Curr Environ Health Rep. Author manuscript; available in PMC 2022 June 01.

Published in final edited form as:
Curr Environ Health Rep. 2021 June ; 8(2): 113–126. doi:10.1007/s40572-021-00310-y.

Fine-scale air pollution models for epidemiologic research: Insights from approaches developed in the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air)

Kipruto Kirwa[1], **Adam A. Szpiro**[2], **Lianne Sheppard**[3], **Paul D. Sampson**[4], **Meng Wang**[1,5], **Joshua P. Keller**[6], **Michael T. Young**[1], **Sun-Young Kim**[1,7], **Timothy V. Larson**[8], **Joel D. Kaufman**[9]

[1]Department of Environmental and Occupational Health Sciences University of Washington School of Public Health, Seattle, WA

[2]Department of Biostatistics University of Washington School of Public Health, Seattle, WA

[3]Departments of Biostatistics and Environmental and Occupational Health Sciences University of Washington School of Public Health, Seattle, WA

[4]Department of Statistics University of Washington School of Public Health, Seattle, WA

[5]Department of Epidemiology and Environmental Health, School of Public Health and Health Professions Research and Education in Energy, Environment and Water Institute University at Buffalo, Buffalo, New York

[6]Department of Statistics, Colorado State University, Fort Collins, Colorado, USA

[7]Institute of Health and Environment, Seoul National University, Seoul, Korea.

[8]Department of Civil and Environmental Engineering University of Washington, Seattle, WA

[9]Departments of Environmental and Occupational Health Sciences, Epidemiology, and Medicine, University of Washington, Seattle, WA

Abstract

Purpose of review—Epidemiological studies of short and long-term health impacts of ambient air pollutants require accurate exposure estimates. We describe the evolution in exposure assessment and assignment in air pollution epidemiology, with a focus on spatiotemporal techniques first developed to meet the needs of the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). Initially designed to capture the substantial variation in pollutant levels and potential health impacts that can occur over small spatial and temporal scales in metropolitan areas, these methods have now matured to permit fine-scale exposure characterization across the contiguous United States, and can be used for understanding long- and short-term health effects of

Compliance with Ethics Guidelines
Conflict of Interest
The authors declare no conflicts of interest in the production of this work.

Human and Animal Rights and Informed Consent
This article does not contain any studies with human or animal subjects performed by any of the authors.

exposure across the lifespan. For context, we highlight how the MESA Air models compare to other available exposure models.

Recent findings—Newer model-based exposure assessment techniques provide predictions of pollutant concentrations with fine spatial and temporal resolution. These validated models can predict concentrations of several pollutants, including particulate matter less than 2.5 μm in diameter ($PM_{2.5}$), oxides of nitrogen, and ozone, at specific locations (such as at residential addresses) over short time intervals (such as two weeks) across the contiguous United States between 1980 and the present. Advances in statistical methods, incorporation of supplemental pollutant monitoring campaigns, improved geographic information systems and integration of more complete satellite and chemical transport model outputs have contributed to the increasing validity and refined spatiotemporal spans of available models.

Summary—Modern models for predicting levels of outdoor concentrations of air pollutants can explain a substantial amount of the spatiotemporal variation in observations and are being used to provide critical insights into effects of air pollutants on the prevalence, incidence, progression, and prognosis of diseases across the lifespan. Additional enhancements in model inputs and model design, such as incorporation of better traffic data, novel monitoring platforms, and deployment of machine learning techniques will allow even further improvements in the performance of pollutant prediction models.

Introduction

Exposure to air pollution is a well-recognized risk factor for numerous adverse health effects, including morbidity and mortality due to multiple disorders and reduction in life expectancy [1–3]. Air pollution is considered the largest environmental cause of disease burden globally [4], and owing to ubiquity of exposure, even modest increases in dose may result in substantial negative health consequences at the population level.

Validity of epidemiological cohort studies of health effects of air pollutants requires accurate estimates of exposure associated with individual study participants. Ideally, researchers would use individual-focused micro-environmental measures of exposure over the period relevant for examining the outcome under consideration. In practice, however, multiple factors render this infeasible, including low reliability and/or high cost of personal exposure measurement devices, the burden of deploying such devices for more than brief periods of time or in large populations, and the inability to focus only on ambient-source pollutants—which are typically the regulated pollutants of primary epidemiological interest—with current personal monitoring technologies. In a prior generation of studies (conducted towards the end of the 20th century) [5, 6], it was common to assign air pollution concentrations from centralized monitors to all individuals residing in a particular region. This approach suffers key shortcomings. Notably, it misses spatial variation in concentration within the coverage area, which can be substantial even across relatively small distances. Also, centralized monitors are available in comparatively few places and some are sited in ways that can make their measurements systematically biased or not reflect community-wide concentrations.

Modern epidemiologic cohort studies of air pollution increasingly rely on statistically-modeled predictions of concentrations. This involves the use of pollutant measurements from relatively few monitored locations in conjunction with multiple types of geographic, atmospheric, and physicochemical information as covariates in statistical models to estimate the concentration of pollutants at a much larger number of unmonitored locations over a given timeframe (Figure). Although this approach addresses some shortcomings of the traditional approaches, statistical prediction of pollutant concentrations has its own challenges, including missingness and irregularity of input data, the need for geographic and other covariates necessary for accurate prediction, and the need for user-friendly statistical techniques with transparent validation procedures that can be used to process data inputs and produce stable predictions.

Reviews have discussed the principles, characteristics, and comparative performance of various methods for assessing and modeling air pollution exposures [7–9]. Here we focus particularly on describing recent developments in statistical prediction of pollutant concentrations through the lens of spatiotemporal modeling approaches developed in the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). MESA Air, a cohort study of more than 7,000 Hispanic, Chinese American, white, and African American adults in six metropolitan regions, was designed to assess the association between ambient air pollutant exposure and cardiovascular disease [10]. We start by outlining the rationale for highly-resolved estimates of pollutant concentrations in cohort studies, introduce the core spatiotemporal model developed for application in the six MESA Air regions, and then discuss how the model has been extended to serve a wider array of epidemiologic use-cases. Along the way, we briefly discuss alternative modeling approaches for each use case and highlight insights learned while developing and deploying the models in health outcome studies. Finally, we outline opportunities for further improving model-based exposure assessment in air pollution studies.

Motivation for estimating fine-scale concentrations of pollutants

The goal of exposure assessment in air pollution epidemiology is to accurately quantify the average concentration of a target pollutant at specific locations, especially where people live or otherwise spend most of their time, over the periods of interest. Air pollutant concentrations vary at fine scale, and the inability to capture this variation risks bias due to exposure misclassification or may result in biased and/or imprecise estimates of health effects [11–13]. Fine-scale variation is important because intra-region variation in concentrations and associated health risks may be larger than inter-region differences [14, 15]. While the magnitude of within-region variation differs by pollutant, all pollutants exhibit important intra-urban variations in concentration, both across space and over time. Important predictors include proximity to roadways and other point sources, meteorology and other seasonal factors, and geographic properties such as predominant land uses, elevation and topography. Structural factors such as policies that influence emissions from industrial or transportation sources may impact secular trends in concentrations, while features like meteorology, traffic patterns, and economic activity result in substantial short-term swings in concentrations. Exposure gradients across even relatively small distances or durations have been shown to be of epidemiological consequence [16], hence estimation and inference about pollutant-outcome relationships requires accurate exposure assessment methods.

Evolution from metropolitan area-wide estimation to nearest monitor to precise residential locations

In early landmark air pollution cohort studies, exposure concentrations were often assigned to all participants in a region based on central "regulatory" monitors [5, 6]. However, this eliminates the opportunity to exploit intra-region spatial variation. An early solution was to assign exposures based on nearest monitors when a study region had more than one monitor [15], but this approach still did not address the core issue of local pollution variation because even in high-income countries, well-calibrated reference-standard monitors are sparsely-located and unevenly-distributed [17].

Exposure assessment using model-based predictions addresses this problem. There are different types of statistical models, but all typically calculate parameters that explain the relationship between pollutant concentrations at measured locations and a wide variety of characteristics thought to influence pollutant formation, distribution and decay, and then deploy those parameters to estimate pollutant levels at unmonitored places and times, conditional on the predictive characteristics. Models are distinguished by the covariate selection and algorithmic approaches employed, the nature of covariates used, whether they account for spatial or temporal variation or both, application of geostatistical techniques to account for spatial structure outside of the mean field, and whether they are designed to take inputs and produce outputs that correspond to "gridded" areas or point locations. The quality of resultant estimates is judged by comparing model predictions to observations where measurements are available. Prediction accuracy can be measured by mean square error, which represents average difference between model predictions for specific locations and their corresponding input observations, or by R^2, a measure of explained variation based on either spatial or temporal comparisons of average predictions versus observations. Spatial R^2 can be operationalized as the cross-validated squared correlation coefficient of the long-term averages of predictions and observations at every monitored location, while temporal R^2 may be calculated as the median of squared correlation coefficients of observations and predictions at each location across the study period.

Model-based methods for predicting air pollutant concentrations include land-use regression (LUR) [18], geostatistical methods such as kriging [19], satellite-derived aerosol optical depth [20, 21], chemical transport modeling (CTM) [22], dispersion models [23], generalized additive models [24], and artificial neural networks [25]. Each has advantages and limitations [7], prompting adoption of fine-scale prediction strategies that tap their collective strengths, including universal kriging-based statistical models that incorporate CTM and satellite-derived variables as covariates [26, 27], and hybrid models that fuse multiple methodologies [28]. LUR has remained the term used for models built using variables obtained from geographic information systems (GIS), although land use is only one of the variable classes used to predict concentrations. Other frequently used classes of predictors include population density, and traffic characteristics, transportation routes, and ground-level characteristics from satellite imaging [18]. In kriging, spatial dependence in the data is used to develop continuous surfaces of pollution, and variability in concentrations is modeled as a factor of spatial trend (local and/or long-range) and non-spatial error components [19]. Simple kriging assumes no global trend in the data and spatially homogenous variation, while universal kriging is the result of combining LUR and simple kriging, whereby LUR predicts aspects of the exposure surface that are related to geography and kriging spatially smooths the residuals [29].

GIS data are a foundation of these models. GIS enhancements continue at a rapid pace, and currently permit compilation of expansive libraries of covariates, ability to accurately calculate covariates for arbitrarily-defined points or buffers around points, frequent covariates updates, improved completeness of covariates, and longer time periods of coverage with more uniform sets of covariates across numerous political jurisdictions. Recent implementations of fine-scale models have used up to 400 geographic covariates, available for the entire United States [30]. These include land use types, vegetation cover, a large array of traffic-related proximity and length measures, emission sources, topography, proximity to sea- and air-ports, and detailed sociodemographic and census data, among others. Still, there are covariates that may be desirable but are not yet incorporated in current models or readily available in a uniform national system, including waterway traffic, national bus and truck routes, and measures of wildfire occurrence and plumes of resultant smoke, and some relevant built environment features. For some urban settings, accurate point-specific residential exposure prediction may also benefit from incorporating elevation estimates (e.g., for multi-story housing), and information regarding “street canyon” effects.

Given increasingly high-dimensional and potentially correlated GIS covariates, pollutant model specification decisions benefit from approaches that both optimally characterize pollutant variation and are parsimonious, to avoid overfitting, collinearity, and computational challenges. Options include variable selection via stepwise regression and/or shrinkage using techniques such as LASSO [29]. A third choice is covariate dimension reduction, which avoids selecting a limited subset of variables and creates composite scores through, for example, partial least squares (PLS) or principal components regression.

The finest feasible spatial scales for model outputs depend partly on the resolutions of model inputs, such as geographic data and satellite or pollutant measurements. Current fine spatial scale models provide predictions either at a point [31] or at grids of 1km x 1km or less [25]. While small grids are a substantial improvement from region-wide assignment, there remains potential for within-grid measurement error—especially in spaces where concentrations can vary over smaller scales as is typically the case in urban areas. Additionally, grids are often defined arbitrarily in ways that make them incompatible with other geographic domains, such as postal codes, at which covariates and outcomes used for health analyses are aggregated. Spatial misalignment may then occur among participant health indicators assessed, for example, at individual level, socioeconomic covariates assessed at yet another geographic scale (e.g., census tract or county level), and exposure estimates evaluated at the grid level, leading to bias that can be difficult to quantify [32].

Point-based predictions may reduce the measurement error at the cost of requiring more elaborate ground-level monitoring and more precisely-calculated model inputs. Point-specific predictions also may provide larger exposure contrasts, boosting the power available to detect health effects. Additionally, point-level predictions at fine temporal resolutions enable calculation of more accurate time-weighted cumulative exposures for study participants who reside in multiple locations with disparate exposure levels during the course of the study.

Developments in model-based pollutant prediction: insights from MESA Air spatiotemporal model

To meet the epidemiological needs of modern cohort studies, advances in model development have tackled a number of pertinent challenges. These include dealing with the complex and irregular nature of input outcome and covariate data, providing finely-resolved estimates at a national scale, and providing accurate predictions for the times preceding availability of reliable, dense regulatory monitoring networks. Exemplifying with a number of epidemiological applications of the MESA Air spatiotemporal model, we discuss how statistical innovations, exposure monitoring improvements, and expanded GIS data availability have helped deal with these challenges. For each application, we also briefly outline alternative modeling approaches. The table shows the performance of a number of statistical models that provide nationwide coverage at fine spatial scale. For non-MESA models, we show performance metrics as reported by the various authors. It is difficult to compare models head-to-head because investigators use different performance statistics and spatial/temporal scales. Moreover, an ideal comparison would involve contrasting predictions from different models at a population-based set of locations across the country, which is not feasible.

Structure of the MESA Air spatiotemporal model

The MESA Air spatiotemporal model is an extension of universal kriging. It is assumed that pollutant concentrations exhibit systematic seasonal and secular trends that vary over space, and that the long-term averages and short-term amplitudes of these trends are predictable based on a combination of geographic characteristics and spatial smoothing [33]. The model can be summarized as

$$C(s,t) = \mu(s,t) + \upsilon(s,t)$$

where $C(s,t)$, the log-transformed two-week average concentration of a pollutant at location s and time t, is a function of a spatiotemporal "mean" air pollution surface [$\mu(s,t)$] and spatially correlated but temporally-independent kriged residuals [$\nu(s,t)$]. The spatiotemporal mean is constructed as a linear combination of temporal trends [$f_i(t)$] of pollution in a given region with spatially varying coefficients [$\beta_i(s)$] that describe the amplitudes of the temporal trends at different locations, and spatiotemporal covariates, if any are used

$$\mu(s,t) = \beta_0(s) + \sum_{i=1}^{m} \beta_i(s) f_i(t) + \sum_{l=1}^{L} \gamma_l \mathscr{M}_l(s,t)$$

where $\beta_0(s)$ represents the long-term mean at location s, while m is a small number of temporal trends derived from singular value decomposition of the long time series of monitored concentrations, and $\mathscr{M}_l$ are any spatiotemporal terms included, with coefficients γ_l. The first temporal basis function $f_0(t)$ is set to 1 and the rest have a mean of 0. The independent spatial surfaces at each time period are modeled with separate kriging models

$$\beta_i \sim N[\boldsymbol{X}_i(s)\boldsymbol{\alpha}_i,\ \boldsymbol{\Sigma}_i(\boldsymbol{\phi}_i)]$$

where $i = 0,\ldots,m$ and $\boldsymbol{X}_i(s)$ are geographic covariates which are typically dimension-reduced in practice using partial least squares (PLS) regression. The model estimates regression parameters of the geographic covariates (α_i), any spatiotemporal covariates used (γ_l), and covariance parameters for the β_i fields [$\Sigma_i(\phi_i)$], depending on the chosen geostatistical covariance structure. $\gamma_l = 0$ if no spatiotemporal covariates are included.

This framework leverages the LUR construct, including the constant time-averaged spatial field $f_0(t)$ to describe trends related to emission sources, population, land use, and near-source concentrations while simultaneously producing smooth, more realistic pollution surfaces owing to spatial smoothing. Validation strategies account for the complexity of the goals in order to, for example, provide an accurate representation of out-of-sample performance in regions that lack adequate monitor coverage [31].

The model's structure also illustrates its ability to provide predictions at fine temporal scales. Reliance on long-term data from reference-grade monitors implies that it accurately captures secular variations in pollutant concentration. While reference data are limited to one-in-three or one-in-six-day availability, diagnostics show that the model also captures fortnightly variations well. Daily data are available from other sources, including some ground monitors and satellite-derived aerosol optical depth (AOD), but these are typically not available at the fine spatial resolution at which this model operates. Overall, the model reliably yields accurate predictions at specific geolocations on an approximately semi-monthly scale.

Metropolitan region-specific spatiotemporal applications incorporating supplemental monitoring data (MESA Air spatiotemporal models)

Many models that use ground-based pollutant observations rely on routine monitors whose spatial sparsity may make them insufficient for fine-scale prediction. Additionally, these monitors are sited primarily for regulatory purposes, not specific epidemiological objectives. Most regulatory monitors for oxides of nitrogen in the US, for example, were until recently sited far from major roads [34]. To support empirically-driven fine-scale prediction, investigators have deployed focused supplemental monitoring schemes to increase the spatial density of available measurements while pursuing specific epidemiological goals

[35]. A study primarily interested in the health effects of traffic-related pollutants can focus supplemental monitoring on locations near to roadways [10]. Investigator-deployed supplemental monitors permit targeting locations of interest that may be underrepresented by regulatory monitors (e.g. homes and schools), capture exposure variations in places with mixed land use characteristics, or uncover exposure dynamics over spatial domains with complex gradients, such as along emission sources or leeward and windward sides of large buildings.

For cost and logistical reasons, supplemental monitoring has generally been constrained to non-synchronous campaigns with short timeframes, resulting in spatiotemporal misalignment with routine monitoring data. To address this, the approach to model fitting and estimation in the spatiotemporal model described above is designed to accommodate irregular pollutant measurements characterized by arbitrary missingness and varying lengths of temporal coverage [33]. This has allowed fine-scale spatiotemporal prediction models to be developed for multiple metropolitan regions where supplemental monitoring has been conducted. In MESA Air, three types of cohort-specific monitors were deployed: 1) "fixed" monitors in participant-dense locations not well-covered by regulatory agency monitors; 2) "home" monitors in about 10% of participant homes for two-week sampling during different seasons (approximately 100 locations in six study regions); and 3) "snapshot" samplers in clusters adjacent to major roadways, on utility poles at 50 m, 100 m, and 350 m from road edges in both directions, and at locales with varying population densities [35, 10]. This set of input data was used to predict concentrations of $PM_{2.5}$ and its chemical components, NO_2, NO_x, and black carbon at each participant's residence [30, 36]. The predictions have been used to gain insights into the role of air pollutants in subclinical cardiovascular disease, among other outcomes [37]. Because the models are easily portable to other pollutants, cohorts, and regions, they have been used in additional epidemiological applications, including assessment of the effect of ozone exposure on the risk of chronic obstructive pulmonary disease [38, 39]. The use of supplemental data to augment regulatory data provided significant improvements in prediction accuracy [40, 30].

Supplemental monitoring can be time-consuming and expensive, and is best deployed with a clear understanding of how it will be used in subsequent analysis. A fraction of supplemental monitors should be co-located with routine monitors for quality control and calibration purposes. Post-modeling validation should account for the complexity of data sources, and disentangle the spatial and temporal contributions of monitor types to overall accuracy in the context of complex spatiotemporal interactions [31].

The MESA Air model is based on spatially-varying temporal pollution processes. An alternative paradigm is to model temporally-varying spatial surfaces, estimating pollutant concentrations as a factor of spatial fixed-effects, smooth functions of time-varying covariates, and time-varying spatial residuals [41, 42]. This approach also allows modeling of complex spatiotemporal interactions and captures fine-scale spatial heterogeneity, although it relies on more complete observations both across time and space, uses a stepwise variable selection approach without dimension-reduction, and does not take advantage of supplemental monitoring. The spatial structure is represented using penalized splines rather than kriging [42].

An emerging alternative source of short-term supplemental data is mobile monitoring, which is especially suitable for ultrafine particles not typically measured by fixed monitors [43, 44]. An advantage of mobile monitoring is spatial extent, making it suitable for spatially heterogeneous pollutants and complex urban terrain. Conversely, mobile monitors do not optimally assess temporal variation, and many aspects of mobile sampling design, including number and location of monitoring sites, length of sampling time per site, and number of repeated samples per site require careful consideration to optimize for good performance and cost-effectiveness [45]. Combining mobile and stationary sources of supplemental data may also yield improved model performance [46]. As application of mobile monitoring-based predictions in epidemiological studies ramps up, additional rigorous evaluation is required to assess whether such exposure estimates bias effects relative to those from fixed monitoring locations.

National spatiotemporal model

While the MESA Air models were developed for specific communities under study, it is desirable to extend these models to the continental scale for application to other populations. Implementation at the national scale can be accomplished by extending statistical infrastructure that is highly effective at the scale of a metropolitan area to a nationwide scale, borrowing information from data-rich areas in order to produce estimates even for sparsely-populated regions with less monitoring infrastructure. To facilitate the extension, investigators divided the country into 9 climatic/topographic regions (for $PM_{2.5}$ model) or 3 regions (for ozone and NO_2), in order to account for sub-national region-specific pollution processes and ensure each region contained supplemental monitors. Simple smoothing was used at regional boundaries to avoid artificial discontinuities. Relative to the city-specific models described above, the national model is parameterized to account for a wider range of concentration time-series and local geographic characteristics by increasing the number of temporal trends and PLS scores. The nationwide version uses data from approximately 940 investigator-deployed non-regulatory monitors and 1,500 regulatory monitors as well as satellite-derived $PM_{2.5}$ and NO_2 measurements, and produces point-wise $PM_{2.5}$, ozone, and NO_2 concentrations between 1999–2017 across the contiguous United States at a two-week temporal time scale [26].

The performance of these models is also excellent, with cross-validated R^2 for $PM_{2.5}$ and NO_2 of 0.89 and 0.87, respectively, and some variation by region (spatial R^2 in the Southeast and Northwest regions were 0.91 and 0.71, respectively). Spatially-clustered cross-validated R^2, a measure of predictive performance at regions with little or no monitor coverage, was satisfactory at 0.77. Greater fine-scale variation was observed in models that additionally incorporated supplemental monitors, rather than relying solely on regulatory data. At best, modest increases in prediction accuracy were gained by including satellite-sourced data, primarily in areas with little ground monitoring coverage. National spatiotemporal models for NO_2 and ozone that employ similar principles are in development, and will additionally include CTM outputs.

Alternative $PM_{2.5}$, PM_{10} and coarse PM models based on spatial trends that vary temporally have also been extended for nationwide application [47]. They also take a regionalized approach, with spatial smoothing using penalized splines and a smaller set of location-specific geographic covariates that is not dimension-reduced, resulting in a more straightforward interpretation of covariate effects on exposure levels. Other recent fine-scale nationwide models are generally characterized by use of hybrid approaches that combine LUR with output from satellite AOD and CTM, rather than supplemental ground monitor data. They are further differentiated by additional analytical features, such as Bayesian Maximum Entropy kriging of residuals from LUR models [48], adjustment of remote sensing inputs with geographically-weighted regression [49], and outputs at finer temporal scales or estimates of urban versus rural model performance [50]. Deep learning neural networks have been proposed to handle complex nonlinear relationships among model inputs and measured concentrations [51]. Key features of another new hybrid implementation are described below [25, 52].

National spatiotemporal model for periods prior to dense regulatory monitoring (Historical national model)

The prediction models discussed above rely on measurements from regulatory monitoring programs and often leverage high resolution satellite data, both of which started being available only in the last two decades. In the US, the Environmental Protection Agency's (EPA) extensive monitoring infrastructure for $PM_{2.5}$ was established in 1999. This poses a challenge when exposure time of interest for many observational studies is before 1999. A common remedy is to use exposure estimates from more recent times relative to time of health outcome assessment. This relies on an assumption that more recent pollutant concentrations reflect concentrations at earlier times. While these sets of concentrations are likely correlated, there is the potential for both bias (given secular trends in several pollutant concentrations) and relative exposure misclassification when different areas change their ranking of concentrations over time. Other alternatives include use of whatever historical (e.g. pre-1999) concentration data are available, however sparse, or back-extrapolation of recently-estimated concentrations. While many LUR-based back-extrapolation efforts have suggested satisfactory temporal transferability of estimates, they have been of relatively limited spatial or temporal scope, focus mostly on oxides of nitrogen, and lack rigorous validation of resulting predictions [53–57]. A few recent national-scale studies suggest that for NO_2, back-extrapolation of estimates from between 2006–2009 explains approximately 73%−83% of the variation in concentrations going back to 1990 [58, 59].

Based on the model structure above, a comprehensive spatiotemporal historical prediction model for $PM_{2.5}$ spanning the contiguous United States for the period 1980 to 2010 was developed by first estimating the concentration trend from data for 1999–2010, then performing linear back-extrapolation [57]. Alternative back-extrapolation approaches were assessed, including estimation of the historical trend based on proxies such as pre-1999 visibility data or $PM_{2.5}$ sulfate data. Validation with numerous contemporaneous external datasets suggested that the linear trend extrapolation technique performed best, with spatial and temporal R^2 ranging between 0.77–0.87 and 0.55–0.58, respectively, depending on type of monitoring site.

Spatiotemporal models incorporating satellite and chemical transport model output as covariates (Spatiotemporal models for Los Angeles, incorporating CTM output)

Nearly 30% of people in the US live more than 20 km from a $PM_{2.5}$ monitor [60], and existing monitors may not be located close enough to pollutant sources of interest to allow precise determination of fine-scale dispersion profiles. Because of sparsity of ground-level monitors, satellite-derived data and outputs from CTM are increasingly incorporated during fine-scale prediction. CTMs such as CMAQ and GEOS-Chem simulate atmospheric physical and chemical processes, taking into account topographic, meteorological and emission source data in order to estimate outdoor pollutant levels [22, 19]. They may facilitate better representation of physicochemical atmospheric characteristics relevant to pollutant concentrations but not encapsulated in land use variables. Satellite-derived AOD is a measure of visibility that, when properly calibrated with regulatory monitors, can serve as a proxy for ground-level pollution.

Because the above model can accommodate spatiotemporal covariates, CTM output was incorporated to predict point-level concentrations of ozone and $PM_{2.5}$ in Los Angeles [27]. The model combined (separately for each pollutant) regulatory data from 25 $PM_{2.5}$ and 37 ozone monitors, more than 100 supplemental ground monitors in the metropolitan region, and 4km x 4km output from the University of California Davis-California Institute of Technology source-oriented CTM [61]. The model also included nearly 200 geographic variables and Caline3QHCR dispersion model output as spatial covariates. It demonstrated better prediction accuracy (lower mean square errors) and precision (higher R^2 values) compared to CTM-only and universal kriging-only alternatives, particularly for ozone in rural locales. The contribution of CTM to improvements in the $PM_{2.5}$ model was more modest. This model incorporates outputs from dispersion, CTM or satellite products as either spatial and spatiotemporal covariates, in contrast to alternatives that fuse estimates from different sources together in hybrid implementations [52].

Supplementing LUR models with CTM and AOD data can improve prediction accuracy by providing a signal where ground-level monitoring is inadequate. Additionally, because land use characteristics typically vary slowly over time and GIS databases are only updated periodically, CTM and AOD inputs provide more variable spatiotemporal coverage. Many contemporary models, especially those covering supra-national scales, combine ground monitor data with CTM and/or AOD measurements [62, 21], although performance improvements have been reported even at sub-national domains [63]. However, CTM and AOD data are typically available at relatively coarse resolutions, suggesting that for purposes of improving fine-scale performance, it may be especially helpful to integrate them with spatially-dense ground monitoring data. Furthermore, recent health-effect studies suggest that using only satellite-derived exposure measures of $PM_{2.5}$ may underestimate risks, supporting incorporation of ground monitoring data for epidemiological studies [64].

While prediction models for $PM_{2.5}$ based on satellite data are appealing, it is not clear that they improve fine-scale long-term prediction of $PM_{2.5}$ in the US [26, 65], because models based on ground observations already perform very well, while satellite observations are only available at coarse spatial scales and are characterized by extensive missingness. Approximately 70% of AOD data are missing due to cloud and snow cover and surface brightness [66]. Much like regulatory ground monitors, satellite AOD is not primarily designed to support epidemiologically-motivated estimation of surface pollutant levels, and its relationship to pollutant concentrations is influenced by a complex of features that is only partially-understood [67, 68], necessitating extensive calibration, ideally on a daily and regional basis [20, 25]. Notwithstanding major progress in AOD satellite data resolution, calibration, and imputation, active ground monitoring will continue to be critical for fine-scale prediction of $PM_{2.5}$. Satellite data has a clear role for understanding $PM_{2.5}$ concentrations in areas of the world with less ground-level monitoring [21]. Similarly, CTM are independently useful, especially in understanding the spatial and temporal relationships between emissions from various sources and subsequent pollutant concentrations [69].

A recent alternative with nationwide application is a hybrid model that relies instead on a neural network to account for complex spatiotemporal nonlinearities and interactions among multiple inputs, including satellite AOD, surface reflectance, CTMs (Geos-CHEM and CMAQ), land use and vegetation cover, and meteorological data [25, 52]. The model uses convolutional layers to aggregate nearby space-time information and account for spatiotemporal autocorrelation. It has recently been used to produce daily national predictions of $PM_{2.5}$ and ozone at a $1km^2$ resolution, for the 2000–2012 period, with overall R^2 values of 0.84 and 0.80, respectively. The model accounts for complex atmospheric processes, has a fine temporal resolution, and the ability to deal naturally with non-linearities. On the other hand, the $1km^2$ resolution introduces possibility of exposure misclassification due to spatial misalignment of participant residence and grid-level exposure. The neural network is trained on data from regulatory monitors, which underrepresent rural expanses and some urban locales, potentially introducing geographically-systematic prediction errors. While overall spatial R^2 is very good, region-specific temporal performance may be more modest, especially in areas with little monitoring coverage.

National spatial models using regulatory monitoring data (National spatial models)

In chronic disease applications, exposure assessment is often focused on spatial contrasts in annual or multi-year average concentrations. A spatial model can be formulated as

$$C(s) = \beta_0(s) + \sum_{i=1}^{n} \beta_i(s) X_i(s) + \sum_{j=1}^{J} \lambda_j(s) \mathscr{F}_j(s) + \varepsilon(s)$$

where $C(s)$ is the log-transformed annual average pollutant concentration at location s, while $X_i(s)$ are PLS components derived from geographic covariates, $\mathscr{F}_j(s)$ are any satellite-derived variables included directly as spatial covariates instead of being fused with geographic variables that generate PLS scores, and $\varepsilon(s)$ are spatially-varying residuals modeled with universal kriging [70, 71].

For national coverage, each region is usually modeled separately. The rationale for regionalized modeling is from findings showing that impressive large-scale spatial performance can mask sub-par predictive characteristics when interrogated at smaller scale [70]. Based on a “pragmatic” approach to dealing with spatial heterogeneity in model structure, modeling $PM_{2.5}$ at the level of sub-national regions characterized by relative geographic homogeneity provides improved prediction accuracy. At a national or continental scale, covariates that seem to define similar characteristics may not have the same implications for exposure levels in all regions. For instance, the impact of traffic density may depend on regionally-varying fleet mixes, and similarly classified roadway types may have far more traffic in urban versus rural areas. In a similarly-designed NO_2 national model, addition of satellite-derived covariates significantly improved prediction performance for locations distant from ground monitors, but only marginally in those with more adequate monitoring coverage [71].

Other models have also shown that satellite data can improve model performance compared to using ground-based measures only [72, 73]. Most national spatial models have focused on predicting NO_2, PM_{10}, and $PM_{2.5}$, with performance ranging from modest to very good (R^2 between 0.22–0.90). A number of other recent fine-scale spatial models provide point-wise predictions for multiple sub-national regions or entire countries based on LUR and universal kriging [74, 24, 75]. Others additionally incorporate satellite-derived data and give predictions resolved to a small grid (up to 100m) [76–78].

Future directions, challenges, and extensions

The next generation of fine-scale exposure models needs to respond to evolving contemporary challenges in air pollution epidemiology, including the need to understand sources of pollutants to better target future exposure control efforts and to incorporate new types of information that can improve prediction accuracy. An additional imperative is the development of openly available, user-friendly modeling platforms to foster reproducible research given the heightened regulatory and public scrutiny to which environmental epidemiology research is now subject.

Methods have recently been developed to assess regional clustering of $PM_{2.5}$ components [79]. Investigators have also published software for implementing comprehensive fine-scale models on publicly accessible venues and in programming languages widely used by environmental scientists and epidemiologists. Examples are *SpatioTemporal* [80] and *airpred* [81], packages for implementing fine-scale national models already used in major epidemiological analyses [37, 2, 82].

Key challenges to optimizing modeling approaches remain. In the US, nationally-standardized long-term high quality traffic volume and congestion data are not yet available, leaving modeling efforts to rely on time-static road-network data or other proxies, which are incomplete or otherwise inadequate. Furthermore, even the best outdoor concentration model will imperfectly reflect an individual's exposure. Most people spend substantial fractions of time indoors, where exposures can be dramatically different from outdoor concentrations.

It is important to note that the current generation of models are trained on data from well-monitored areas. The result is potentially large prediction errors in data-poor areas and little applicability to rural areas in high-income countries and to low- and middle-income countries. The advent of “low-cost” monitors may play a big role in expanding the spatial density of sources for fine-scale modeling. Work is ongoing to address a number of well-recognized limitations of low-cost monitors, including reliability and accuracy of measurements [83, 84]

Conclusions

The continued evolution of models for fine-scale prediction of pollutant concentrations has improved the scientific evaluation of health effects of air pollutants. Models now exist that leverage a wide array of land use, satellite, meteorological, and physicochemical inputs to provide small-grid or pointwise predictions of key air pollutants across the contiguous United States and elsewhere, taking advantage of supplemental monitoring campaigns while accounting for complex spatiotemporal dynamics. Machine learning algorithms may offer an alternative approach for complex spatiotemporal correlations. Major challenges remain, including the computational burden of increasingly sophisticated models, worse prediction accuracy at places with little available ground-level input data, and a lack of comprehensive traffic data. High quality, fine-scale, nation-wide modeling also needs to be expanded to more countries, more pollutants — including ultrafine particulate and non-criteria hazardous air pollutants — and an understanding of the sources and components of major pollutants.

Sources of Support

This publication was developed under a STAR research assistance agreements RD831697 (MESA Air), RD-83830001 (MESA Air Next Stage), and RD83479601 (UW Center for Clean Air Research), awarded by the U.S. Environmental Protection Agency. It has not been formally reviewed by the EPA. The views expressed in this document are solely those of the authors and the EPA does not endorse any products or commercial services mentioned in this publication. Research reported in this publication was also supported by the University of Washington EDGE Center of the NIA under award number: P30ES007033, by ECHO PATHWAYS (NIH grants: 1UG3OD023271-01 and 4UH3OD023271-03), and by grants R56ES026528 and P30ES007033 from NIEHS and R01ES026187 from NIA and NIEHS. This work was supported in part by the UW NIEHS sponsored Biostatistics, Epidemiologic and Bioinformatics Training in Environmental Health (BEBTEH) Training Grant, Grant #: NIEHS T32ES015459.

References

1. Brook RD, Rajagopalan S, Pope CAI, Brook JR, Bhatnagar A, Diez-Roux AV et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. Circulation. 2010;121(21):2331–78. doi:10.1161/CIR.0b013e3181dbece1. [PubMed: 20458016]
2. Di Q, Wang Y, Zanobetti A, Wang Y, Koutrakis P, Choirat C et al. Air Pollution and Mortality in the Medicare Population. N Engl J Med 2017;376(26):2513–22. doi:10.1056/NEJMoa1702747. [PubMed: 28657878]
3. Beelen R, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, Hoffmann B et al. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. Lancet. 2014;383(9919):785–95. doi:10.1016/S0140-6736(13)62158-3. [PubMed: 24332274]
4. Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. Lancet. 2017;389(10082):1907–18. doi:10.1016/S0140-6736(17)30505-6. [PubMed: 28408086]
5. Dockery DW, Pope CA 3rd, Xu X, Spengler JD, Ware JH, Fay ME et al. An association between air pollution and mortality in six U.S. cities. N Engl J Med 1993;329(24):1753–9. doi:10.1056/nejm199312093292401. [PubMed: 8179653]

6. Pope CA 3rd, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE et al. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. Am J Respir Crit Care Med 1995;151(3 Pt 1):669–74. doi:10.1164/ajrccm/151.3_Pt_1.669. [PubMed: 7881654]
7. Hoek G. Methods for Assessing Long-Term Exposures to Outdoor Air Pollutants. Curr Environ Health Rep 2017;4(4):450–62. doi:10.1007/s40572-017-0169-5. [PubMed: 29064065]
8. Berrocal VJ, Guan Y, Muyskens A, Wang H, Reich BJ, Mulholland JA et al. A comparison of statistical and machine learning methods for creating national daily maps of ambient PM2.5 concentration. Atmospheric Environment. 2020;222:117130. doi:10.1016/j.atmosenv.2019.117130. [PubMed: 32863727]
9. Sorek-Hamer M, Just AC, Kloog I. Satellite remote sensing in epidemiological studies. Current opinion in pediatrics. 2016;28(2):228–34. doi:10.1097/mop.0000000000000326. [PubMed: 26859287]
10. Kaufman JD, Adar SD, Allen RW, Barr RG, Budoff MJ, Burke GL et al. Prospective study of particulate air pollution exposures, subclinical atherosclerosis, and clinical cardiovascular disease: The Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). Am J Epidemiol 2012;176(9):825–37. doi:10.1093/aje/kws169. [PubMed: 23043127]
11. Szpiro AA, Paciorek CJ. Measurement error in two-stage analyses, with application to air pollution epidemiology. Environmetrics. 2013;24(8):501–17. doi:10.1002/env.2233. [PubMed: 24764691]
12. Vlaanderen J, Portengen L, Chadeau-Hyam M, Szpiro A, Gehring U, Brunekreef B et al. Error in air pollution exposure model determinants and bias in health estimates. Journal of Exposure Science & Environmental Epidemiology. 2019;29(2):258–66. doi:10.1038/s41370-018-0045-x. [PubMed: 29880834]
13. Butland BK, Samoli E, Atkinson RW, Barratt B, Katsouyanni K. Measurement error in a multi-level analysis of air pollution and health: a simulation study. Environmental Health. 2019;18(1):13. doi:10.1186/s12940-018-0432-8. [PubMed: 30764837]
14. Eeftens M, Tsai M-Y, Ampe C, Anwander B, Beelen R, Bellander T et al. Spatial variation of PM2.5, PM10, PM2.5 absorbance and PM coarse concentrations between and within 20 European study areas and the relationship with NO2 – Results of the ESCAPE project. Atmospheric Environment. 2012;62:303–17. doi:10.1016/j.atmosenv.2012.08.038.
15. Miller KA, Siscovick DS, Sheppard L, Shepherd K, Sullivan JH, Anderson GL et al. Long-term exposure to air pollution and incidence of cardiovascular events in women. N Engl J Med 2007;356(5):447–58. doi:10.1056/NEJMoa054409. [PubMed: 17267905]
16. Brunekreef B, Hoffmann B. Air pollution and heart disease. The Lancet. 2016;388(10045):640–2. doi:10.1016/S0140-6736(16)30375-0.
17. United States Environmental Protection Agency. SLAMS/NAMS/PAMS Network Review Guidance: Regulatory Requirements and Other Data Needs. Research Triangle Park, DC: US EPA; 1998.
18. Hoek G, Beelen R, de Hoogh K, Vienneau D, Gulliver J, Fischer P et al. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmos Environ 2008;42:7561–78.
19. Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahsuvaroglu T et al. A review and evaluation of intraurban air pollution exposure models. J Expo Anal Environ Epidemiol 2005;15(2):185–204. doi:10.1038/sj.jea.7500388. [PubMed: 15292906]
20. Streets DG, Canty T, Carmichael GR, de Foy B, Dickerson RR, Duncan BN et al. Emissions estimation from satellite retrievals: A review of current capability. Atmospheric Environment. 2013;77:1011–42. doi:10.1016/j.atmosenv.2013.05.051.
21. van Donkelaar A, Martin RV, Brauer M, Boys BL. Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter. Environmental health perspectives. 2015;123(2):135–43. doi:10.1289/ehp.1408646. [PubMed: 25343779]
22. Bey I, Jacob DJ, Yantosca RM, Logan JA, Field BD, Fiore AM et al. Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. Journal of Geophysical Research: Atmospheres. 2001;106(D19):23073–95. doi:10.1029/2001jd000807.

23. Beevers SD, Kitwiroon N, Williams ML, Kelly FJ, Ross Anderson H, Carslaw DC. Air pollution dispersion models for human exposure predictions in London. J Expo Sci Environ Epidemiol 2013;23(6):647–53. doi:10.1038/jes.2013.6. [PubMed: 23443237]
24. Hart JE, Yanosky JD, Puett RC, Ryan L, Dockery DW, Smith TJ et al. Spatial modeling of PM10 and NO2 in the continental United States, 1985–2000. Environ Health Perspect 2009;117(11):1690–6. doi:10.1289/ehp.0900840. [PubMed: 20049118]
25. Di Q, Kloog I, Koutrakis P, Lyapustin A, Wang Y, Schwartz J. Assessing PM2.5 Exposures with High Spatiotemporal Resolution across the Continental United States. Environ Sci Technol 2016;50(9):4712–21. doi:10.1021/acs.est.5b06121. [PubMed: 27023334]
26. Wang M, Sampson P, Bechle M, Marshall J, Vedal S, Kaufman J. National PM2. 5 and NO2 Spatiotemporal Models Integrating Intensive Monitoring Data and Satellite-Derived Land Use Regression in a Universal Kriging Framework in the United States: 1999–2016. ISEE Conference Abstracts; Ottawa, Canada2018.
27. Wang M, Sampson PD, Hu J, Kleeman M, Keller JP, Olives C et al. Combining Land-Use Regression and Chemical Transport Modeling in a Spatiotemporal Geostatistical Model for Ozone and PM2.5. Environ Sci Technol 2016;50(10):5111–8. doi:10.1021/acs.est.5b06001. [PubMed: 27074524]
28. Di Q, Koutrakis P, Schwartz J. A hybrid prediction model for PM2.5 mass and components using a chemical transport model and land use regression. Atmospheric Environment. 2016;131:390–9. doi:10.1016/j.atmosenv.2016.02.002.
29. Mercer LD, Szpiro AA, Sheppard L, Lindström J, Adar SD, Allen RW et al. Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NOx) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). Atmospheric environment (Oxford, England : 1994). 2011;45(26):4412–20. doi:10.1016/j.atmosenv.2011.05.043.
30. Keller JP, Olives C, Kim SY, Sheppard L, Sampson PD, Szpiro AA et al. A unified spatiotemporal modeling approach for predicting concentrations of multiple air pollutants in the multi-ethnic study of atherosclerosis and air pollution. Environ Health Perspect 2015;123(4):301–9. doi:10.1289/ehp.1408145. [PubMed: 25398188]
31. Lindstrom J, Szpiro AA, Sampson PD, Oron AP, Richards M, Larson TV et al. A Flexible Spatio-Temporal Model for Air Pollution with Spatial and Spatio-Temporal Covariates. Environ Ecol Stat 2014;21(3):411–33. doi:10.1007/s10651-013-0261-4. [PubMed: 25264424]
32. Keller JP, Peng RD. Error in estimating area-level air pollution exposures for epidemiology. Environmetrics. 2019;30(8):e2573. doi:10.1002/env.2573.
33. Sampson PD, Szpiro AA, Sheppard L, Lindström J, Kaufman JD. Pragmatic estimation of a spatio-temporal air quality model with irregular monitoring data. Atmos Environ 2011;Vol. 45(No. 36):6593–606.
34. Federal Register. Revision to the near-road NO2: A rule by the Environmental protection Agency (Document Citation: 81 FR 96381). Washington, DC. 2016. https://www.federalregister.gov/documents/2016/12/30/2016-31645/revision-to-the-near-road-no2. Accessed May 23, 2020.
35. Cohen MA, Adar SD, Allen RW, Avol E, Curl CL, Gould T et al. Approach to estimating participant pollutant exposures in the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). Environ Sci Technol 2009;43(13):4687–93. [PubMed: 19673252]
36. Kim S-Y, Sheppard L, Bergen S, Szpiro AA, Sampson PD, Kaufman JD et al. Prediction of fine particulate matter chemical components with a spatio-temporal model for the Multi-Ethnic Study of Atherosclerosis cohort. Journal of Exposure Science & Environmental Epidemiology. 2016;26(5):520–8. doi:10.1038/jes.2016.29. [PubMed: 27189258]
37. Kaufman JD, Adar SD, Barr RG, Budoff M, Burke GL, Curl CL et al. Association between air pollution and coronary artery calcification within six metropolitan areas in the USA (the Multi-Ethnic Study of Atherosclerosis and Air Pollution): a longitudinal cohort study. The Lancet. 2016;388(10045):696–704. doi:10.1016/S0140-6736(16)00378-0.
38. Hansel NN, Paulin LM, Gassett AJ, Peng RD, Alexis N, Fan VS et al. Design of the Subpopulations and Intermediate Outcome Measures in COPD (SPIROMICS) AIR Study. BMJ Open Respir Res 2017;4(1):e000186. doi:10.1136/bmjresp-2017-000186.

39. Paulin LM, Gassett AJ, Alexis NE, Kirwa K, Kanner RE, Peters S et al. Association of Long-term Ambient Ozone Exposure With Respiratory Morbidity in Smokers. JAMA Internal Medicine. 2020;180(1):106–15. doi:10.1001/jamainternmed.2019.5498. [PubMed: 31816012]
40. Szpiro AA, Sampson PD, Sheppard L, Lumley T, Adar SD, Kaufman J. Predicting Intra-Urban Variation in Air Pollution Concentrations with Complex Spatio-Temporal Dependencies. Environmetrics. 2009;21(6):606–31. doi:10.1002/env.1014. [PubMed: 24860253]
41. Paciorek CJ, Yanosky JD, Puett RC, Laden F, Suh HH. Practical Large-Scale Spatio-Temporal Modeling of Particulate Matter Concentrations. Ann Appl Stat 2009;3(1):370–97.
42. Yanosky JD, Paciorek CJ, Suh HH. Predicting chronic fine and coarse particulate exposures using spatiotemporal models for the Northeastern and Midwestern United States. Environ Health Perspect 2009;117(4):522–9. doi:10.1289/ehp.11692. [PubMed: 19440489]
43. Hatzopoulou M, Valois MF, Levy I, Mihele C, Lu G, Bagg S et al. Robustness of Land-Use Regression Models Developed from Mobile Air Pollutant Measurements. Environmental Science & Technology. 2017;51(7):3938–47. doi:10.1021/acs.est.7b00366. [PubMed: 28241115]
44. Jones RR, Hoek G, Fisher JA, Hasheminassab S, Wang D, Ward MH et al. Land use regression models for ultrafine particles, fine particles, and black carbon in Southern California. Science of The Total Environment. 2020;699:134234. doi:10.1016/j.scitotenv.2019.134234.
45. Saha PK, Li HZ, Apte JS, Robinson AL, Presto AA. Urban Ultrafine Particle Exposure Assessment with Land-Use Regression: Influence of Sampling Strategy. Environmental Science & Technology. 2019;53(13):7326–36. doi:10.1021/acs.est.9b02086.
46. Tessum MW, Larson T, Gould TR, Simpson CD, Yost MG, Vedal S. Mobile and Fixed-Site Measurements To Identify Spatial Distributions of Traffic-Related Pollution Sources in Los Angeles. Environmental Science & Technology. 2018;52(5):2844–53. doi:10.1021/acs.est.7b04889. [PubMed: 29382190]
47. Yanosky JD, Paciorek CJ, Laden F, Hart JE, Puett RC, Liao D et al. Spatio-temporal modeling of particulate air pollution in the conterminous United States using geographic and meteorological predictors. Environmental Health. 2014;13(1):63. doi:10.1186/1476-069X-13-63. [PubMed: 25097007]
48. Beckerman BS, Jerrett M, Serre M, Martin RV, Lee S-J, van Donkelaar A et al. A Hybrid Approach to Estimating National Scale Spatiotemporal Variability of PM2.5 in the Contiguous United States. Environmental Science & Technology. 2013;47(13):7233–41. doi:10.1021/es400039u. [PubMed: 23701364]
49. van Donkelaar A, Martin RV, Spurr RJ, Burnett RT. High-Resolution Satellite-Derived PM2.5 from Optimal Estimation and Geographically Weighted Regression over North America. Environ Sci Technol 2015;49(17):10482–91. doi:10.1021/acs.est.5b02076. [PubMed: 26261937]
50. Bechle MJ, Millet DB, Marshall JD. National Spatiotemporal Exposure Surface for NO2: Monthly Scaling of a Satellite-Derived Land-Use Regression, 2000–2010. Environmental Science & Technology. 2015;49(20):12297–305. doi:10.1021/acs.est.5b02882. [PubMed: 26397123]
51. Li T, Shen H, Yuan Q, Zhang X, Zhang L. Estimating Ground-Level PM2.5 by Fusing Satellite and Station Observations: A Geo-Intelligent Deep Learning Approach. Geophysical Research Letters. 2017;44(23):11,985–11,93. doi:10.1002/2017gl075710.
52. Di Q, Rowland S, Koutrakis P, Schwartz J. A hybrid model for spatially and temporally resolved ozone exposures in the continental United States. Journal of the Air & Waste Management Association (1995). 2017;67(1):39–52. doi:10.1080/10962247.2016.1200159. [PubMed: 27332675]
53. Chen H, Goldberg MS, Crouse DL, Burnett RT, Jerrett M, Villeneuve PJ et al. Back-extrapolation of estimates of exposure from current land-use regression models. Atmospheric Environment. 2010;44(35):4346–54. doi:10.1016/j.atmosenv.2010.07.061.
54. Wang R, Henderson SB, Sbihi H, Allen RW, Brauer M. Temporal stability of land use regression models for traffic-related air pollution. Atmospheric Environment. 2013;64:312–9. doi:10.1016/j.atmosenv.2012.09.056.
55. Eeftens M, Beelen R, Fischer P, Brunekreef B, Meliefste K, Hoek G. Stability of measured and modelled spatial contrasts in NO₂ over time. Occupational and Environmental Medicine. 2011;68(10):765. doi:10.1136/oem.2010.061135. [PubMed: 21285243]

56. Cesaroni G, Porta D, Badaloni C, Stafoggia M, Eeftens M, Meliefste K et al. Nitrogen dioxide levels estimated from land use regression models several years apart and association with mortality in a large cohort study. Environmental Health. 2012;11(1):48. doi:10.1186/1476-069X-11-48. [PubMed: 22808928]

57. Kim S-Y, Olives C, Sheppard L, Sampson PD, Larson TV, Keller JP et al. Historical Prediction Modeling Approach for Estimating Long-Term Concentrations of PM2.5 in Cohort Studies before the 1999 Implementation of Widespread Monitoring. Environmental Health Perspectives. 2017;125(1):38–46. doi:doi:10.1289/EHP131. [PubMed: 27340825]

58. Gulliver J, de Hoogh K, Hoek G, Vienneau D, Fecht D, Hansell A. Back-extrapolated and year-specific NO2 land use regression models for Great Britain - Do they yield different exposure assessment? Environment International. 2016;92–93:202–9. doi:10.1016/j.envint.2016.03.037. [PubMed: 26562560]

59. Knibbs LD, Coorey CP, Bechle MJ, Marshall JD, Hewson MG, Jalaludin B et al. Long-term nitrogen dioxide exposure assessment using back-extrapolation of satellite-based land-use regression models for Australia. Environmental Research. 2018;163:16–25. doi:10.1016/j.envres.2018.01.046. [PubMed: 29421169]

60. Dominici F, Schwartz J, Di Q, Braun D, Choirat C, Zanobetti A. Assessing Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution: Phase 1. Research Report 200. Boston, MA: Health Effects Institute; 2019.

61. Hu J, Zhang H, Ying Q, Chen SH, Vandenberghe F, Kleeman MJ. Long-term particulate matter modeling for health effect studies in California – Part 1: Model performance on temporal and spatial variations. Atmos Chem Phys 2015;15(6):3445–61. doi:10.5194/acp-15-3445-2015.

62. Meng J, Li C, Martin RV, van Donkelaar A, Hystad P, Brauer M. Estimated Long-Term (1981–2016) Concentrations of Ambient Fine Particulate Matter across North America from Chemical Transport Modeling, Satellite Remote Sensing, and Ground-Based Measurements. Environmental Science & Technology. 2019;53(9):5071–9. doi:10.1021/acs.est.8b06875. [PubMed: 30995030]

63. Akita Y, Baldasano JM, Beelen R, Cirach M, de Hoogh K, Hoek G et al. Large Scale Air Pollution Estimation Method Combining Land Use Regression and Chemical Transport Modeling in a Geostatistical Framework. Environmental Science & Technology. 2014;48(8):4452–9. doi:10.1021/es405390e. [PubMed: 24621302]

64. Jerrett M, Turner MC, Beckerman BS, Pope CA, Donkelaar Av, Martin RV et al. Comparing the Health Effects of Ambient Particulate Matter Estimated Using Ground-Based versus Remote Sensing Exposure Estimates. Environmental Health Perspectives. 2017;125(4):552–9. doi:doi:10.1289/EHP575. [PubMed: 27611476]

65. Paciorek CJ, Liu Y. Limitations of remotely sensed aerosol as a spatial proxy for fine particulate matter. Environ Health Perspect 2009;117(6):904–9. doi:10.1289/ehp.0800360. [PubMed: 19590681]

66. Belle JH, Liu Y. Evaluation of Aqua MODIS Collection 6 AOD Parameters for Air Quality Research over the Continental United States. Remote Sensing. 2016;8(10):815.

67. Belle JH, Chang HH, Wang Y, Hu X, Lyapustin A, Liu Y. The Potential Impact of Satellite-Retrieved Cloud Parameters on Ground-Level PM2.5 Mass and Composition. International Journal of Environmental Research and Public Health. 2017;14(10):1244.

68. Kumar N. What Can Affect AOD–PM2.5 Association? Environmental Health Perspectives. 2010;118(3):A109–A10. doi:doi:10.1289/ehp.0901732.

69. Matthias V, Arndt JA, Aulinger A, Bieser J, Denier van der Gon H, Kranenburg R et al. Modeling emissions for three-dimensional atmospheric chemistry transport models. Journal of the Air & Waste Management Association. 2018;68(8):763–800. doi:10.1080/10962247.2018.1424057. [PubMed: 29364776]

70. Sampson PD, Richards M, Szpiro AA, Bergen S, Sheppard L, Larson TV et al. A regionalized national universal kriging model using Partial Least Squares regression for estimating annual PM2.5 concentrations in epidemiology. Atmospheric Environment. 2013;75:383–92. doi:10.1016/j.atmosenv.2013.04.015. [PubMed: 24015108]

71. Young MT, Bechle MJ, Sampson PD, Szpiro AA, Marshall JD, Sheppard L et al. Satellite-Based NO2 and Model Validation in a National Prediction Model Based on Universal Kriging and Land-

Use Regression. Environ Sci Technol 2016;50(7):3686–94. doi:10.1021/acs.est.5b05099. [PubMed: 26927327]

72. Hoek G, Eeftens M, Beelen R, Fischer P, Brunekreef B, Boersma KF et al. Satellite NO2 data improve national land use regression models for ambient NO2 in a small densely populated country. Atmospheric Environment. 2015;105:173–80. doi:10.1016/j.atmosenv.2015.01.053.
73. Bechle MJ, Millet DB, Marshall JD. Remote sensing of exposure to NO2: Satellite versus ground-based measurement in a large urban area. Atmospheric Environment. 2013;69:345–53. doi:10.1016/j.atmosenv.2012.11.046.
74. Kim S-Y, Song I. National-scale exposure prediction for long-term concentrations of particulate matter and nitrogen dioxide in South Korea. Environmental Pollution. 2017;226:21–9. doi:10.1016/j.envpol.2017.03.056. [PubMed: 28399503]
75. Kim S-Y, Bechle M, Hankey S, Sheppard L, Szpiro AA, Marshall JD. Concentrations of criteria pollutants in the contiguous U.S., 1979 – 2015: Role of prediction model parsimony in integrated empirical geographic regression. PLOS ONE. 2020;15(2):e0228535. doi:10.1371/journal.pone.0228535. [PubMed: 32069301]
76. Hystad P, Setton E, Cervantes A, Poplawski K, Deschenes S, Brauer M et al. Creating national air pollution models for population exposure assessment in Canada. Environ Health Perspect 2011;119(8):1123–9. doi:10.1289/ehp.1002976. [PubMed: 21454147]
77. Vienneau D, de Hoogh K, Bechle MJ, Beelen R, van Donkelaar A, Martin RV et al. Western European Land Use Regression Incorporating Satellite- and Ground-Based Measurements of NO2 and PM10. Environmental Science & Technology. 2013;47(23):13555–64. doi:10.1021/es403089q. [PubMed: 24156783]
78. Novotny EV, Bechle MJ, Millet DB, Marshall JD. National Satellite-Based Land-Use Regression: NO2 in the United States. Environmental Science & Technology. 2011;45(10):4407–14. doi:10.1021/es103578x. [PubMed: 21520942]
79. Keller JP, Drton M, Larson T, Kaufman JD, Sandler DP, Szpiro AA. COVARIATE-ADAPTIVE CLUSTERING OF EXPOSURES FOR AIR POLLUTION EPIDEMIOLOGY COHORTS. Ann Appl Stat 2017;11(1):93–113. doi:10.1214/16-AOAS992. [PubMed: 28572869]
80. Lindstrom J, Szpiro AA, Sampson PD, Bergen S, Sheppard L. SpatioTemporal: An R Package for Spatio-Temporal Modeling of Air Pollution. R Package Version 1.1.7. 2013. http://cran.r-project.org/web/packages/SpatioTemporal/index.html. Accessed November 19, 2018.
81. Sabath M, Di Q, Schwartz J, Braun D, Dominici F, Choirat C, editors. aipred: A flexible R package implementing methods for predicting air pollution. IEEE International Conference on Data Science and Advanced Analytics; 2018; Turin, Italy: IEEE.
82. Wang M, Aaron CP, Madrigano J, Hoffman EA, Angelini E, Yang J et al. Association Between Long-term Exposure to Ambient Air Pollution and Change in Quantitatively Assessed Emphysema and Lung Function. JAMA. 2019;322(6):546–56. doi:10.1001/jama.2019.10255. [PubMed: 31408135]
83. Bi J, Wildani A, Chang HH, Liu Y. Incorporating Low-Cost Sensor Measurements into High-Resolution PM2.5 Modeling at a Large Spatial Scale. Environmental Science & Technology. 2020;54(4):2152–62. doi:10.1021/acs.est.9b06046. [PubMed: 31927908]
84. Zusman M, Schumacher CS, Gassett AJ, Spalt EW, Austin E, Larson TV et al. Calibration of low-cost particulate matter sensors: Model development for a multi-city epidemiological study. Environment International. 2020;134:105329. doi:10.1016/j.envint.2019.105329. [PubMed: 31783241]

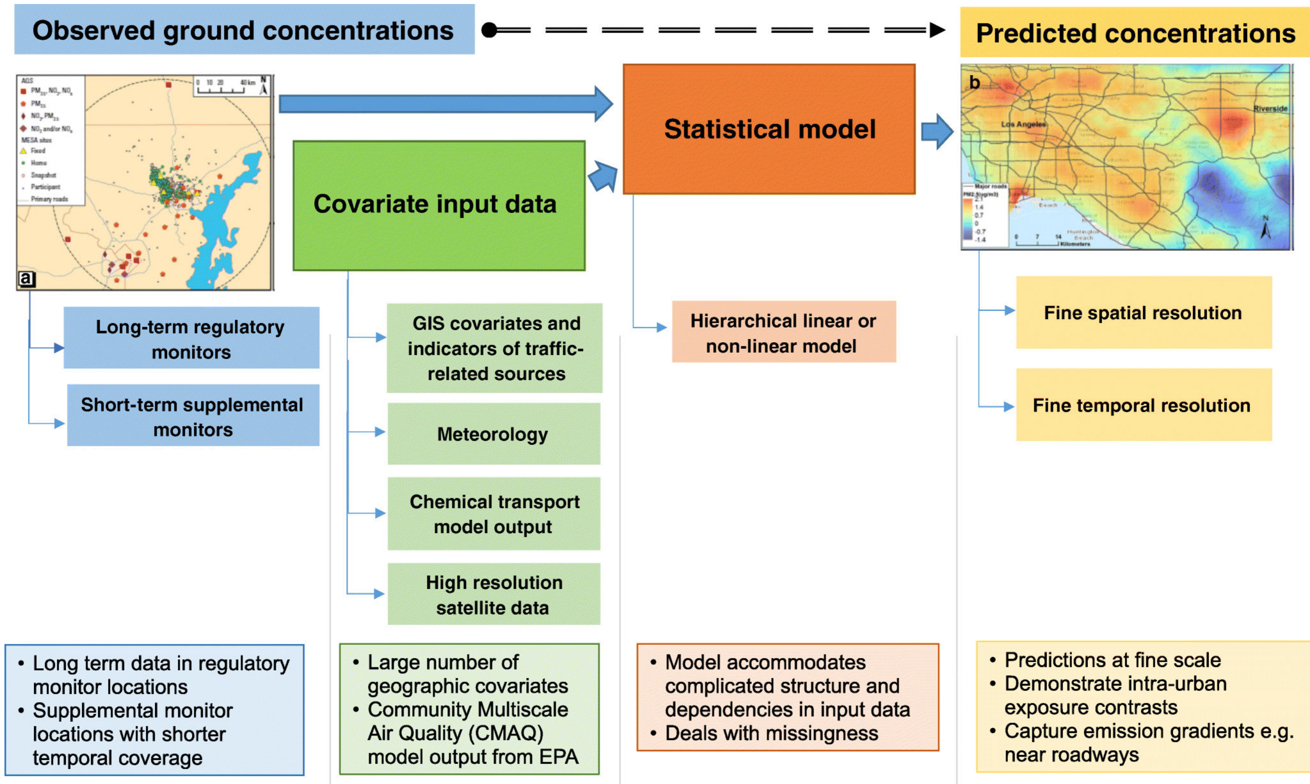

Figure.

Schematic of spatiotemporal air pollution model.

Table.

Characteristics and performance statistics of selected fine scale air pollution models. A) Models developed in MESA-Air B) Selected models with regional and national coverage

Model	Spatial resolution	Temporal resolution	Key characteristics	Pollutants	Correlation-based R^2 [a]	Mean square error-based R^2 [a]	RMSE	Reference
A. Models developed in MESA-Air								
MESA Air and SPIROMICS Air city-specific spatiotemporal models	Point	Two weeks	Incorporate regulatory data supplemented by investigator-deployed monitors. Developed for cities with MESA and SPIROMICS cohort participants [c]	$PM_{2.5}$	0.78–0.93	0.80–0.93	1.20–2.82 $\mu g/m^3$	Keller et al. [30] [d]
				NOx	0.59–0.92	0.61–0.93	3.43–15.92 ppb	
				NO_2	0.75–0.91	0.74–0.89	1.24–3.82 ppb	
				Ozone	0.75–0.92	0.72–0.90	2.45–4.15 ppb	
				Black carbon		0.51–0.78	0.074–0.329 $10^{-5}/m$	
Spatiotemporal models for Los Angeles	Point	Two weeks	Incorporate CTM data	$PM_{2.5}$	0.80		3.10 $\mu g/m^3$	Wang et al. [27] [f]
				Ozone	0.84		3.62 ppb	Wang et al. [27] [f]
National spatial models	Point	Annual	Based on regionalized universal kriging	$PM_{2.5}$		0.88		Sampson et al. [70]
			Incorporates satellite data	NO_2		0.85 [b]		Young et al. [71]
National spatiotemporal models	Point	Two weeks	Incorporate national network of regulatory monitors and investigator-deployed monitors	$PM_{2.5}$		0.89	1.10 $\mu g/m^3$	Wang et al. [26]
				NO_2		0.87	2.76 ppb	
Historical national spatiotemporal model	Point	Annual	Covers 1980 – 2010	$PM_{2.5}$		0.91	1.14 $\mu g/m^3$	Kim et al. [57] [e]
B. Other selected models with regional and national coverage								
Region-specific spatiotemporal models	Point	Monthly	Based on general additive modeling of geographic and meteorological covariates and regulatory monitoring data	$PM_{2.5}$	0.62–0.89			Yanosky et al. [47] [g]
				PM_{10}	0.44–0.72			
				$PM_{2.5–10}$	0.33–0.64			
National spatiotemporal models	Point	Monthly		$PM_{2.5}$	0.89			Yanosky et al. [47]

Model	Spatial resolution	Temporal resolution	Key characteristics	Pollutants	Correlation-based R^{2a}	Mean square error-based R^{2a}	RMSE	Reference
				PM_{10}	0.69			
				$PM_{2.5–10}$	0.61			
National machine learning-based models	1 × 1 km	Daily	Neural network-based model incorporating satellite, CTM, geographic and meteorological covariates	$PM_{2.5}$	0.83		1.29 $\mu g/m^3$	Di et al. [25] [h]
		Annual		Ozone	0.76		7.36 ppb	Di et al. [52] [h]
National spatiotemporal model based on temporally-scaled LUR surface	0.1 × 0.1 km	Monthly	Incorporates regulatory monitor data and satellite-derived ground-level observations. Monthly scaling of LUR surface	NO_2	0.82			Bechle et al. [50]
National spatial integrated empirical geographic models	Point	Annual	Incorporate land use and satellite-derived pollution data in a universal kriging framework	$PM_{2.5}$		0.85	0.13 $\mu g/m^3$	Kim et al. [75] [i]
				NO_2		0.84	0.23 ppb	
				PM_{10}		0.57	0.27 $\mu g/m^3$	
				O_3		0.82	0.06 ppb	
National spatiotemporal model based on Bayesian Maximum Entropy interpolation	30 x 30 m	Annual	Combines LUR and Bayesian Maximum Entropy interpolation (kriging of LUR residuals)	$PM_{2.5}$	0.79		2.15 $\mu g/m^3$	Beckerman et al. [48]
North American model based on regression adjustment of satellite data	1 × 1 km	Annual	Uses geographically-weighted regression adjustment of satellite AOD data	$PM_{2.5}$	0.82		1.5 $\mu g/m^3$	van Donkelaar et al. [49]

Abbreviations: CTM – chemical transport model, RMSE – root mean square error. SPIROMICS – Sub-Populations and Intermediate Outcome Measures in COPD study

[a] Correlation-based R^2 is the square of the correlation coefficient between observations and cross-validated predictions, it captures precision. Mean square error-based R^2 is calculated a 1 minus the ratio of MSE to data variability to additionally account for bias.

[b] The R^2 without satellite-derived estimate of ground-level NO_2 as covariate was similar (0.84).

[c] These included New York City, NY; Baltimore, MD; Chicago, IL; Saint Paul, MN; Ann Arbor, MI; Salt Lake City, UT; Winston Salem, NC; Los Angeles, CA; and San Francisco, CA.

[d] R^2 represents range of long-term MSE-based R^2 at home locations across the modeled metro regions. Model output has been updated since it was originally published.

[e] Based on cross-validation with Federal Reference Method/Interagency Monitoring of Protected Visual Environment network monitors. Other datasets were also used for external validation.

[f] CTM data are output from UC-Davis/Caltech Air Quality Model.

[g] R^2 values represent range observed across 7 US regions.

[h] CTM data from GEOS-Chem.

[i] RMSE values are standardized, i.e. divided by mean RMSE across all monitoring sites.