# Global, high-resolution, reduced-complexity air quality modeling using InMAP (Intervention Model for Air Pollution)

**Short title: Global, reduced-complexity air quality modeling**

Sumil K. Thakrar[1,2], Christopher W. Tessum[3], Joshua S. Apte[4,5], Srinidhi Balasubramanian[1], Dylan B. Millet[6], Spyros N. Pandis[7,8], Julian D. Marshall[9], Jason D. Hill[1*].

[1]Department of Bioproducts & Biosystems Engineering, University of Minnesota, St Paul, MN, USA.

[2]Department of Applied Economics, University of Minnesota, St Paul, MN, USA.

[3]Department of Civil and Environmental Engineering, University of Illinois at Urbana–Champaign, Urbana, IL, USA.

[4]Department of Civil and Environmental Engineering, University of California, Berkeley, Berkeley, CA, USA.

[5]School of Public Health, University of California, Berkeley, Berkeley, CA, USA.

[6]Department of Soil, Water, and Climate, University of Minnesota, St Paul, MN, USA.

[7]Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA.

[8]Department of Chemical Engineering, University of Patras, Patras, Greece.

[9]Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, USA.

**Corresponding author:** Sumil K. Thakrar

**Email:** sthakrar@umn.edu

ORCIDs:

| | |
|---|---|
| Sumil K. Thakrar: | 0000-0003-2205-3333 |
| Christopher W. Tessum: | 0000-0002-8864-7436 |
| Joshua S. Apte: | 0000-0002-2796-3478 |
| Srinidhi Balasubramanian: | 0000-0002-6561-5984 |
| Dylan B. Millet: | 0000-0003-3076-125X |
| Spyros N. Pandis | 0000-0001-8085-9795 |
| Julian D. Marshall: | 0000-0003-4087-1209 |
| Jason D. Hill: | 0000-0001-7609-6713 |

## Abstract

Each year, millions of premature deaths worldwide are caused by exposure to outdoor air pollution, especially fine particulate matter ($PM_{2.5}$). Designing policies to reduce these deaths relies on air quality modeling for estimating changes in $PM_{2.5}$ concentrations from many policy scenarios at high spatial resolution. However, air quality modeling typically has high requirements for computation and expertise, which limits policy design, especially in countries where most $PM_{2.5}$-related deaths occur. Lower requirement reduced-complexity models exist but are generally unavailable worldwide. Here, we adapt InMAP, a reduced-complexity model originally developed for the United States, to simulate annual-average primary and secondary $PM_{2.5}$ concentrations across a global-through-urban spatial domain: "Global InMAP". Global InMAP uses a variable resolution grid, with 4 km horizontal grid cell widths in cities. We evaluate Global InMAP performance both against measurements and a state-of-the-science chemical transport model, GEOS-Chem. For the emission scenarios considered, Global InMAP reproduced GEOS-Chem pollutant concentrations with a normalized mean bias of 59%–121%, which is sufficient for initial policy assessment and scoping. Global InMAP can be run on a desktop computer; simulations here took 2.6–4.4 hours. This work presents a global, open-source, reduced-complexity air quality model to facilitate air pollution policy assessment worldwide, providing a screening tool for reducing the deaths where they occur most.

## Introduction

Exposure to outdoor air pollution is the largest environmental health risk factor worldwide, associated with millions of excess deaths each year[1,2]. The deaths are mostly attributable to fine particulate matter ($PM_{2.5}$), which can either be emitted directly, or can form indirectly from precursor pollutants that are emitted from a wide variety of natural and anthropogenic emission sources, including transportation, agriculture, and electricity generation[3,4]. Designing strategies to reduce mortality relies on understanding how specific emission sources affect ambient $PM_{2.5}$ concentrations, and thereby, human health, across a range of possible technology or policy scenarios.

InMAP[5] (Intervention Model for Air Pollution) is a reduced-complexity, open-source air quality model that has been used to inform strategies to reduce $PM_{2.5}$-related mortality from specific emission sources. For example, InMAP has been used to estimate fine-scale pollution impacts across distances[6], measures of pollution inequity across racial-ethnic and socioeconomic groups[7], the health impacts of specific sectors under different policy scenarios[8,9], and individual impacts of commodities[10]. However, as with other widely used reduced-complexity air quality models such as EASIUR[11], AP2[12], and COBRA[13], InMAP previously has only been configured and evaluated for the United States, a country with just 4% of the world's population and 2% of the world's air quality-related deaths[2,4].

Chemical transport models (CTMs) are employed for estimating the effects of emission sources on pollutant concentrations and health impacts and are considered state-of-the-science for air quality modeling. However, they require substantial time, expertise, and computational resources (*e.g.*, several computation days per simulation month), limiting the use cases and therefore the extent to which they can inform many multidimensional policy decisions[5,14], especially when hundreds of policy scenarios are being considered. Although GEOS-Chem is one of the most widely used CTMs, 60% of deaths from outdoor air pollution occur in countries where there are no known users or institutions using GEOS-Chem[15,16]. Thus, researchers and practitioners would benefit from additional models and tools beyond CTMs to investigate air pollution and emission-control strategies. Such tools would be useful even though the uncertainty may be higher than with a CTM. For example, because damages per tonne emitted varies by orders of magnitude, for

many analyses an uncertainty of a factor of 2 or 3, or higher (e.g., an order of magnitude estimate), can provide scientifically relevant results that can usefully inform policy decisions.

Some global air quality models are available with a lower operational difficulty than CTMs, including TM5-FASST[17], source-receptor relationships built from GEOS-Chem adjoint[18], and EMEP[19]. Compared to the existing global air quality models with lower operational difficulty than CTMs, InMAP has higher spatial resolution, is easier to use, and has lower computational costs. A recent notable effort[20] to build a monthly life cycle assessment model for $PM_{2.5}$ has not yet been tested against measurements or compared with results from a CTM. A diversity of independently evaluated reduced-complexity models will increase their applicability and the robustness of policy assessments worldwide[21].

Here, we developed and configured InMAP for use on a global spatial domain ("Global InMAP"). We ran a year-long, global CTM simulation using GEOS-Chem[22], and used its outputs to globally parameterize the chemistry, physics, and meteorology of InMAP. We then ran InMAP on global emission inventories to predict total $PM_{2.5}$ concentrations as well as changes in concentrations from three specific scenarios of emission changes. We compared the results to a global dataset of ground observations, as well as to $PM_{2.5}$ concentrations and changes in concentrations predicted by GEOS-Chem. Lastly, we compared Global InMAP to the United States versions of InMAP for two emission scenarios.

## Materials and Methods

The InMAP model, fully described in Tessum *et al.*[5], estimates annual-average concentrations of fine particulate matter ($PM_{2.5}$), including both primary (*i.e.*, directly emitted) and secondary (*i.e.*, formed in the atmosphere) components, to guide research and policy. As with other reduced-complexity models, InMAP is designed to be faster and easier to use than CTMs, and will typically have lower accuracy and precision than CTMs as a tradeoff for the greater speed and ease of use.

InMAP explicitly tracks secondary $PM_{2.5}$ contributions from particulate ammonium ($pNH_4$), particulate sulfate ($pSO_4$), particulate nitrate ($pNO_3$), and secondary organic aerosol (SOA), from emissions of $PM_{2.5}$ precursors (sulfur oxides ($SO_x$), nitrogen oxides ($NO_x$), ammonia ($NH_3$), and non-methane volatile organic compounds (NMVOCs)). InMAP estimates pollutant concentrations by approximating the steady-state solution to a set of differential equations governing pollutant emissions, reaction, advection, diffusion, and removal. It solves the equations by discretizing over space and time, using a variable resolution grid, and spatially varying parameterizations that simplify the reaction, advection, and removal terms in the equations. Whereas CTMs simulate chemistry and physics (reaction, advection, removal) using first principles and mechanistic or empirical representations for specific processes, InMAP simulates chemistry and physics using simplified representations that are parameterized by the outputs of a CTM simulation.

InMAP as configured over the United States ("US InMAP") was parameterized using outputs from WRF-Chem[23,24]. However, WRF-Chem is not commonly used for global simulations. Instead, InMAP was parameterized here using outputs from GEOS-Chem[22], a global CTM. The full list of equations used in InMAP is given in Tessum *et al.*[5] Details of the model configuration, GEOS-Chem simulation inputs, global emission inventories, and performance evaluation are provided below.

### *Global InMAP computational grid*

As with previous InMAP configurations for the US[5-10], the horizontal resolution of the Global InMAP computational grid varies across space and is higher in places with larger population or population density. Here, we used 2020 projected population data at 0.01° resolution[25] to create the computational grid. We employed a population density threshold of $5.5 \times 10^8$ deg$^{-2}$ and a population threshold of 100,000. For any grid cell, if either threshold was exceeded, then the model subdivided it into smaller cells until the smallest cell size was reached.

The resulting computational grid (Figure S1) has ~1.5 million grid cells (ground-level: 170,358 grid cells), whose horizontal resolution at ground-level ranges from 5° × 4° (which corresponds to ~500 km length at the equator) in remote locations to 0.04° × 0.03° (~4 km length at the equator) in urban locations. The spatial domain encompasses the vast majority of the Earth's surface: latitudes from -87.0° to +81.0° and longitudes from -178.0° to +172.0°. Global InMAP does not track pollution across the poles or antimeridian[26]. The resulting grid covers all but ~5 million people (< 0.1% of the total population) in parts of New Zealand and other islands in the Pacific Ocean. The population-weighted average grid-cell size is 1,000 km$^2$ (for comparison, ~39,000 km$^2$ for GEOS-Chem). The resulting pre-processed gridded input data file is ~700 MB and is provided in a freely available dataset (doi:10.5281/zenodo.4641948).

*GEOS-Chem simulation*

Chemical and physical atmospheric parameters used in Global InMAP, such as annual-average gas/particle-phase partitioning coefficients, were derived from the outputs of an annual GEOS-Chem "Classic" (version 11-01f) simulation (2016-01-01 until 2017-01-01), with meteorology provided by MERRA-2[27]. The GEOS-Chem outputs were used in the same way as the corresponding WRF-Chem variables were used for US InMAP (see Tessum *et al.*[5]). The full list of GEOS-Chem variables used in Global InMAP, and descriptions of how they are used, are in Table S1.

The GEOS-Chem model code and configuration were derived from a simulation performed by Hammer *et al.*,[28] where the chemical mechanism included complex secondary organic aerosol (SOA) formation with semi-volatile primary organic aerosol[29,30]. We used the standard horizontal spatial resolution for global simulations in GEOS-Chem, 2° × 2.5°, (~ 220 km × 275 km at the equator) with 47 vertical levels, following the configuration described in Hammer *et al.*[28]

GEOS-Chem also allows for higher resolution grids nested within a larger domain[31]. Again following Hammer *et al.*[28], we ran GEOS-Chem nested grid simulations over the same time period for Asia, Europe, and North America, at 0.5° × 0.625° resolution, which covers 75% of the world's population. First, boundary conditions for the nested grid simulations were recorded every 180 minutes of simulation time, at 2° × 2.5° resolution, during the global simulation. In our application, emergent properties extracted for use in Global InMAP, such as the annual-average temperature and wind velocity vectors, are only specified up to this coarse resolution. However, Global InMAP can still be used on a higher resolution (variable) grid, and the resolution of the emission inventory is also not limited by the resolution of the GEOS-Chem output.

*Emission inputs*

To estimate concentrations of total $PM_{2.5}$ and speciated components using Global InMAP, we compiled a global emission inventory of $NH_3$, primary $PM_{2.5}$, $NO_x$, $SO_x$, and NMVOC. For consistency, we chose the same emission inventories as those used in the GEOS-Chem simulation, but, where possible, processed to a higher spatial resolution as described below for the Global InMAP computational grid. Total annual emissions fluxes for the emission inventories used in the Global InMAP simulation are given in Table 1.

Where possible, the total emission inventories used for the Global InMAP simulation were compiled using the standalone version of HEMCO[32], using the same configuration as used in the GEOS-Chem simulation except at 0.25° × 0.25° horizontal resolution.

Differences in grid resolutions, time steps, and environmental fields can result in small differences when the same emission inventories are processed. HEMCO standalone provides both high resolution emissions and consistency with the GEOS-Chem simulation but cannot be used for some emission inventories that require detailed chemical or meteorological inputs. For those, we instead saved out emissions ("diagnostics") from the GEOS-Chem simulation, gridded at 2° × 2.5°, and used these in the global InMAP simulation.

Table 1 gives the total annual emissions for Global InMAP inputs, and the data source for each group of emissions used. Global and regional emission inventories used for anthropogenic sources of $PM_{2.5}$ and precursors include: EDGAR[33] v.4.3.2, the National Emissions Inventory (NEI) 2011 for the United States, BRAVO[34] (Big Bend Regional Aerosol and Visibility Observational study) for Mexico, the Criteria Air Contaminant (CAC) emission inventory for Canada, EMEP[35] for Europe, MIX[36] v1.1 for Asia, MEIC[37] v1.2 for China, Lu et al.[38] for $SO_x$ emissions in China and India, AEIC[39] for aircraft emissions, PARANOX[40] for ship emissions, and RETRO[41] for biofuel emissions. Biomass burning emissions are from the RETRO[41] and GFED-4[42] emission inventories. Natural emission inventories used here include Ge et al.[43] for volcanic emissions, Hudman et al.[44] for soil $NO_x$, MEGAN[45] for biogenic emissions, and DEAD[46] for dust emissions.

Only a subset of NMVOC emissions is likely to form SOA[47,48]. For Global InMAP anthropogenic emissions, we included isoprene, monoterpenes, benzene, toluene, xylenes, trimethylbenzenes, alkanes with more than 4 carbon atoms, and other aromatics, from the EDGAR[33] v4.3.2 emission inventory. For biogenic emissions, we included limonene, isoprene, alpha-pinene, beta-pinene, sabinene, carene, and monoterpenes from the global GEOS-Chem simulation. For biomass burning, we include benzene, toluene, xylenes, alkenes with more than 3 carbon atoms, and alkanes with more than 4 carbon atoms, from the RETRO biomass burning emission inventory[41].

Although Global InMAP has the functionality to include vertically elevated emissions, there is a lack of global information on emission heights for many sources[33]. HEMCO processed emissions were thus derived at the lowest vertical layer, except for aircraft emissions, lightning $NO_x$ and volcanic $SO_x$. For simplicity in configuring the Global InMAP emissions, here we only used the emissions from these sources in the lowest vertical layer, which excluded 8% of global $NO_x$ emissions and 16% of global $SO_x$ emissions.

$PM_{2.5}$ concentrations are not directly tracked in GEOS-Chem, but rather are calculated from its underlying components that are grouped in such a way as to facilitate chemical and atmospheric modeling. For example, dust is grouped by several size classes that do not perfectly map onto $PM_{2.5}$. HEMCO and GEOS-Chem diagnostic outputs also typically report emissions in these groups, requiring some conversions for use in Global InMAP. Here, we did so in accordance with the standard GEOS-Chem recommendations (see Table S1 for the $PM_{2.5}$ equation used). Following Hammer *et al.*[28] and Li *et al.*[49], irreversible aqueous formation of SOA from isoprene was included in total $PM_{2.5}$ mass, whereas reversible formation was excluded.

InMAP data inputs for pollutant removal through deposition likewise required modification for Global InMAP simulations. Specifically, Global InMAP requires land cover data to calculate dry deposition rates for gases and particles in each ground-level grid cell. For the United States, InMAP used land cover data from the United States Geological Survey National Land Cover

Database[50]. For Global InMAP, we instead used the Olson 2001 Land Use Map at 0.025° × 0.025° resolution[51], which is also used in GEOS-Chem.

*Comparison with other air quality models and measurements*

Using the global emission inventories described in the previous section, we generated Global InMAP results and compared them against other models and measurements (1) for total concentrations; (2) for three perturbation scenarios wherein we modified global emissions from a specific sector and predicted the resulting concentration changes; and (3) for United States electricity and transportation emissions, to compare Global InMAP with US InMAP.

First, we evaluated Global InMAP predictions of $PM_{2.5}$ (total and speciated) against annual-average ground-level measurements, as is commonly done for air quality models[52,53]. To this end, we compiled and vetted a global measurement dataset for total and speciated $PM_{2.5}$ (see Supplementary Text and Table S2 for additional details). We reported metrics commonly used for evaluating model performance: normalized mean error and bias (NME and NMB), the squared linear correlation coefficient, $R^2$, and the slope of the best-fit line, S (see Supplementary Information for equations used)[54]. Using this approach, model-measurement comparisons were generated for Global InMAP and (separately) for the GEOS-Chem simulation (described above).

To provide context for the model-measurement comparison results, we reported model criteria published by Emery *et al.*[54] (see Supplementary Information). Performance criteria were provided as a general reference point, not as "pass/fail" criteria. The criteria are intended for evaluating $PM_{2.5}$ concentrations over sub-annual lengths of time[55], or for daily average measurements within 1000 km, where there are more than 10 measurements[54]. Here, we used the criteria more broadly to identify the stronger and weaker aspects of model performance.

Second, we simulated the effects of three emissions perturbations with Global InMAP and GEOS-Chem simulations and compared their predicted pollutant concentration increments. The perturbations chosen were: (i) a 100% increase (4.9 Tg) in global $SO_2$ emissions from power generation for 2 months (2016-01-01 until 2016-03-01); (ii) a 100% increase (7.5 Tg) in global $NH_3$ emissions from agricultural soils for 3 months (2016-01-01 until 2016-04-01); (iii) a 100% increase (1.4 Tg) in global $NO_x$ emissions from road transport for 1 month (2016-01-01 until 2016-02-01). All emissions changes were from the EDGAR emissions database (v.4.2, 0.1° × 0.1° resolution) as described above. For each of the scenarios chosen, we ran global, annual 2° × 2.5° GEOS-Chem simulations similar to those described above, with the change in emissions implemented using a constant temporal profile over the timescale of the perturbation. As InMAP is an "intervention" model (designed to model changes in emissions directly), for Global InMAP we ran the changes in emissions from the EDGAR emission inventories at native resolution.

Lastly, because InMAP has already been configured and evaluated over the contiguous United States, we performed two simulations for United States emission changes using Global InMAP and US InMAP. To this end, we compiled emission inventories over the United States using the National Emissions Inventory (NEI) 2014v.1, processed exactly as in Thakrar *et al.*[8] We investigated two sources of $PM_{2.5}$ and precursor emissions: coal-powered electricity generation (NEI Source Classification Code: 10100212) and gasoline passenger vehicles (NEI Source Classification Code: 2201210080).

## Results

*Computational requirements*

The annual, global simulations described above (system: 98 processors on 1 node of a supercomputing cluster; 36 GB memory) required 4 hours for Global InMAP (1.5 million grid cells) and 100 hours for GEOS-Chem (2° × 2.5° grid resolution, 0.6 million grid cells). The perturbation simulations, when run on the same system, took 2.6–4.4 hours.

Other GEOS-Chem simulations require comparably high resources[56]. The variable resolution InMAP grid allows for much higher spatial resolution over areas with high population density than is possible with the GEOS-Chem uniform grid, while only requiring 4% of the computational time.

*Model-to-measurement comparisons*

The Global InMAP simulation using total emissions was able to predict total $PM_{2.5}$ concentrations against measurements globally with NMB = –60%; NME = 63%; and $R^2$ = 0.35 (see Figures 1–2, S2). As with the GEOS-Chem simulation, the performance of the Global InMAP simulation varied by region (see Figures S3–S9). The Global InMAP simulation was generally most accurate in Oceana (NMB: -45%; $R^2$: 0.64; see Figure S7), North America (NMB: -54%; $R^2$: 0.59; see Figure S6), and Europe (NMB: -64%; $R^2$: 0.28; see Figure S5), and least accurate in South America (NMB: -74%; $R^2$: 0.05; see Figure S8). Across many heavily polluted regions in Asia, the Global InMAP simulation predicted much lower $PM_{2.5}$ concentrations than are measured (difference: > 30 μg $m^{-3}$) (Figure S5), in particular across the Indo-Gangetic Plain. The underprediction may have arisen because of potentially low emissions inputs, *e.g.* from industrial and agricultural $NH_3$ emissions[57] or missing NMVOC species from biomass burning[58]. The Global InMAP simulation may have underpredicted pollution from episodic events, such as biomass burning in the Indo-Gangetic Plain, because Global InMAP assumes that emissions occur at an annual-average rate. Furthermore, the chemistry that is included in Global InMAP may not be sufficiently complex to predict $PM_{2.5}$ with high accuracy in heavily polluted areas[59].

We also compared annual-average predicted concentrations from the Global InMAP simulation to annual-average measurements of $pSO_4$, $pNO_3$, and $pNH_4$ globally (Figures 3–6). The Global InMAP simulation predicted these components well (NME: 50%–67%; $R^2$: 0.24–0.38) and was generally biased low against measurements for $pNO_3$ (especially in areas with $pNO_3$ >2 μg $m^{-3}$), and high for $pSO_4$ and $pNH_4$. Because the Global InMAP simulation did not have a strong low bias against secondary inorganic $PM_{2.5}$ measurements, it is likely that much of the low bias of the Global InMAP simulation against total $PM_{2.5}$ measurements arose from its prediction of primary $PM_{2.5}$ concentrations (see Figure 7). However, measurement data for SOA and primary $PM_{2.5}$ concentrations were not available at the evaluation sites (see Figure 8 for ground-level concentrations of these species).

We also compared the GEOS-Chem simulation against the same measurement data, to contextualize the Global InMAP results. The GEOS-Chem simulation predicted total $PM_{2.5}$ measurements with an $R^2$ of 0.55. For comparison, a GEOS-Chem simulation that used the same code and emissions[28] reported an $R^2$ of 0.61 when using a more comprehensive measurement dataset and averaging results across years 2010–2018 instead of just 2016.

Both the Global InMAP and the GEOS-Chem simulations predicted lower annual-average total $PM_{2.5}$ concentrations than were observed. For all species and regions, the direction of bias against measurements was the same for the Global InMAP simulation as for the GEOS-Chem simulation. This suggests that the Global InMAP simulation was inheriting the bias from the GEOS-Chem simulation inputs to some extent. If that was the case, then future improvements to the GEOS-Chem model and to the emission inventories used here could further reduce Global InMAP biases.

The Global InMAP simulation broadly reproduced spatial patterns of pollutant concentrations predicted by GEOS-Chem. However, there were some features present in the GEOS-Chem simulation that were not captured by the Global InMAP simulation. Such features included high annual-average $PM_{2.5}$ concentrations from biomass burning, including the Alberta fires, crop burning in the Indo-Gangetic Plain, peatland fires in Singapore and Malaysia, and burning in Siberia. InMAP may have underpredicted $PM_{2.5}$ concentrations from biomass burning relative to the GEOS-Chem simulation because it assumes emissions happen at an annual-average rate. Across Western China, the Global InMAP simulation tended to misrepresent the spatial patterns provided by the parent GEOS-Chem simulation for both primary and secondary $PM_{2.5}$, including high concentrations over the Himalayas and Sichuan Basin, and low concentrations in surrounding areas. This may suggest that the annual-average advection scheme used by InMAP does not yet adequately capture complex air flows over steep terrain.

*Evaluation of predicted responses to changes in emissions*

The major intended use of InMAP is to estimate the changes in human exposure to $PM_{2.5}$ concentrations for given scenarios of emission changes. Therefore, its ability to reproduce the changes predicted by the original CTM could be considered its most important attribute, more important than its ability to reproduce current absolute concentrations. However, InMAP is designed to predict human exposure with high spatial resolution in urban areas, while GEOS-Chem is designed to predict global chemical transport and runs at comparatively low resolution. Directly comparing the two models requires re-gridding the higher-resolution Global InMAP results to match the lower-resolution GEOS-Chem results, which cancels out predictive advantages Global InMAP might gain from its use of higher spatial resolution. Therefore, results in this section could be considered a conservative evaluation of Global InMAP's predictive performance.

Figures 9–11 show annual-average pollutant concentration increments predicted by the GEOS-Chem and Global InMAP simulations for increases in $SO_x$ emissions from power generation, $NH_3$ emissions from agricultural soils, and $NO_x$ emissions from road transportation. When regridding Global InMAP predictions to the GEOS-Chem grid, we found that Global InMAP reproduced the GEOS-Chem results with an average area-weighted NME of 118–182% and an average area-weighted NMB of 59–121% (see Table 2). For the $NO_x$ and $NH_3$ emissions scenarios, Global InMAP exhibited better performance against GEOS-Chem on a population-weighted basis than on an area-weighted basis. For the $SO_x$ emissions scenario, Global InMAP exhibited the lowest performance against the GEOS-Chem simulation, having overpredicted changes in $pSO_4$ concentrations especially in populated areas.

The Global InMAP simulations predicted greater variability in concentration changes over urban areas than the 2° × 2.5° GEOS-Chem simulations for the same emissions scenarios, owing to its higher resolution computational grid. Figure 12 compares the $pNO_3$ concentration changes over Cairo, São Paulo, and Tokyo (the largest cities in Africa, South America, and Asia[60]) for the $NO_x$ perturbation scenario as predicted by Global InMAP and GEOS-Chem. Changes in concentrations predicted by Global InMAP correlated with changes in emissions at the urban scale. Higher resolution GEOS-Chem simulations that resolve intra-urban gradients would be even more computationally expensive than the GEOS-Chem simulations performed here[56].

Global InMAP predicted similar spatial patterns and magnitudes of changes in pollutant concentrations as did US InMAP for a given emissions perturbation (see Figure S10), with NME and NMB within ± 50% for both scenarios considered (see Table 2). This demonstrated consistency between the InMAP versions derived from WRF-Chem and GEOS-Chem inputs, suggesting that no major errors were introduced in the Global InMAP model development (see

Table S1; Tessum *et al.*[5]). For InMAP applications focusing only on the United States, continued use of US InMAP is warranted, as the WRF-Chem simulation used to parameterize US InMAP provides higher spatial resolution than does the nested GEOS-Chem simulation employed for Global InMAP.

**Discussion**

Here, we extended InMAP, a reduced-complexity air quality model originally developed for use in the United States, to simulate a global-through-urban spatial domain. InMAP is designed to supplement rather than supplant state-of-the-science tools such as GEOS-Chem or other global models, *e.g.*, for cases in which (i) resources to implement a CTM are unavailable, (ii) numerous simulations are needed to evaluate a large variety of policy scenarios, or (iii) the primary need is initial assessment and screening. The accuracy of InMAP is not as good as with a CTM (e.g., here, a normalized mean error of 63% (InMAP) versus 41% (GEOS-Chem)), yet for many scientific and policy questions lacking readily-available CTM-quality results, InMAP provides useful information.

Global InMAP requires relatively low computational resources, allowing it to be run on a desktop computer rather than a supercomputer. Simulations predicting annual-average concentrations take several hours rather than days. For example, compared to the global GEOS-Chem simulation described here, the Global InMAP simulation was 25× faster at predicting total annual-average $PM_{2.5}$ concentrations, despite the Global InMAP simulation having 39× higher population-weighted average spatial resolution (down to ~4km in urban areas).

As expected, the expedience of Global InMAP comes at the expense of lower predictive accuracy compared to a comprehensive CTM. This Global InMAP simulation is biased low against measurements for total $PM_{2.5}$ across all regions. Among species, it is biased high against measurements of $pSO_4$ and $pNH_4$, and low against measurements of $pNO_3$. The low computational resource requirements make Global InMAP particularly well-suited to applications where hundreds of policy scenarios are evaluated, as is often done using reduced-complexity models for the United States[8,10,61], or when no other air quality models are available at the urban scale. In places with higher population and pollution exposure than the United States, there is even more potential for a reduced-complexity model such as Global InMAP to inform impactful policy decisions. Global InMAP may be important for informing preliminary hypotheses about policy decisions in its early stages (*e.g.*, “What is the best location to site a new facility that may be a major pollution source?”), allowing computational resources to be used instead for CTMs at a later stage to check consistency with the findings.

Global InMAP performance varies regionally, and it tends to perform worse against measurements in places where GEOS-Chem also performs poorly (*e.g.*, South America). This suggests that predictive accuracy in those areas is generally lower across models, so that Global InMAP may provide a comparative advantage. Further, this suggests that Global InMAP performance in those regions may improve based on future advancements in emission inventories or GEOS-Chem model inputs.

By directly estimating annual-average $PM_{2.5}$ concentrations at high spatial resolution, Global InMAP is configured to easily estimate changes in human exposure and health impacts. When estimating human health effects of emissions changes, there will also be sizeable uncertainties from estimating the emissions changes themselves and from the concentration-response function employed[62]; Global InMAP errors should thus be contextualized with those in mind. For the United States, a previous study[6] found that the largest source of uncertainty in estimating monetized $PM_{2.5}$ health impacts was the economic valuation of premature mortality, followed by

the concentration response function, whereas uncertainty in $PM_{2.5}$ concentrations from the choice of air quality model was the smallest source of uncertainty considered. Since uncertainty in the air pollution model is not the largest source of uncertainty in many contexts, a reduced-complexity model (RCM) can deliver useful information; that conclusion especially applies to the many cases where resources may exist to run an RCM but not to establish and run a conventional CTM. Indeed, there are many cases in which a CTM simulation is infeasible, yet an RCM or other approach could provide some information. As mentioned above, InMAP is not a replacement for a CTM; instead, it provides screening-level information, results for questions that would involve too many model runs to use a CTM, or results that would be otherwise infeasible. By providing a global, open source, air quality model with high spatial resolution and low computational requirements, we hope to facilitate the wide practice of air pollution policy assessment worldwide.

## Acknowledgments

The authors thank Peter Adams, Susan Anenberg, Zoe Chafe, Xin Chen, Kimberly Colgan, Nina Domingo, Madisen Gittlin, Melanie Hammer, Daven Henze, Iyad Kheirbek, Gary Kleiman, Chi Li, Randall Martin, Dev Millstein, Shayak Sengupta, Kelley Wells, Ruili Wu, and Xueying Yu. The authors acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing resources that contributed to the research results reported within this paper.

This project was funded by the US EPA (Assistance Agreement No. R835873 to CACES; J.S.A., S.B., J.D.H, J.D.M, D.B.M., S.N.P., C.W.T, S.K.T.), the Clean Air Fund via c40 (J.D.M., C.W.T.), the Wellcome Trust (Grant 216075/Z/19/Z; J.D.M., C.W.T.), the William F. Wilcke Fellowship Fund (S.K.T.), and USDA/NIFA (Project MIN-12-110; J.D.H.). The manuscript has not been formally reviewed by the funders; the views expressed in this document are solely those of authors and do not necessarily reflect those of the funders. Funders do not endorse any products or commercial services mentioned in this publication, and no endorsement of the funders or their goals or affiliates is implied.

The authors declare no competing financial interest.

## Author contributions

Conceptualization: SKT, CWT, DBM, JDM, JDH. Data Curation: SKT, CWT, SB, DBM, JDM, JDH. Methodology: SKT, CWT, DBM, JDM, JDH. Resources: SKT, CWT, SB, DBM, JDM, JDH. Software: CWT, SKT. Supervision: CWT, JDM, JDH. Vizualisation: SKT. Writing – original draft: SKT; Writing – review & editing: SKT, CWT, JSA, SB, DBM, SNP, JDM, JDH.

## References

1. Landrigan PJ, Fuller R, Acosta NJ, Adeyi O, Arnold R, Baldé AB, et al. The Lancet Commission on pollution and health. The lancet. 2018 Feb 3;391(10119):462-512.
2. Burnett R, Chen H, Szyszkowicz M, Fann N, Hubbell B, Pope CA, et al. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proceedings of the National Academy of Sciences. 2018 Sep 18;115(38):9592-7.
3. Silva RA, Adelman Z, Fry MM, West JJ. The impact of individual anthropogenic emissions sectors on the global burden of human mortality due to ambient air pollution. Environmental health perspectives. 2016 Nov;124(11):1776-84.
4. Lelieveld J, Evans JS, Fnais M, Giannadaki D, Pozzer A. The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature. 2015 Sep;525(7569):367-71.
5. Tessum CW, Hill JD, Marshall JD. InMAP: A model for air pollution interventions. PloS one. 2017 Apr 19;12(4):e0176131.
6. Goodkind AL, Tessum CW, Coggins JS, Hill JD, Marshall JD. Fine-scale damage estimates of particulate matter air pollution reveal opportunities for location-specific mitigation of emissions. Proceedings of the National Academy of Sciences. 2019 Apr 30;116(18):8775-80.
7. Tessum CW, Apte JS, Goodkind AL, Muller NZ, Mullins KA, Paolella DA, et al. Inequity in consumption of goods and services adds to racial–ethnic disparities in air pollution exposure. Proceedings of the National Academy of Sciences. 2019 Mar 26;116(13):6001-6.

8. Thakrar SK, Balasubramanian S, Adams PJ, Azevedo IM, Muller NZ, Pandis SN, et al. Reducing mortality from air pollution in the United States by targeting specific emission sources. Environmental Science & Technology Letters. 2020 Jul 15;7(9):639-45.
9. Liu L, Hwang T, Lee S, Ouyang Y, Lee B, Smith SJ, Tessum CW, Marshall JD, Yan F, Daenzer K, Bond TC. Health and climate impacts of future United States land freight modelled with global-to-urban models. Nature Sustainability. 2019 Feb;2(2):105-12.
10. Hill J, Goodkind A, Tessum C, Thakrar S, Tilman D, Polasky S, Smith T, Hunt N, Mullins K, Clark M, Marshall J. Air-quality-related health damages of maize. Nature Sustainability. 2019 May;2(5):397-403.
11. Heo J, Adams PJ, Gao HO. Reduced-form modeling of public health impacts of inorganic PM2. 5 and precursor emissions. Atmospheric Environment. 2016 Jul 1;137:80-9.
12. Muller NZ. Boosting GDP growth by accounting for the environment. Science. 2014 Aug 22;345(6199):873-4.
13. USEPA. User's manual for the co-benefits risk assessment (COBRA), 2018.
14. Lee CJ, Martin RV, Henze DK, Brauer M, Cohen A, Donkelaar AV. Response of global particulate-matter-related mortality to changes in local precursor emissions. Environmental Science & Technology. 2015 Apr 7;49(7):4335-44.
15. GBD Results Tool. Accessed: 25th January 2021. http://ghdx.healthdata.org/gbd-results-tool
16. Maps of GEOS-Chem User Groups. Accessed: 25th January 2021. http://acmg.seas.harvard.edu/geos/geos_people.html
17. Van Dingenen R, Dentener F, Crippa M, Leitao J, Marmer E, Rao S, et al. TM5-FASST: a global atmospheric source–receptor model for rapid impact analysis of emission changes on air quality and short-lived climate pollutants. Atmospheric Chemistry and Physics. 2018 Nov 13;18(21):16173-211.
18. Henze DK, Hakami A, Seinfeld JH. Development of the adjoint of GEOS-Chem. Atmospheric Chemistry and Physics. 2007 May 11;7(9):2413-33.
19. Amann M, Bertok I, Borken-Kleefeld J, Cofala J, Heyes C, Höglund-Isaksson L, Klimont Z, Nguyen B, Posch M, Rafaj P, Sandler R. Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications. Environmental Modelling & Software. 2011 Dec 1;26(12):1489-501.
20. Oberschelp C, Pfister S, Hellweg S. Globally regionalized monthly life cycle impact assessment of particulate matter. Environmental Science & Technology. 2020 Nov 23;54(24):16028-38.
21. Gilmore EA, Heo J, Muller NZ, Tessum CW, Hill JD, Marshall JD, et al. An inter-comparison of the social costs of air quality from reduced-complexity models. Environmental Research Letters. 2019 Jul 9;14(7):074016.
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 Oct 16;106(D19):23073-95.
23. Grell GA, Peckham SE, Schmitz R, McKeen SA, Frost G, Skamarock WC, Eder B. Fully coupled “online” chemistry within the WRF model. Atmospheric Environment. 2005 Dec 1;39(37):6957-75.
24. Tessum CW, Hill JD, Marshall JD. Twelve-month, 12 km resolution North American WRF-Chem v3. 4 air quality simulation: performance evaluation. Geoscientific Model Development. 2015 Apr 7;8(4):957-73.
25. Gridded Population of the World, Version 4 (GPWv4): National Identifier Grid. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H41V5BX1.
26. Park RJ, Jacob DJ, Field BD, Yantosca RM, Chin M. Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy. Journal of Geophysical Research: Atmospheres. 2004 Aug 16;109(D15).

27. Gelaro R, McCarty W, Suárez MJ, Todling R, Molod A, Takacs L, et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). Journal of climate. 2017 Jul 15;30(14):5419-54.
28. Hammer MS, van Donkelaar A, Li C, Lyapustin A, Sayer AM, Hsu NC, Levy RC, Garay MJ, Kalashnikova OV, Kahn RA, Brauer M. Global estimates and long-term trends of fine particulate matter concentrations (1998–2018). Environmental Science & Technology. 2020 Jun 3;54(13):7879-90.
29. Pye, H. O. T; Seinfeld, J. H. A global perspective on aerosol from low-volatility organic compounds. Atmospheric Chemistry and Physics **2010**, *10* (9), 4377-4401.
30. Pye HO, Chan AW, Barkley MP, Seinfeld JH. Global modeling of organic aerosol: the importance of reactive nitrogen (NOx and NO3). Atmospheric Chemistry and Physics. 2010 Nov 30;10(22):11261-76.
31. Wang YX, McElroy MB, Jacob DJ, Yantosca RM. A nested grid formulation for chemical transport over Asia: Applications to CO. Journal of Geophysical Research: Atmospheres. 2004 Nov 27;109(D22).
32. Keller CA, Long MS, Yantosca RM, Da Silva AM, Pawson S, Jacob DJ. HEMCO v1. 0: a versatile, ESMF-compliant component for calculating emissions in atmospheric models. Geoscientific Model Development. 2014 Jul 14;7(4):1409-17.
33. Crippa M, Guizzardi D, Muntean M, Schaaf E, Dentener F, Van Aardenne JA, et al. Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4. 3.2. Earth Syst. Sci. Data. 2018 Oct 26;10(4):1987-2013.
34. Kuhns H, Green M, Etyemezian V, Watson J, Pitchford M. Big bend regional aerosol and visibility observational (BRAVO) study emissions inventory. Report prepared for BRAVO Steering Committee, Desert Research Institute, Las Vegas, Nevada. 2003.
35. Vestreng V. Emission data reported to UNECE/EMEP: Quality assurance and trend analysis & presentation of WebDab: MSC-W status report 2002.
36. Li M, Zhang Q, Kurokawa JI, Woo JH, He K, Lu Z, Ohara T, Song Y, Streets DG, Carmichael GR, Cheng Y. MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP. Atmospheric Chemistry and Physics. 2017 Jan 20;17(2):935-63.
37. Zheng B, Huo H, Zhang Q, Yao ZL, Wang XT, Yang XF, Liu H, He KB. High-resolution mapping of vehicle emissions in China in 2008. Atmospheric Chemistry and Physics. 2014 Sep 17;14(18):9787-805.
38. Lu Z, Zhang Q, Streets DG. Sulfur dioxide and primary carbonaceous aerosol emissions in China and India, 1996–2010. Atmospheric Chemistry and Physics. 2011 Sep 23;11(18):9839-64.
39. Stettler ME, Eastham S, Barrett SR. Air quality and public health impacts of UK airports. Part I: Emissions. Atmospheric environment. 2011 Oct 1;45(31):5415-24.
40. Holmes CD, Prather MJ, Vinken GC. The climate impact of ship NOx emissions: an improved estimate accounting for plume chemistry. Atmospheric Chemistry and Physics. 2014 Jul 4;14(13):6801-12.
41. Schultz MG, Heil A, Hoelzemann JJ, Spessa A, Thonicke K, Goldammer JG, Held AC, Pereira JM, van Het Bolscher M. Global wildland fire emissions from 1960 to 2000. Global Biogeochemical Cycles. 2008 Jun;22(2).
42. Giglio L, Randerson JT, Van Der Werf GR. Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). Journal of Geophysical Research: Biogeosciences. 2013 Mar;118(1):317-28.
43. Ge C, Wang J, Carn S, Yang K, Ginoux P, Krotkov N. Satellite-based global volcanic SO2 emissions and sulfate direct radiative forcing during 2005–2012. Journal of Geophysical Research: Atmospheres. 2016 Apr 16;121(7):3446-64.

44. RC Hudman, NE Moore, AK Mebust, RV Martin, AR Russell, LC Valin, and RC Cohen. Steps towards a mechanistic model of global soil nitric oxide emissions: implementation and space based-constraints. Atmospheric Chemistry & Physics, 12(16), 2012.
45. AB Guenther, Xiaoyan Jiang, CL Heald, T Sakulyanontvittaya, Ti any Duhl, LK Emmons, and X Wang. The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2. 1): an extended and updated framework for modeling biogenic emissions, 2012.
46. Zender CS, Bian H, Newman D. Mineral Dust Entrainment and Deposition (DEAD) model: Description and 1990s dust climatology. Journal of Geophysical Research: Atmospheres. 2003 Jul 27;108(D14).
47. Ahmadov R, McKeen SA, Robinson AL, Bahreini R, Middlebrook AM, De Gouw JA, et al. A volatility basis set model for summertime secondary organic aerosols over the eastern United States in 2006. Journal of Geophysical Research: Atmospheres. 2012 Mar 27;117(D6).
48. Philip S, Martin RV, Pierce JR, Jimenez JL, Zhang Q, Canagaratna MR, Spracklen DV, Nowlan CR, Lamsal LN, Cooper MJ, Krotkov NA. Spatially and seasonally resolved estimate of the ratio of organic mass to organic carbon. Atmospheric Environment. 2014 Apr 1;87:34-40.
49. Li C, Martin RV, van Donkelaar A, Boys BL, Hammer MS, Xu JW, Marais EA, Reff A, Strum M, Ridley DA, Crippa M. Trends in chemical composition of global and regional population-weighted fine particulate matter estimated for 25 years. Environmental science & technology. 2017 Oct 3;51(19):11185-95.
50. Homer C, Huang C, Yang L, Wylie B, Coan M. Development of a 2001 national land-cover database for the United States. Photogrammetric Engineering & Remote Sensing. 2004 Jul 1;70(7):829-40.
51. GEOS-Chem Wiki. Olson land map. Accessed: 25th January 2021. http://wiki.seas.harvard.edu/geos-chem/index.php/Olson_land_map
52. Dennis R, Fox T, Fuentes M, Gilliland A, Hanna S, Hogrefe C, Irwin J, Rao ST, Scheffe R, Schere K, Steyn D. A framework for evaluating regional-scale numerical photochemical modeling systems. Environmental Fluid Mechanics. 2010 Aug;10(4):471-89.
53. Diao M, Holloway T, Choi S, O'Neill SM, Al-Hamdan MZ, Van Donkelaar A, et al. Methods, availability, and applications of PM2. 5 exposure estimates derived from ground measurements, satellite, and atmospheric models. Journal of the Air & Waste Management Association. 2019 Dec 2;69(12):1391-414.
54. Emery C, Liu Z, Russell AG, Odman MT, Yarwood G, Kumar N. Recommendations on statistics and benchmarks to assess photochemical model performance. Journal of the Air & Waste Management Association. 2017 Apr 27;67(5):582-98.
55. Boylan JW, Russell AG. PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models. Atmospheric environment. 2006 Aug 1;40(26):4946-59.
56. Clune L, Kouatchou J, Putman WM, Thompson MA, Trayanov AL, Molod AM, Martin RV, Jacob DJ. GEOS-Chem High Performance (GCHP): A next-generation implementation of the GEOS-Chem chemical transport model for massively parallel applications.
57. Van Damme M, Clarisse L, Whitburn S, Hadji-Lazaro J, Hurtmans D, Clerbaux C, et al. Industrial and agricultural ammonia point sources exposed. Nature. 2018 Dec;564(7734):99-103.
58. Akherati A, He Y, Coggon MM, Koss AR, Hodshire AL, Sekimoto K, et al. Oxygenated aromatic compounds are important precursors of secondary organic aerosol in biomass-burning emissions. Environmental Science & Technology. 2020 Jun 19;54(14):8568-79.
59. Zheng B, Zhang Q, Zhang Y, He KB, Wang K, Zheng GJ, et al. Heterogeneous chemistry: a mechanism missing in current models to explain secondary inorganic

aerosol formation during the January 2013 haze episode in North China. Atmospheric Chemistry and Physics. 2015 Feb 25;15(4):2031-49.
60. United Nations, Department of Economic and Social Affairs, Population Division (2019). World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420). New York: United Nations.
61. Muller NZ, Mendelsohn R. Efficient pollution regulation: getting the prices right. American Economic Review. 2009 Dec;99(5):1714-39.
62. Coffman E, Burnett RT, Sacks JD. Quantitative Characterization of Uncertainty in the Concentration–Response Relationship between Long-Term PM2. 5 Exposure and Mortality at Low Concentrations. Environmental Science & Technology. 2020 Jul 23;54(16):10191-200.

## Figures and Tables

**Table 1.** $PM_{2.5}$ and precursor emissions inputs into GEOS-Chem and Global InMAP.

| Pollutant | GEOS-Chem (Tg/yr) | Global InMAP (Tg/yr) | Global InMAP data sources | Maximum resolution |
|---|---|---|---|---|
| *Anthropogenic* | | | | |
| $PM_{2.5}$ | 24.45 | 32.93 | EDGAR, NEI, CAC, MEIC | 0.25° × 0.25° |
| $NH_3$ | 51.52 | 47.39 | EDGAR, CAC, NEI, MIX, MEIC | 0.25° × 0.25° |
| $SO_x$ | 84.33 | 84.33 | EDGAR, BRAVO, EMEP, NEI, CAC, MIX, MEIC, Lu et al. | 2° × 2.5° |
| $NO_x$ | 64.85 | 76.28 | EDGAR, BRAVO, EMEP, NEI, CAC, MIX, MEIC, AEIC | 0.25° × 0.25° |
| NMVOC | - [b] | 58.15 | EDGAR | 0.1° × 0.1° |
| *Natural* | | | | |
| $PM_{2.5}$ | 244.53 | 244.53 | DEAD, GEOS-Chem diagnostics | 2° × 2.5° |
| $NH_3$ | 17.38 | 15.97 | GEIA | 0.25° × 0.25° |
| $SO_x$ | 28.32 | 0.42[a] | Ge et al., GEOS-Chem diagnostics | 2° × 2.5° |
| $NO_x$ | 28.02 | 16.60[a] | Hudman et al., GEOS-Chem diagnostics | 2° × 2.5° |
| NMVOC | - [b] | 553.14 | MEGAN, GEOS-Chem diagnostics | 2° × 2.5° |
| *Biomass burning* | | | | |
| $PM_{2.5}$ | 35.30 | 35.30 | GFED-4 | 0.25° × 0.25° |
| $NH_3$ | 4.24 | 4.24 | GFED-4 | 0.25° × 0.25° |
| $SO_x$ | 2.25 | 2.25 | GFED-4 | 0.25° × 0.25° |
| $NO_x$ | 20.28 | 20.28 | GFED-4 | 0.25° × 0.25° |
| NMVOC | - [b] | 5.10 | RETRO | 0.5° × 0.5° |

[a]Only $NO_x$ and $SO_x$ emissions in the lowest vertical layer were used in Global InMAP, yet the majority of natural $NO_x$ and $SO_x$ emissions are emitted from lightning and volcanoes at higher levels. [b]Not all NMVOC emissions from GEOS-Chem simulation are reported.

**Table 2.** Area- and population-weighted normalized mean bias (NMB) and error (NME) for Global InMAP predicted changes in concentrations against changes in concentrations from GEOS-Chem or US InMAP, arising from scenarios of changes in emissions. Positive bias indicates that Global InMAP has higher average concentration changes than the other model.

| Model comparison | Scenario | Weighting | NME (%) | NMB (%) |
|---|---|---|---|---|
| Global InMAP against GEOS-Chem | $NH_3$ increase from agricultural soils | area-wtd. | 118.2 | 58.7 |
| | | population-wtd. | 88.6 | 52.5 |
| | $NO_x$ increase from road transportation | area-wtd. | 180.7 | 96.2 |
| | | population-wtd. | 102.2 | 40.8 |
| | $SO_x$ increase from power generation | area-wtd. | 181.3 | 120.7 |
| | | population-wtd. | 273.4 | 259.3 |
| Global InMAP against US InMAP | Coal-powered electricity | area-wtd. | 38.4 | -18.8 |
| | | population-wtd. | 38.7 | -10.5 |
| | Gasoline passenger vehicles | area-wtd. | 48.4 | -23.0 |
| | | population-wtd. | 48.8 | -46.7 |

**Figure 1.** Annual-average ground-level total $PM_{2.5}$ concentrations from the Global InMAP and GEOS-Chem simulations for year 2016.

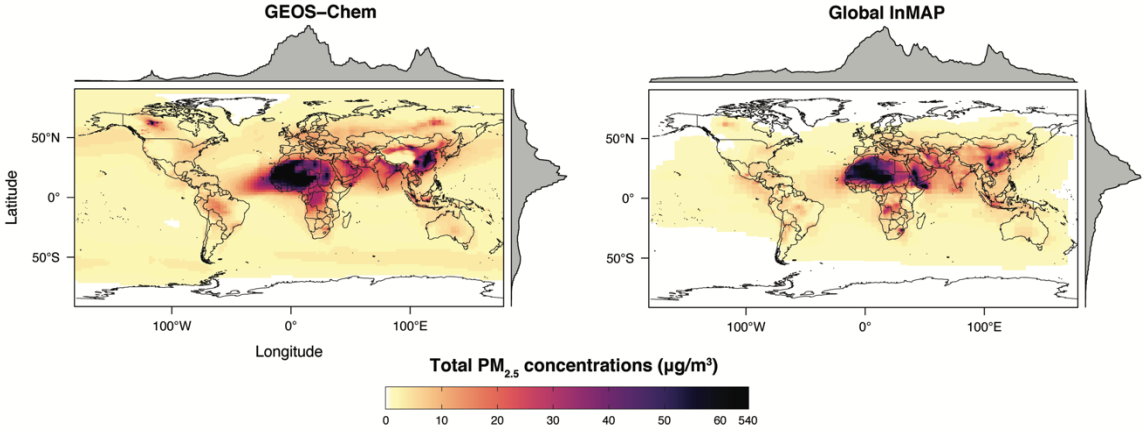
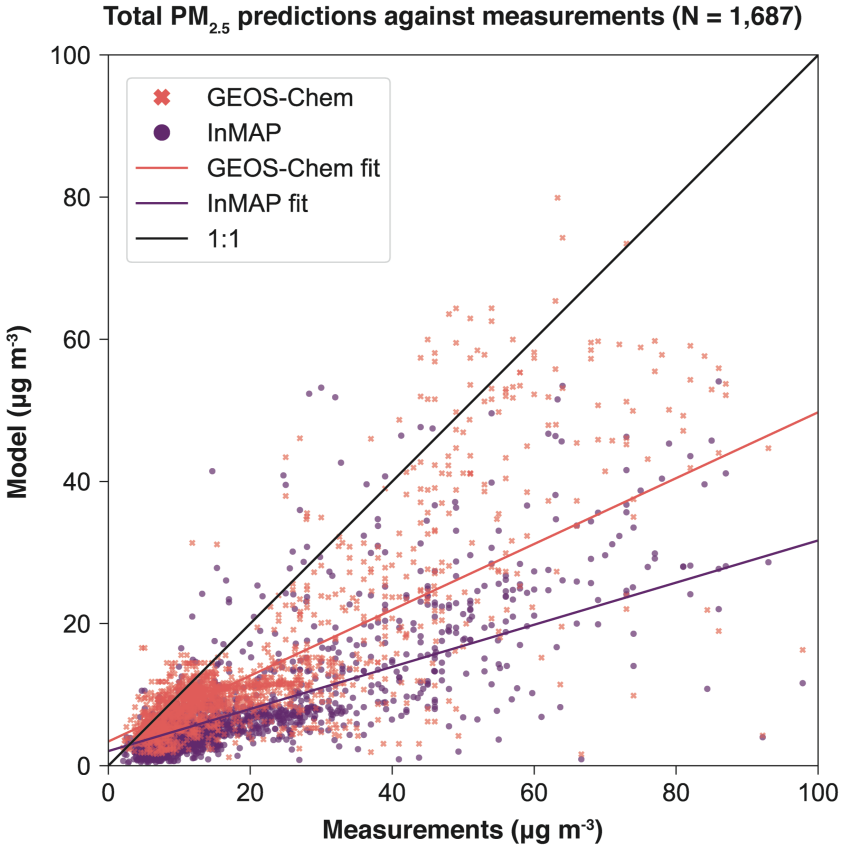
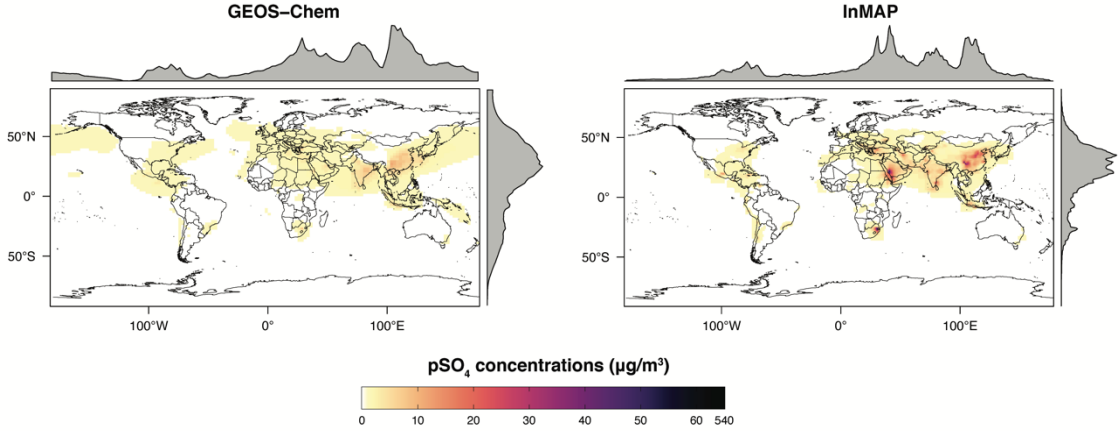


**Figure 2.** Annual-average total $PM_{2.5}$ concentrations from the Global InMAP and GEOS-Chem simulations against measurements. Only values ≤100 μg $m^{-3}$ are plotted here, excluding 25 (1.5%) model-measurement pairs (full figure shown in Supplementary Information, Figure S2).



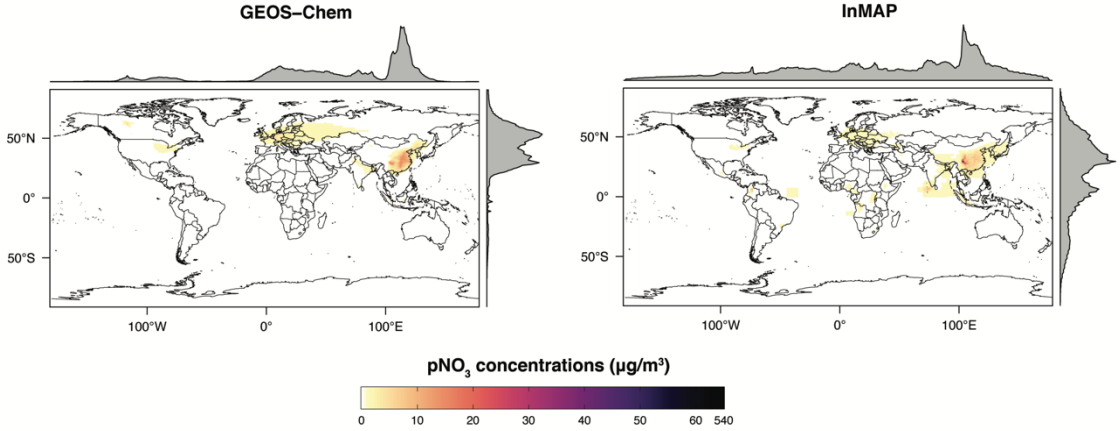

| | NMB | NME | Slope | $R^2$ |
|---|---|---|---|---|
| **GEOS-Chem** | -37% | 41% | 0.46 | 0.55 |
| *- pop-wtd* | -37% | 40% | 0.47 | 0.75 |
| **Global InMAP** | -60% | 63% | 0.30 | 0.35 |
| *- pop-wtd* | -61% | 64% | 0.31 | 0.56 |

690 **Figure 3.** Global InMAP and GEOS-Chem annual-average ground-level $pSO_4$ concentrations.



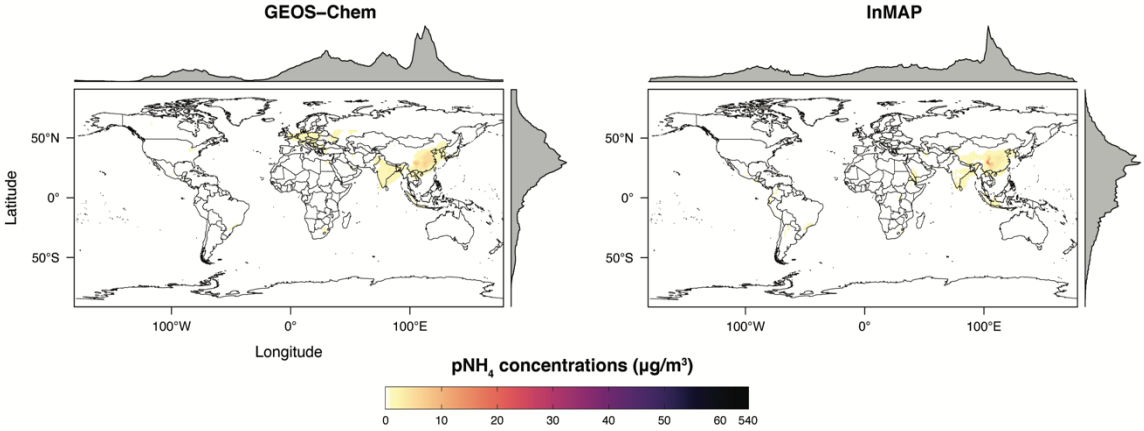

691
692

693 **Figure 4.** Global InMAP and GEOS-Chem annual-average ground-level $pNO_3$ concentrations.



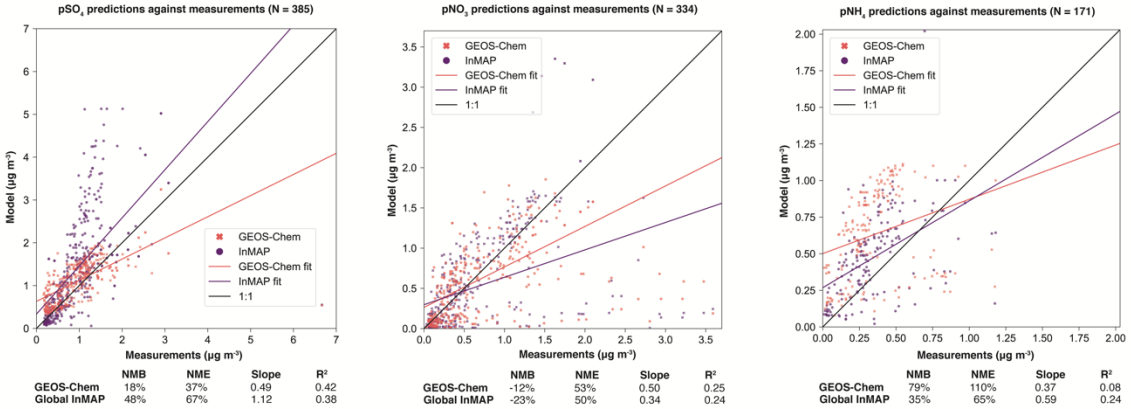

694
695

696 **Figure 5.** Global InMAP and GEOS-Chem annual-average ground-level $pNH_4$ concentrations.




697
698

**Figure 6.** Global InMAP and GEOS-Chem annual-average $pSO_4$, $pNO_3$, and $pNH_4$ concentrations against measurements.

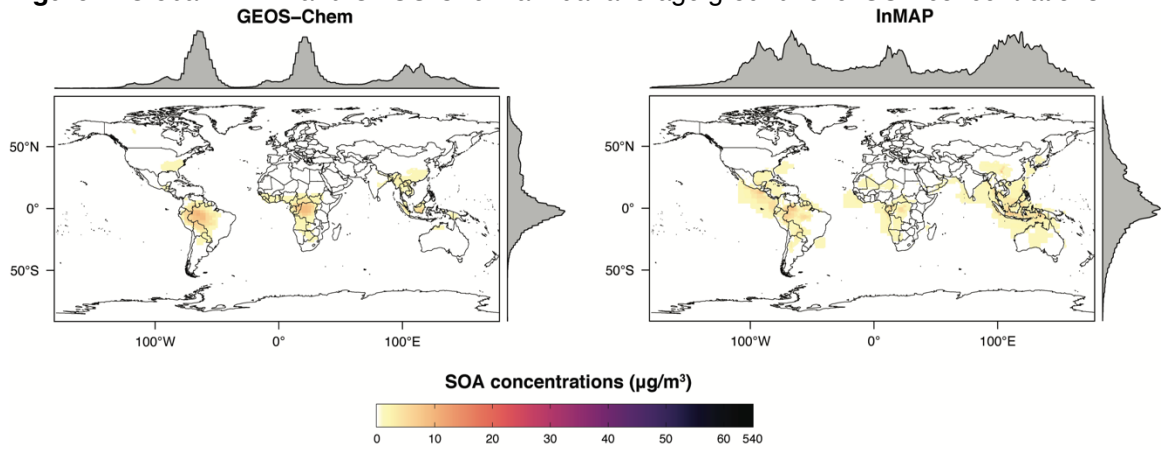



| | NMB | NME | Slope | $R^2$ |
|---|---|---|---|---|
| GEOS-Chem | 18% | 37% | 0.49 | 0.42 |
| Global InMAP | 48% | 67% | 1.12 | 0.38 |

| | NMB | NME | Slope | $R^2$ |
|---|---|---|---|---|
| GEOS-Chem | -12% | 53% | 0.50 | 0.25 |
| Global InMAP | -23% | 50% | 0.34 | 0.24 |

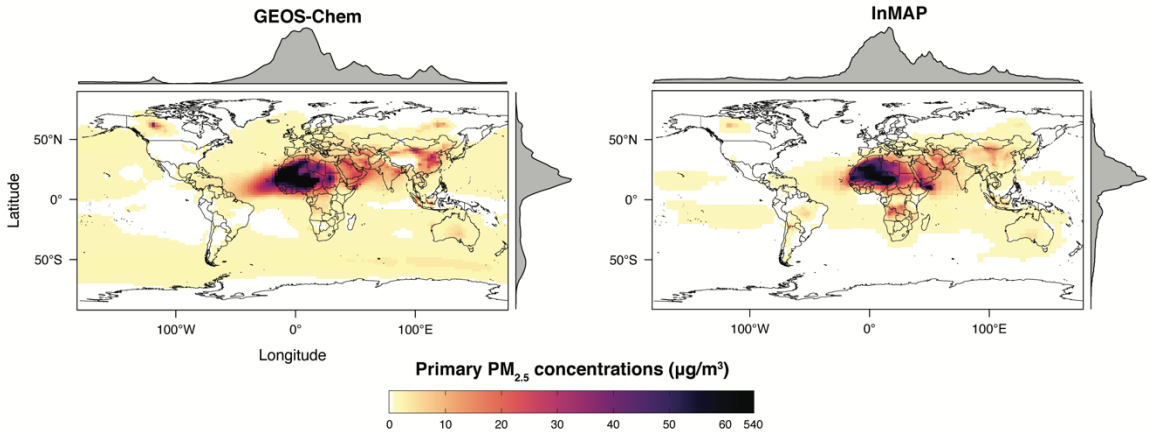
| | NMB | NME | Slope | $R^2$ |
|---|---|---|---|---|
| GEOS-Chem | 79% | 110% | 0.37 | 0.08 |
| Global InMAP | 35% | 65% | 0.59 | 0.24 |

703 **Figure 7.** Global InMAP and GEOS-Chem annual-average ground-level SOA concentrations.




704
705

706 **Figure 8.** Global InMAP and GEOS-Chem annual-average ground-level primary $PM_{2.5}$
707 concentrations.




708
709

**Figure 9.** Comparison between Global InMAP and GEOS-Chem for predicting changes in $pNO_3$ concentrations from a 100% increase in $NO_x$ emissions from road transportation.

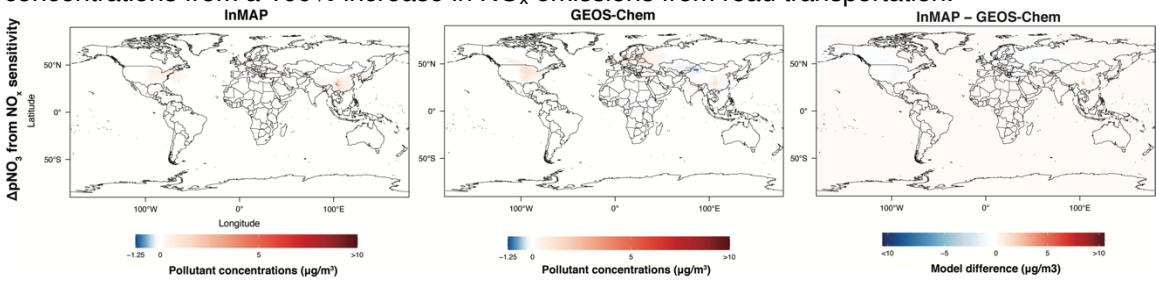


**Figure 10.** Comparison between Global InMAP and GEOS-Chem for predicting changes in $pSO_4$ concentrations from a 100% increase in $SO_x$ emissions from power generation.

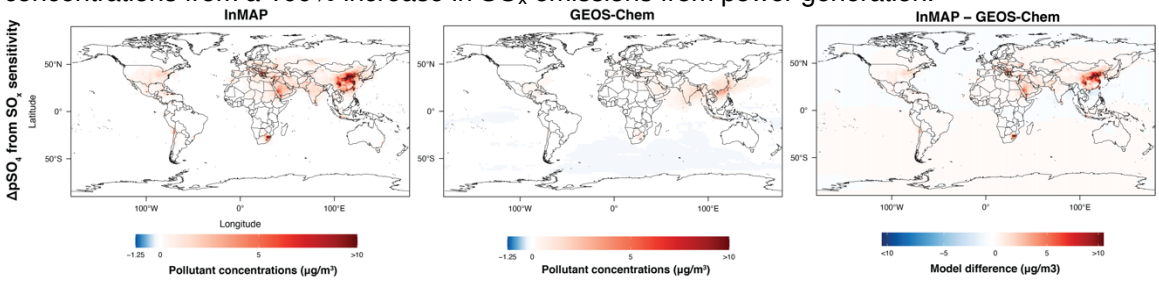


**Figure 11.** Comparison between Global InMAP and GEOS-Chem for predicting changes in $pNH_4$ concentrations from a 100% increase in $NH_3$ emissions from agricultural soils.

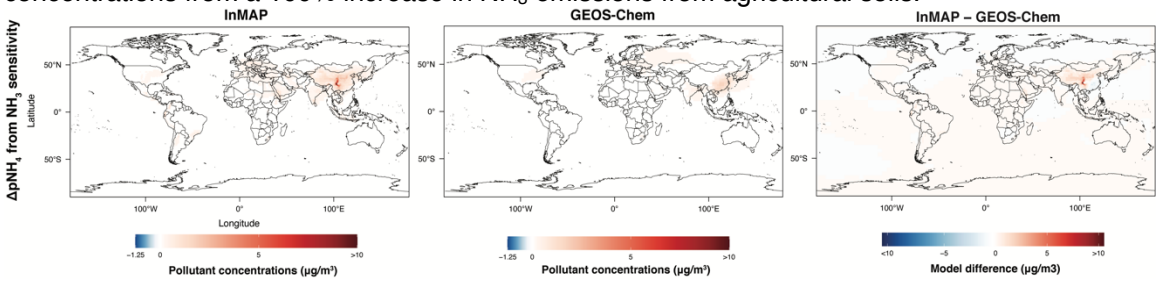
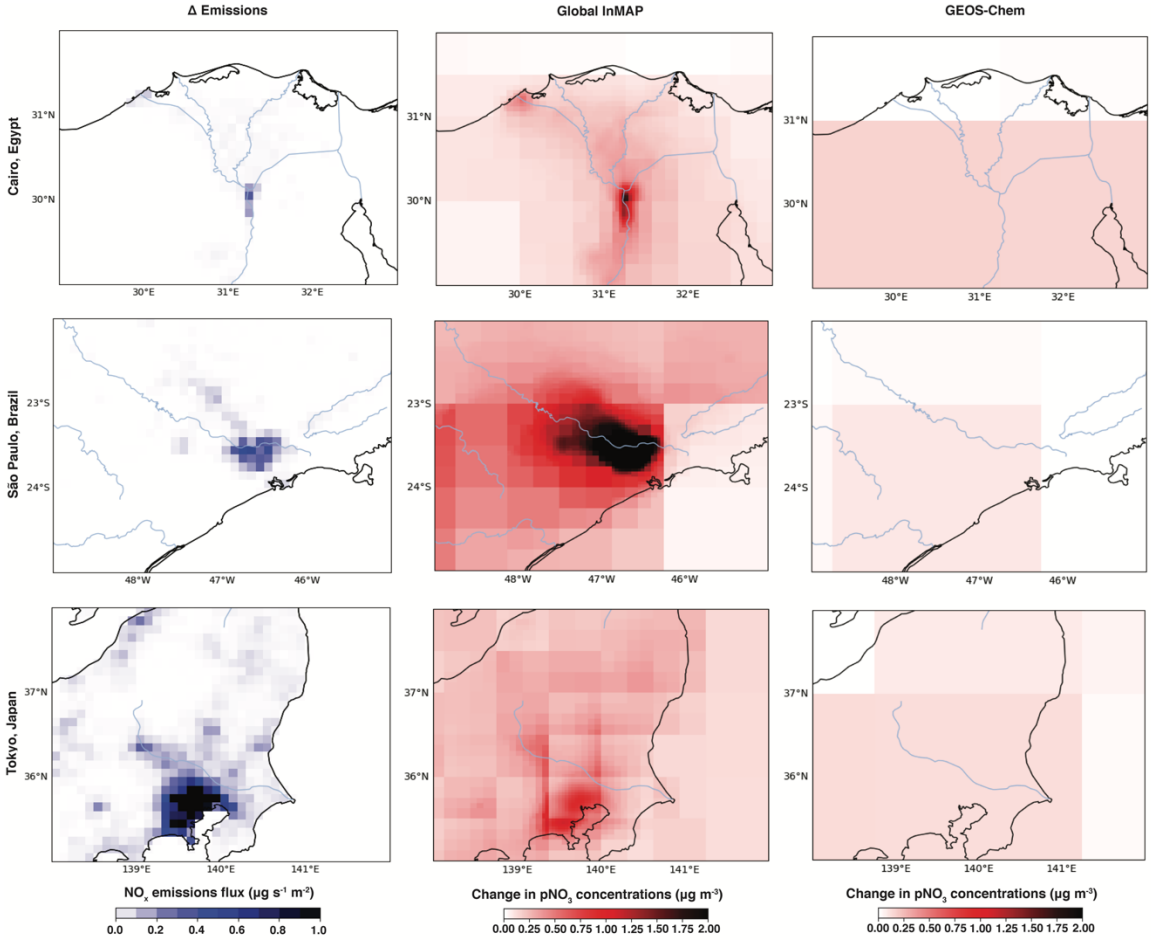


**Figure 12.** *First column:* 100% increase in $NO_x$ emissions from road transport across Cairo, São Paulo, and Tokyo. *Second and third column:* resulting changes in $pNO_3$ concentrations predicted by the Global InMAP and the GEOS-Chem simulations. For each map, blue lines indicate rivers and black lines indicate land borders.



**Global, high-resolution reduced-complexity air quality modeling using InMAP (Intervention Model for Air Pollution)**

Sumil K. Thakrar, Christopher W. Tessum, Joshua S. Apte, Srinidhi Balasubramanian, Dylan B. Millet, Spyros N. Pandis, Julian D. Marshall, Jason D. Hill.

**Global, high-resolution, reduced-complexity air quality modeling using InMAP (Intervention Model for Air Pollution)**

Sumil K. Thakrar[1,2], Christopher W. Tessum[3], Joshua S. Apte[4,5], Srinidhi Balasubramanian[1], Dylan B. Millet[6], Spyros N. Pandis[7,8], Julian D. Marshall[9], Jason D. Hill[1*].

[1]Department of Bioproducts & Biosystems Engineering, University of Minnesota, St Paul, MN, USA.

[2]Department of Applied Economics, University of Minnesota, St Paul, MN, USA.

[3]Department of Civil and Environmental Engineering, University of Illinois at Urbana–Champaign, Urbana, IL, USA.

[4]Department of Civil and Environmental Engineering, University of California, Berkeley, Berkeley, CA, USA.

[5]School of Public Health, University of California, Berkeley, Berkeley, CA, USA.

[6]Department of Soil, Water, and Climate, University of Minnesota, St Paul, MN, USA.

[7]Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA, USA.

[8]Department of Chemical Engineering, University of Patras, Greece.

[9]Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, USA.

**Corresponding author:** Sumil K. Thakrar

**Email:** sthakrar@umn.edu

**Supplementary Information Text**

*Measurement data description*

Ground-level measurements of total $PM_{2.5}$, $pNH_4$, $pNO_3$, and $pSO_4$ concentrations across year 2016 were compiled from the World Health Organization database and supplemented with additional measurements from other official channels such as governmental and non-governmental agencies (see Table S2). Included measurements were vetted according to quality control criteria, including those used by the 2012 United States National Ambient Air Quality Standards. Only measurements that directly measured $PM_{2.5}$ were included; $PM_{10}$ measurements that were converted to $PM_{2.5}$ were excluded. Further, data without correct latitude and longitude were excluded. Global InMAP directly estimates annual-average pollutant concentrations, so measurement data from each monitoring site were averaged across the year. To avoid temporal biases across the day, all measurement data were averaged daily values of pollutant concentrations. To avoid seasonal biases, measurements had to be reported for at least 75% of days in the year from each monitoring site included in our dataset.

After vetting, the final dataset of annual-average pollutant concentrations included ~1,700 total $PM_{2.5}$ data points across 62 countries; 171 $pNH_4$ data points across 1 country (the US); 334 $pNO_3$ data points across 4 countries, and 385 $pSO_4$ data points across 12 countries. The final dataset is provided in Dataset S1.

*Performance metric and criteria descriptions*

Normalized mean bias and error (NMB and NME), are given by:

$$NMB = \frac{\sum_i P_i - O_i}{\sum_i O_i} \times 100$$

$$NME = \frac{\sum_i |P_i - O_i|}{\sum_i O_i} \times 100$$

where, for monitor location $i$, $P_i$ are the model predictions and $O_i$ are the observations of annual-average pollutant concentrations.

Model criteria for $PM_{2.5}$, $pSO_4$, and $pNH_4$ concentrations, are $R^2 \geq 0.16$, NME ≤ 50%, and |NMB| ≤ 30%. For $pNO_3$, model criteria are NME ≤ 115%, |NMB| ≤ 65% (with no criteria for $R^2$).

For model-to-model comparisons, weighted NMB and NME are given by:

$$NMB_{weighted} = \frac{\sum_i (GI_i - M_i) \times w_i}{\sum_i M_i \times w_i} \times 100$$

802 $$NME_{weighted} = \frac{\sum_i |GI_i - M_i| \times w_i}{\sum_i M_i \times w_i} \times 100$$

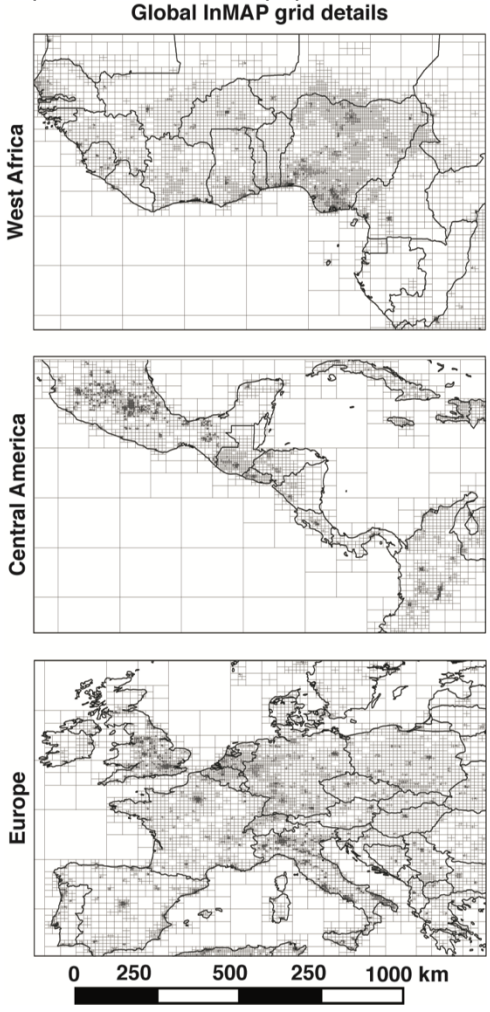
803

804 Where $w_i$ are the weights (areas or population counts) for each grid cell $i$, $GI$ are the Global
805 InMAP predictions, and $M$ are the predictions from the other model (GEOS-Chem or US InMAP).

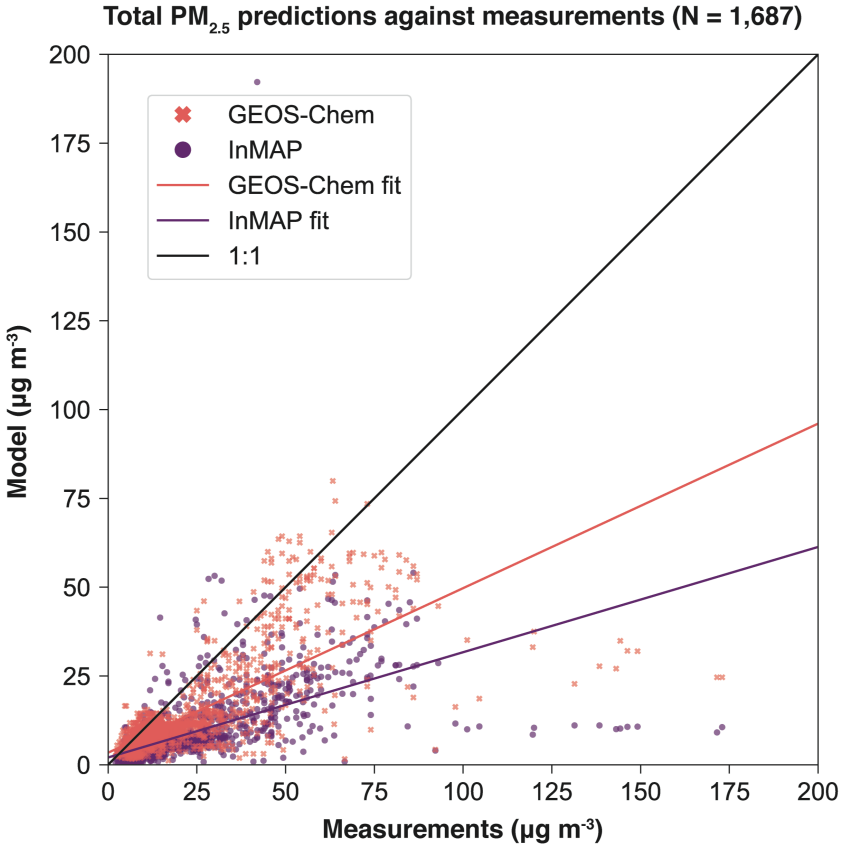
806

807

**Supplementary Figure 1.** Detail of the Global InMAP horizontal computational grid over West Africa, Central America, and Europe for illustration. Grid cells are as small as 0.04° × 0.03° (~4 km length) in areas with a higher population such as Lagos in Nigeria, San Salvador in El Salvador, and London in the United Kingdom. Grid cells are as large as 5° × 4° (~500 km length) in places with a lower population, such as across the Atlantic Ocean.

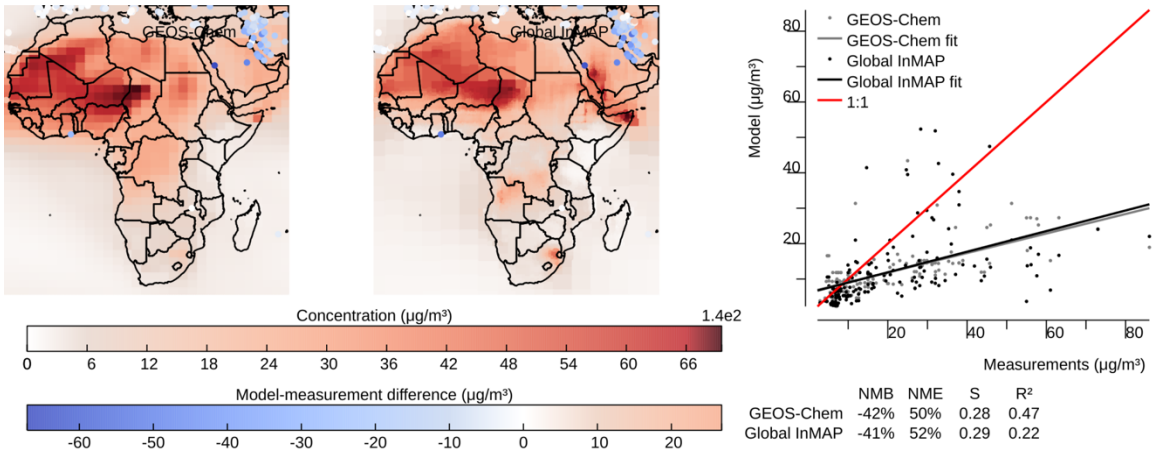


**Supplementary Figure 2.** InMAP and GEOS-Chem annual-average primary $PM_{2.5}$ concentrations against measurements, including outliers (above 100 μg $m^{-3}$). Pop-wtd: population-weighted metrics.



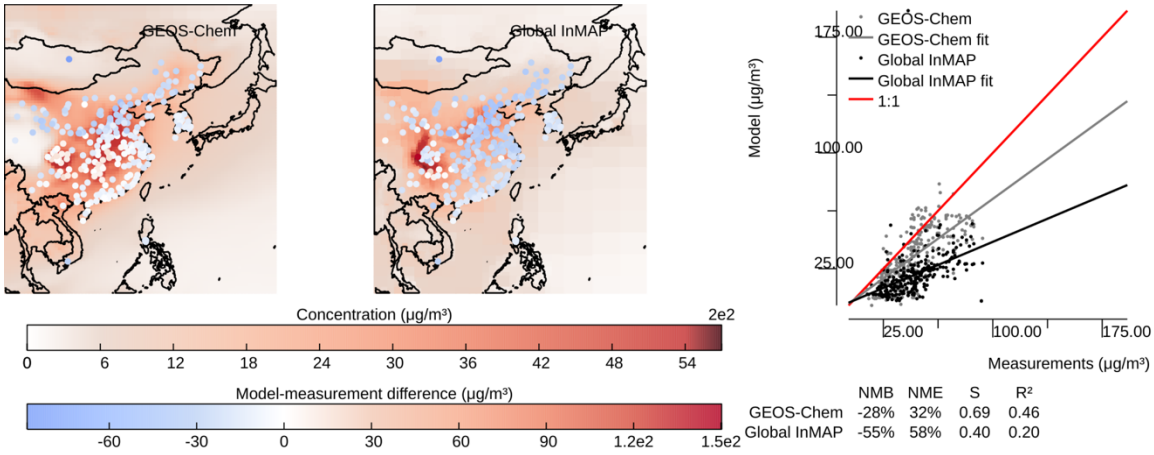

| | NMB | NME | Slope | $R^2$ |
|---|---|---|---|---|
| **GEOS-Chem** | -37% | 41% | 0.46 | 0.55 |
| *- pop-wtd* | -37% | 40% | 0.47 | 0.75 |
| **Global InMAP** | -60% | 63% | 0.30 | 0.35 |
| *- pop-wtd* | -61% | 64% | 0.31 | 0.56 |

**Supplementary Figure 3.** Performance of Global InMAP and GEOS-Chem simulations against total annual-average $PM_{2.5}$ measurements for Africa. Dots on each map show measurement site locations, whose color corresponds to the model-measurement difference in $PM_{2.5}$ concentrations.

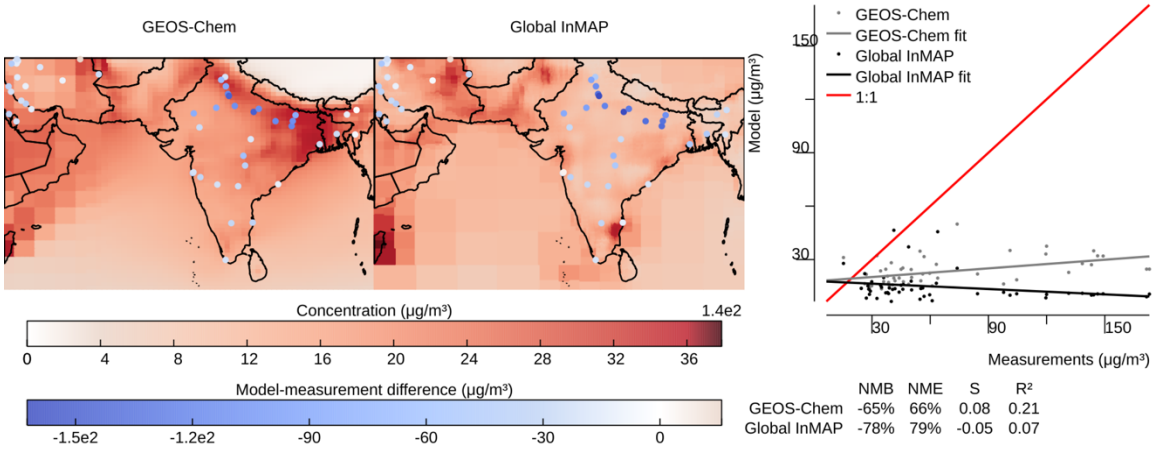



| | NMB | NME | S | $R^2$ |
|---|---|---|---|---|
| GEOS-Chem | -42% | 50% | 0.28 | 0.47 |
| Global InMAP | -41% | 52% | 0.29 | 0.22 |

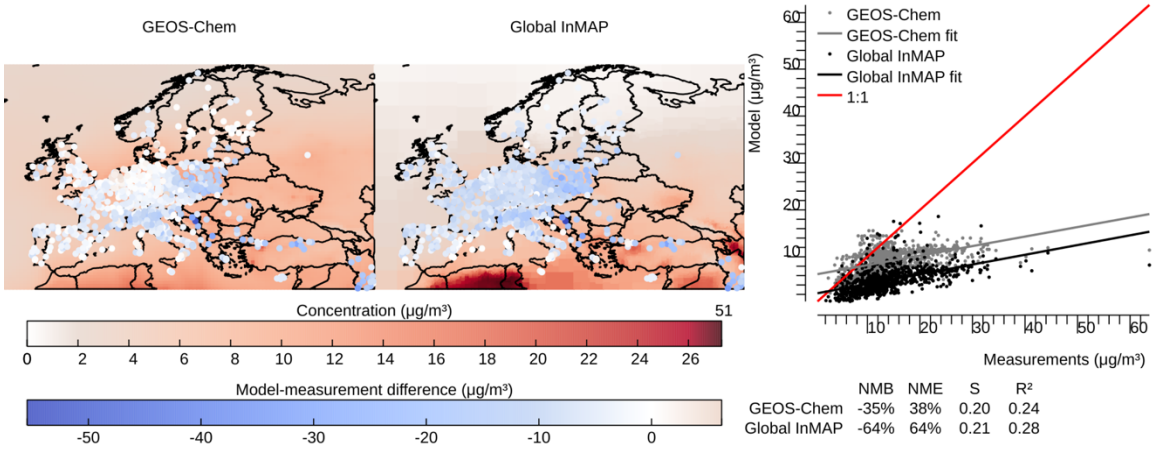
**Supplementary Figure 4.** Performance of Global InMAP and GEOS-Chem simulations against total annual-average $PM_{2.5}$ measurements for East Asia. Dots on each map show measurement site locations, whose color corresponds to the model-measurement difference in $PM_{2.5}$ concentrations.



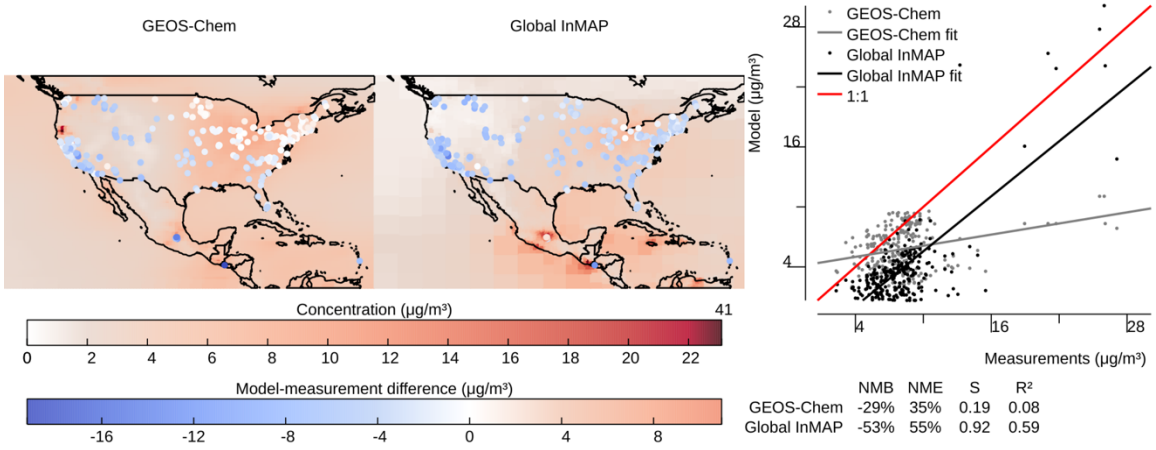
**Supplementary Figure 5.** Performance of Global InMAP and GEOS-Chem simulations against total annual-average $PM_{2.5}$ measurements for South Asia. Dots on each map show measurement site locations, whose color corresponds to the model-measurement difference in $PM_{2.5}$ concentrations.



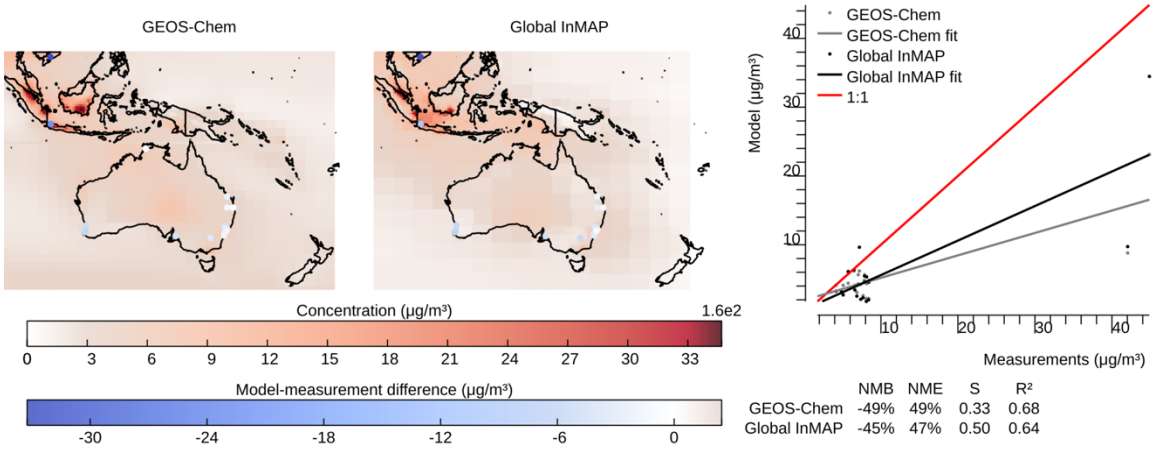
**Supplementary Figure 6.** Performance of Global InMAP and GEOS-Chem simulations against total annual-average $PM_{2.5}$ measurements for Europe. Dots on each map show measurement site locations, whose color corresponds to the model-measurement difference in $PM_{2.5}$ concentrations.



**Supplementary Figure 7.** Performance of Global InMAP and GEOS-Chem simulations against total annual-average $PM_{2.5}$ measurements for North and Central America. Dots on each map show measurement site locations, whose color corresponds to the model-measurement difference in $PM_{2.5}$ concentrations.

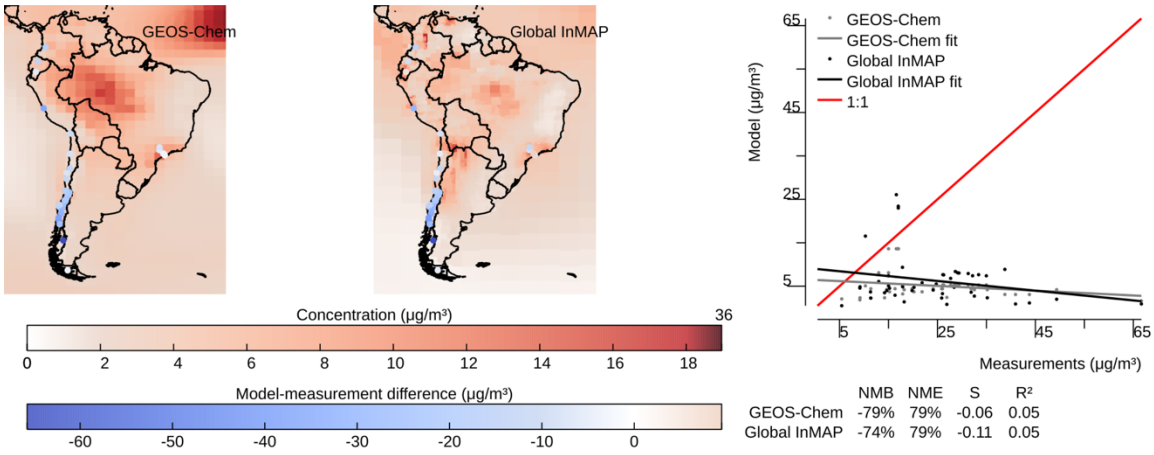


**Supplementary Figure 8.** Performance of Global InMAP and GEOS-Chem simulations against total annual-average $PM_{2.5}$ measurements for Oceana. Dots on each map show measurement site locations, whose color corresponds to the model-measurement difference in $PM_{2.5}$ concentrations.

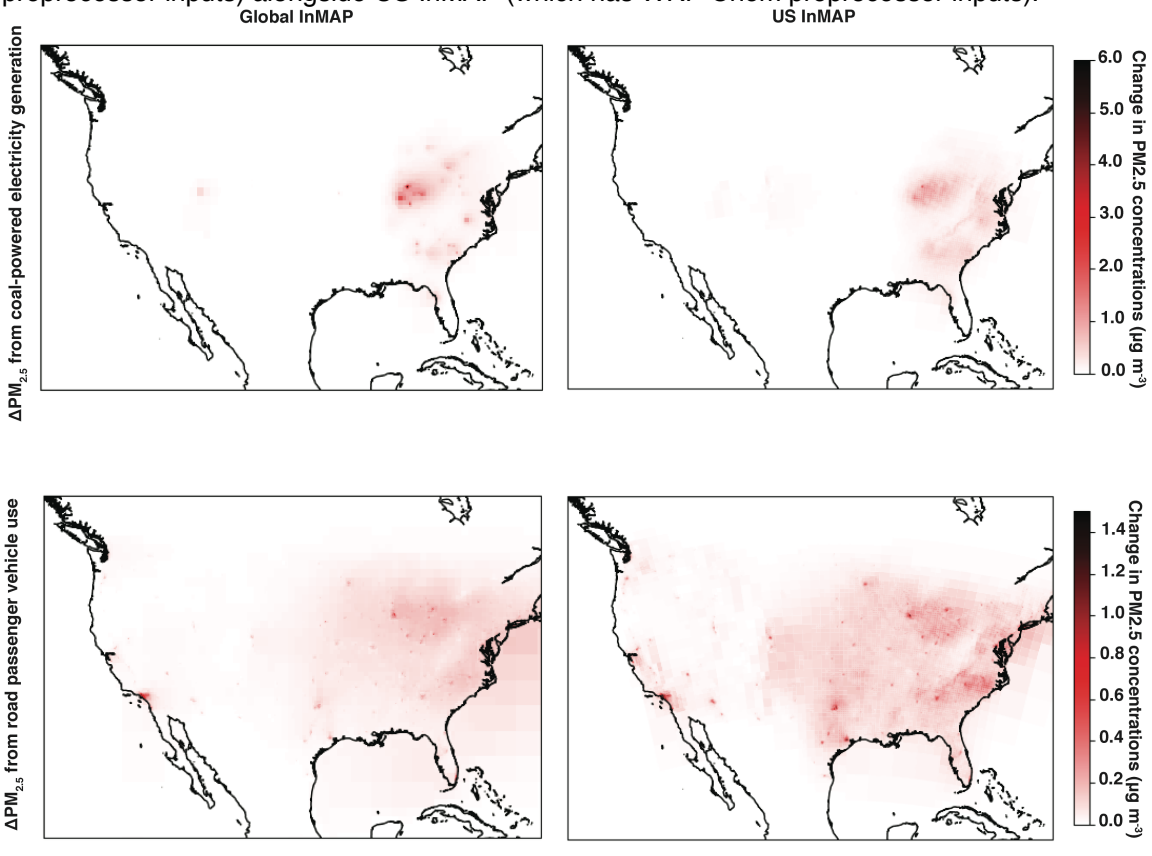



| | NMB | NME | S | $R^2$ |
|---|---|---|---|---|
| GEOS-Chem | -49% | 49% | 0.33 | 0.68 |
| Global InMAP | -45% | 47% | 0.50 | 0.64 |

**Supplementary Figure 9.** Performance of Global InMAP and GEOS-Chem simulations against total annual-average $PM_{2.5}$ measurements for South America. Dots on each map show measurement site locations, whose color corresponds to the model-measurement difference in $PM_{2.5}$ concentrations.



**Supplementary Figure 10.** Changes in Total $PM_{2.5}$ concentrations from road vehicle emissions and from power generation emissions as predicted by Global InMAP (which has GEOS-Chem preprocessor inputs) alongside US InMAP (which has WRF-Chem preprocessor inputs).



**Supplementary Table 1.** Names and descriptions of GEOS-Chem outputs used to calculate Global InMAP parameters.

| Name(s) | Description and use in Global InMAP preprocessor |
|---|---|
| BENZ, TOLU, XYLE, NAP, POG1, POG2 | Anthropogenic VOCs that are SOA precursors; used to determine VOC/SOA partitioning |
| ASOA1, ASOA2, ASOA3, ASOAN | Anthropogenic SOA; used to determine VOC/SOA partitioning |
| ISOP, LIMO, MTPA, MTPO | Biogenic VOCs that are SOA precursors; used for model evaluation |
| TSOA0, TSOA1, TSOA2, TSOA3, SOAGX, SOAMG, SOAIE, SOAME, LVOCOA, ISN1OA | Biogenic SOA; used for model evaluation |
| NO, NO2 | Components of $NO_x$; used to determine $NO_x$ /$pNO_3$ partitioning |
| NIT, NITs | Components of $pNO_3$; used to determine $NO_x$ /$pNO_3$ partitioning |
| SO2 | Gaseous $SO_2$ and sulfate; used to determine $SO_x$ /$pSO_4$ partitioning |
| SO4, SO4s, DMS | Particulate $SO_4$; used to determine $SO_x$ /$pSO_4$ partitioning |
| NH3 | Ammonia; used to determine $NH_3$/$pNH_4$ partitioning |
| NH4 | Particulate Ammonium; used to determine $NH_3$/$pNH_4$ partitioning |
| 1.33×(NH4 + NIT + SO4) + BCPI + BCPO + 1.4×(POA1 + POA2) + 2.1×(OPOA1 + OPOA2) + 1.16×(TSOA1 + TSOA2 + TSOA3 + ASOAN + ASOA1 + ASOA2 + ASOA3 + SOAGX + INDIOL + SOAMG + SOAIE + SOAME + LVOCOA + ISN1OA) + DST1 + 0.38×DST2 + 1.86×SALA | Total $PM_{2.5}$ concentration in the baseline simulation; used for model evaluation |
| Z0M | Momentum roughness length |
| U, V, OMEGA | Wind fields; used to determine advection and mixing coefficients |
| PBLH | Planetary boundary layer height; used to determine mixing coefficients |
| HFLUX | Surface heat flux; used to determine mixing and dry deposition |
| USTAR | Friction velocity; used to determine mixing and dry deposition |
| T | Temperature; used to calculate chemical reaction rates and plume rise |
| PS, P | Base state pressure plus perturbation pressure; used to calculate chemical reaction rates and plume rise |
| OH, H2O2 | Hydroxyl radical and hydrogen peroxide concentrations; used to calculate chemical reaction rates |
| FRSNO | Fraction of land covered by snow; used to calculate dry deposition |
| PFLCU, PFLLSAN | Mixing ratio of rain; used to calculate wet |

| | |
|---|---|
| | deposition |
| CLOUD | Fraction of grid cell covered by clouds; used to calculate wet deposition |
| QL | Cloud mixing ratio; used to calculate aqueous-phase chemical reaction rates |
| AIRDEN | Inverse air density; used to calculate mixing and to convert between mixing ratio and mass concentration |
| PARDF, PARDR | Downward shortwave and longwave radiative flux at ground level; used to calculate dry deposition |

869
870

**Supplementary Table 2.** Measurement data sources for 2016 used in evaluating Global InMAP and GEOS-Chem annual-average predictions of pollutant concentrations. The World Health Organization data includes data from other regulatory sources and monitoring networks globally.

| Region | Data | Source |
|---|---|---|
| Global | $PM_{2.5}$ | World Health Organization |
| Europe | $PM_{2.5}$, $pNO_3$, $pSO_4$ | European Environment Agency |
| Canada | $PM_{2.5}$ | National Air Pollution Surveillance Program |
| United States of America | $PM_{2.5}$, $pNO_3$, $pSO_4$, $pNH_4$ | Environmental Protection Agency |
| India | $PM_{2.5}$ | Central Pollution Control Board |
| Australia | $PM_{2.5}$ | Australian Government State of the Environment |

**Supplementary Table 3.** Global InMAP and GEOS-Chem performance metrics for total $PM_{2.5}$ concentrations globally, speciated $PM_{2.5}$ concentrations globally, and total $PM_{2.5}$ concentrations regionally. Bold values do not meet the performance criteria (see Supplementary Text). NMB: normalized mean bias (%); NME: normalized mean error (%). Pop. wtd.: population-weighted metrics.

| | Global InMAP | | | GEOS-Chem | | |
|---|---|---|---|---|---|---|
| | NMB (%) | NME (%) | $R^2$ | NMB (%) | NME (%) | $R^2$ |
| Total $PM_{2.5}$ | **-60** | **63** | 0.35 | **-37** | 41 | 0.55 |
| - *pop. wtd.* | -61 | 64 | 0.56 | -37 | 40 | 0.75 |
| | | | | | | |
| $pSO_4$ | 48 | **67** | 0.38 | 18 | 37 | 0.42 |
| $pNO_3$ | -24 | 50 | 0.24 | -12 | 53 | 0.25 |
| $pNH_4$ | 35 | **65** | 0.24 | **79** | **110** | **0.08** |
| | | | | | | |
| Africa | **-41** | **52** | 0.22 | **-42** | 50 | 0.47 |
| - *pop. wtd.* | -45 | 55 | 0.46 | -46 | 51 | 0.82 |
| East Asia | **-55** | **58** | 0.20 | -28 | 32 | 0.46 |
| - *pop. wtd.* | -55 | 58 | 0.18 | -28 | 31 | 0.80 |
| South Asia | **-78** | **79** | **0.07** | **-65** | **66** | 0.21 |
| - *pop. wtd.* | -79 | 80 | 0.00 | -66 | 66 | 0.68 |
| Europe | **-64** | **64** | 0.28 | **-35** | 38 | 0.24 |
| - *pop. wtd.* | -63 | 63 | 0.71 | -29 | 34 | 0.85 |
| North & Central America | **-53** | **55** | 0.59 | -29 | 35 | **0.08** |
| - *pop. wtd.* | -55 | 58 | 0.45 | -34 | 41 | 0.91 |
| Oceana | **-45** | **47** | 0.64 | **-49** | **49** | 0.68 |
| - *pop. wtd.* | -43 | 46 | 0.58 | -58 | 58 | 0.87 |
| South America | **-74** | **79** | **0.05** | **-79** | **79** | **0.05** |
| - *pop. wtd.* | -76 | 80 | 0.07 | -73 | 73 | 0.87 |