et al. (2019): 63.5 m stack height, 4.1 m stack diameter, 519.2 K emission temperature and 24.7 m/s emission velocity. (There is no global database of EGU stack properties.)

For *surface transportation*, we create a spatial surrogate using a weighted average of roadway lengths of OpenStreetMap (OSM) roadways in each CEDS grid cell. We use the following weighted average of OSM roadway types: 36% motorways, 21% trunk roads, 18% primary roads, 9% secondary roads, 1% tertiary, unclassified, and service roads, and 14% residential roads. This weighting is derived from U.S. data on urban road uses (US DOT FHA, 2017); the taxonomy of roadway types is described in Table S2 (OSM, 2019).

For *international shipping*, we create a spatial surrogate from the combined length of OSM features tagged as river, riverbank, pier, ferry, ferry terminal, boat, and mooring. For *non-combustion agriculture*, we create a spatial surrogate from the combined length of OSM features tagged as farm, farmland, or vineyard (OSM, 2017). For *industrial combustion and processes* and *solvents*, we create a spatial surrogate from the combined area of OSM features tagged as industrial buildings or “industrial” or “quarry” land use. For the remaining categories (*residential, commercial, and other, and waste disposal and handling*), we assume a spatial distribution similar to population density, which we represent using the year-2020 projected population from the WorldPop database (Tatem, 2017).

2.3 InMAP Air Quality Modeling

Air quality model description: The Intervention Model for Air Pollution (InMAP) is a mechanistic reduced-complexity air quality model (RCM) that estimates annual-average changes

in primary and secondary PM_{2.5} concentrations attributable to annual changes in emissions of PM_{2.5} and its precursors. In addition to the emissions and population data described in Section 2.1, InMAP requires information on meteorological characteristics and on chemical transport and reaction rates. We use meteorological and background chemistry inputs generated from the outputs of the GEOS-Chem global atmospheric chemical transport model (CTM) simulation for the year 2016, with a base spatial resolution of 2×2.5 degrees and regional nests over Asia, Europe, and North America at 0.5×0.625 degree spatial resolution. The GEOS-Chem simulation uses the SOA_SVPOA chemical mechanism with standard emissions inputs as processed by the HEMCO emissions processor. Further details regarding the GEOS-Chem configuration are described by Thakrar et al. (2021).

The InMAP configuration used here leverages pre-processed physical and chemical information from the output of GEOS-Chem and uses a variable spatial-resolution computational grid to perform simulations that are several orders of magnitude less computationally intensive than conventional CTMs, yet with spatial resolution that is higher than is typically possible using a conventional CTM for a given domain. Conventional CTMs create a three-dimensional Eulerian grid and simulate changes in pollutant concentration in each cell at a high temporal resolution (<1 hour) based on physical transport via wind flow and plume rise, emissions, physical removal (e.g., deposition), and interdependent non-linear physico-chemical transformation pathways. In contrast, InMAP uses time-averaged transport and reaction rates in its algorithms for emission, plume rise, transport, transformation, and removal of atmospheric pollution. To reduce computational intensity, the algorithms are in some cases simplified relative to similar algorithms in a conventional CTM; these simplified representations are calibrated using output from a conventional CTM (GEOS-Chem). This is the first detailed application of InMAP to global cities, but it has previously been used to study PM_{2.5} air pollution in the US, including racial-ethnic disparities in exposure (Tessum et al., 2019; Tessum et al., 2021), and exposure to air pollution from agriculture (Hill et al., 2019), electricity (Thind et al. 2019), and freight (Liu et al. 2019). Further details regarding the InMAP model, including model formulation and performance evaluation, for the US and globally, are described in detail elsewhere (Tessum et al., 2017; Thakrar et al. 2021).

Air quality model application: First, we estimate, for each city, the total PM_{2.5} concentrations and the contributions from each of the 12 source sectors. Here, our spatial domain is global, so we refer to these as the “global” simulations. We configure InMAP to use a variable-resolution grid with 2×2.5 degree resolution for the largest cells, each of which are allowed to split into 4 smaller cells up to 6 times recursively, for a minimum grid cell size of 0.031×0.039 degrees (about 3×4 km² at the equator). The resulting grid (Figure S1) was created by recursively splitting any grid cell containing more than 100,000 people or containing more than 55 million people per square degree in any part of the cell. These 12 simulations (one per source sector), each require ~36 hours on a circa-2019 computer with 32 CPU cores.

Next, we estimate, for each city, the same two parameters as in the first step (total PM_{2.5} and contributions from each sector) but in this case only considering within-city emissions (i.e. emissions originating within a city boundary developed in partnership with city officials for each of the 96 cities). For these city simulations, we use a different InMAP configuration designed to

take advantage of the fact that all emissions are in the relatively small area of one city; this allows us to reduce computational expense while maintaining high spatial resolution in areas where it is beneficial. Specifically, for each city we use a modeling spatial domain of 20×25 degrees ($\sim 2220 \times 2775 \text{ km}^2$, which is $\sim 1\%$ of the surface area of the earth), centered at the center of the city in question. This size was chosen with the expectation of capturing $>90\%$ of the total exposure to $\text{PM}_{2.5}$ pollution resulting from emissions originating in the city (Goodkind et al., 2019). We use maximum and minimum grid cell sizes of 2×2.5 degrees and 0.01×0.013 degrees ($\sim 220 \times 280 \text{ km}^2$ and $\sim 1.2 \times 1.4 \text{ km}^2$ at the equator). As expected, these individual city simulations often end up with higher spatial resolution in the city in question (relative to the global simulations) owing to the decreased overall population in the model spatial domain as compared to the global simulation. The resulting 1,152 individual-city simulations (i.e., 96 cities \times 12 source sectors) each take ~ 2 -30 hours on a circa-2019 computer with 8 CPU cores, depending on the population density of the region in question.

2.4 Data Analysis

For each source sector, we estimate exposure impacts from emissions originating outside of each city by subtracting population-weighted concentrations caused by emissions within the city from total population-weighted concentrations. In this manner we obtain the fraction of total $\text{PM}_{2.5}$ concentrations caused by within-city emissions as well as the fraction of total $\text{PM}_{2.5}$ concentrations caused by different emission sources located either within or outside of the city. However, because the individual city simulations are typically at higher spatial resolution (as described above), in $\sim 3\%$ of the 1152 city-sector combinations the simulated concentrations caused by within-city emissions are greater than the simulated concentrations caused by all emissions, resulting in a negative contribution from out-of-city emissions. The median of those negative values is $-0.0098 \mu\text{g m}^{-3}$, and the largest value is $-1.7 \mu\text{g m}^{-3}$. In these cases, we set the concentrations caused by all emissions equal to the concentrations caused by within-city emissions, so that the contribution from out-of-city emissions equals zero.

Because the air quality model simulations we perform here require a substantial amount of expertise, time and computational resources, we investigate whether patterns exist in the underlying results that could potentially allow extrapolation beyond the 96 cities we studied. To do so, we analyze the relationship between the fraction of $\text{PM}_{2.5}$ originating from within city sources and various city characteristics such as city population, gross domestic product (GDP), city area, and “population buffer fraction”. We define population buffer fraction as the city population divided by the total population within a radius of 200-km from the city centroid. The goal of this supplementary analysis is to explore whether there might exist a straightforward method to reproduce the results shown here without extensive air quality modeling.

2.5 Model Evaluation

We evaluate InMAP model prediction accuracy on total $\text{PM}_{2.5}$ concentration by comparing against measurements from the WHO ambient (outdoor) air pollution database (WHO, 2016), which provides $\text{PM}_{2.5}$ measurements for 53 of the 96 cities studied here. We assess performance using metrics including mean bias (MB), mean error (ME), mean fractional bias (MFB), mean fractional error (MFE), and coefficient of determination (r^2). Definitions of these metrics are in

Table S3. We also use the same metrics to evaluate the global GEOS-Chem model predictions as configured and run by Thakrar *et al.* (2021) against the same PM_{2.5} measurements in the same cities. In addition, we compare InMAP model predictions of total PM_{2.5} concentrations against satellite-based predictions of PM_{2.5} concentrations in 91 global cities (Southerland *et al.* 2021), and predictions by McDuffie *et al.* (2021) in 43 global cities.

Although there do not exist city-specific estimates of PM_{2.5} source apportionment for all of the cities studied here, we compare our estimates for fractional contributions of six similar emission sources with those reported by McDuffie *et al.* (2021) among 43 global cities. We also compare our estimates for fractions of total PM_{2.5} and fractions caused by four emission sources generated by within-city emissions with those reported by Thunis *et al.* (2017) among 17 European cities estimated by the European Commission.

3. Results

3.1. InMAP PM_{2.5} concentration and sources analysis

InMAP model results provide estimates (population-weighted concentrations) for each city of primary and secondary PM_{2.5}, chemically-speciated by source type for within-city versus outside-city emissions. Results are summarized next, with full results for each city in supporting dataset S1.

The median (interquartile range [IQR]) predicted concentration among the 96 cities is 17 (7-43) µg/m³, of which we estimate 57% (37%-73%) is primary (the rest is secondary), and 49% (33%-70%) comes from within-city emissions (the rest comes from sources outside the urban boundary; Table S4).

Concentration estimates for each city (Figure 1) demonstrate substantial variability among urban areas, in terms of concentrations as well as the proportion that is primary vs secondary particulate matter. The highest levels of total PM_{2.5} as predicted here are mainly in Asian cities. As described in Section 3.3, InMAP and other mechanistic model predictions of total PM_{2.5} concentrations in global cities are often substantially different from measured concentrations. Some studies (for example McDuffie *et al.* (2021)) calibrate their mechanistic model predictions to measurement and remote sensing data, but we do not do that here to provide a more accurate sense of the level of uncertainty surrounding our predictions.

The largest contributors to both total PM_{2.5} concentrations (Figure 2) and PM_{2.5} concentrations caused by within-city emissions are most commonly industry, energy transformation and extraction and residential and commercial activities. For example, 49% and 38% of cities have industry ranked the largest for total PM_{2.5} and PM_{2.5} concentrations caused by within-city emissions, respectively; 26% of cities have energy transformation and extraction ranked the largest for both total PM_{2.5} and PM_{2.5} concentrations caused by within-city emissions; and 9% and 22% of cities have residential and commercial activities ranked the largest total PM_{2.5} and PM_{2.5} concentrations caused by within-city emissions, respectively. The two largest sources of PM_{2.5} caused by out-of-city emissions are the same as total PM_{2.5} and PM_{2.5} caused by within-city emissions; however, the third largest source of PM_{2.5} caused by out-of-city emissions is dust, which is the largest out-of-city source in 21% of cities. Only 9% of cities have surface

transportation listed as the largest PM_{2.5} source caused by within-city emissions. Although industrial combustion/processing and energy transformation/extraction are the top PM_{2.5} sources in many cities, there is substantial variability in which emission sources contribute the most across all 96 cities. Thus, an important implication of these findings is that one-size-fits-all approaches to air quality management are unlikely to work across urban areas. Instead, management practice likely needs approaches tuned to local context, including which sources dominate for that city.

3.2 Fraction of within-city emitted PM_{2.5} and city characteristics

As described in Section 2.4, we tested correlations between the fraction of PM_{2.5} caused by within-city emissions and city characteristics such as city size and urban GDP (Figure S2). The results suggest low correlation between fraction of within-city emitted PM_{2.5} and city population, GDP and city area for total, primary and secondary PM_{2.5}. However, there are positive correlations between the fraction of within-city emitted PM_{2.5} and population buffer fraction for total and secondary PM_{2.5}.

There is no statistically significant difference in total PM_{2.5} concentration, city population, GDP or area among cities that have different top ranking emission sources (ANOVA $p=0.29$, 0.69 , 0.10 , and 0.68 , respectively). However, there is a statistical difference in fraction of within-city generated PM_{2.5} among cities that have different top sources ($p=0.00$), where cities with high fraction of within-city generated PM_{2.5} have top PM_{2.5} sources as residential and commercial, industrial and energy, indicating these sources are likely generated locally. Additionally, there is a statistical difference in population buffer fraction among cities having different top sources ($p=0.03$), where residential and commercial, dust, industrial and energy sources tend to be the top PM_{2.5} sources among cities with higher population buffer fractions, while agricultural, international shipping and surface transportation tend to be the top PM_{2.5} sources among low population buffer fraction cities. Along with the results shown in Figure S2, our findings suggest if a city does not have other densely-populated areas nearby, it tends to have the most locally-generated PM_{2.5}, whereas if a city has other densely populated areas nearby, it tends to have a smaller proportion of PM_{2.5} generated locally. The level of population buffer fraction is strongly associated with certain sources, as well as with total and secondary PM_{2.5} concentration. None of the urban parameters we investigated are well-correlated with the top PM_{2.5} source for a city. This finding suggests that atmospheric modeling holds value for understanding the local context of which sources contribute the most to local pollution.

3.3 Comparison of InMAP results with measurements and other studies

We evaluate the InMAP predicted total PM_{2.5} concentrations against measured total PM_{2.5} concentrations. There are 53 global cities that have both InMAP predictions and measured data collected by WHO (WHO, 2016); results are shown in the first panel of Figure 3. The r^2 is 0.47 and bias and error are listed in Table S5. For the same cities, the model-measurement agreement is better for GEOS-Chem than for InMAP ($r^2=0.57$; Figure 3, second panel) and GEOS-Chem has lower error but larger bias (Table S5).

We additionally compare the InMAP predicted total PM_{2.5} concentrations using satellite derived PM_{2.5} concentrations (Southerland et al. 2021) among 91 cities in Figure S3 ($r^2=0.22$),

as well as using modeled total PM_{2.5} concentrations reported by McDuffie *et al.* (2021) in the first panel of Figure S4 ($r^2 = 0.27$). (Note that total PM_{2.5} concentrations reported by McDuffie *et al.* (2021) are calibrated to measurement and remote sensing data, which explains their good agreement with measurement data.) We compare our estimates for fractions of total PM_{2.5} caused by six common emission sources with those reported by McDuffie *et al.* (2021) in Figure 4. Most fractions of these emission sources show good correlations (r^2 : 0.21-0.53) between the two studies except the fraction of PM_{2.5} from residential, commercial and other ($r^2 = 0.01$). We also compare InMAP fractions of total PM_{2.5} concentrations caused by within-city emissions with these fractions from 17 European cities (Thunis *et al.* 2017) as shown in Figure S5 ($r^2 = 0.17$). Additionally, we compare the fractions of total PM_{2.5} caused by four common emission sources, as well as the fractions of total PM_{2.5} concentrations caused by these four emission sources generated by within-city emissions to that study in Figures S6 and S7. There are no correlations between InMAP predictions and predictions by Thunis *et al.* (2017), except the fractions of total PM_{2.5} caused by agricultural and residential sources (the differences in emission source categories are listed in Table S6). We also evaluate the InMAP within-city fractions for total PM_{2.5} against zero-out simulations with GEOS-Chem conducted in five cities and find good agreement ($r^2 = 0.67$) when InMAP is run at the coarser resolution used by GEOS-Chem. Model comparison is described in more detail in the supporting information.

4. Discussion

In this study, we have produced estimates of PM_{2.5} concentration, its source composition, and contribution of in-city vs. out-of-city sources—the latter of which was not previously available in many global cities—that can be useful as a screening tool and in many cases may be robust enough to inform policy action to enact more effective strategies for improving public health. We find that although industrial- and energy-related sources are the largest contributors to PM_{2.5} in a plurality of cities, there is considerable diversity among cities in which source types are most important (Figure 2). Because there is considerable uncertainty inherent in the results presented here, in some cases it may be advisable to leverage these results to obtain more precise estimates using a more mechanistically detailed air quality model and locally produced emission inventories before taking policy action, to the extent that resources are available to support the additional analysis.

Beyond the specific results for individual cities archived in the supporting information and summarized above, we would like to call the reader's attention to several salient points that emerge from this analysis.

The first point is that we did not find any strong patterns among cities that could predict which emission sources contribute most to a city's PM_{2.5} concentrations, or how much of a city's ambient PM_{2.5} concentration originates from emissions within the city boundary (Figure S2). To some extent, this contradicts previous findings by Apte *et al.* (2012); this discrepancy may be explained by the more detailed model of the atmosphere used here. The implication of this point is that it is necessary to perform atmospheric modeling in a given city to get a realistic estimate of the sources of that city's pollution—there don't seem to be any shortcuts. Since it may not be practical to carry out urban-level air quality simulations in a large number of cities using a

comprehensive model like GEOS-Chem, this underscores the utility of reduced-complexity models such as InMAP.

The second point is that urban air quality analyses require urban emissions inventories. As described above, we use spatial surrogates—mainly based on OpenStreetMap data—to allocate 0.5×0.5 degree CEDS emissions to much smaller InMAP grid cells. This spatial downscaling is important: in an analysis comparing InMAP and GEOS-Chem predicted contributions of within-city emissions for a subset of five of the 96 cities studied here (i.e. Johannesburg, Buenos Aires, Addis Ababa, Chengdu, and Guadalajara), we found that the r^2 value between GEOS-Chem and InMAP predictions when InMAP used emissions at their native resolution was 0.67, but when InMAP used the same emissions downscaled with the spatial surrogates described above, the r^2 value decreased to 0.2 (Figure S8, methods in supporting text). This implies that the use of high-resolution emissions provides information that couldn't be reproduced by—for example—applying a correction factor to simulation results based on low-resolution emissions.

Building on the second point, the third point is that in this analysis the emission inventory appears to be a larger source of potential error than the choice of air quality model. For example, the r^2 value for total concentration predictions in five cities between GEOS-Chem and InMAP when using the same emission inventory is 0.98 (Figure S8), but the r^2 value between GEOS-Chem and InMAP for the fractional contribution of Residential, Commercial, and Other emissions in 43 cities when using different versions of the CEDS inventory ([Hoesly et al. 2018](#) vs. [McDuffie et al. 2021](#)) at different spatial resolutions is 0.01 (Figure 4).

As described above, high-resolution emissions estimates are important for urban-scale analysis. The fourth point is that downscaling existing inventories using spatial surrogates can only yield improvements up to a certain point. The global CEDS inventory used here is mainly based on national emissions estimates that are themselves downscaled to a 0.5×0.5 degree grid using mainly population density estimates. This can lead to spatial misallocations that cannot be fixed by further downscaling. For example, using the national-population-based spatial allocation method above in a country with substantial residential coal emissions could allocate a plurality of those emissions to the cosmopolitan capital city, where in reality there were relatively few residential coal emissions owing to the capital's relatively high level of affluence. The next generation of global emissions inventories may benefit from the emergence of new streams of relevant data—for example from smart phones and satellites—in combination with local expertise facilitated by networks of cities like C40.

Acknowledgements

We acknowledge funding from the Clean Air Fund via C40 and from the Wellcome Trust (Grant 216075/Z/19/Z). All authors were funded by both sources. Further, we acknowledge contributions from C40 employees Tibebe Assefa, Andrea Bizberg, Toby Coombes, Helen Ho, Seneca Naidoo, Nwabisa Potwana, Nadia Shah, and Culley Thomas. Finally, we acknowledge C40 for providing geographical city boundaries and access to feedback from local city administrators.

Table 1. Sectors of anthropogenic emissions from the Community Emissions Data System (Hoesly et al. 2018)

Sector	Specification	Spatial Surrogate
Non-combustion agricultural sector (AGR)	manure management, soil emissions, rice cultivation, enteric fermentation, and other	Agricultural sector
Energy transformation and extraction (ENE)	electricity production, heat production, other energy transformation, related fugitive emissions, and fossil fuel fires	Energy generation
Industrial combustion and processes (IND)	combustion for manufacturing of goods and minerals and for construction, production of cement, lime, and “other minerals”, mining, chemical production, paint application, wood, pulp, and paper products	Industrial sector
Surface transportation (road, rail, other) (TRA)	air, road, rail, and water transportation	Roadways
Residential, commercial, and other (RCO)	commercial-institutional, residential, agriculture-forestry-fishing, and other-unspecified emissions	Population
Solvents (SLV)	used in degreasing and cleaning	Industrial sector
Waste disposal and handling (WST)	solid waste disposal, waste combustion, wastewater handling, and other	Population
International shipping (SHP)	VOCs from oil tanker loading/leakage	Waterways

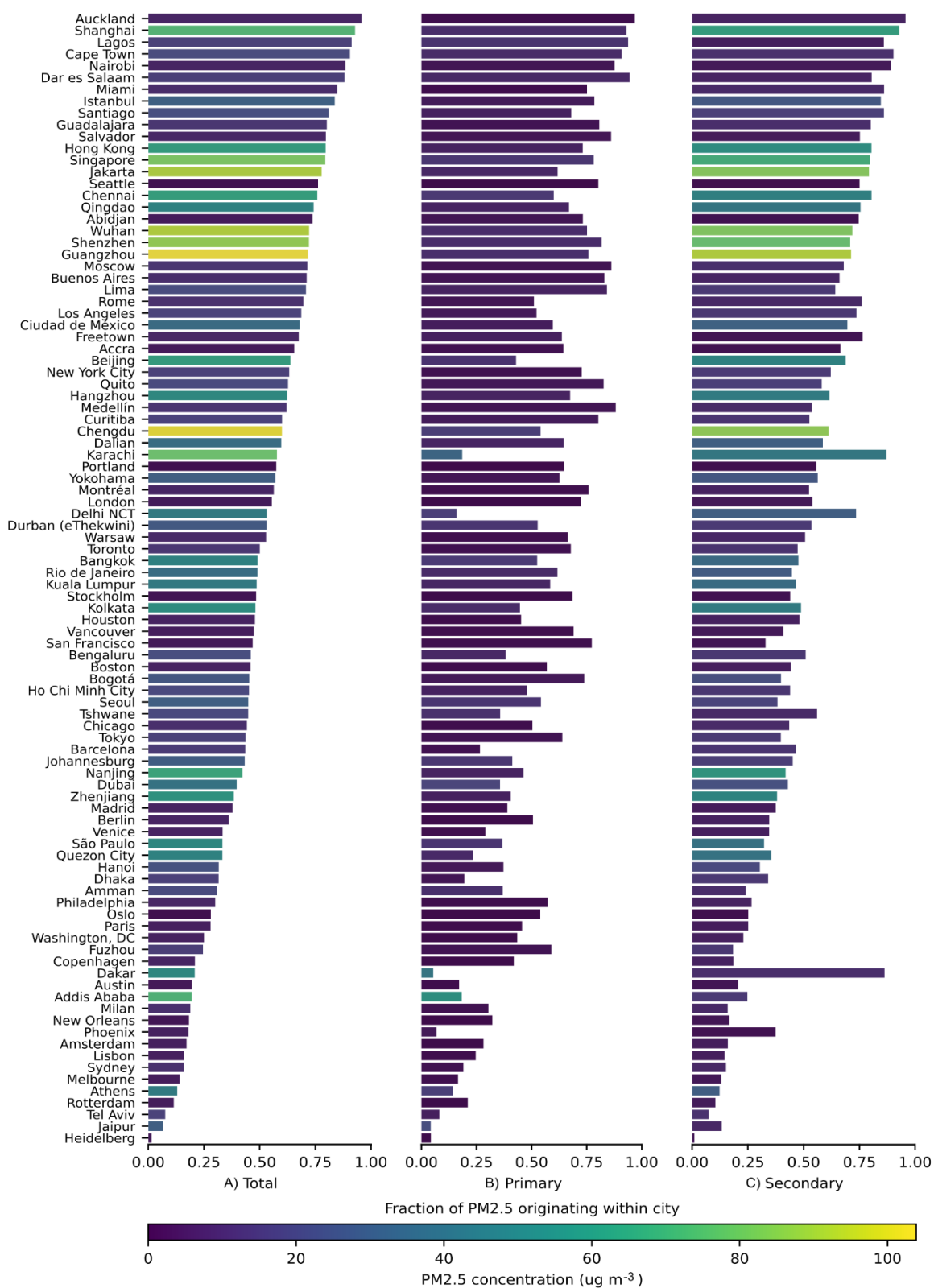


Figure 1. Fractions of PM_{2.5} originating from within city sources for A) total PM_{2.5}, B) primary PM_{2.5} and C) secondary PM_{2.5} among 96 global cities. Color scales represent PM_{2.5} concentration ($\mu\text{g m}^{-3}$).

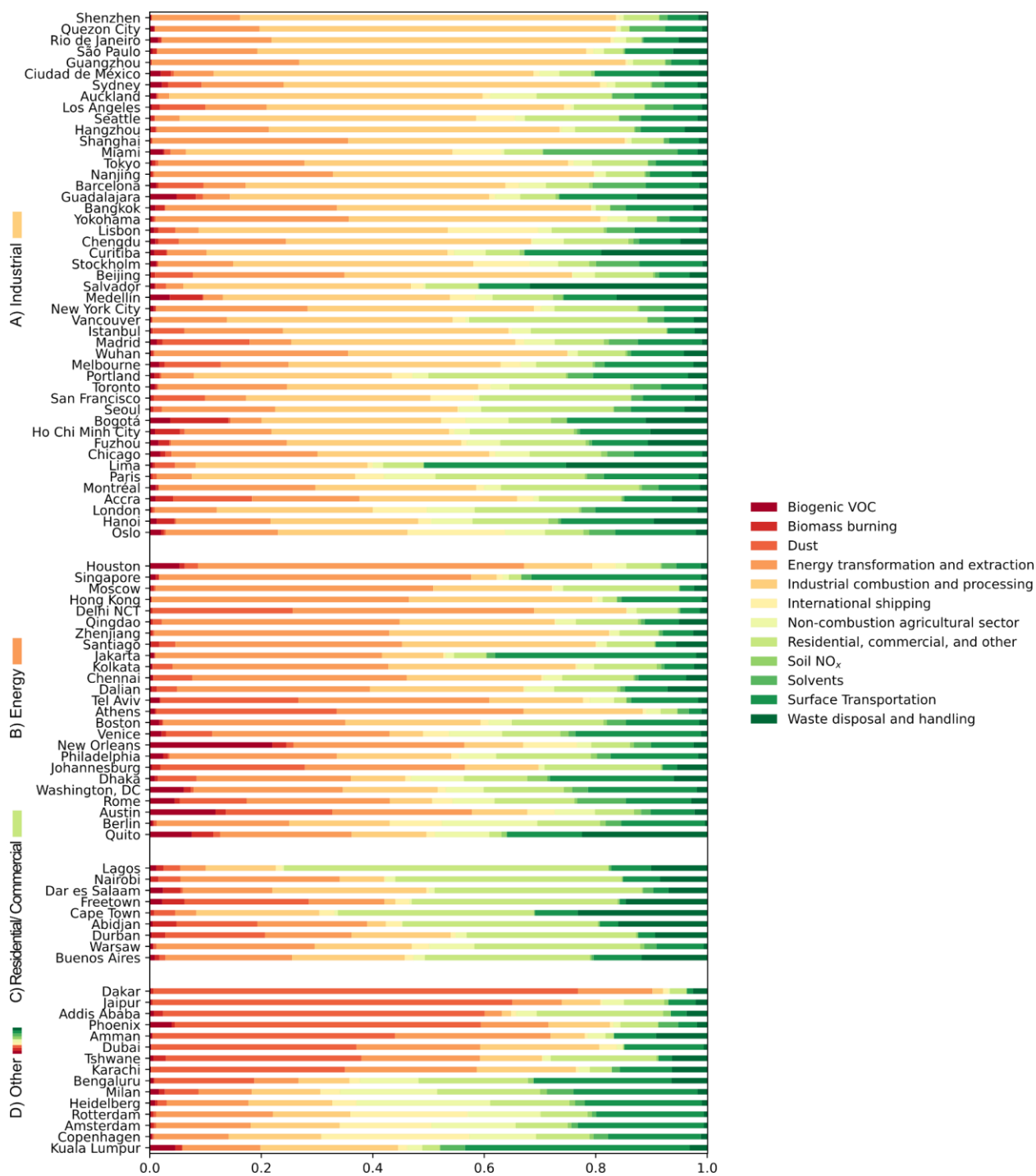


Figure 2. Proportions of total $PM_{2.5}$ from 12 sources among 96 global cities, grouped by the largest sources: A) Industrial combustion and processing, B) Energy transformation and extraction, C) Residential, commercial, and other, and D) Other sectors.

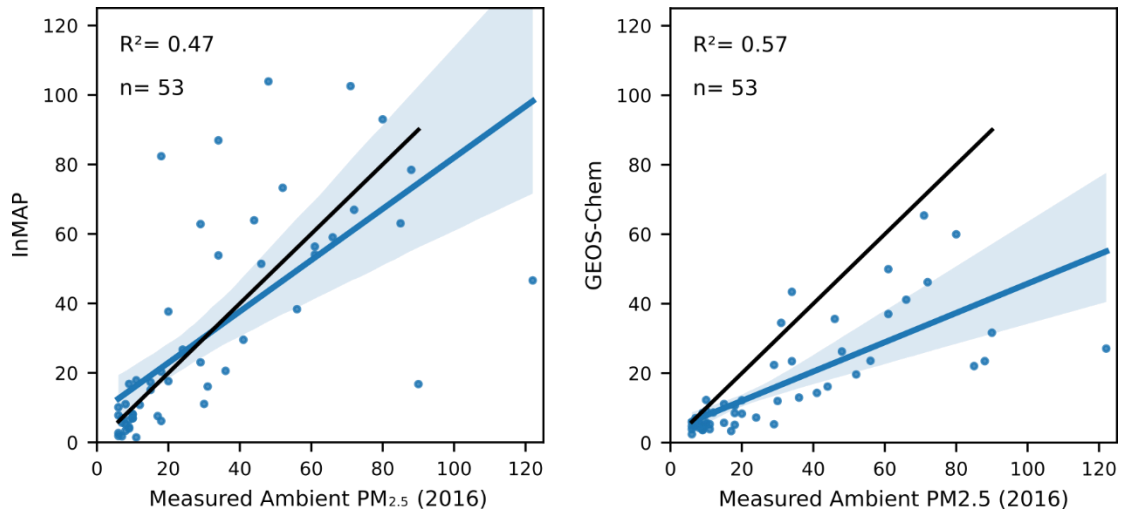


Figure 3. Comparison of InMAP predicted total PM_{2.5} concentrations and measured total ambient PM_{2.5} concentrations (WHO 2016, left) and GEOS-Chem predicted total PM_{2.5} concentrations and measured total ambient PM_{2.5} concentrations (right) among 53 global cities. The blue line is a least-squares model fit and blue shaded areas indicate the 95% confidence interval of a least squares fit. The black line represents a 1:1 relationship.

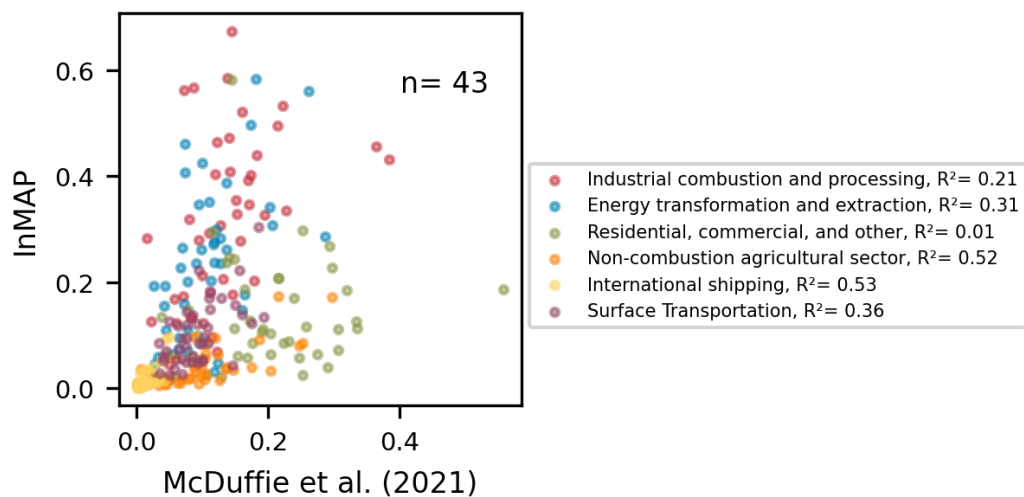


Figure 4. Comparison of fractions of total $PM_{2.5}$ caused by different emission sources between InMAP (this study) and McDuffie *et al.* (2021) among 43 global cities.

References

- World Health Organization (WHO) Website; <https://www.who.int/health-topics/air-pollution>. (Accessed: 2021-8-16).
- Myllyvirta, L. *Quantifying the Economic Costs of Air Pollution from Fossil Fuels*; Centre for Research on Energy and Clean Air, 2020, <https://energyandcleanair.org/wp/wp-content/uploads/2020/02/Cost-of-fossil-fuels-briefing.pdf>. (Accessed: 2021-8-16)
- State of Global Air/2019—A Special Report on Global Exposure to Air Pollution and Its Disease Burden*; Health Effects Institute (HEI): Boston, MA, 2019. https://www.stateofglobalair.org/sites/default/files/soga_2019_report.pdf. (Accessed: 2021-8-16)
- United Nation Environmental Programme (UNEP) Website: <https://wesi.unep.org/airpollution>.
- Levy, J.I.; Baxter, L.K.; Schwartz, J. Uncertainty and variability in health-related damages from coal-fired power plants in the United States. *Risk Anal.* 2009, 29, 1000–1014.
- Heo, J.; Adams, P.J.; Gao, H.O. Public health costs accounting of inorganic PM 2.5 pollution in metropolitan areas of the United States using a risk-based source-receptor model. *Environ Int.* 2017, 106, 119–126.
- Casey, J.A.; Karasek, D.; Ogburn, E.L.; Goin, D.E.; Dang, K.; Braveman, P.A. et al. Coal and oil power plant retirements in California associated with reduced preterm birth among populations nearby. *Am J Epidemiol.* 2018, 187(8), 1586–1594. <https://doi.org/10.1093/aje/kwy110/4996680>.
- Mikati, I.; Benson, A.F.; Luben, T.J.; Sacks, J.D.; Richmond-Bryant, J. Disparities in distribution of particulate matter emission sources by race and poverty status. *Am J Public Health.* 2018, 108, 480–485.
- Tessum, C.W.; Hill, J.D.; Marshall, J.D. InMAP: A model for air pollution interventions, *PLoS ONE* 2017, 12(4), e0176131. DOI 10.1371/journal.pone.0176131.
- Goodkind, A. L.; Tessum, C. W.; Coggins, J. S.; Hill, J. D.; Marshall, J. D. Fine-scale damage estimates of particulate matter air pollution reveal opportunities for location-specific mitigation of emissions. *Proc. Natl. Acad. Sci. USA.* 2019, 116, 8775–8780.
- Tessum, C.W.; Apte, J.S.; Goodkind, A.L.; Muller, N.Z.; Mullins, K.A.; Paoletta, D.A.; Polasky, S.; Springer, N.P.; Thakrar, S.K.; Marshall, J.D.; Hill, J.D. Inequity in consumption of goods and services adds to racial-ethnic disparities in air pollution exposure. *Proc. Natl. Acad. Sci. USA.* 2019, 116, 6001–6006.
- Tschofen, P.; Azevedo, I.L.; Muller, N.Z. Fine particulate matter damages and value added in the US economy. *Proc. Natl. Acad. Sci. USA.* 2019, 116 (40), 19857–19862
- Hill, J.D.; Goodkind, A.L.; Tessum, C.W.; Thakrar, S.K.; Tilman, D.; Polasky, S.; Smith, T.; Hunt, N.; Mullins, K.A.; Clark, M.; Marshall, J.D. Air-quality-related health damages of maize. *Nat. Sustain.* 2019, 2, 397.
- Thind, M.P.S.; Tessum, C.W.; Azevedo, I.L.; Marshall, J.D. Fine Particulate Air Pollution from Electricity Generation in the US: Health Impacts by Race, Income, and Geography. *Environmental Science & Technology* 2019, 53 (23), 14010–14019. DOI 10.1021/acs.est.9b02527.

Liu, L.; Hwang, T.; Lee, S.; Ouyang, Y.; Lee, B.; Smith, S.J.; Tessum, C.W.; Marshall, J.D.; Yan, F.; Daenzer, K.; Bond, T.C., Health and climate impacts of future united states land freight modelled with global-to-urban models. *Nat. Sustain.* 2019, 2, 105–112.

Tessum, C.W.; Paoletta, D.A.; Chambliss, S.E.; Apte, J.S.; Hill, J.D.; Marshall, J.D. PM_{2.5} polluters disproportionately and systemically affect people of color in the United States. *Science Advances*, 2021, 7(18), eabf4491.

Thakrar, S.; Tessum, C.W.; Apte, J.; Balasubramanian, S.; Millet, D. B.; Pandis, S.; Marshall, J.D.; Hill, J.D. Global, High-Resolution, Reduced-Complexity Air Quality Modeling Using InMAP (Intervention Model for Air Pollution), ChemRxiv. Preprint. 2021. DOI 10.26434/chemrxiv.14330375.v1.

Friedman, M.S.; Powell, K.E.; Hutwagner, L.; Graham, L.M.; Teague, W.G. Impact of changes in transportation and commuting behaviors during the 1996 Summer Olympic Games in Atlanta on air quality and childhood asthma. *JAMA*, 2021, 285(7), 897-905.

Streets, D.G.; Fu, J.S.; Jang, C.J.; Hao, J.; He, K.; Tang, X.; Zhang, Y.; Wang, Z.; Li, Z.; Zhang, Q.; Wang, L. Air quality during the 2008 Beijing Olympic games. *Atmospheric Environment*, 2007, 41(3), 480-492.

Tonne, C.; Beevers, S.; Armstrong, B.; Kelly, F.; Wilkinson, P. Air pollution and mortality benefits of the London Congestion Charge: spatial and socioeconomic inequalities. *Occupational and Environmental Medicine*, 2008, 65(9), 620-627.

Slovic, A.D. and Ribeiro, H. Policy instruments surrounding urban air quality: The cases of São Paulo, New York City and Paris. *Environmental science & policy*, 2018, 81, 1-9.

Izquierdo, R.; Dos Santos, S.G.; Borge, R.; De La Paz, D.; Sarigiannis, D.; Gotti, A.; Boldo, E. Health impact assessment by the implementation of Madrid City air-quality plan in 2020. *Environmental Research*, 2020, 183, 109021.

Lelieveld, J.; Evans, J.S.; Fnais, M.; Giannadaki, D.; Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale, *Nature* 2015, 525, 367–371. DOI 10.1038/nature15371.

Anenberg, S.C.; Miller, J.; Henze, D.K.; Minjares, R.; Achakulwisut, P. The global burden of transportation tailpipe emissions on air pollution-related mortality in 2010 and 2015. *Environmental Research Letters*, 2019, 14(9), p.094012.

McDuffie, E.E.; Martin, R.V.; Spadaro, J.V.; Burnett, R.; Smith, S.J.; O'Rourke, P.; Hammer, M.S.; van Donkelaar, A.; Bindle, L.; Shah, V.; Jaeglé, L. Source sector and fuel contributions to ambient PM_{2.5} and attributable mortality across multiple spatial scales. *Nature Communications*, 2021, 12(1), 1-12.

C40 Annual Report 2020; C40 Cities. 2021. https://c40-production-images.s3.amazonaws.com/other_uploads/images/2827_C40_Annual_Report_2020_vMay2021_lightfile.original.pdf. (Accessed: 2021-11-5)

Hoesly, R.M.; Smith, S.J.; Feng, L.; Klimont, Z.; Janssens-Maenhout, G.; Pitkanen, T.; Seibert, J.J.; Vu, L.; Andres, R.J.; Bolt, R.M.; Bond, T.C. Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the community emissions data system (CEDS). *Geosci. Model Dev.* 2018, 11, 369–408.

Guenther, A.; Karl, T.; Harley, P.; Wiedinmyer, C.; Palmer, P.I.; Geron, C. Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmos. Chem. and Phys.* 2006, 6, 3181–3210.

Zender, C.S.; Bian, H.; Newman, D. Mineral Dust Entrainment and Deposition (DEAD) model: Description and 1990s dust climatology. *J. Geophys. Res.: Atmos.* 2003, 108.

Hudman R. C.; Moore, N.E.; Mebust, A.K.; Martin, R.V.; Russell, A.R.; Valin, L.C.; Cohen, R.C. Steps towards a mechanistic model of global soil nitric oxide emissions: Implementation and space based-constraints. *Atmos. Chem. Phys.* 2012, 12, 7779–7795.

GEOS-Chem authors, GEOS-Chem emissions. 2019
<http://geoschemdata.computecanada.ca/ExtData/HEMCO/> (Accessed: 2021-7-12)

Giglio, L.; Randerson, J. T.; van der Werf, G. R. Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). *J. of Geophys. Res.: Biogeosci.* 2013, 118, 317–328.

Tong, D.; Zhang, Q.; Davis, S.J.; Liu, F.; Zheng, B.; Geng, G.; Xue, T.; Li, M.; Hong, C.; Lu, Z.; Streets, D.G. Targeted emission reductions from global super-polluting power plant units. *Nat. Sustain.* 2018, 1, 59.

Fann, N.; Baker, K. R.; Fulcher, C. M. Characterizing the PM_{2.5}-related health benefits of emission reductions for 17 industrial, area and mobile emission sectors across the US. *Environ. Intl.* 2012, 49, 141–151.

U.S. Environmental Protection Agency. 2014 National Emissions Inventory, version 1 Technical Support Document. Research Triangle Park, North Carolina, 2016.
https://www.epa.gov/sites/production/files/2016-12/documents/nei2014v1_tsd.pdf. (Accessed: 2021-7-13)

U.S. Department of Transportation Federal Highway Administration, Annual Vehicle-Miles of Travel, 1980–2017 by Functional System National Summary, 2017,
<https://www.fhwa.dot.gov/policyinformation/statistics/2017/vm202.cfm> (Accessed: 2021-7-13).

OpenStreetMap contributors, Planet dump retrieved from <https://planet.osm.org> 2017,
<https://www.openstreetmap.org> (Accessed: 2021-7-13).

OpenStreetMap contributors, OpenStreetMap Wiki: Key: highway 2019,
<https://wiki.openstreetmap.org/wiki/Key:highway> (Accessed: 2021-09-08).

Tatem, A.J. WorldPop, open data for spatial demography. *Scientific Data.* 2017, 4, 170004.

World Health Organization (WHO). Air quality database: Update 2016.
<https://www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database/2016>
 (Accessed: 2021-8-10).

Southerland, V.A.; Anenberg, S.C.; Brauer, M.; Mohegh, A.; Hammer, M.; Martin, R.; van Donkelaar, A.; Apte, J. Global urban temporal trends in fine particulate matter (PM_{2.5}) and attributable health burdens. 2021. Under review.

Thunis, P.; Degraeuwe, B.; Pisoni, E.; Trombetti, M.; Peduzzi, E.; Belis, C.A.; Wilson, J.; and Vignati, E. Urban PM_{2.5} Atlas – Air Quality in European Cities, EUR 28804 EN, Publication

Office of the European Union, Luxembourg, 2017.

<https://publications.jrc.ec.europa.eu/repository/handle/JRC108595> (Accessed: 2021-8-10).

Stohl, A.; Aamaas, B.; Amann, M.; Baker, L.H.; Klimont, Z.; Kupiainen, K.; Heyes, C.

Evaluating the climate and air quality impacts of short-lived pollutants. *Atmospheric Chemistry and Physics* 2015 15 (18) 10529-10566. 10.5194/acp-15-10529-2015.