https://doi.org/10.5194/egusphere-2025-2663
Preprint. Discussion started: 14 July 2025

A Geographically Weighted Gaussian Process Regression Emulator of the GCHP 13.0.0 Global Air Quality Model

Anthony Y. H. Wong[1], Sebastian D. Eastham[2], Erwan Monier[3], Noelle E. Selin[1,4,5]

[1]Center for Sustainability Science and Strategy, Massachusetts Institute of Technology, Cambridge, MA, USA
[2]Brahmal Vasudevan Institute for Sustainable Aviation, Department of Aeronautics, Faculty of Engineering, Imperial College London, South Kensington, London, UK
[3]Department of Land, Air and Water Resources, University of California, Davis, CA, USA
[4]Department of Earth, Atmospheric and Planetary Science, Massachusetts Institute of Technology, Cambridge, MA, USA
[5]Institute for Data, Systems and Society, Massachusetts Institute of Technology, Cambridge, MA, USA

Correspondence to: Anthony Y. H. Wong (ayhwong@mit.edu) and Noelle E. Selin (selin@mit.edu)

Abstract. Air quality modelling has been an essential tool to study the impacts of socio-economic changes and policies on air quality and associated social costs due to human health impacts. However, high computational and human resource demands limit the use of state-of-the-art air quality models outside of the atmospheric science community. We address this limitation by training Geographically Weighted Gaussian Process Regressors (GW-GPR) on the outputs of a series of perturbation experiments from the high-fidelity GEOS-Chem High Performance global chemical transport model (GCHP 13.0.0). The Gaussian Process Regressor relates changes in annual mean surface anthropogenic $PM_{2.5}$ to changes in short-lived air pollutant emissions and atmospheric CH_4 and CO_2 levels for each GCHP model grid cell. In comparison to existing widely adopted linearized and regionalized approaches, our method can account for sub-regional changes in air pollutant emission patterns and incorporates the non-linear response of secondary air pollutants to precursor and greenhouse gas emissions. We evaluate and demonstrate the utility of our model by predicting the global distribution of $PM_{2.5}$ in 2050 (relative to 2014) under 4 sets of climate and air pollution control policy scenarios. The emulator reproduces grid cell-level changes in anthropogenic $PM_{2.5}$ ($R^2 = 0.94 – 0.99$ over the 4 scenarios tested), and associated global changes in premature mortalities at 95% confidence level, while requiring < 10 seconds of CPU time (vs. ~3000 CPU hours for GCHP) for each scenario. The emulator is also able to capture projected global trends of population-weighted $PM_{2.5}$ from the AerChemMIP ensemble within the ensemble range. To our knowledge, the GW-GPR emulator is the first global-scale emulator operating at grid cell level with explicit consideration of non-linearities in atmospheric chemistry, climate change, and uncertainties resulting from both chemistry and climate variability. The accuracy, speed and simplicity of the emulator also show the capability of machine learning algorithms in emulating global air quality models, and make air quality modelling accessible for global climate/air pollution scenario analysis and integrated assessment.

1 Introduction

Fine (aerodynamic diameter ≤ 2.5 µm) particulate matter ($PM_{2.5}$) is among the most important air pollutants at global level, threatening both human and ecosystem health. Globally, $PM_{2.5}$ exposure was estimated to be responsible for ~4 million deaths in 2019 (Sang et al., 2022), and addressing health and environmental impacts from ambient air pollution has been explicitly stated as part of the Sustainable Development Goals (United Nations, 2015). The conventional approach to evaluate the impacts of socio-economic changes and policy interventions on air quality involves producing the projected air pollutant emission inventories (and meteorological fields if direct impacts of climate change are considered) and feeding them as inputs to a chemical transport model to simulate the impacts on air pollutant concentration. This process is highly demanding in terms of human and computational resources, which limits its usage for policy analysis and integrated assessments.

To increase the accessibility of air quality modelling for the broader scientific and stakeholder communities, strategies have been developed to reduce the complexity of air quality modelling by drawing from full chemical transport model experiments, resulting in various reduced-form air quality models that are faster and easier to run while retaining a reasonable level of accuracy. One approach involves dividing the world into regions. By assuming a linear relationship between air pollutant emissions in one region (source) and the air pollutant concentrations in other regions (receptor), source-receptor (SR) matrices are constructed by running a series of chemical transport model experiments with emission perturbed individually at each region. The SR matrices can then be used as a linearized global air quality model. This approach is also useful in spatial attribution of air pollution, which is applied in the Task Force on Hemispheric Transport of Air Pollutants (HTAP) (Galmarini et al., 2017; Liang et al., 2018). Designed to be a useful tool for science-policy analysis, the TM5-FASST model (Van Dingenen et al., 2018) computes SR matrices for 56 regions of the world, and subsequently process the output into public health, agriculture and climate impact metrics. Another approach involves using output of full complexity chemical transport model to parameterize some physical and chemical processes, resulting in a reduced-order chemical transport model that can be run faster and in higher resolution (Tessum et al., 2017; Thakrar et al., 2022). These SR (Huang et al., 2023; Reis et al., 2022) and reduced-order (Camilleri et al., 2023) models have been frequently applied in recent science and policy studies, showing the utility and demand for these outputs.

However, both SR and reduced-order models rely on several simplifying assumptions, which do not always hold. Many methods rely on linearizing the relationship between emissions and concentrations, which has been shown to be a reasonable approximation when the emission change is relatively small (Van Dingenen et al., 2018), but the formation rate of major secondary air pollutants such as inorganic $PM_{2.5}$ (Ansari and Pandis, 1998) are known to respond non-linearly to precursor emissions, especially when there are shifts in chemical regimes. This issue can be relevant when exploring a wide range of climate and air quality scenarios given the large range of possible air pollutant emissions and the discrepancy in the rates of change of different precursor emissions (NO_x vs NH_3 vs SO_2 for $PM_{2.5}$) (Atkinson et al., 2022; Rao et al., 2017; Turnock et al., 2020). Existing SR matrices and reduced-order models also often ignore the direct effects of climate change on air pollution

(Jacob and Winner, 2009). Garcia-Menendez et al. (2015) find that under a high-warming scenario, climate change alone can increase population-weighted annual average $PM_{2.5}$ by 1.5 µg m^{-3} between 2000 and 2100 over contiguous United States. Recent innovations in regional reduced-form air quality models have moved beyond simple linear scaling, by applying non-linear regression techniques. Conibear et al., (2022) and Vander Hoorn et al. (2022) successfully trained Gaussian Process Regression to emulate the grid cell-level response of annual mean $PM_{2.5}$ to a large range (-100% to +50%) of sectoral emission perturbations over China and the Perth greater metropolitan region respectively, with regional chemical transport model perturbation experiments as training data. Colette et al. (2022) applied multivariate quadratic regressions to emulate the simulated $PM_{2.5}$ response to emission control policies over Europe, achieving an accuracy of <2% relative error in 95% of grid cells. Meanwhile, a geographically-weighted linear regression emulator was shown to reproduce $PM_{2.5}$ response to precursor emissions from the parent chemical transport model within 10% accuracy over Europe (Pisoni et al., 2017).

Building on these regional scale applications, we combine Geographic Weighting and Gaussian Process regression (GW-GPR) techniques to emulate the output of a high-fidelity global chemical transport model, GEOS-Chem High Performance (GCHP) driven by meteorological data from multiple climate simulations with the Community Atmosphere Model (CAM, collectively GCHP-CAM) (Eastham et al., 2023) This results in a global reduced-form air quality model that can account for spatially heterogenous pollutant emission changes and non-linearity in atmospheric chemistry under multiple climate scenarios, and provide robust uncertainty estimates, without drastically increasing the computational cost.

2 Method

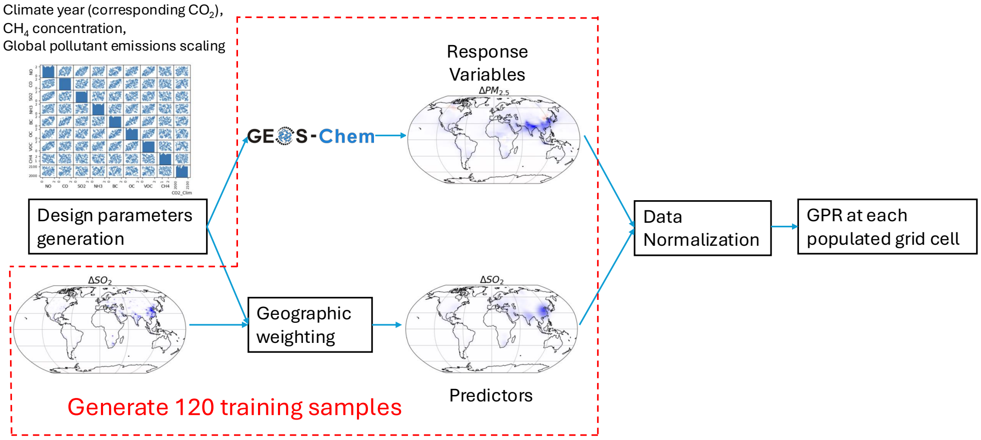


Fig. 1 Schematic showing the steps involved in the development of the emulator

We emulate the response of annual mean surface $PM_{2.5}$ to changes in air pollutant emissions and climate modelled by a high-fidelity global chemical transport model, GEOS-Chem High Performance (GCHP). Here, we provide descriptions for our GCHP model setup and experiments, and the process of constructing the emulator (Fig. 1).

2.1 GEOS-Chem High Performance Model

We use the GCHP-CAM modelling system described and evaluated in Eastham et al. (2023). The modelling system is based on a customized version of GCHP 13.0.0 (The International GEOS-Chem User Community, 2024) that can be driven by the modelled meteorological fields of the Community Atmosphere Model (CAM) version 3.1. Here we provide a brief description of the modelling system and specific setups for our work.

GCHP (Eastham et al., 2018) simulates $PM_{2.5}$ by resolving the chemistry, transport, emission and deposition of relevant chemical species. Oxidant chemistry is simulated using a coupled VOC-CO-NO_x-O_3-aerosol-halogen chemical mechanism (Sherwen et al., 2016). $PM_{2.5}$ includes contribution from nitrate, sulphate, ammonium, black carbon (BC), organic carbon (OC), fine dust, sea salt and secondary organic aerosols (SOA). The formation of secondary inorganic aerosols is simulated by considering the thermodynamic equilibrium of the NH_4^+–Na^+–SO_4^{2-}–NO_3^-–Cl^-–H_2O system through ISORROPIA II (Fountoukis and Nenes, 2007). Organic aerosols are assumed to be non-volatile.

The model is driven with climate projections simulated by the IGSM-CAM (Monier et al., 2013) a modelling framework that links the MIT Integrated Global System Model (IGSM, Monier et al., 2018) to the National Center for Atmospheric Research (NCAR) Community Atmosphere Model (CAM) 3.1 (Collins et al., 2006). The simulations are described in detail in Monier et al. (2015). The global climate model is run from 2000 – 2100 with a horizontal resolution of 2° × 2.5° on 26 vertical layers up to 2.2 hPa. We choose the high-warming "REF" scenario (10 W/m^2 in 2100, resulting in 4.3 °C warming in 2080 – 2100 versus 1990 – 2009) to provide samples across a wide range of warming and CO_2 concentration. The meteorological data from this climate projection is processed into the format of the Modern Era Retrospective for Research and Analysis version 2 (MERRA-2) (Gelaro et al., 2017) meteorological fields used by native GCHP. GCHP is run at C48 (~200km) horizontal resolution with the same vertical layers with the CAM simulations. The model output is remapped into a 2° latitude × 2.5° longitude horizontal grid conservatively (Jones, 1999).

Anthropogenic emissions of non-greenhouse gas (GHG) air pollutants are from the Community Emission Data System (Hoesly et al., 2018). Biogenic volatile organic compounds (BVOC) emissions follow Guenther et al. (2012) with isoprene inhibition by CO_2 (Possell and Hewitt, 2011; Tai et al., 2013) included. Soil NO_x emissions follows Hudman et al. (2012). While BVOC and soil NO_x emissions are both calculated online (and therefore respond to climate and atmospheric CO_2 level), mineral dust (Meng et al., 2021) and lightning NO_x (Murray et al., 2012) emissions are held at 2014 level. The monthly surface CH_4 concentration is prescribed by spatially kriging the observations from National Oceanic and Atmospheric Administration Global Monitoring Laboratory Cooperative Air Sampling Network at 2014 level. Scaling of anthropogenic emissions and atmospheric CH_4 concentration in the training sets are described in the next section.

Aerosol concentrations are archived at daily time resolution, and subsequently processed into annual mean surface total anthropogenic $PM_{2.5}$. Total $PM_{2.5}$ mass is calculated from the aerosol concentration by considering aerosol hygroscopic growth at 35% relative humidity, aligning with the $PM_{2.5}$ measurement standard of the United States Environmental Protection Agency (USEPA) (Latimer and Martin, 2019). Anthropogenic $PM_{2.5}$ mass is calculated by the above method, but only summing a subset of aerosol species (sulphate, nitrate, ammonium, BC and OC) while leaving out the aerosol species that are mostly from natural sources (dust, sea salt).

2.2 GCHP-CAM experiments

2.2.1 Generating the training set through perturbation experiments

To effectively sample the sensitivity of $PM_{2.5}$ over a wide range of climate and air pollution emissions, we generate the training set from a series of GCHP-CAM perturbation experiments by manipulating 9 input variables that affect $PM_{2.5}$ and oxidant levels: 7 air pollutant emissions (NO_x, SO_2, NH_3, NMVOC, BC, OC, carbon monoxide (CO)) that are commonly provided by integrated assessment models (Gidden et al., 2019), CH_4 concentration, and global warming level. Global warming level cannot be directly implemented in GCHP as a scaling factor; the global warming scaling is implemented by driving the model with simulated meteorological fields at the year with the closest CO_2 level under the "REF" scenario.

120 sets of scaling factors (0 to 1) for the 9 input variables are generated following a Latin Hypercube Sampling (LHS) (McKay et al., 1979) strategy. However, the changes in emissions of different pollutants are correlated, since different air pollutants often share similar emission sources (e.g. combustion). To account for co-emissions of air pollutants, we calculate the spatial correlations of the emissions of the 7 air pollutants from CEDS between 2000 and 2017. The correlation of CH_4 and global warming level with other variables are set to be 0 to provide independence between climate and air pollution control policies. This correlation matrix is then used to rearrange the LHS scaling factors using an Iman-Conover Transform (Conover and Iman, 1982).

We first run GCHP-CAM with meteorological fields from 1st Oct 2013 to 31st Dec 2014 and 2014 anthropogenic emissions, CH_4 and CO_2 concentration under the "REF" scenario to generate the baseline for comparison, with the first 3 months of model run discarded as spin-up (output not used). The rearranged LHS scaling factors sets are then linearly mapped to the range of inputs (Table 1), and each set corresponds to a 1-year perturbation simulation, again with an extra 3 months before as spin-up. Each of the perturbation simulations are driven by the globally scaled 2014 anthropogenic air pollutant emissions and surface CH_4 levels. The scaling factor for CH_4 (0.5 to 2.5) is chosen to enclose the range of CH_4 concentration in 2100 projected by Meinshausen et al. (2020) over all scenarios. Instead of total radiative forcing, the global warming level is parameterized as atmospheric CO_2 level as CO_2 is projected to dominate (68 – 85%) total radiative forcing in the 21st century (Meinshausen et al., 2020). In addition, atmospheric CO_2 level also directly affects isoprene emission, which could affect atmospheric oxidant (e.g. OH, O_3) (e.g. Tai et al., 2013), and therefore potentially secondary inorganic aerosol formation.

Variables	Range
Air pollutant emission scaling factor	0 – 2
Surface CH_4 concentration scaling factor	0.5 – 2.5
Atmospheric CO_2 level	369.9 – 813.5 ppm

Table 1. Range of scaling factors and CO_2 concentration of the perturbation experiments.

2.2.1 IGSM-GAINS-TAPS combined air quality and climate legislation scenarios

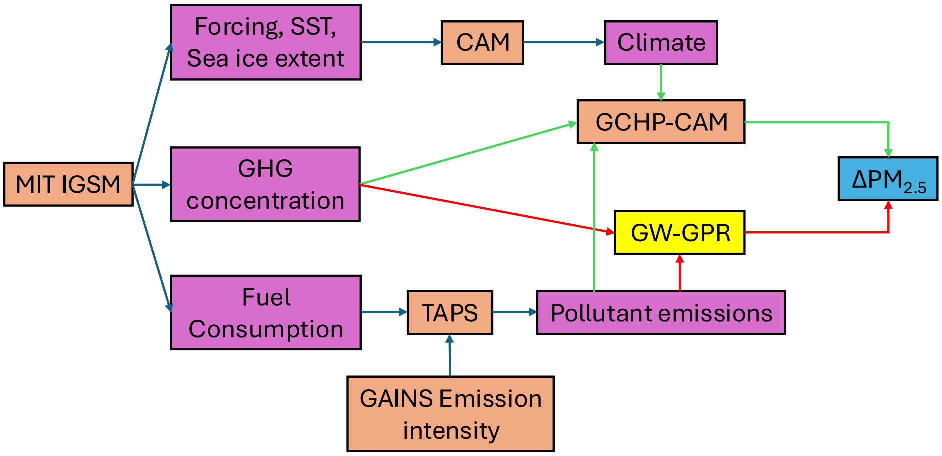


Fig. 2 Schematic of GW-GPR model evaluation using IGSM-GAIN-TAPS climate and air pollution scenarios. Orange boxes represent existing modelling systems, purple boxes represent data sets. Yellow represents our newly developed emulator (GW-GPR). Green and red arrows represent how GCHP-CAM and GW-GPR predicts changes in $PM_{2.5}$, respectively.

To assess the utility of the emulator within the context of integrated assessment modelling, we evaluate the ability of the emulator in reproducing GCHP output anthropogenic $PM_{2.5}$ over 2 climate (Current Trend (CT) and Accelerated Action (AA)) × 2 air pollution control (Current Legislation (CLE) and Maximum Feasible Reduction (MFR)), in total 4 scenarios (CT_CLE, CT_MFR, AA_CLE, AA_MFR) in 2050. CT assumes the implementation of Nationally Determined Contributions (NDCs) from Paris Agreement through 2030. Despite such effort, climate is not stabilized, and global mean temperature continues to increase. AA assumes the extension of these initial NDCs to align with the long-term goal of Paris Agreement, therefore the ability of limit and stabilize anthropogenic warming to 1.5 °C at 2100 with at least 50% probability. The CLE scenario assumes complying with existing region- and source-specific air pollutant emission limits, while the MFR scenario assumes increasing deployments of currently available lowest-emitting technologies.

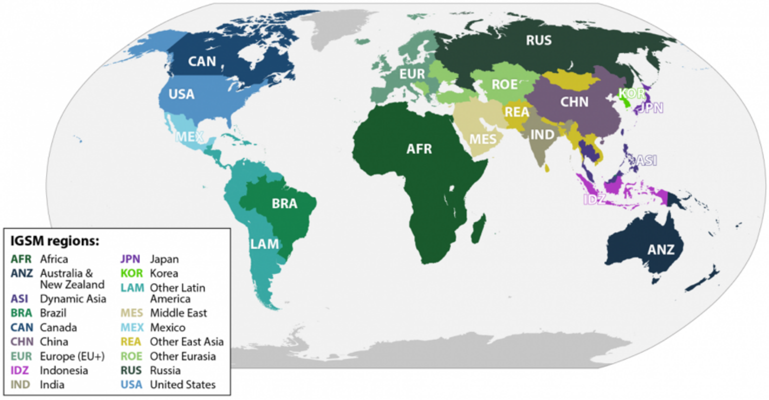


Fig. 3 IGSM region definition (retrieved from https://globalchange.mit.edu/research/research-tools/eppa, date accessed: 31st March 2025)

https://doi.org/10.5194/egusphere-2025-2663
Preprint. Discussion started: 14 July 2025

The climate scenarios are generated from the MIT IGSM framework. The human system component of IGSM, Economic Projection and Policy Analysis model version 7 (EPPA7) (Chen et al., 2022), is a global multi-sector (22 sectors) multi-region (18 regions, Fig. 3) recursive-dynamic computable general equilibrium model. As a part of climate/economic scenario projection, EPPA provides regionalized and sectorized consumptions of different fuel types, and populations. The yearly global average atmospheric GHG concentrations is then derived by driving the MIT Earth System Model (MESM) (Sokolov et al., 2018) with corresponding EPPA output.

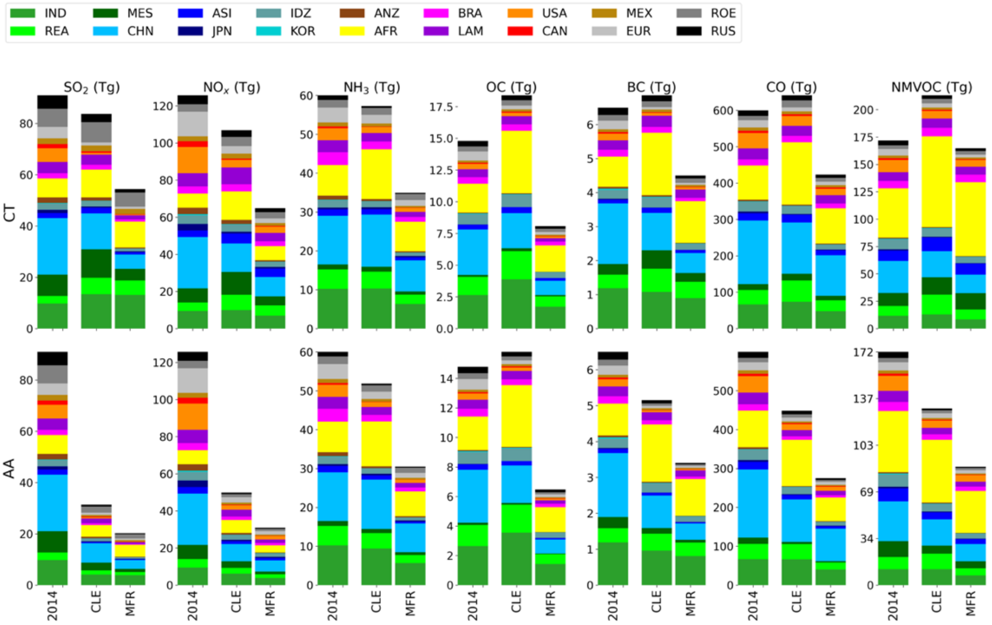


Fig. 4 Total and regional emissions for the four IGSM-GAINS-TAPS scenarios considered in this study at base year and 2050.

The projected trends of air pollutant emission intensities under CLE and MFR scenarios are from Greenhouse gas-Air pollution INteractions and Synergies (GAINS) (Amann et al., 2011), based on GAINS4/ECLIPSE (Evaluating the Climate and Air Quality Impacts of Short-lived Pollutants) v6b data (GAINS Developer Team, 2021; Klimont et al., 2017; Smith et al., 2020; Stohl et al., 2015). Future air pollutant emissions for each scenarios can be derived through the Tool for Air Pollution Scenarios (TAPS) (Atkinson et al., 2022) by considering climate and air pollution policies independently. The resulting air pollutant emissions of each of the 4 combined scenarios are shown in Fig. 4.

We perform 10 years of GCHP-CAM simulations for each IGSM-GAINS-TAPS scenarios with their respective anthropogenic emissions. The meteorological years (2031 – 2041 for AA and 2040 – 2050 for CT) are chosen to match the level of warming of our meteorological fields (“REF” scenario).

2.3 Geographically Weighted Gaussian Process Regression (GW-GPR) Emulator

We use Gaussian Process Regression (GPR) (Williams and Rasmussen, 1995) to relate the changes in pollutant emissions and climate with the corresponding changes in annual mean $PM_{2.5}$ level at grid cell level. GPR is a non-linear and non-parametric regression algorithm that does not require prior assumptions of the functional relationship between input and output variables. Instead, predictions are made by assuming the training and prediction sets follow a joint multivariate normal distribution:

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = N\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad (1)$$

Where Y_1 is the random variable representing the prediction, and Y_2 is the random variable representing the response from training set. μ_1 and μ_2 are their means, and Σ_{11} ,Σ_{12} ,Σ_{21} ,Σ_{22} are the covariance matrix blocks. The prediction process can be viewed as finding $(Y_1|Y_2 = \boldsymbol{a}) \sim N(\mu', \Sigma')$, where $\boldsymbol{a}$ is the model output from training set. Thus, the mean (μ') and variance (Σ') of the prediction can be calculated as:

$$\mu' = \mu_1 + \Sigma_{12}{\Sigma_{22}}^{-1}(\boldsymbol{a} - \mu_2) \quad (2)$$

$$\Sigma' = \Sigma_{11} - \Sigma_{12}{\Sigma_{22}}^{-1}\Sigma_{21} \quad (3)$$

By setting the prior mean of prediction as 0 and proper normalization during training process, μ_1 and $\mu_2 = 0$. Then μ' is essentially a sum of **a** weighted by the correlations between the training and prediction vectors ($\Sigma_{12}\Sigma_{22}^{-1}$). The elements of the correlation matrix take the form:

$$\sigma_{\mathrm{ij}} = k(x_{\mathrm{i}}, x_{\mathrm{j}}) \quad (4)$$

Where x_i and x_j are the input vectors at each training or prediction points, and k is the covariance function (kernel) that can be chosen to control the shape and smoothness of the prediction. We use a sum of anisotropic kernels to represent the nature of our problem (smooth functions (rational quadratic function) with unknown points of chemical regime change + local interactions among variables (Matern 3/2 function) + noise from climate variability (white noise)). We use the GPR as implemented in Scikit-learn version 1.3.2 (Pedregosa et al., 2011), and only train a GPR model for each populated (population density > 1 person/km^2) model grid cell. As the predictions are random variables, the uncertainty and confidence interval of each prediction can be derived by its standard deviation.

We apply geographic weighting to the emission fields of individual pollutant species to emulate the process of chemical transport of emitted species. An isotropic 2D Gaussian dispersion kernel is applied to calculate the effective air pollutant emission changes ($\Delta E_{\mathrm{weighted,x}}$) at each grid cell x:

$$\Delta E_{\mathrm{weighted,x}} = \sum_{\mathrm{y}} e^{-d_{\mathrm{y,x}}^2/2L^2}\, \Delta E_{\mathrm{y}} \quad (5)$$

Where ΔE_{y} is the emission change all individual grid cells considered within the dispersion kernel (y), $d_{\mathrm{y,x}}$ is the distance between grid cell y and x, and L is the dispersion length scale. The length scales (in the unit of grid cell, 1 grid cell = 2° latitude × 2.5° longitude) of each species are chosen to match their atmospheric lifetime. The dispersion kernel is implemented by the Gaussian Blurring algorithm as in Scipy version 1.10.1 (Virtanen et al., 2020). We assign $L = 1$ grid cell for relatively short-lived species (NO_x, NH_3, NMVOC), $L = 2$ grid cells for longer-lived species (BC, OC, SO_2), and $L = 3$ grid cells for CO. We note that some previous regional studies (e.g. Pisoni et al., 2017) have treated the parameters of the dispersion kernel as optimizable hyperparameters. However, our GCHP-CAM experiments are conducted with uniform global scaling factors for emission fields. After the variable normalization procedure, training configurations with different dispersion kernels would effectively collapse to the same 120 sets of global scaling factors prescribed in the GCHP-CAM experiments. Therefore, our training set cannot be used to directly optimize the dispersion kernel.

While we do not perform rigorous optimization of the dispersion kernel, we explore the associated uncertainties by sensitivity simulations of training the GW-GPR emulator with halved and doubled dispersion length scales of the GW-GPR emulator. The result is discussed in Section 3.3.2. The input variables are normalized by their corresponding global maximum value after the geographic weighting. Since the output variables are not blurred, and $\mu_2 = 0$ simplifies computation, the output variables are normalized by local mean and maximum at each grid cell.

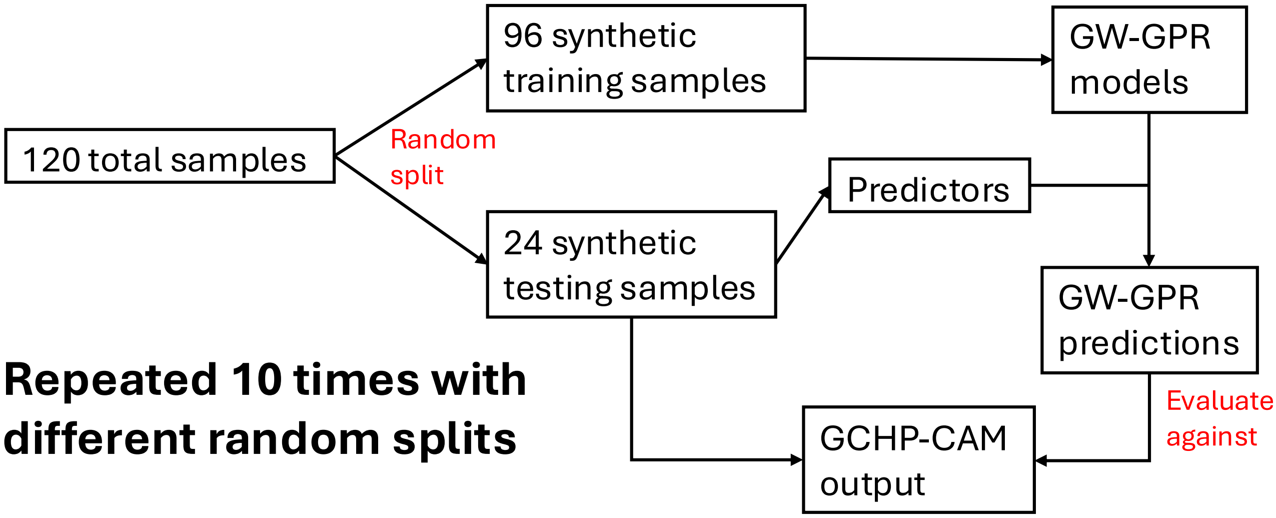


Fig. 5 Schematic of the cross-validation procedure

We evaluate the generalization ability of our models using the repeated random sub-sampling technique. For each repetition (Figure 5), we randomly split the data into training (80%) and testing (20%) sets. New GW-GPR models are built from the synthetic training set and predictions are mode over the synthetic testing set at grid cell level. This process is repeated 10 times, and the performance metrics are calculated using all 10 synthetic testing sets and corresponding predictions. In addition, we perform Sobol global sensitivity analysis (Sobol′, 2001) by drawing 1024 × (2 × number of input variables + 2) samples to compute the total sensitivity indices for each input variables using Saltelli sampling (Saltelli et al., 2010) at each grid cell to using SALib 1.4.8 (Herman and Usher, 2017; Iwanaga et al., 2022), which helps us identify the importance of each input variable for different locations.

https://doi.org/10.5194/egusphere-2025-2663
Preprint. Discussion started: 14 July 2025

2.4 AerChemMIP data

	SSP3-7.0 Ensemble members	SSP3-7.0-lowNTCF ensemble members	Model references	Data references
EC-Earth3	2	2	(Döscher et al., 2022)	(Consortium (EC-Earth), 2023)
GFDL-ESM4	1	1	(Dunne et al., 2020)	(Horowitz et al., 2023)
GISS-E2.1-G	16	4	(Kelley et al., 2020)	(NASA Goddard Institute For Space Studies (NASA/GISS), 2023a, b)
GISS-E2.1-H	1	1	(Kelley et al., 2020)	(NASA Goddard Institute For Space Studies (NASA/GISS), 2023c)
IPSL-CM5A2-INCA	1	1	(Sepulchre et al., 2020)	(Boucher et al., 2023)
MIROC-ES2H	1	0	(Kawamiya et al., 2020)	(Watanabe et al., 2023)

Table 2. The numbers of ensemble members with all 5 anthropogenic $PM_{2.5}$ components (sulphate, nitrate, ammonium, BC, OC) available for SSP3-7.0 and SSP3-7.0-lowNTCF scenarios from the AerChemMIP archive.

To further assess and demonstrate the utility of the emulator, particularly under global change scenarios, we also compare the output of the GW-GPR emulator with the output from the Aerosol Chemistry Intercomparison Project (AerChemMIP) under 2 Shared Socio-economic Pathways (SSP)-based scenarios: 1) the standard SSP3-7.0 "Regional Rivalry" scenario (radiative forcing = 7.0 W m^{-2} at 2100); and 2) a variant of SSP3-7.0 with same socio-economic assumptions as the standard SSP3-7.0, but stronger air quality control measures, resulting in lower emissions of Near Term Climate Forcers (SSP3-7.0-lowNTCF) (Fujimori et al., 2017). AerChemMIP (Collins et al., 2017) is endorsed by the Coupled-Model Intercomparison Project 6 (CMIP 6) to quantify the impacts of aerosols and chemically reactive gases on climate. There are minimum model complexity requirements (atmosphere-ocean general circulation model with tropospheric aerosols driven by pollutant emission fluxes) to participate in AerChemMIP ensemble.

We calculate the anthropogenic $PM_{2.5}$ (sum of sulphate, nitrate, ammonium, BC and OC) from AerChemMIP archive in an identical manner as for GCHP (summing contributions from individual components after applying the hygroscopic growth factor at 35% RH). We only include ensemble members with the output of all five anthropogenic $PM_{2.5}$ components available. The models selected and numbers of realizations included for each model are summarized in table 2.

The emulator calculates the decadal changes in anthropogenic $PM_{2.5}$ concentrations relative to 2020 over 2030 – 2090, using surface air pollutant emissions (calculated as anthropogenic + open burning emissions) and GHG concentration (Meinshausen et al., 2020) provided by the Input4MIP repository. As air pollutant emissions are provided every 10 years, one emulator prediction is done per decade. For each ensemble member in the AerChemMIP archive, the corresponding decadal changes in

https://doi.org/10.5194/egusphere-2025-2663
Preprint. Discussion started: 14 July 2025

anthropogenic $PM_{2.5}$ are calculated by comparing the decadal average anthropogenic $PM_{2.5}$ with that of the first decade (2015 – 2024). The model-specific decadal changes in anthropogenic $PM_{2.5}$ is then calculated by averaging the result from all the ensemble members from the corresponding model. AerChemMIP and Input4MIP data are retrieved via the search engine of Earth System Grid Federation (ESGF) (https://aims2.llnl.gov, last access: 26th Aug 2024).

2.5 Health Impact Calculation

For the four climate and air pollution control scenarios, we also estimate the impacts of changes in anthropogenic $PM_{2.5}$ on public health through premature mortalities. GCHP and emulator output are upsampled from 2° × 2.5° to 0.5° × 0.5° using the nearest neighbour algorithm, which matches the horizontal resolution of the age-specific population data we use (Gridded Population of the World version 4.11, last access: 19th Apr, 2024) (Center For International Earth Science Information Network-CIESIN-Columbia University, 2018). Country-level baseline age- and cause-specific mortality rates are provided by the World Health Organization (WHO) (WHO, 2018). The age- and cause-specific changes in the annual mortality due to chronic $PM_{2.5}$ exposure for scenario *i* ($\Delta Mort_i$) is calculated from the relative mortality risks under the baseline (total $PM_{2.5}$ from the 2014 baseline run) (RR_{base}) and each scenario *i* (RR_i):

$$\Delta Mort_i = Mort_{\mathrm{base}}\left(\frac{RR_i - RR_{\mathrm{base}}}{RR_{\mathrm{base}}}\right) \quad (6)$$

where $Mort_{\mathrm{base}}$ is the age- and cause-specific mortalities from the WHO.

We use the age-specific non-linear Concentration Response Functions from the Global Exposure Mortality Model (Burnett et al., 2018) to calculate RR_i and RR_{base} for non-communicable diseases and lower respiratory infections attributable to outdoor $PM_{2.5}$ pollution. Since our emulator focuses on the changes and corresponding impacts in anthropogenic $PM_{2.5}$, the synthetic $PM_{2.5}$ level for scenario *i* at each grid cell is calculated as $PM_{2.5,\mathrm{base}} + \Delta PM_{2.5,i}$, where $PM_{2.5,\mathrm{base}}$ is the modelled total $PM_{2.5}$ in year 2014, and $\Delta PM_{2.5,i}$ is the modelled/emulated change in anthropogenic $PM_{2.5}$ for scenario *i*.

3 Comparisons with GCHP-CAM

In this section we discuss and explain the performance of our GW-GPR emulators against GCHP-CAM simulations, measured by the response of anthropogenic $PM_{2.5}$ pollution level and associated premature mortalities.

3.1 Computing resource requirement

GCHP-CAM requires 2400 – 3000 CPU hours (Intel Xeon Processor E5-2679A v4, processor base frequency = 2.6 GHz)) to simulate $PM_{2.5}$ for each 1-year run. The one-time operation of fitting the GW-GPR emulator requires 280 CPU hours (Intel Xeon Processor E5-2670, processor base frequency = 2.6 GHz). Once trained, the emulator requires approximately 10 CPU seconds to generate global $PM_{2.5}$ predictions for one scenario (Intel Xeon Processor Silver 4214R, processor base frequency = 2.4 GHz). This demonstrates the magnitude of the speed up offered by the emulator.

285 **3.2 Emulator validation and sensitivity**

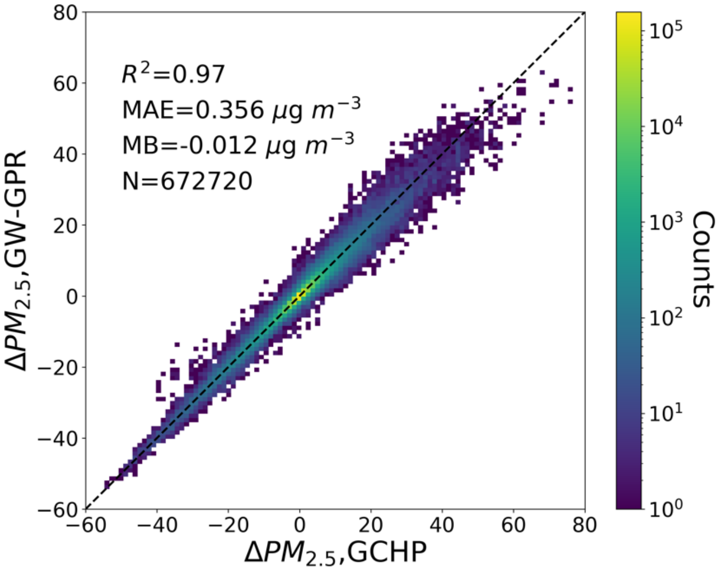
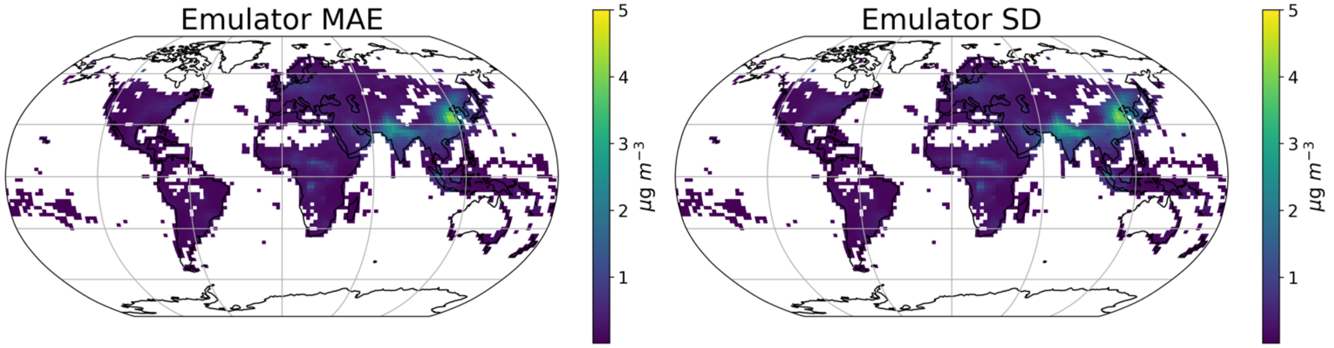


Fig 6. 2D histogram from the grid cell by grid cell comparison between $\Delta PM_{2.5}$ predicted by the GW-GPR emulator ($\Delta PM_{2.5}$, GW-GPR) and that simulated by GCHP-CAM ($\Delta PM_{2.5}$, GCHP) from the 10-fold random sub-sampling cross-validation.



290 **Fig. 7 The mean absolute error (MAE) of emulator prediction against the parent model (GCHP-CAM), and the average standard deviation of emulator predictions (indicative of uncertainties from climate variability and chemistry) at grid cell level (240 data points at each grid cell)**

Fig. 6 shows the result grid cell by grid cell comparison of changes in annual mean anthropogenic $PM_{2.5}$ ($\Delta PM_{2.5}$) predicted by the GW-GPR emulator against that simulated by its parent model (GCHP-CAM) across all the data points generated by the 295 random subsampling. The GW-GPR emulator can predict $\Delta PM_{2.5}$ GCHP-CAM with reasonable accuracy (R^2 = 0.97, mean absolute error (MAE) = 0.356 μg m^{-3}) and minimal overall bias (mean bias (MB) = -0.012 μg m^{-3}). Fig.7 shows the spatial distribution of grid cell level MAE of the GW-GPR emulator, and the emulator output standard deviation (which can characterize the uncertainty of emulator output). The largest MAE is found over Northern China and Northern India (up to 5 μg m^{-3}), where the anthropogenic $PM_{2.5}$ and precursor emissions are very high in the base year of 2014. The emulator output standard deviation have similar magnitudes and spatial distributions (spatial $R^2 = 0.99$) as MAE, indicating that emulator output standard deviation is an appropriate measure of chemical and climate uncertainties of emulator predictions relative to the parent model.

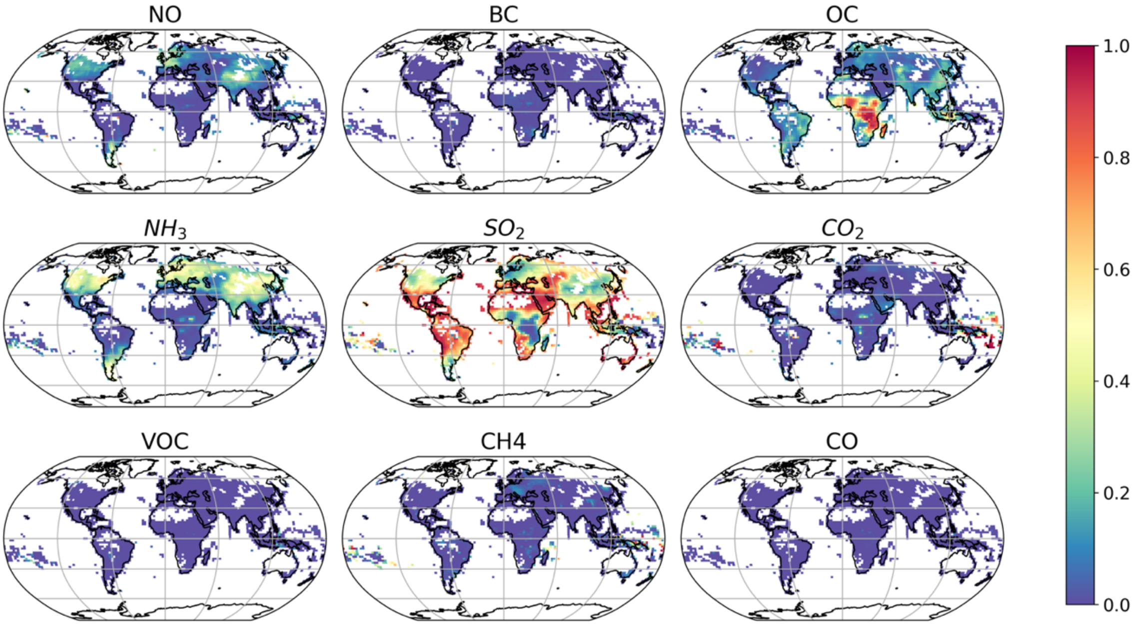


Fig. 8 Spatial patterns of Sobol Total Sensitivity Indices for each predictor for $\Delta PM_{2.5}$.

Fig. 8 shows the normalized Sobol Total Sensitivity Indices of the GW-GPW emulator to each of the input variables (in the unit of fractional rather than absolute changes), which measure how much each input variable is responsible for the variance in the output over the whole domain of input data, including the interaction among variables. In other words, the Sensitivity Indices indicate how important the specific input variable is in controlling the output. The importance of input variables is spatially heterogenous. Over North America, Europe, $\Delta PM_{2.5}$ is mostly sensitive to SO_2 (46 – 57% of total sensitivity index) and NH_3 (28 – 32%), and to a lesser extent NO emissions (10 – 14%). The pattern of total sensitivity indices over India and China are similar (9 – 13% for NO, 21 – 37% for NH_3 and 36 – 57% for SO_2), but the sensitivity of $\Delta PM_{2.5}$ to OC (13% vs 3 – 7% over North America and Europe) is higher over these regions. For most of the rest of the northern hemisphere, $\Delta PM_{2.5}$ is primarily sensitive to SO_2 emissions (e.g. >80% over Mexico and Middle East). Over the southern hemisphere, $\Delta PM_{2.5}$ remains highly sensitive to SO_2 emissions (47% over Indonesia – 76% over Brazil). In Brazil, Indonesia, and Africa, $\Delta PM_{2.5}$ is also sensitive to OC emissions (15% - 38%). Sensitivity of $\Delta PM_{2.5}$ to BC is relatively low globally (mean = 0.006 over the globe). This is because BC is largely co-emitted with OC, while the OC emissions are always around 1 – 2 times larger than BC emission by mass. Thus, the variance attributable to BC is mostly captured by the variance attributable to OC. The sensitivity index of CO_2, CH_4, VOC and CO are also relatively low (<3%) globally, except over the certain regions with low anthropogenic emissions (tropical Pacific islands, edge of the Amazon and central African rainforests), reflecting the fact that our emulator does not account for secondary organic aerosols.

3.3 Emulator performance over IGSM-GAINS-TAPS scenarios

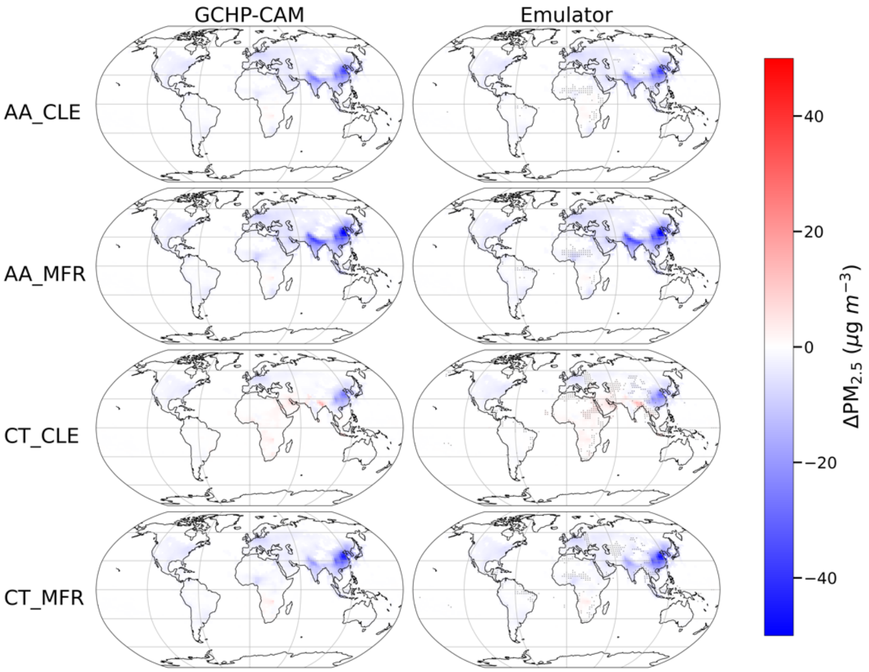


Fig. 9 Spatial patterns of GCHP-CAM and emulator predicted $\Delta PM_{2.5}$ for each of the 4 IGSM-GAINS-TAPS scenarios at 2050 (relative to 2014). Only results in grid cells with population density > 1 person km^{-2} are shown. The dots show where GCHP-CAM output does not fall within the 95% confidence interval of emulator prediction.

Scenario	Regression technique	R^2	MAE ($\mu g\ m^{-3}$)	MB ($\mu g\ m^{-3}$)	% of grid cells agreeing within 1 (2) SD
AA_CLE	GRP	0.99	0.25	+0.04	76.2 (94.3)
	MLR	0.98	0.35	+0.02	
AA_MFR	GPR	1.00	0.20	+0.05	84.8 (96.0)
	MLR	0.99	0.29	-0.08	
CT_CLE	GPR	0.94	0.42	+0.10	58.7 (82.9)
	MLR	0.93	0.47	+0.04	
CT_MFR	GPR	0.98	0.34	+0.07	68.3 (87.6)
	MLR	0.98	0.39	-0.01	

Table 3. Gaussian Process Regression (GPR) and multilinear regression (MLR) emulator performance metrics (spatial coefficient of determination (R^2), mean absolute error (MAE), mean bias (MB), computed at grid cell level, N = 2803) for each IGSM-GAINS-TAPS scenarios, relative to GCHP-CAM output. The rightmost column indicates the percentage of gridc ells that the GPR emulator prediction agrees with GCHP-CAM output within 1 (2) standard deviation (prediction uncertainty of emulator).

Fig. 9 shows the GCHP-CAM and GW-GPR (emulator) output $\Delta PM_{2.5}$ (2045 – 2054 mean vs 2014) over each IGSM-GAINS-TAPS scenario at grid cell level. The global performance metrics are shown in table 3. Generally, the emulator performs comparably to that in the random subsampling evaluation (R^2 = 0.94 – 0.99, MAE = 0.20 – 0.42 μg m^{-3}). 58.7 – 84.8% and 82.9% (96%) of the grid cells have emulator predictions agreeing with GCHP-CAM within 1 (2) standard deviation of emulator output respectively (computed with eq. 3).

When the predicted spatial distributions of $\Delta PM_{2.5}$ are converted into premature mortalities using the GEMM CRF, we find that the GCHP-CAM and emulator output produce similar impacts on global premature mortalities over the 4 IGSM-GAINS-TAPS scenarios tested (differences within 1.2%) that agree within the range of uncertainty due to the GEMM CRF parameters (Fig. 10). This shows the emulator's ability to reproduce both the magnitudes and spatial distributions of $\Delta PM_{2.5}$ from GCHP-CAM, and the suitability of emulator output for public health impact calculation at global level.

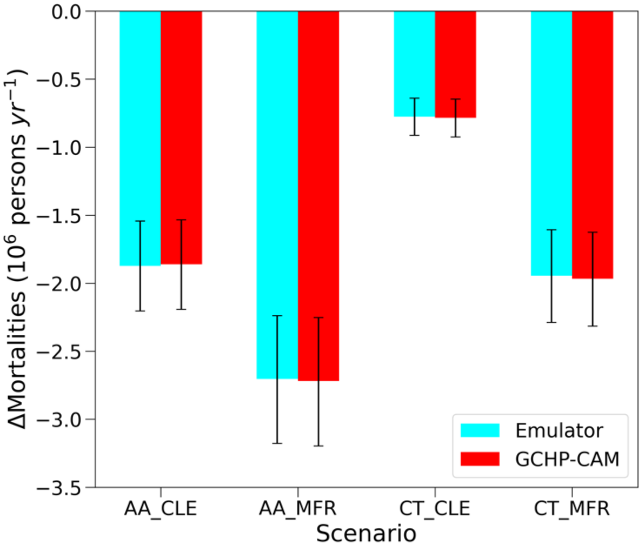


Fig. 10 Changes in global annual premature mortality attributable to $PM_{2.5}$ exposure under each of the four scenarios between 2050 and 2014, calculated from the emulator and GCHP-Chem output $\Delta PM_{2.5}$. The error bars represent the uncertainties due to GEMM CRF parameters, calculating by applying the 2.5 and 97.5 percentile estimate of the GEMM CRF parameters.

In general, the emulator performs the best over the western hemisphere (longitude < -20°), where the emulator error is within 2 μg m^{-3} (MAE = 0.1 μg m^{-3}), and 77.4% of emulator predictions agrees with GCHP-CAM output within 1 emulator output standard deviation. In contrast, the emulator output shows consistent high bias of up to 2 μg m^{-3} over the 4 scenarios over the Sahel. Around the Bay of Bengal, emulator output does not agree with GCHP-CAM output with 1 emulator output standard deviation in a large portion of grid cells. In the subsections below, we will explore the potential sources of error by comparing the result presented above with that from alternative emulator architectures.

350 ### 3.3.1 Comparison with linear model

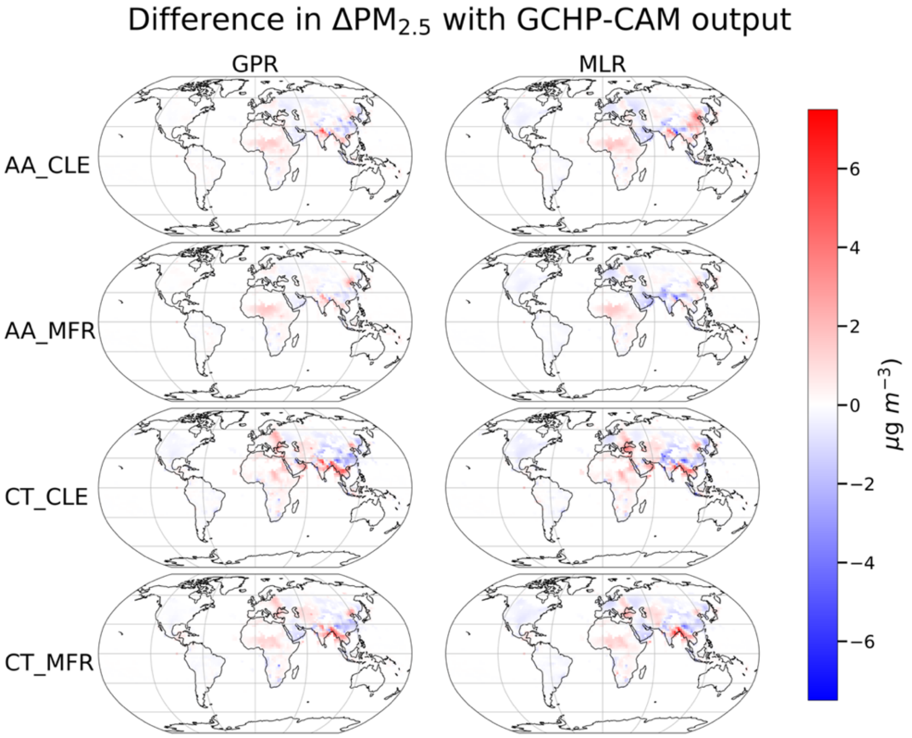
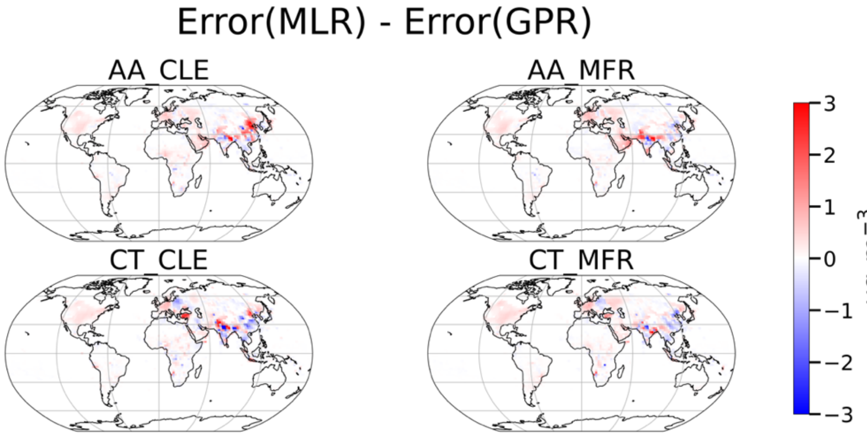


Fig. 11 GPR and MLR emulator errors relative to GCHP-CAM simulated $\Delta PM_{2.5}$ over the 4 IGSM-GAINS-TAPS scenarios at 2050 (relative to 2014).



355 **Fig. 12 Difference in the absolute error (relative to GCHP-CAM output) between MLR and GPR emulation. Red (positive) indicate GPR is more accurate than MLR at the given grid cell, while blue (negative) indicates the opposite.**

For comparison, we train a multilinear regression (MLR) emulator with identical variables, geographic weighting and normalization schemes, and the performance metrics of the multilinear emulator is also shown in table 3. In all scenarios, the MLR estimator has a larger global MAE than GPR (by 0.05 μg m^{-3} (19%) in CT_CLE to 0.10 μg m^{-3} (40%) in AA_CLE).

However, in 3 out of 4 scenarios tested (expect AA_MFR), the GPR emulator has higher MB than the MLR emulator, though the overall magnitudes of MB remain relatively small (within 0.1 μg m^{-3}).

Fig. 11 shows the spatial distribution of MLR and GPR error relative to their parent model (GCHP-CAM), and Fig. 12 shows the difference in absolute values of such errors between MLR and GPR. In the 4 IGSM-GAINS-TAPS scenarios, GPR predictions have less absolute error relative to the parent model than MLR in 57.6 % (CT_CLE) to 66.4% (AA_MFR) of the grid cells. In all 4 scenarios, GPR outperforms MLR over the US (MAE = 0.05 – 0.14 μg m^{-3} for GPR vs 0.17 – 0.29 μg m^{-3} for MLR), western and southern Europe (MAE = 0.05 – 0.06 μg m^{-3} for GPR vs 0.20 – 0.26 μg m^{-3} for MLR), Middle East (MAE = 0.16 – 0.56 μg m^{-3} for GPR vs 0.40 – 0.72 μg m^{-3} for MLR), South America (MAE = 0.08 – 0.14 μg m^{-3} for GPR vs 0.14 – 0.16 μg m^{-3} for MLR), and South Asia.

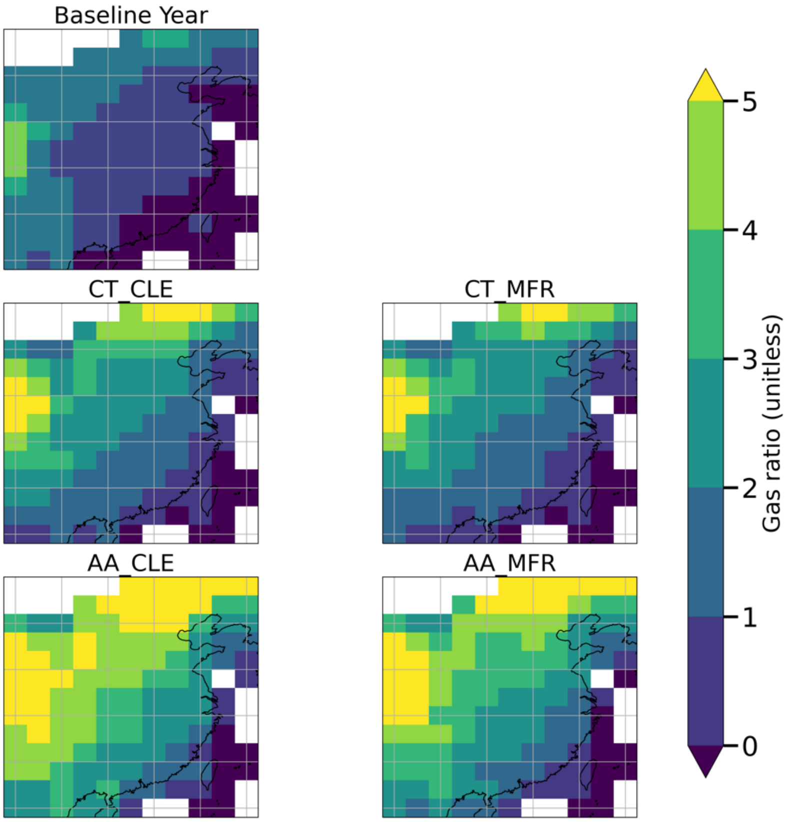


Fig. 13. Gas ratio (GR) over China, which indicate secondary inorganic $PM_{2.5}$ sensitivity to NH_3 emissions. Secondary inorganic $PM_{2.5}$ is weakly sensitive to NH_3 when GR < 0. 0 < GR < 1 indicate stronger sensitivity of secondary inorganic $PM_{2.5}$ to NH_3 emissions. When GR > 1, sensitivity of secondary inorganic $PM_{2.5}$ to NH_3 emissions decreases as GR increases.

In northern China, GPR is slightly less accurate than MLR on average under the MFR scenarios (by 0.31 μg m^{-3} under AA_MFR and 0.38 μg m^{-3} under CT_MFR, measured by regional MAE), but considerably more accurate under CLE scenarios (by 2.01 μg m^{-3} under AA_CLE and 0.81 μg m^{-3} under CT_CLE. We analyse the shifts in the chemical regime of secondary inorganic aerosol formation calculating the Gas Ratio (*GR*) (Paulot and Jacob, 2014) over China at baseline year and under all 4 scenarios (Fig. 13):

$$GR = \frac{[NH_3]+[NH_4^+]-2[SO_4^{2-}]}{[HNO_3]+[NO_3^+]} \quad (7)$$

GR < 0 indicates that secondary inorganic $PM_{2.5}$ is weakly sensitive to NH_3 emissions through adding NH_4^+ to existing SO_4^{2-} and HSO_4^- ions. 0 < *GR* < 1 indicates that there is enough NH_3 to react with SO_4^{2-}, such that NH_3 and HNO_3 start partitioning into NH_4NO_3 aerosol, leading to strong sensitivity of secondary inorganic $PM_{2.5}$ to NH_3 emissions. In this regime, secondary inorganic $PM_{2.5}$ is more sensitive to NH_3 emissions. *GR* > 1 indicates that there is more than enough NH_3 to react with both SO_4^{2-} and HNO_3, and $PM_{2.5}$ sensitivity to NH_3 emissions will weaken continuously as *GR* keep increasing beyond 1 (Ansari and Pandis, 1998). At the baseline year, *GR* over Norther China is largely between 0 – 1. Under all four scenarios, *GR* increases beyond 1 over northern China. However, the increases in GR are the strongest under AA_CLE, followed by AA_MFR, while CT_CLE and CT_MFR have lower *GR* than the two AA scenarios. This indicates stronger shifts in secondary inorganic $PM_{2.5}$ sensitivity to precursor emissions relative to the baseline year (and therefore more non-linearity) under the two AA scenarios (especially AA_CLE) than the two CT scenarios, which is more well-captured by GPR than MLR.

The results in this sub-section show that GPR generally outperforms MLR. When emission changes could potentially trigger non-linear aerosol chemistry, non-linear emulators can be significantly more accurate than linear emulators. This justifies the use of non-linear regression techniques (e.g. GPR) in developing air quality emulators.

3.3.2 Sensitivity to dispersion length scales

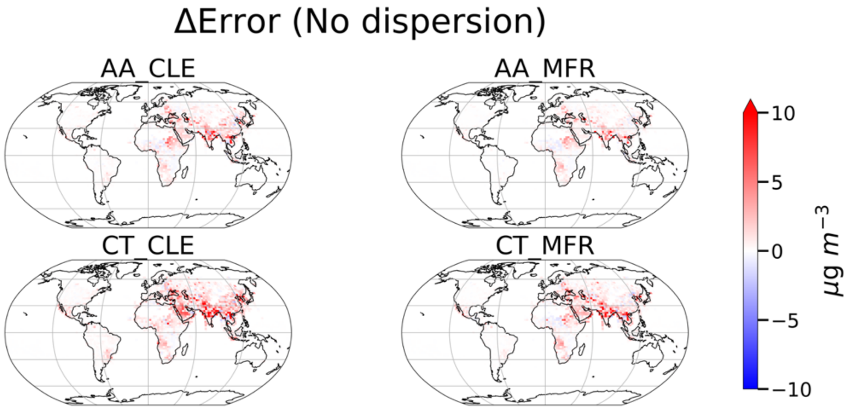


Fig. 14 Changes in absolute error (relative to GCHP-CAM output) when no dispersion kernel is implemented. Red (positive) indicates that turning off dispersion worsens the performance (increasing error), blue (negative) indicates the opposite.

In addition to regression techniques, we also conduct 3 sensitivity tests of altering the dispersion length scales: 1) no dispersion; 2) halving the dispersion length scale, and 3) doubling the dispersion length scale. Fig. 14 shows the changes in absolute error of emulator prediction when the dispersion kernel is disabled, i.e. no geographic weighting is done. Turning off the geographic weighting scheme worsens the performance of the emulator, increasing the global MAE by between 0.31 (AA_MFR) and 0.88

https://doi.org/10.5194/egusphere-2025-2663
Preprint. Discussion started: 14 July 2025

(CT_CLE) μg m^{-3}, and locally absolute error by up to 29.1 μg m^{-3}. This decline in model performance is much larger than that by switching from GPR to MLR, indicating the necessity of the geographic weighting scheme in our emulator.

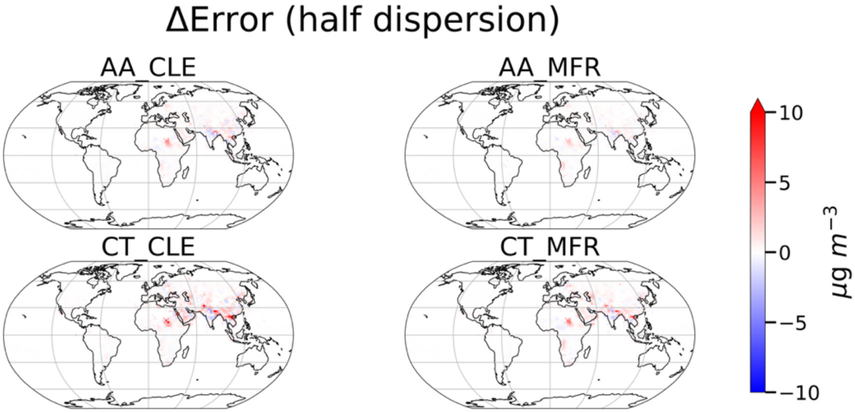


Fig. 15 Changes in absolute error (relative to GCHP-CAM output) when the dispersion length scale is halved. Red (positive) indicates that turning off dispersion worsens the performance (increasing error), blue (negative) indicates the opposite.

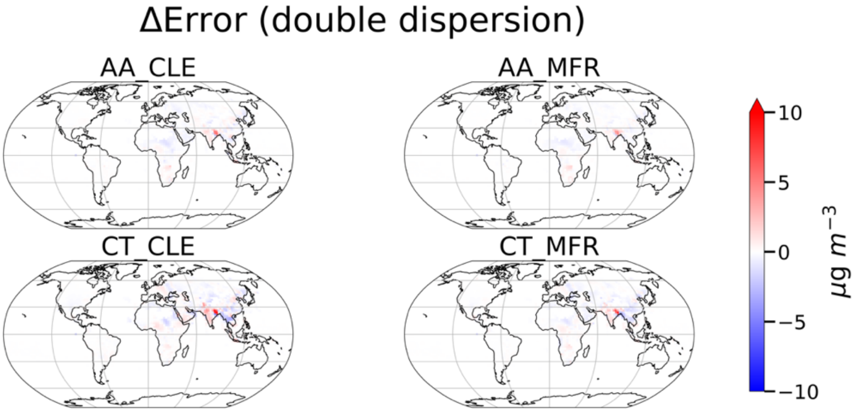


Fig. 16 Changes in absolute error (relative to GCHP-CAM output) when the dispersion length scale is doubled. Red (positive) indicates that turning off dispersion worsens the performance (increasing error), blue (negative) indicates the opposite.

Fig. 15 shows the changes in absolute error of emulator predictions when the dispersion length scale is halved. Halving the dispersion length scale increases global MAE at all the scenarios by 0.05 (AA_MFR) to 0.21 (CT_CLE) μg m^{-3}, especially over the Sahel and northern India, where absolute error increases by up to 13.2 μg m^{-3}. Fig. 16 shows changes in absolute error of emulator predictions when the dispersion length scale is doubled. Doubling the dispersion length scale leads to minor changes in global MAE across the 4 scenarios (-0.015 μg m^{-3} under CT_CLE to +0.015 μg m^{-3} under CT_MFR). The geographic pattern of emulator performance changes is similar across different scenarios. After doubling the dispersion length scale, the emulator performs better over the Sahel, China and Indochina by up to 4.8 μg m^{-3} locally, but worse over India, where the regional MAE increases by 0.28 (AA_MFR) to 1.01 μg m^{-3} (CT_CLE), and local error increases by up to 4.05 to 9.67 μg m^{-3} locally over the 4 scenarios.

The results of these sensitivity tests illustrate that the accuracy of our emulator is sensitive to the choice of dispersion kernel. While our choice of a set of globally uniform dispersion length scales provides a reasonable first-order approximation to emulate pollutant dispersion, performance of the emulator could conceivably be improved by regional, or even grid cell specific dispersion kernels. However, this will greatly increase the computing power required to train the model, and potentially require many additional global change scenarios to train and benchmark the model.

4 Comparison with AerChemMIP ensemble

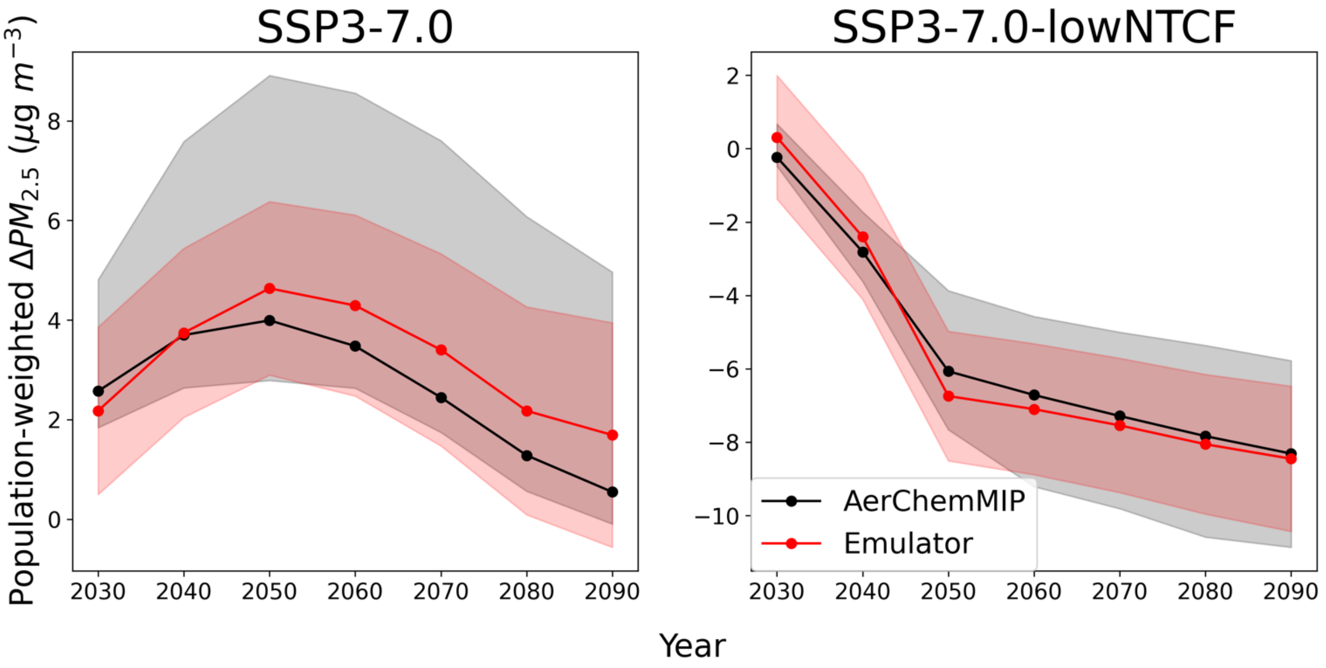


Fig. 17 Changes in global decadal mean population-weighted anthropogenic $PM_{2.5}$ relative to 2015 – 2024 mean, predicted by the models in AerChemMIP ensemble and emulator with the shades indicate the range of uncertainty (ensemble range for AerChemMIP, 1 standard deviation for emulator), under standard SSP3-7.0 and the SSP3-7.0 low Near-Term Climate Forcer (SSP3-7.0-lowNTCF) scenario

Fig. 17 shows the changes in of global population-weighted average anthropogenic $PM_{2.5}$ simulated by the models in the AerChemMIP ensemble and GW-GPR emulator over 2030 – 2090 under SSP3-7.0 and SSP3-7.0-lowNTCF scenarios, relative to 2015 – 2024 average. The solid lines represent the mean predictions, and the shaded areas represent the ranges of uncertainty (min/max for AerChemMIP ensemble, 1 standard deviation for emulator). Global population-weighted average $\Delta PM_{2.5}$ from the emulator is within the range of AerChemMIP ensemble for both SSP3-7.0 and SSP3-7.0-lowNTCF scenarios over 2030 – 2090. The emulator predicted decadal mean global population-weighted average $\Delta PM_{2.5}$ falls well within the range and differs by less than 1.15 μg m^{-3} with the mean of AerChemMIP ensemble for all decades under both scenarios.

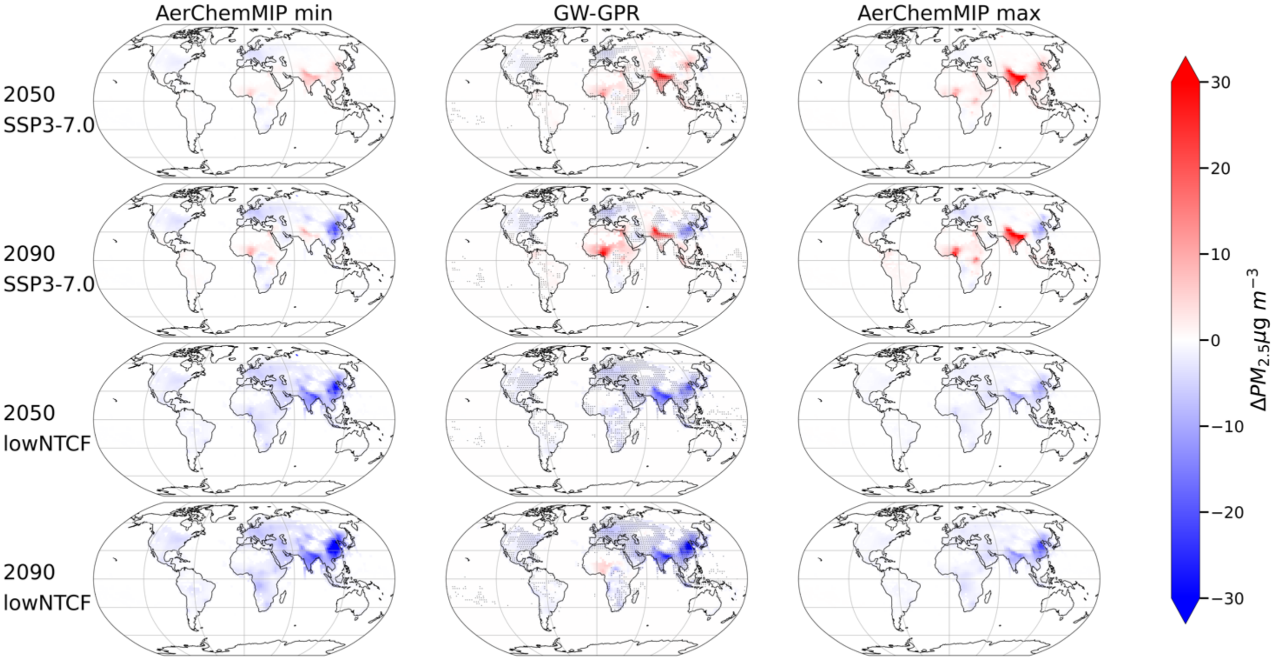


Fig. 18 Multimodel minimum and maximum magnitude of $\Delta PM_{2.5}$ simulated by models in AerChemMIP, and GW-GPR emulator predicted $\Delta PM_{2.5}$ in 2050 and 2090 under SSP3-7.0 and SSP3-7.0-lowNTCF (abbreviated as lowNTCF in plot labels). Dots in the middle (GW-GPR) column indicate the grid cells where prediction of GW-GPR do not fall between the minimum and maximum of AerChemMIP simulation output.

Fig. 18 show the multimodel minimum and maximum $\Delta PM_{2.5}$ simulated by models in AerChemMIP and GW-GPR emulator predicted $\Delta PM_{2.5}$ in 2050 and 2090 under SSP3-7.0 and SSP3-7.0-lowNTCF scenarios. Output from models in AerChemMIP is conservatively regridded to the same horizontal resolution for comparison. GW-GPR produces similar spatial patterns of $\Delta PM_{2.5}$ as AerChemMIP models (e.g. large increases and decreases in $PM_{2.5}$ over northern China and northern India) in all four scenario-year combinations shown. Over major population centres in the northern Hemisphere (eastern North America, Europe, northern India), GW-GPR emulator predictions of $\Delta PM_{2.5}$ largely fall within the range of AerChemMIP model output. This contributes to the agreement of global population-weighted average $\Delta PM_{2.5}$ between GW-GPR emulator and models in AerChemMIP (Fig.17).

However, there is systematic disagreement between GW-GPR and AerChemMIP model output over western Africa. Under SSP3-7.0, GW-GPR predicts 13.8 μg m^{-3}, while models in AerChemMIP predicts a 4.4 – 10.7 μg m^{-3} increase in anthropogenic $PM_{2.5}$ in 2090 over Nigeria. Under SSP3-7.0-lowNTCF, GW-GPR predicts a 3.0 μg m^{-3} increase, while models in AerChemMIP predict a 1.8 – 3.0 μg m^{-3} decrease in anthropogenic $PM_{2.5}$ in 2090 over Nigeria.

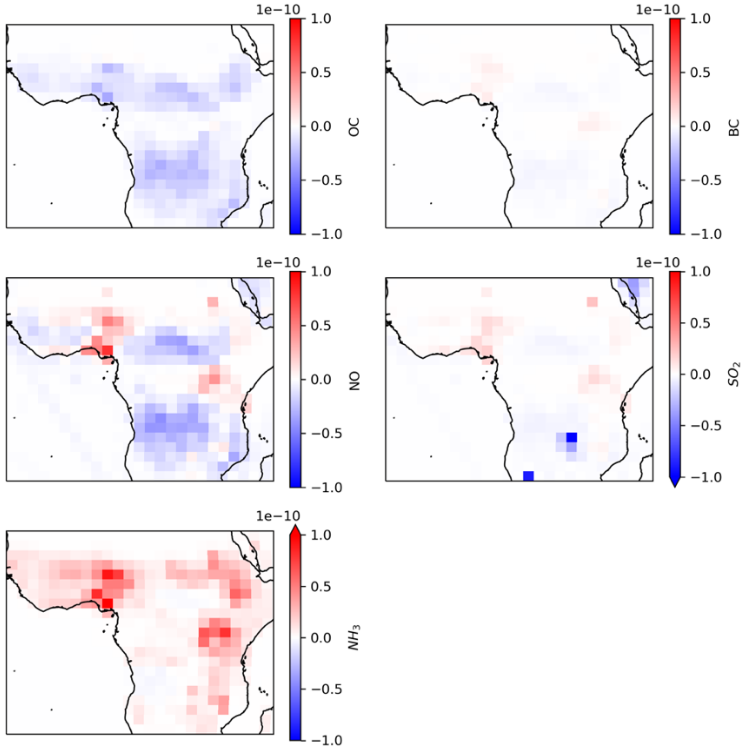


Fig. 19 Changes in air pollutant emissions (kg m^{-2} s^{-1}) in 2090 relative to 2020 over equatorial Africa under SSP3-7.0-lowNTCF

To explore the potential sources of error and bias of GW-GPR emulator, we conduct more detailed analysis of $\Delta PM_{2.5}$ over equatorial Africa in 2090 under SSP3-7.0-lowNTCF. Fig. 19 shows the changes in OC, BC, NO, SO_2 and NH_3 emissions over equatorial Africa. Over Nigeria, the magnitudes of NO (+145%), SO_2 (+180%) and NH_3 (+140%) emission changes are well beyond the range prescribed in our training set (±100%), which could lead to failure of machine learning algorithms. We also recognize GW-GPR has consistent positive biases over GCHP-CAM (Fig. 10) over equatorial Africa that cannot be effectively eliminated by switching to MLR (Fig. 11). This hints that the mismatch between regional pollutant transport patterns and prescribed dispersion kernel could be another possible source of error of GW-GPR over equatorial Africa. Theoretically, the isotropic geographic weighting scheme only emulates of primary pollutants and precursors of secondary pollutants as a purely diffusive process, but not directly emulating the dispersion of secondary pollutants nor the advective components of pollutant dispersion. This might also contribute to the failure of the emulator to predict the changes in regional aerosol background, especially when the emission changes are highly spatially heterogenous.

The comparison between GW-GPR and AerChemMIP output further confirms ability of GW-GPR in predicting the spatiotemporal changes in anthropogenic $PM_{2.5}$ exposure under global change scenarios. However, GW-GPR predictions must be interpreted cautiously when the changes in emissions are well beyond that prescribed in the training set (±100%), or there is high level of spatial heterogeneity in pollutant emission changes within a region.

5 Discussion

In this work, we apply a classic emulator building workflow (carefully sampling the input space to create samples for training machine learning models) that has been widely applied in engineering (Alizadeh et al., 2020) to build a reduced-form global air quality model from a high-fidelity global 3-D chemical transport model, GCHP-CAM. Similar techniques (also often choosing Gaussian Process Regression as the machine learning algorithm) have been used for uncertainty analysis and parameter calibration in atmospheric chemistry modelling (Reyes-Villegas et al., 2023; Ryan and Wild, 2021; Wild et al., 2020), and directly emulate air quality models at local and regional scales (Conibear et al., 2021; Vander Hoorn et al., 2022). Our work applies this approach for global change scenarios, where climate change, inter-regional chemical transport and discrepancies in chemical regimes pose another layer of challenges.

To address these challenges, there are a few unique features of the emulator architectures in comparison to other reduced-form global air quality models. We design the emulator to be usable for a wide range of integrated assessment modelling and policy evaluation, where new scenarios or greenhouse gas concentration and pollutant emissions are routinely generated, but complementary atmospheric simulations are not always available. Therefore, rather than directly using the meteorological fields as input (e.g. Chen et al., 2023), we parameterize anthropogenic climate change intensity as a function of atmospheric CO_2 concentration, and use geographic weighting (Pisoni et al., 2018) to approximate the effect of chemical transport. Rather than exploring the source-receptor relationships between pre-defined regions, the emulator is trained at individual grid cell level. Therefore, the accuracy of the emulator is not affected by different definitions of regions and sub-regional changes in spatial patterns pollutant emissions, which is important for application across different integrated assessment frameworks. This also allow us to tackle the non-linearity in the atmospheric chemical system by exploring more combinations of pollutant emission changes (more efficiently via Latin Hypercube Sampling) and machine learning (via Gaussian Process Regression). By analysing emulator performance at grid cell level, we find the GW-GPR emulator successful in reproducing the global and regional changes in $PM_{2.5}$ simulated by GCHP-CAM under 4 climate and air quality legislation (IGSM-GAINS-TAPS) scenarios and that from the AerChemMIP archive under SSP3-7.0 and SSP3-7.0-lowNTCF scenarios. We also find that the emulator may underperform when 1) the magnitude of pollutant emission changes is well beyond that prescribed in the training set (±100%); 2) the dispersion kernel ignores the advective component of pollutant transport, therefore misrepresents region-specific directional pollutant transport patterns; 3) the spatial pattern of pollutant emission changes is highly heterogenous within a region. This points to some potential ways of improving the accuracy of the GW-GPR framework (e.g. fitting anisotropic dispersion kernels for each grid cells, expanding the training set), which can be explored in the future.

In addition to the mean, GPR also calculates the standard deviation of the prediction, which can be interpreted to characterize the statistical uncertainty of emulator output. As unforced climate variability directly contributes to the interannual variabilities of $PM_{2.5}$, previous studies recommend 10 – 20 years of averaging to robustly detect the changes in $PM_{2.5}$ over contiguous United States (Brown-Steiner et al., 2018; Garcia-Menendez et al., 2017; Pienkosz et al., 2019), especially when the signal to be detected is smaller or comparable in magnitude to the underlying unforced climate variability. Due to limitations in computing time, we focus on exploring a wider range pollutant emission and climate change by generating each sample in the training set using one year of simulation, rather than running multiple years of simulation to generating robust signals amidst unforced climate variabilities for each set of perturbation experiment. This significant source of uncertainty, however, is captured by the uncertainty quantification algorithm of the GPR. Therefore, in addition to the magnitude, our emulator also provides the uncertainties in $\Delta PM_{2.5}$, which can be important in quantifying the overall uncertainties of health impacts of future air pollution (Saari et al., 2019).

In combination with the emission intensity projections from GAINS, TAPS can translate integrated assessment model output to spatially explicit air pollutant emission inventories. Combining with the GW-GPR emulator (Fig. 2), we can potentially produce gridcell-level projection of anthropogenic $PM_{2.5}$ changes for any climate and air quality integrated assessment scenarios within seconds, as demonstrated by our IGSM-GAINS-TAPS emulation exercise. This opens up the possibility for including air quality impacts within climate and sustainability decision making and scientific analysis. As climate projections move towards including scenario design as part of the uncertainty (Guivarch et al., 2022; Lamontagne et al., 2018; O'Neill et al., 2016, 2020), climate and global change scenarios generated will increase by orders of magnitude (Lamontagne et al., 2018; Shindell and Smith, 2019). Tools with proper balance between accuracy and computing resource requirements become more important in enabling uncertainty analysis and impact assessments. Our work shows the potential of machine learning techniques in enabling rapid and accurate global air quality assessment. Future work includes applying the $PM_{2.5}$ emulator to study more global change scenarios, improving and extending the emulator to calculate changes in other pollutants (e.g. O_3) and local climate forcing, and building software package and web interface to increase the accessibility of the emulator.

Author contributions

AYHW, NES and SDE conceptualized the study. EM developed the underlying CAM variant and produced the meteorological data. SDE developed and evaluated the GCHP-CAM model. AYHW conducted GCHP-CAM simulations, preparing air pollutant emissions for the IGSM-GAINS-TAPS scenarios, GW-GPR modelling, acquiring and processing AerChemMIP, Input4MIP data. AYHW analyzed the results and wrote the manuscript with input from NES and SDE. NES provided supervision and acquired funding for this project.

Competing interests

The authors have no competing interests

Acknowledgement

This study is part of the Bring Computation to the Climate Challenge (BC3) project at MIT, which is supported by Schmidt Sciences, LLC. and the MIT Climate Grand Challenges. The computation of this study was performed on the Svante computing cluster, which is supported by MIT Center for Sustainability Science and Strategy. We acknowledge the climate modelling groups participating in AerChemMIP and make the model output available through ESGF, and the funding agencies supporting the above work. We thank Will Atkinson for developing TAPS, Emmie Le Roy for configuring GCHP-CAM, and other members of the BC3 team and Selin group for their valuable insight.

Code and data availability

The GW-GPR software is publicly available at https://github.com/ayhwong/GW-GPR, with a brief user guide in the repository. A frozen version of the repository, which in addition contains the associated datasets required to reproduce the study and the figures presented in this manuscript, is available on Zenodo (Wong, 2025) (https://doi.org/10.5281/zenodo.15484655). TAPS is publicly available at https://github.com/watkin-mit/TAPS. The GAINS data used in this study were processed by Atkinson et al. (2022). The raw GAINS scenario data are available at https://gains.iiasa.ac.at/models. The source code of GCHP 13.0.0 is available at https://github.com/geoschem/GCHP/releases/tag/13.0.0. Detailed descriptions and greenhouse gas emissions for CT and AA scenarios can be found within the MIT Joint Program on the Science and Policy for Global Change 2023 Global Change Outlook (Paltsev et al., 2023). The unprocessed CAM and GCHP output are available by contacting the corresponding authors.

References

Alizadeh, R., Allen, J. K., and Mistree, F.: Managing computational complexity using surrogate models: a critical review, Res Eng Design, 31, 275–298, https://doi.org/10.1007/s00163-020-00336-7, 2020.

Amann, M., Bertok, I., Borken-Kleefeld, J., Cofala, J., Heyes, C., Höglund-Isaksson, L., Klimont, Z., Nguyen, B., Posch, M., Rafaj, P., Sandler, R., Schöpp, W., Wagner, F., and Winiwarter, W.: Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications, Environmental Modelling & Software, 26, 1489–1501, https://doi.org/10.1016/j.envsoft.2011.07.012, 2011.

Ansari, A. S. and Pandis, S. N.: Response of inorganic PM to precursor concentrations, Environmental Science and Technology, 32, 2706–2714, https://doi.org/10.1021/es971130j, 1998.

Atkinson, W., Eastham, S. D., Chen, Y.-H. H., Morris, J., Paltsev, S., Schlosser, C. A., and Selin, N. E.: A tool for air pollution scenarios (TAPS v1.0) to enable global, long-term, and flexible study of climate and air quality policies, Geosci. Model Dev., 15, 7767–7789, https://doi.org/10.5194/gmd-15-7767-2022, 2022.

Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., Meurdesoif, Y., Cadule, P., Devilliers, M., Dupont, E., and Lurton, T.: IPCC DDC: IPSL IPSL-CM5A2-INCA model output prepared for CMIP6 ScenarioMIP (1), https://doi.org/10.26050/WDCC/AR6.C6SPIPICI, 2023.

Brown-Steiner, B., Selin, N. E., Prinn, R. G., Monier, E., Tilmes, S., Emmons, L., and Garcia-Menendez, F.: Maximizing ozone signals among chemical, meteorological, and climatological variability, Atmospheric Chemistry and Physics, 18, 8373–8388, https://doi.org/10.5194/acp-18-8373-2018, 2018.

Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C. A., Apte, J. S., Brauer, M., Cohen, A., Weichenthal, S., Coggins, J., Di, Q., Brunekreef, B., Frostad, J., Lim, S. S., Kan, H., Walker, K. D., Thurston, G. D., Hayes, R. B., Lim, C. C., Turner, M. C., Jerrett, M., Krewski, D., Gapstur, S. M., Diver, W. R., Ostro, B., Goldberg, D., Crouse, D. L., Martin, R. V., Peters, P., Pinault, L., Tjepkema, M., Van Donkelaar, A., Villeneuve, P. J., Miller, A. B., Yin, P., Zhou, M., Wang, L., Janssen, N. A. H., Marra, M., Atkinson, R. W., Tsang, H., Thach, T. Q., Cannon, J. B., Allen, R. T., Hart, J. E., Laden, F., Cesaroni, G., Forastiere, F., Weinmayr, G., Jaensch, A., Nagel, G., Concin, H., and Spadaro, J. V.: Global estimates of mortality associated with longterm exposure to outdoor fine particulate matter, Proceedings of the National Academy of Sciences of the United States of America, 115, 9592–9597, https://doi.org/10.1073/pnas.1803222115, 2018.

Camilleri, S. F., Montgomery, A., Visa, M. A., Schnell, J. L., Adelman, Z. E., Janssen, M., Grubert, E. A., Anenberg, S. C., and Horton, D. E.: Air quality, health and equity implications of electrifying heavy-duty vehicles, Nat Sustain, 6, 1643–1653, https://doi.org/10.1038/s41893-023-01219-0, 2023.

Center For International Earth Science Information Network-CIESIN-Columbia University: Gridded Population of the World, Version 4 (GPWv4): National Identifier Grid, Revision 11, https://doi.org/10.7927/H4TD9VDP, 2018.

Chen, W., Lu, X., Yuan, D., Chen, Y., Li, Z., Huang, Y., Fung, T., Sun, H., and Fung, J. C. H.: Global PM $_{2.5}$ Prediction and Associated Mortality to 2100 under Different Climate Change Scenarios, Environ. Sci. Technol., 57, 10039–10052, https://doi.org/10.1021/acs.est.3c03804, 2023.

Chen, Y.-H. H., Paltsev, S., Gurgel, A., Reilly, J. M., and Morris, J.: A Multisectoral Dynamic Model for Energy, Economic, and Climate Scenario Analysis, LCE, 13, 70–111, https://doi.org/10.4236/lce.2022.132005, 2022.

Colette, A., Rouïl, L., Meleux, F., Lemaire, V., and Raux, B.: Air Control Toolbox (ACT_v1.0): a flexible surrogate model to explore mitigation scenarios in air quality forecasts, Geosci. Model Dev., 15, 1441–1465, https://doi.org/10.5194/gmd-15-1441-2022, 2022.

Collins, W. D., Rasch, P. J., Boville, B. A., Hack, J. J., McCaa, J. R., Williamson, D. L., Briegleb, B. P., Bitz, C. M., Lin, S.-J., and Zhang, M.: The Formulation and Atmospheric Simulation of the Community Atmosphere Model Version 3 (CAM3), Journal of Climate, 19, 2144–2161, https://doi.org/10.1175/JCLI3760.1, 2006.

Collins, W. J., Lamarque, J.-F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., Maycock, A., Myhre, G., Prather, M., Shindell, D., and Smith, S. J.: AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6, Geosci. Model Dev., 10, 585–607, https://doi.org/10.5194/gmd-10-585-2017, 2017.

Conibear, L., Reddington, C. L., Silver, B. J., Chen, Y., Knote, C., Arnold, S. R., and Spracklen, D. V.: Statistical Emulation of Winter Ambient Fine Particulate Matter Concentrations From Emission Changes in China, GeoHealth, 5, e2021GH000391, https://doi.org/10.1029/2021GH000391, 2021.

Conibear, L., Reddington, C. L., Silver, B. J., Chen, Y., Knote, C., Arnold, S. R., and Spracklen, D. V.: Sensitivity of Air Pollution Exposure and Disease Burden to Emission Changes in China Using Machine Learning Emulation, GeoHealth, 6, e2021GH000570, https://doi.org/10.1029/2021GH000570, 2022.

Conover, W. J. and Iman, R. L.: Analysis of Covariance Using the Rank Transformation, Biometrics, 38, 715, https://doi.org/10.2307/2530051, 1982.

Consortium (EC-Earth), E.-E.: IPCC DDC: EC-Earth-Consortium EC-Earth3-AerChem model output prepared for CMIP6 AerChemMIP, https://doi.org/10.26050/WDCC/AR6.C6ACEEEEA, 2023.

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R., Boussetta, S., Caron, L.-P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., Fuentes-Franco, R., Gröger, M., V. Hardenberg, J., Hieronymus, J., Karami, M. P., Keskinen, J.-P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégoz, M., Miller, P. A., Moreno-Chamarro, E., Nieradzik, L., Van Noije, T., Nolan, P., O'Donnell, D., Ollinaho, P., Van Den Oord, G., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Le Sager, P., Schmith, T., Schrödner, R., Serva, F., Sicardi, V., Sloth Madsen, M., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., and Zhang, Q.: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6, Geosci. Model Dev., 15, 2973–3020, https://doi.org/10.5194/gmd-15-2973-2022, 2022.

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P. P. G., Griffies, S. M., Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P. C. D., Nikonov, S., Paynter, D. J., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T., Schwarzkopf, D. M., Sentman, L. T., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A. T., Wyman, B., Zeng, Y., and Zhao, M.: The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics, J Adv Model Earth Syst, 12, e2019MS002015, https://doi.org/10.1029/2019MS002015, 2020.

Eastham, S. D., Long, M. S., Keller, C. A., Lundgren, E., Yantosca, R. M., Zhuang, J., Li, C., Lee, C. J., Yannetti, M., Auer, B. M., Clune, T. L., Kouatchou, J., Putman, W. M., Thompson, M. A., Trayanov, A. L., Molod, A. M., Martin, R. V., and Jacob, D. J.: GEOS-Chem High Performance (GCHP v11-02c): a next-generation implementation of the GEOS-Chem chemical transport model for massively parallel applications, Geosci. Model Dev., 11, 2941–2953, https://doi.org/10.5194/gmd-11-2941-2018, 2018.

Eastham, S. D., Monier, E., Rothenberg, D., Paltsev, S., and Selin, N. E.: Rapid Estimation of Climate–Air Quality Interactions in Integrated Assessment Using a Response Surface Model, ACS Environ. Au, 3, 153–163, https://doi.org/10.1021/acsenvironau.2c00054, 2023.

Fountoukis, C. and Nenes, A.: ISORROPIAII: A computationally efficient thermodynamic equilibrium model for K+-Ca2+-Mg2+-NH4+-Na+-SO42--NO3--Cl--H2O aerosols, Atmospheric Chemistry and Physics, 7, 4639–4659, https://doi.org/10.5194/acp-7-4639-2007, 2007.

Fujimori, S., Hasegawa, T., Masui, T., Takahashi, K., Herran, D. S., Dai, H., Hijioka, Y., and Kainuma, M.: SSP3: AIM implementation of Shared Socioeconomic Pathways, Global Environmental Change, 42, 268–283, https://doi.org/10.1016/j.gloenvcha.2016.06.009, 2017.

GAINS Developer Team: GAINS 4.01 IAM model outputs, IIASA, 2021.

Galmarini, S., Koffi, B., Solazzo, E., Keating, T., Hogrefe, C., Schulz, M., Benedictow, A., Griesfeller, J. J., Janssens-Maenhout, G., Carmichael, G., Fu, J., and Dentener, F.: Technical note: Coordination and harmonization of the multi-scale, multi-model activities HTAP2, AQMEII3, and MICS-Asia3: simulations, emission inventories, boundary conditions, and model output formats, Atmos. Chem. Phys., 17, 1543–1555, https://doi.org/10.5194/acp-17-1543-2017, 2017.

Garcia-Menendez, F., Saari, R. K., Monier, E., and Selin, N. E.: U.S. Air Quality and Health Benefits from Avoided Climate Change under Greenhouse Gas Mitigation, Environ. Sci. Technol., 49, 7580–7588, https://doi.org/10.1021/acs.est.5b01324, 2015.

Garcia-Menendez, F., Monier, E., and Selin, N. E.: The role of natural variability in projections of climate change impacts on U.S. ozone pollution, Geophysical Research Letters, 44, 2911–2921, https://doi.org/10.1002/2016GL071565, 2017.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G. K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), Journal of Climate, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S., Luderer, G., Kriegler, E., Van Vuuren, D. P., Van Den Berg, M., Feng, L., Klein, D., Calvin, K., Doelman, J. C., Frank, S., Fricko, O., Harmsen, M., Hasegawa, T., Havlik, P., Hilaire, J., Hoesly, R., Horing, J., Popp, A., Stehfest, E., and Takahashi, K.: Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century, Geosci. Model Dev., 12, 1443–1475, https://doi.org/10.5194/gmd-12-1443-2019, 2019.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions, Geoscientific Model Development, 5, 1471–1492, https://doi.org/10.5194/gmd-5-1471-2012, 2012.

Guivarch, C., Le Gallic, T., Bauer, N., Fragkos, P., Huppmann, D., Jaxa-Rozen, M., Keppo, I., Kriegler, E., Krisztin, T., Marangoni, G., Pye, S., Riahi, K., Schaeffer, R., Tavoni, M., Trutnevyte, E., Van Vuuren, D., and Wagner, F.: Using large ensembles of climate change mitigation scenarios for robust insights, Nat. Clim. Chang., 12, 428–435, https://doi.org/10.1038/s41558-022-01349-x, 2022.

Herman, J. and Usher, W.: SALib: An open-source Python library for Sensitivity Analysis, JOSS, 2, 97, https://doi.org/10.21105/joss.00097, 2017.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geosci. Model Dev., 11, 369–408, https://doi.org/10.5194/gmd-11-369-2018, 2018.

Horowitz, L. W., Naik, V., Sentman, L., Paulot, F., Blanton, C., McHugh, C., Radhakrishnan, A., Rand, K., Vahlenkamp, H., Zadeh, N. T., Wilson, C., Ginoux, P., He, J., John, J. G., Lin, M., Paynter, D. J., Ploshay, J., Zhang, A., and Zeng, Y.: IPCC DDC: NOAA-GFDL GFDL-ESM4 model output prepared for CMIP6 AerChemMIP (1), https://doi.org/10.26050/WDCC/AR6.C6ACNGGFE, 2023.

Huang, X., Srikrishnan, V., Lamontagne, J., Keller, K., and Peng, W.: Effects of global climate mitigation on regional air quality and health, Nat Sustain, 6, 1054–1066, https://doi.org/10.1038/s41893-023-01133-5, 2023.

Hudman, R. C., Moore, N. E., Mebust, A. K., Martin, R. V., Russell, A. R., Valin, L. C., and Cohen, R. C.: Steps towards a mechanistic model of global soil nitric oxide emissions: Implementation and space based-constraints, Atmospheric Chemistry and Physics, 12, 7779–7795, https://doi.org/10.5194/acp-12-7779-2012, 2012.

Iwanaga, T., Usher, W., and Herman, J.: Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses, SESMO, 4, 18155, https://doi.org/10.18174/sesmo.18155, 2022.

Jacob, D. J. and Winner, D. A.: Effect of climate change on air quality, Atmospheric Environment, 43, 51–63, https://doi.org/10.1016/j.atmosenv.2008.09.051, 2009.

Jones, P. W.: First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates, Mon. Wea. Rev., 127, 2204–2210, https://doi.org/10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2, 1999.

Kawamiya, M., Hajima, T., Tachiiri, K., Watanabe, S., and Yokohata, T.: Two decades of Earth system modeling with an emphasis on Model for Interdisciplinary Research on Climate (MIROC), Prog Earth Planet Sci, 7, 64, https://doi.org/10.1186/s40645-020-00369-5, 2020.

Kelley, M., Schmidt, G. A., Nazarenko, L. S., Bauer, S. E., Ruedy, R., Russell, G. L., Ackerman, A. S., Aleinov, I., Bauer, M., Bleck, R., Canuto, V., Cesana, G., Cheng, Y., Clune, T. L., Cook, B. I., Cruz, C. A., Del Genio, A. D., Elsaesser, G. S., Faluvegi, G., Kiang, N. Y., Kim, D., Lacis, A. A., Leboissetier, A., LeGrande, A. N., Lo, K. K., Marshall, J., Matthews, E. E., McDermid, S., Mezuman, K., Miller, R. L., Murray, L. T., Oinas, V., Orbe, C., García-Pando, C. P., Perlwitz, J. P., Puma, M. J., Rind, D., Romanou, A., Shindell, D. T., Sun, S., Tausnev, N., Tsigaridis, K., Tselioudis, G., Weng, E., Wu, J., and Yao, M.: GISS-E2.1: Configurations and Climatology, J Adv Model Earth Syst, 12, e2019MS002025, https://doi.org/10.1029/2019MS002025, 2020.

Klimont, Z., Kupiainen, K., Heyes, C., Purohit, P., Cofala, J., Rafaj, P., Borken-Kleefeld, J., and Schöpp, W.: Global anthropogenic emissions of particulate matter including black carbon, Atmos. Chem. Phys., 17, 8681–8723, https://doi.org/10.5194/acp-17-8681-2017, 2017.

Lamontagne, J. R., Reed, P. M., Link, R., Calvin, K. V., Clarke, L. E., and Edmonds, J. A.: Large Ensemble Analytic Framework for Consequence-Driven Discovery of Climate Change Scenarios, Earth's Future, 6, 488–504, https://doi.org/10.1002/2017EF000701, 2018.

Latimer, R. N. C. and Martin, R. V.: Interpretation of measured aerosol mass scattering efficiency over North America using a chemical transport model, Atmos. Chem. Phys., 19, 2635–2653, https://doi.org/10.5194/acp-19-2635-2019, 2019.

Liang, C.-K., West, J. J., Silva, R. A., Bian, H., Chin, M., Davila, Y., Dentener, F. J., Emmons, L., Flemming, J., Folberth, G., Henze, D., Im, U., Jonson, J. E., Keating, T. J., Kucsera, T., Lenzen, A., Lin, M., Lund, M. T., Pan, X., Park, R. J., Pierce, R. B., Sekiya, T., Sudo, K., and Takemura, T.: HTAP2 multi-model estimates of premature human mortality due to intercontinental transport of air pollution and emission sectors, Atmos. Chem. Phys., 18, 10497–10520, https://doi.org/10.5194/acp-18-10497-2018, 2018.

McKay, M. D., Beckman, R. J., and Conover, W. J.: Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, Technometrics, 21, 239–245, https://doi.org/10.1080/00401706.1979.10489755, 1979.

Meinshausen, M., Nicholls, Z. R. J., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., Canadell, J. G., Daniel, J. S., John, A., Krummel, P. B., Luderer, G., Meinshausen, N., Montzka, S. A., Rayner, P. J., Reimann, S., Smith, S. J., Van Den Berg, M., Velders, G. J. M., Vollmer, M. K., and Wang, R. H. J.: The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500, Geosci. Model Dev., 13, 3571–3605, https://doi.org/10.5194/gmd-13-3571-2020, 2020.

Meng, J., Martin, R. V., Ginoux, P., Hammer, M., Sulprizio, M. P., Ridley, D. A., and Van Donkelaar, A.: Grid-independent high-resolution dust emissions (v1.0) for chemical transport models: application to GEOS-Chem (12.5.0), Geosci. Model Dev., 14, 4249–4260, https://doi.org/10.5194/gmd-14-4249-2021, 2021.

Monier, E., Scott, J. R., Sokolov, A. P., Forest, C. E., and Schlosser, C. A.: An integrated assessment modeling framework for uncertainty studies in global and regional climate change: the MIT IGSM-CAM (version 1.0), Geosci. Model Dev., 6, 2063–2085, https://doi.org/10.5194/gmd-6-2063-2013, 2013.

Monier, E., Gao, X., Scott, J. R., Sokolov, A. P., and Schlosser, C. A.: A framework for modeling uncertainty in regional climate change, Climatic Change, 131, 51–66, https://doi.org/10.1007/s10584-014-1112-5, 2015.

Monier, E., Paltsev, S., Sokolov, A., Chen, Y.-H. H., Gao, X., Ejaz, Q., Couzo, E., Schlosser, C. A., Dutkiewicz, S., Fant, C., Scott, J., Kicklighter, D., Morris, J., Jacoby, H., Prinn, R., and Haigh, M.: Toward a consistent modeling framework to assess multi-sectoral climate impacts, Nat Commun, 9, 660, https://doi.org/10.1038/s41467-018-02984-9, 2018.

Murray, L. T., Jacob, D. J., Logan, J. A., Hudman, R. C., and Koshak, W. J.: Optimized regional and interannual variability of lightning in a global chemical transport model constrained by LIS/OTD satellite data, J. Geophys. Res., 117, 2012JD017934, https://doi.org/10.1029/2012JD017934, 2012.

NASA Goddard Institute For Space Studies (NASA/GISS): IPCC DDC: NASA-GISS GISS-E2.1G model output prepared for CMIP6 AerChemMIP (1), https://doi.org/10.26050/WDCC/AR6.C6ACGIGE1, 2023a.

NASA Goddard Institute For Space Studies (NASA/GISS): IPCC DDC: NASA-GISS GISS-E2.1G model output prepared for CMIP6 AerChemMIP ssp370-lowNTCF (1), https://doi.org/10.26050/WDCC/AR6.C6ACGIGE1SL, 2023b.

NASA Goddard Institute For Space Studies (NASA/GISS): IPCC DDC: NASA-GISS GISS-E2.1H model output prepared for CMIP6 CMIP (1), https://doi.org/10.26050/WDCC/AR6.C6CMGIGH1, 2023c.

Nations, U.: Transforming our world: the 2030 Agenda for Sustainable Development, 2015.

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, Geosci. Model Dev., 9, 3461–3482, https://doi.org/10.5194/gmd-9-3461-2016, 2016.

O'Neill, B. C., Carter, T. R., Ebi, K., Harrison, P. A., Kemp-Benedict, E., Kok, K., Kriegler, E., Preston, B. L., Riahi, K., Sillmann, J., Van Ruijven, B. J., Van Vuuren, D., Carlisle, D., Conde, C., Fuglestvedt, J., Green, C., Hasegawa, T., Leininger, J., Monteith, S., and Pichs-Madruga, R.: Achievements and needs for the climate change scenario framework, Nat. Clim. Chang., 10, 1074–1084, https://doi.org/10.1038/s41558-020-00952-0, 2020.

Paltsev, S., Schlosser, C. A., Chen, H., Gao, X., Gurgel, A., Jacoby, H., Morris, J., Prinn, R. G., Reilly, J. M., Salunke, P., and Sokolov, A. P.: 2023 Global Change Outlook, MIT Joint Program on the Science and Policy of Global Change, 2023.

Paulot, F. and Jacob, D. J.: Hidden cost of U.S. agricultural exports: Particulate matter from ammonia emissions, Environmental Science and Technology, 48, 903–908, https://doi.org/10.1021/es4034793, 2014.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., and others: Scikit-learn: Machine learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Pienkosz, B. D., Saari, R. K., Monier, E., and Garcia-Menendez, F.: Natural Variability in Projections of Climate Change Impacts on Fine Particulate Matter Pollution, Earth's Future, 7, 762–770, https://doi.org/10.1029/2019EF001195, 2019.

Pisoni, E., Clappier, A., Degraeuwe, B., and Thunis, P.: Adding spatial flexibility to source-receptor relationships for air quality modeling, Environmental Modelling & Software, 90, 68–77, https://doi.org/10.1016/j.envsoft.2017.01.001, 2017.

Pisoni, E., Albrecht, D., Mara, T. A., Rosati, R., Tarantola, S., and Thunis, P.: Application of uncertainty and sensitivity analysis to the air quality SHERPA modelling tool, Atmospheric Environment, 183, 84–93, https://doi.org/10.1016/j.atmosenv.2018.04.006, 2018.

Possell, M. and Hewitt, C. N.: Isoprene emissions from plants are mediated by atmospheric CO2 concentrations, Global Change Biology, 17, 1595–1610, https://doi.org/10.1111/j.1365-2486.2010.02306.x, 2011.

Rao, S., Klimont, Z., Smith, S. J., Van Dingenen, R., Dentener, F., Bouwman, L., Riahi, K., Amann, M., Bodirsky, B. L., Van Vuuren, D. P., Aleluia Reis, L., Calvin, K., Drouet, L., Fricko, O., Fujimori, S., Gernaat, D., Havlik, P., Harmsen, M., Hasegawa, T., Heyes, C., Hilaire, J., Luderer, G., Masui, T., Stehfest, E., Strefler, J., Van Der Sluis, S., and Tavoni, M.: Future air pollution in the Shared Socio-economic Pathways, Global Environmental Change, 42, 346–358, https://doi.org/10.1016/j.gloenvcha.2016.05.012, 2017.

Reis, L. A., Drouet, L., and Tavoni, M.: Internalising health-economic impacts of air pollution into climate policy: a global modelling study, The Lancet Planetary Health, 6, e40–e48, https://doi.org/10.1016/S2542-5196(21)00259-X, 2022.

Reyes-Villegas, E., Lowe, D., Johnson, J. S., Carslaw, K. S., Darbyshire, E., Flynn, M., Allan, J. D., Coe, H., Chen, Y., Wild, O., Archer-Nicholls, S., Archibald, A., Singh, S., Shrivastava, M., Zaveri, R. A., Singh, V., Beig, G., Sokhi, R., and McFiggans, G.: Simulating organic aerosol in Delhi with WRF-Chem using the volatility-basis-set approach: exploring model uncertainty with a Gaussian process emulator, Atmos. Chem. Phys., 23, 5763–5782, https://doi.org/10.5194/acp-23-5763-2023, 2023.

Ryan, E. and Wild, O.: Calibrating a global atmospheric chemistry transport model using Gaussian process emulation and ground-level concentrations of ozone and carbon monoxide, Geosci. Model Dev., 14, 5373–5391, https://doi.org/10.5194/gmd-14-5373-2021, 2021.

Saari, R. K., Mei, Y., Monier, E., and Garcia-Menendez, F.: Effect of Health-Related Uncertainty and Natural Variability on Health Impacts and Cobenefits of Climate Policy, Environ. Sci. Technol., 53, 1098–1108, https://doi.org/10.1021/acs.est.8b05094, 2019.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S.: Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, Computer Physics Communications, 181, 259–270, https://doi.org/10.1016/j.cpc.2009.09.018, 2010.

Sang, S., Chu, C., Zhang, T., Chen, H., and Yang, X.: The global burden of disease attributable to ambient fine particulate matter in 204 countries and territories, 1990–2019: A systematic analysis of the Global Burden of Disease Study 2019, Ecotoxicology and Environmental Safety, 238, 113588, https://doi.org/10.1016/j.ecoenv.2022.113588, 2022.

Sepulchre, P., Caubel, A., Ladant, J.-B., Bopp, L., Boucher, O., Braconnot, P., Brockmann, P., Cozic, A., Donnadieu, Y., Dufresne, J.-L., Estella-Perez, V., Ethé, C., Fluteau, F., Foujols, M.-A., Gastineau, G., Ghattas, J., Hauglustaine, D., Hourdin, F., Kageyama, M., Khodri, M., Marti, O., Meurdesoif, Y., Mignot, J., Sarr, A.-C., Servonnat, J., Swingedouw, D., Szopa, S., and Tardif, D.: IPSL-CM5A2 – an Earth system model designed for multi-millennial climate simulations, Geosci. Model Dev., 13, 3011–3053, https://doi.org/10.5194/gmd-13-3011-2020, 2020.

Sherwen, T., Schmidt, J. A., Evans, M. J., Carpenter, L. J., Großmann, K., Eastham, S. D., Jacob, D. J., Dix, B., Koenig, T. K., Sinreich, R., Ortega, I., Volkamer, R., Saiz-Lopez, A., Prados-Roman, C., Mahajan, A. S., and Ordóñez, C.: Global impacts of tropospheric halogens (Cl, Br, I) on oxidants and composition in GEOS-Chem, Atmos. Chem. Phys., 16, 12239–12271, https://doi.org/10.5194/acp-16-12239-2016, 2016.

Shindell, D. and Smith, C. J.: Climate and air-quality benefits of a realistic phase-out of fossil fuels, Nature, 573, 408–411, https://doi.org/10.1038/s41586-019-1554-z, 2019.

https://doi.org/10.5194/egusphere-2025-2663
Preprint. Discussion started: 14 July 2025

Smith, S. J., Klimont, Z., Drouet, L., Harmsen, M., Luderer, G., Riahi, K., Van Vuuren, D. P., and Weyant, J. P.: The Energy Modeling Forum (EMF)-30 study on short-lived climate forcers: introduction and overview, Climatic Change, 163, 1399–1408, https://doi.org/10.1007/s10584-020-02938-5, 2020.

Sobol′, I. M.: Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Mathematics and Computers in Simulation, 55, 271–280, https://doi.org/10.1016/S0378-4754(00)00270-6, 2001.

Sokolov, A., Kicklighter, D., Schlosser, A., Wang, C., Monier, E., Brown-Steiner, B., Prinn, R., Forest, C., Gao, X., Libardoni, A., and Eastham, S.: Description and Evaluation of the MIT Earth System Model (MESM), J Adv Model Earth Syst, 10, 1759–1789, https://doi.org/10.1029/2018MS001277, 2018.

Stohl, A., Aamaas, B., Amann, M., Baker, L. H., Bellouin, N., Berntsen, T. K., Boucher, O., Cherian, R., Collins, W., Daskalakis, N., Dusinska, M., Eckhardt, S., Fuglestvedt, J. S., Harju, M., Heyes, C., Hodnebrog, Ø., Hao, J., Im, U., Kanakidou, M., Klimont, Z., Kupiainen, K., Law, K. S., Lund, M. T., Maas, R., MacIntosh, C. R., Myhre, G., Myriokefalitakis, S., Olivié, D., Quaas, J., Quennehen, B., Raut, J.-C., Rumbold, S. T., Samset, B. H., Schulz, M., Seland, Ø., Shine, K. P., Skeie, R. B., Wang, S., Yttri, K. E., and Zhu, T.: Evaluating the climate and air quality impacts of short-lived pollutants, Atmos. Chem. Phys., 15, 10529–10566, https://doi.org/10.5194/acp-15-10529-2015, 2015.

Tai, A. P. K., Mickley, L. J., Heald, C. L., and Wu, S.: Effect of CO2 inhibition on biogenic isoprene emission: Implications for air quality under 2000 to 2050 changes in climate, vegetation, and land use, Geophysical Research Letters, 40, 3479–3483, https://doi.org/10.1002/grl.50650, 2013.

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

Thakrar, S. K., Tessum, C. W., Apte, J. S., Balasubramanian, S., Millet, D. B., Pandis, S. N., Marshall, J. D., and Hill, J. D.: Global, high-resolution, reduced-complexity air quality modeling for PM2.5 using InMAP (Intervention Model for Air Pollution), PLoS ONE, 17, e0268714, https://doi.org/10.1371/journal.pone.0268714, 2022.

The International GEOS-Chem User Community: geoschem/GCHP: GCHP 14.4.3, , https://doi.org/10.5281/ZENODO.4428926, 2024.

Turnock, S. T., Allen, R. J., Andrews, M., Bauer, S. E., Deushi, M., Emmons, L., Good, P., Horowitz, L., John, J. G., Michou, M., Nabat, P., Naik, V., Neubauer, D., O'Connor, F. M., Olivié, D., Oshima, N., Schulz, M., Sellar, A., Shim, S., Takemura, T., Tilmes, S., Tsigaridis, K., Wu, T., and Zhang, J.: Historical and future changes in air pollutants from CMIP6 models, Atmos. Chem. Phys., 20, 14547–14579, https://doi.org/10.5194/acp-20-14547-2020, 2020.

Van Dingenen, R., Dentener, F., Crippa, M., Leitao, J., Marmer, E., Rao, S., Solazzo, E., and Valentini, L.: TM5-FASST: a global atmospheric source–receptor model for rapid impact analysis of emission changes on air quality and short-lived climate pollutants, Atmos. Chem. Phys., 18, 16173–16211, https://doi.org/10.5194/acp-18-16173-2018, 2018.

Vander Hoorn, S., Johnson, J. S., Murray, K., Smit, R., Heyworth, J., Lam, S., and Cope, M.: Emulation of a Chemical Transport Model to Assess Air Quality under Future Emission Scenarios for the Southwest of Western Australia, Atmosphere, 13, 2009, https://doi.org/10.3390/atmos13122009, 2022.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., Van Der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., Van Mulbregt, P., SciPy 1.0 Contributors, Vijaykumar, A., Bardelli, A. P., Rothberg, A., Hilboll, A., Kloeckner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C. N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D. A., Hagen, D. R., Pasechnik, D. V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G. A., Ingold, G.-L., Allen, G. E., Lee, G. R., Audren, H., Probst, I., Dietrich, J. P., Silterra, J., Webber, J. T., Slavič, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J. L., De Miranda Cardoso, J. V., Reimer, J., Harrington, J., Rodríguez, J. L. C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmerer, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N. J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P. A., Lee, P., McGibbon, R. T., Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., et al.: SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Watanabe, S., Hajima, T., Sudo, K., Abe, M., Arakawa, O., Ogochi, K., Arakawa, T., Tatebe, H., Ito, A., Ito, A., Komuro, Y., Nitta, T., Noguchi, M. A., Ogura, T., Ohgaito, R., Sekiguchi, M., Suzuki, T., Tachiiri, K., Takata, K., Takemura, T., Watanabe, M., Yamamoto, A., Yamazaki, D., Yoshimura, K., and Kawamiya, M.: IPCC DDC: MIROC MIROC-ES2H model output prepared for CMIP6 CMIP (1), https://doi.org/10.26050/WDCC/AR6.C6CMMIMIH, 2023.

WHO: WHO methods and data sources for country-level causes of death 2000-2016, 2018.

Wild, O., Voulgarakis, A., O'Connor, F., Lamarque, J.-F., Ryan, E. M., and Lee, L.: Global sensitivity analysis of chemistry–climate model budgets of tropospheric ozone and OH: exploring model diversity, Atmos. Chem. Phys., 20, 4047–4058, https://doi.org/10.5194/acp-20-4047-2020, 2020.

Williams, C. and Rasmussen, C.: Gaussian Processes for Regression, in: Advances in Neural Information Processing Systems, 1995.

Wong, A. Y. H.: Models and Data for Reproducing the Result in "A Geographically Weighted Gaussian Process Regression Emulator of the GCHP 13.0.0 Global Air Quality Model," https://doi.org/10.5281/ZENODO.15484655, 21 May 2025.