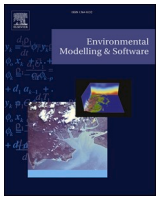
Environmental Modelling and Software 197 (2026) 106827

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: www.elsevier.com/locate/envsoft



# A generalized user-friendly method for fusing observational data and chemical transport model (Gen-Friberg V1.0: GF-1)

Zongrun Li [a], Abiola S. Lawal [a,b], Bingqing Zhang [c], Kamal J. Maji [a], Pengfei Liu [c], Yongtao Hu [a], Armistead G. Russell [a], M. Talat Odman [a,*]

[a] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, 30332, USA
[b] School of Civil and Environmental Engineering, University of Connecticut, Storrs, CT, 06269, USA
[c] School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, 30332, USA

ARTICLE INFO

*Keywords:*
Chemical transport model
Data fusion
Health assessments
Software
User-friendly

ABSTRACT

A generalized, user-friendly data fusion method (Gen-Friberg) to reduce differences between chemical transport models (CTMs) and observational data is implemented to be compatible with widely used CTMs such as CMAQ, GEOS-Chem, and WRF-Chem. Key source code improvements included encapsulating the data fusion algorithm within a single function and enabling parallel processing to minimize runtime for long simulations. We applied the data fusion method to CMAQ outputs and observations from 2010 to 2019 to evaluate the method's performance. After data fusion, pollutant concentration fields showed improved performance. Additionally, we assessed the generalizability of the data fusion method by demonstrating its effectiveness in reducing bias in the GEOS-Chem and WRF-Chem concentration fields using evaluations based on 2017 simulations. Comparisons across CMAQ, GEOS-Chem, and WRF-Chem with and without data fusion demonstrate that data fusion reduces inter-model discrepancies, yielding more consistent concentration fields for use in health and policy assessments.

## 1. Introduction

Accurate concentration fields of ambient air species are critical to performing epidemiologic analyses and understanding inequalities in potential exposures among various socioeconomic communities. Such fields are central to environmental justice and public health studies for a thorough understanding of morbidity and mortality impacts from air pollution exposure.

Data from observational networks (CSN; CASTNET; Ng et al., 2022; Hand et al., 2011; EMEP, 2018; CNEMC), consisting of monitors managed by national, state, local, and tribal agencies, are often used to assess air pollution levels and their spatiotemporal patterns. However, due to high operation costs, monitors in these networks tend to be sparse in their spatial coverage and are also limited in their measurement frequencies, as some particulate matter (PM) species are not reported or recorded for each day but instead once every three to six days (CSN). The limited spatial and temporal coverage of monitoring data adds uncertainties when directly used in epidemiological studies, particularly if spatial gradients are significant.

To address the problem, mathematical interpolation methods can be applied to estimate the concentration field from observation data (Lin et al., 2018; Li et al., 2016). However, the interpolation methods do not consider the physics and chemistry behind the pollution transport and evolution, adding to uncertainty, especially in areas with limited monitoring sites. Chemical transport models (CTMs), such as GEOS-Chem (GC) (Bey et al., 2001), the Community Multiscale Air Quality Modeling System (CMAQ) (Byun and Schere, 2006), and the Weather Research and Forecasting (WRF) model coupled with Chemistry (WRF-Chem) (Grell et al., 2005), take into account emissions, meteorology, and chemistry and can provide complete spatiotemporal concentration fields over their domains of application. GEOS-Chem is a global CTM to simulate atmospheric composition, currently developed and maintained by Harvard University and Washington University in St. Louis. It is driven by reanalysis meteorological data from NASA's Global Modeling and Assimilation Office (GMAO). In contrast, CMAQ and WRF-Chem are primarily used for regional air quality simulations. CMAQ is developed and maintained by the U.S. Environmental Protection Agency (EPA) and uses preprocessed meteorological fields from models, such as WRF. Once the meteorological simulations are completed, CMAQ can be run independently, which allows for faster air

* Corresponding author.
*E-mail address:* odman@gatech.edu (M.T. Odman).

https://doi.org/10.1016/j.envsoft.2025.106827
Received 11 September 2025; Received in revised form 10 November 2025; Accepted 5 December 2025
Available online 6 December 2025

quality simulations and sensitivity analyses without rerunning the meteorology. So, the model is commonly used for evaluating emission control strategies and conducting attainment tests. WRF-Chem, developed by the National Center for Atmospheric Research (NCAR) and the National Oceanic and Atmospheric Administration (NOAA), simulates atmospheric composition along with meteorology, explicitly representing two-way interactions between chemical and meteorological processes. The model is advantageous for research on chemical-meteorological interactions, such as aerosol-radiation-cloud interactions (Jerez et al., 2021; Yang et al., 2020; Archer-Nicholls et al., 2016). In addition to providing continuous spatiotemporal concentration fields, CTMs can also simulate counterfactual scenarios to understand health impacts from specific pollutant sources or emission control policies (Skipper et al., 2023; Henneman et al., 2017). Unfortunately, CTM outputs are uncertain due to uncertainties in meteorological inputs (Gilliam et al., 2015; Garcia-Menendez et al., 2013), simplifications in emissions' vertical and temporal profiles (Lawal et al., 2022; Li et al., 2023a), inaccuracies in emission intensity estimates (Hanna et al., 2005; Zhao et al., 2017), and reduced complexity of chemical species and mechanisms (Dodge, 2000; Cao et al., 2021; Huijnen et al., 2019).

To address the limitations and leverage the strengths of both observations and CTMs, data fusion methods are employed to enhance the accuracy of concentration field estimations. Traditional data fusion methods often use statistical and geographical interpolation to fuse different data sources, including satellite retrievals, CTM simulations, and observations. Van Donkelaar et al. (van Donkelaar et al., 2019) generated surface concentration fields of $PM_{2.5}$ chemical components by integrating aerosol optical depth (AOD) from multiple satellite products and GEOS-Chem simulations. Geographically Weighted Regression (GWR) was then applied to merge observational data with those concentration fields, reducing the bias between observations and the modeled concentrations. Xue et al. (2017) developed a three-step method to fuse satellite, CMAQ, and observation data. The linear mixed effect model was applied to predict surface $PM_{2.5}$ from AOD retrieval and to calibrate the CMAQ $PM_{2.5}$ concentration field using observations as the reference. Then, a maximum likelihood estimator integrated the AOD-derived and the calibrated-CMAQ fields to derive a combined concentration field. The bias of the combined concentration field was additionally reduced by adding a Kriging-interpolated residual term. Senthilkumar et al. (2019) calculated a normalized dimensionless ratio between observations and their corresponding CMAQ simulations at monitoring sites. They used the Inverse Distance Weighting method to spatially interpolate the ratio, which they then multiplied with the intensity-adjusted CMAQ simulations to generate a fused concentration field.

Recently, machine learning and deep learning have been increasingly applied to integrate observational data with chemical transport models and other data sources. These approaches are often regarded as "black boxes" due to their complex nature. After selecting the model architecture, such as the random forest or convolutional neural network, the model is trained by providing input features, which typically include simulation outputs from CTMs, meteorological indicators, and geographical static data, along with the corresponding target values, generally observations. For instance, Lyu et al. (2019) conducted a stacking ensemble learning framework that included random forest, neural network, gradient boosting machine, and general linear models to fuse CTM and observation data. Simulated data from CMAQ and other supporting variables derived from meteorological fields and land-use data were used as input features, while observations were used as response variables. Input features and observations were collected to train these machine learning models, and the models' predictions were ensembled to predict the fused concentration field. The field's bias was additionally reduced by adding a Kriging interpolated residual term as in Xue et al. (2017). Tang et al. (2024) applied a random forest model to fuse satellite AOD retrievals and WRF-Chem outputs with observation data. The satellite retrievals, meteorological data, elevation data, emission data, and concentration outputs are re-gridded at the same resolution and utilized as random forest input features. The observational data were used as target values for the random forest model to reduce the bias between model predictions and observations. Li et al. (2023b) trained a recurrent spatiotemporal deep-learning model using WRF and CMAQ outputs as input features and monitored ozone ($O_3$) as target values. The deep learning model was trained by reducing the bias between observations and corresponding concentration predictions.

Despite the growing and important uses of data fusion-generated concentration fields for health and public policy research, few offer open-source, well-documented code. Additionally, there is a lack of shared model architectures and trained weights for machine learning and deep learning approaches. Since the software for data fusion methods is generally not publicly available, researchers can only access the provided fused data, making it difficult to apply the data fusion methods to their own simulations. A good example of an open-sourced data fusion software is the tool developed by Li et al. (2019), which is available in the Air Benefit and Cost and Attainment Assessment System. The software is wrapped with a user-friendly graphical user interface (GUI) and provides different data fusion methods, including enhanced Voronoi Neighbor Averaging (eVNA) (Ding et al., 2016) and Downscaler (DS) (Berrocal et al., 2010). Its effectiveness in reducing the CMAQ model's bias has been demonstrated in previous studies (Yang et al., 2020; Yuan et al., 2023). The eVNA approach uses Voronoi Neighbor Averaging (VNA) (Kreveld et al., 1997) interpolation method to produce concentration fields from observations. These interpolated concentrations are then scaled by the ratio of the modeled concentration in the CMAQ grid cell to that in the CMAQ grid cell containing the monitor. The DS method employs a Bayesian hierarchical model to correct biases in CMAQ outputs using observations. It assumes prior distributions for time-varying additive and multiplicative biases. It also models spatially and temporally varying additive and multiplicative adjustment terms using Gaussian processes. Markov chain Monte Carlo, including forward-filtering and backward-sampling, estimates model parameters by updating priors with observations. The biases estimated by the model are then used to correct CMAQ simulations both spatially and temporally. However, this software, like most other examples, has limited generalizability because it focuses on integrating only a single CTM with observations rather than being adaptable to other CTMs. CTMs generally share a similar output data structure, often in Network Common Data Form (NetCDF) format, and typically provide hourly grid-based surface concentration fields. Standard daily pollutant metrics can then be derived from these hourly outputs and saved in NetCDF format through post-processing. Thus, migrating a data fusion method from one model to another can be relatively straightforward if the input data format is well-defined, with each variable stored as a tensor (high-dimensional matrix) along the spatial and temporal dimensions and accompanied by metadata specifying the map projection used and time information such as beginning date and time, frequency and ending date and time.

To improve the transparency and efficiency for researchers who are interested in implementing data fusion methods or using the data fusion outputs from various CTMs for their research, we develop and implement a general data fusion framework based on the Friberg et al. (2016) data fusion method (Gen-Friberg version 1.0). Friberg et al.'s data fusion method has shown better performance than different monitor-interpolation methods, air quality models, and hybrid modeling methods that reduced CTM bias using receptor model outputs (Yu et al., 2018). The data fusion method by Frieberg et al. has been widely used as part of several ensemble data fusion approaches (Bates et al., 2018; Huang et al., 2018) and health studies (Maji et al., 2024a, 2024b; Picciotto et al., 2024). In this study, we developed a framework that unifies the input data format, offering a user-friendly solution for data fusion. To ensure most users can easily use the data fusion method, we encapsulated the algorithm into one function that supports multiple CTMs, including CMAQ, WRF-Chem, and GEOS-Chem. For computational

efficiency, the data fusion model can be executed in parallel to speed up the process for long simulations. Relevant post-analysis tools were developed for bare CTM and post data fusion performance evaluation. This study demonstrated the ease and broad applicability of this proposed framework and software tool for data fusion using several CTMs and monitoring networks. The model can be downloaded from the GitHub repository along with its auxiliary tools.

## 2. Material and methods

The new data fusion tool, Gen-Friberg version 1.0 (GF-1), provides an effective and efficient method to fuse daily observational and CTM-simulated data. Given that the data may come from various sources or undergo different processing methods, we standardized the input data formats. Users are responsible for preparing their data in the defined format to ensure the model can be successfully executed.

### 2.1. Input data (observation and CTM) and format requirements

For observational data, users should calculate the daily observational concentrations based on aggregating methods such as mean, maximum, or daily maximum 8-h. Then, these data to be utilized in data fusion should be combined as the observational inputs in one comma-separated values (CSV) format file. For each observational record, the unique ID for the monitoring site and its coordinates (latitude and longitude) need to be included to provide spatial information. The observation time should be recorded in YYYY-MM-DD format (e.g., 2021-01-01 for January 1st, 2021). The time zone in the observational data must be the same as the one in the CTM or post-processed CTM outputs. For example, both datasets can use the local standard time. The data fusion outputs will inherit the time zone of the input data.

The CTM data must be in NetCDF format. Hourly (or sub-hourly) CTM data must be converted to daily data using the same aggregation method as the observational data. For CMAQ, the (hr2day) post-processing program provided by the CMAQ code, which generates gridded daily concentrations from hourly data, is an effective tool for this process. We provided similar utilities for processing GEOS-Chem and WRF-Chem hourly data. The daily CTM data covering the study period may exist in multiple files due to discontinuous CTM simulation or post-processing. We required combining all the CTM data in one NetCDF file to standardize CTM data input. A utility for combining multiple CTM files into a single NetCDF file is provided (data format details are on the GitHub page).

In this study, we applied GF-1 to fuse CMAQ simulation results and observations for daily average $PM_{2.5}$, daily maximum 8-h average (MDA8) $O_3$, and daily average $NO_2$ in the contiguous United States (CONUS) from 2010 to 2019 and evaluated its performance. We collected the observational data for $PM_{2.5}$, $O_3$, and $NO_2$ from EPA's Air Quality System (AQS) monitoring sites (Fig. 1) (AQS). CMAQ data was obtained from the Air QUAlity TimE Series Project (EQUATES) (U.S. EPA, 2021) data repository over the same time frame and spatial domain. The EQUATES data is constructed using WRF v4.1.1 and CMAQ v5.3.2 and provides 12-km resolution gridded fields of pollutant concentrations over CONUS. Also, we included a one-year (2017) CONUS GEOS-Chem simulation with 0.5° (latitude) × 0.625° (longitude) resolution and a one-year (2017) WRF-Chem simulation with a 36-km resolution over CONUS. The use of GEOS-Chem and WRF-Chem demonstrates the ease of applicability of GF-1 on other CTMs with a simple function overwrite, details of which are provided in Supplementary Information of Text S2, Text S3, and the GitHub page.

### 2.2. Data fusion method and implementation optimization

GF-1 is an advancement of the Friberg et al. (2016) approach, which includes three general steps to fuse observational and simulation data to reduce spatiotemporal biases of CTM outputs. In summary, the first two steps focus on minimizing spatial and temporal biases respectively, while the third step integrates their outputs to generate the final fused fields.

The fused concentration field in the first step is obtained using either a zero-intercept linear or an exponential equation, derived by comparing annual mean observational data with CTM data. The selected equation aims to minimize yearly bias. Then, Kriging is conducted to reduce the spatial bias in the simulation. Since the zero-intercept linear regression is a special case of exponential regression (the exponent is 1), we combine these two cases and mathematically express the first step as follows.

$$\overline{OBS_m} = \alpha \times \overline{CTM(s)_m}^{\beta} + \varepsilon \tag{1}$$

$$\overline{FC(s)} = \alpha_{year} \times \overline{CTM(s)}^{\beta} + \varepsilon \tag{2}$$

$$FC_1(s,t) = \left(\frac{OBS_m(t)}{\overline{OBS_m}}\right)_{krig} \times \overline{FC(s)} \tag{3}$$

Equation (1) is a regression equation for adjusting the annual mean of CTM predictions, where $\overline{OBS_m}$ is the monitor's annual mean value, $\overline{CTM(s)_m}$ is the annual mean CTM concentrations at corresponding monitor locations. $\varepsilon$ denotes the regression residual, minimized through the least squares method during the estimation of regression parameters $\alpha$ and $\beta$. This regression method is expected to capture the linear (when $\beta = 1$) or non-linear relations between observations and corresponding co-located simulations. Friberg et al.'s study did not propose a selection algorithm for linear or exponential regression. For GF-1, we implement both modes of regression (i.e., linear and exponential). An algorithm that uses 10-fold cross-validation for selecting the optimal regression is provided (default mode). The regression with the lowest root mean square error (RMSE) is automatically selected for CTM adjustment. In Equation (2), $\overline{FC(s)}$ is the adjusted annual mean gridded CTM concentration ($\overline{CTM(s)}$) over the study domain for each grid cell at location $s$, preserving the annual spatial distributions from CTM with adjusted intensity. The regression parameter $\beta$ (note for linear regression $\beta = 1$) is derived based on all averaged observations at different monitors and

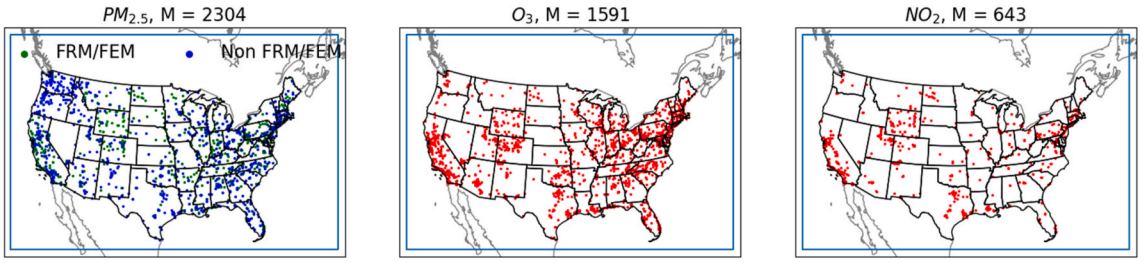



**Fig. 1.** EPA AQS $PM_{2.5}$, $O_3$, and $NO_2$ monitoring sites number (M) and locations through the study period (2010–2019). The blue rectangle indicates the EQUATES domain. For $PM_{2.5}$, the green and blue dots show the locations of Federal Reference Method (FRM)/Federal Equivalent Method (FEM) monitors and non-FRM/FEM monitors, respectively.

their co-located annual average CTM simulations (Equation (1)) over all studied years. Then, the regression parameters $\alpha_{year}$ are estimated based on annual averaged observations and their co-located annual average CTM simulations for each year (Equation (2)). In Equation (3), $OBS_m(t)$ are the daily observations at monitor $m$ on day $t$ and the normalized observations are interpolated by the Kriging interpolation to derive the $FC_1$. We implemented the Kriging interpolation using the 2D ordinary Kriging function (OrdinaryKriging) in the PyKrige package (Murphy et al., 2024) with the exponential variogram model. For the grid cells where monitors are located, this function ensures that interpolated concentrations remain highly correlated with the observations. For locations without monitors, it estimates concentrations based on the spatial correlation with nearby observations, producing a smoothed surface that reflects the spatial patterns of the observations. Then, the interpolated normalized observations are multiplied by the annual intensity-adjusted CTM concentration field ($\overline{FC(s)}$) to reduce the spatial distribution difference between simulations and observations.

The second step of the data fusion method mainly focuses on reducing the seasonal bias in CTM simulation results. First, the daily CTM results are adjusted using Equation (4), where $CTM(s,t)$ is the daily CTM data. $\overline{FC(s)}$ and $\overline{CTM(s)}$ are the annual spatial averages as before.

$$CTM(s,t)_{adj} = CTM(s,t) \times \frac{\overline{FC(s)}}{\overline{CTM(s)}} \tag{4}$$

$$\overline{CTM_{adj}(t)} = \frac{\sum_{m=1}^{M} CTM_{adj}(m,t)}{M} \tag{5}$$

$$\overline{OBS(t)} = \frac{\sum_{m=1}^{M} OBS_m(t)}{M} \tag{6}$$

A seasonal ratio $\beta_{season}$, which is the ratio of $\overline{CTM_{adj}(t)}$ to $\overline{OBS(t)}$ (calculated by Equations (5) and (6), respectively) for each Julian day $jt$, is used to train the following trigonometric sinusoidal function (Equation (7)):

$$\frac{\overline{CTM_{adj}(t)}}{\overline{OBS(t)}} = \beta_{season}(jt) = e^{A \times cos\left[\frac{2\pi}{365.25}(jt - jt_{max})\right]} + \varepsilon \tag{7}$$

The ratio captures the seasonal variation of both modeled and measured data sets. The period of the trigonometric function is 365.25 days, which is derived from averaging the total number of days over a four-year cycle. $A$ and $jt_{max}$ are parameters derived from the regression by minimizing the regression residual $\varepsilon$. $jt$ is the Julian date of day $t$.

After adjusting the daily CTM results and developing the temporal regression function, the fused-concentration field in the second step ($FC_2$) is calculated as follows (Equation (8)):

$$FC_2(s,t) = CTM(s,t)_{adj} \times \beta_{season}(jt) \tag{8}$$

In the third step, the method optimally integrates the results from the first and second steps to produce the final data fusion output. For grid cell location $s$, the method estimates the observation's correlations to $FC_1$ and $FC_2$ separately. Then, the method uses these correlations to derive weights for calculating the weighted average of $FC_1$ and $FC_2$ as the final combined data fusion result. To calculate the weighting factor for $FC_1$, which has spatial information provided by interpolated observations, the method uses an exponential correlogram equation to estimate spatial correlations for each CTM grid cell. First, the method calculates the distance $d$ and Pearson correlation value, $R_{obs}$, between each monitor pairing. Parameters $R_{coll}$ and $r$ are then determined based on these values using the exponential correlogram equation as follows (Equation (9)):

$$R_{obs}(d) = R_{coll}e^{-\frac{d}{r}} + \varepsilon \tag{9}$$

$R_{coll}$ is the intercept, which represents instrumentation error, while $r$ is the distance at which the $R_{obs}$ has an e-fold decrease. $\varepsilon$ is the regression residual, which can be minimized by fitting given values of $R_{obs}(d)$ and their corresponding $d$. Then, for any CTM grid location denoted by $s$ on day $t$, the correlation weighting factor $R_1(s,t)$ for $FC_1$ is estimated by Equation (10):

$$R_1(s,t) = R_{coll}e^{-\frac{x(s,t)}{r}} \tag{10}$$

where $x(s,t)$ is the distance between grid cell location $s$ and its closest monitor with data collected on day $t$. For $FC_2$, the spatial information is provided by the CTM. The method evaluates the CTM's performance in predicting spatial concentration distribution by calculating the mean Pearson correlation between the observation and CTM prediction at included monitors (Equation (11)).

$$R_2 = \frac{1}{M}\sum_{m=1}^{M} corr(OBS_m, CTM_m) \tag{11}$$

where $OBS_m$ is the time series of observational data at monitor $m$ over the entire study period, $CTM_m$ is the time series of associated CTM predictions, and $corr$ is the Pearson correlation coefficient defined in Text S1. Then, a combined weight factor is calculated as follows (Equation (12)).

$$W(s,t) = \frac{R_1(s,t) \times (1 - R_2)}{R_1(s,t) \times (1 - R_2) + R_2 \times (1 - R_1(s,t))} \tag{12}$$

The weight factor evaluates the confidence in using the $FC_1$ results at grid cell location $s$ on day $t$. When the area is near the monitor, $R_1$ is dominant, leading to a higher weight $W$. It suggests placing more trust in $FC_1$, which includes the interpolated observations for spatial concentration distribution estimation. For locations far from the monitors, $FC_2$, which represents the CTM with seasonal correction, plays a dominant role in the data fusion. With the weight factor $W$, the method uses the weighted average on $FC_1$ and $FC_2$ to calculate the ultimate data fusion results $FC_{opt}$ (Equation (13)).

$$FC_{opt} = W(s,t) \times FC_1(s,t) + (1 - W(s,t)) \times FC_2(s,t) \tag{13}$$

Global regression parameters are calculated first (i.e., β, $R_{coll}$, r, A, and $jt_{max}$), before proceeding with the steps of fusing the CTM simulation data with the corresponding yearly observation data set. After determining the parameters, the yearly data fusion processes for different years were independent from each other, so it was implemented in parallel to speed up long-period data fusion.

### 2.3. Performance evaluation

We first evaluated the CMAQ and data fused-CMAQ performances by comparing results to observational data using the metrics suggested by Emery et al. (2017) (the details of the criteria are in Text S1) to show the effectiveness of the data fusion algorithm. Improved performance is expected after data fusion. To estimate the uncertainty of the data fusion method, we conducted a 5-fold cross-validation. For each fold, we withheld 20 % of the observational data and then fused the remaining 80 % of observations with CMAQ data. The performance was evaluated based on the 20 % of observations not used in data fusion. Two withholding strategies were employed: random withholding and site-wise withholding. The random withholding withheld a portion of observations from each monitor, assessing how well the resulting fields capture temporal variations. The site-wise method withheld observations from 20 % of the monitoring sites and fused observations from the remaining monitors with CMAQ data. The performance of the method was evaluated on the observations from 20 % withheld monitors to assess how well the method captures spatial patterns. Cross-validation and data withholding only require pre-processing of observational data and post-analysis of data fusion outputs, without any modifications to the

data fusion code itself. Therefore, users can easily implement scripts for uncertainty evaluation by calling the data fusion functions with different observational data inputs in each fold, following their own withholding strategies. Example scripts for these two withholding strategies used in the study are provided to assist with these uncertainty evaluations.

### 2.4. Application of Gen-Friberg to other CTMs

We applied GF-1 to 2017 simulations with GEOS-Chem and WRF-Chem over CONUS using 0.5° (latitude) × 0.625° (longitude) and 36-km grid resolutions, respectively, to demonstrate its generalizability (the domain boundaries are shown in Fig. S1). We conducted a 5-fold cross-validation using both random and site-wise withholding to assess differences in performance when using GF-1 with different CTMs. Additionally, we compared the GEOS-Chem and WRF-Chem results, both before and after data fusion, to 2017 CMAQ simulations matching the nearest grid cells. For each GEOS-Chem or WRF-Chem grid cell, we identified the closest CMAQ grid cell and compared their simulated pollutant concentrations. To understand air pollutant exposure discrepancies among different models, we focused our comparisons on grid cells within the CONUS region. This comparison allowed us to assess the differences in pollutant simulations among different CTMs or data fused-CTMs results, which are essential for better understanding the confidence and uncertainty in concentration fields used for health impact assessments.

## 3. Results

Here, we analyzed and discussed the global parameters derived by analyzing CMAQ and observational data and showed how they affect the data fusion process. Then, we evaluated the GF-1's effectiveness in reducing concentration biases and the model uncertainty when applied to CMAQ data. We also discussed GF-1's performance when applied to GEOS-Chem and WRF-Chem and compared the results with those from its application to CMAQ.

### 3.1. Global parameters used in data fusion

We used the "default mode" and found that the linear regression had a lower RMSE than the exponential regression in step 1. Notice that the implementation of exponential regression involved numerically determining the optimal regression parameters, while the linear regression model relied on an analytical solution. Therefore, in our implementation, linear regression was treated separately from exponential regression, despite mathematical derivations suggesting a subset relationship. The regression parameters developed from the combined 10-year data set of CMAQ and observations showed that the CMAQ simulated the magnitude of annual averaged $O_3$ and $PM_{2.5}$ with slopes ($\alpha$) of 1 and 0.98, respectively (Table 1). However, the slope of linear regression for $NO_2$ was 1.12, which indicated that CMAQ underestimated $NO_2$ by 12 %. For each yearly slope derived by linear regressions of annual CMAQ and observations (Table S1), we also found that $O_3$ and $PM_{2.5}$ simulation from CMAQ showed a good performance on capturing the intensity with slopes close to 1 in all selected years, while CMAQ tended to underestimate $NO_2$ from 2014 to 2019.

For seasonal ratio correction, we derived the following function (Equation (14)) by taking the logarithm of Equation (7):

$$\log \frac{\overline{CMAQ_{adj}(jt)}}{\overline{OBS(jt)}} = A \times cos\left[\frac{2\pi}{365.25}(jt - jt_{max})\right] \quad (14)$$

$A$, the amplitude indicates the relative ratios between CMAQ and observation in the seasonal adjustment (Table 1). $jt_{max}$, is the Julian day when the ratio difference between observations and CMAQ reaches its maximum. All three pollutants showed large differences in the summer or earlier fall season ($jt_{max}$ of $PM_{2.5}$, $O_3$, and $NO_2$ are in July, September, and August separately) by analyzing CMAQ simulations from 2010 to 2019. Around $jt_{max}$, positive $A$ for $PM_{2.5}$ indicated that the daily spatial-averaged CMAQ had overestimations, while negative $A$ for $O_3$ and $NO_2$ showed that CMAQ predictions were slightly lower than the observations.

$R_{coll}$ and r represent the correlation between observations from different monitors. $PM_{2.5}$ had the highest $R_{coll}$, as its monitoring network is denser than the networks for $O_3$ and $NO_2$. With more monitors close to each other, providing additional training data points in the exponential correlogram Equation (9), $R_{coll}$ was more accurately estimated. $NO_2$ had the lowest $R_{coll}$, indicating that $NO_2$ correlations, even between closely located monitors, could be low. The spatial variation of $NO_2$ was high, especially for the near-road monitors. For example, near-road monitors upwind or downwind from the road can have significant differences in the $NO_2$ concentration even if they are close to each other. Parameter r indicates the e-folding distance of Pearson correlation compared to $R_{coll}$. The e-folding distance of $O_3$ was nearly double that of $PM_{2.5}$ despite similar $R_{coll}$ values, indicating that $PM_{2.5}$ exhibits greater spatial heterogeneity than $O_3$. $NO_2$ has the largest e-folding distance, which results from its low initial correlation ($R_{coll}$). Because the correlation starts at a small value, the decline over distance occurs more gradually.

The data fusion method utilizes the weights $W$ to fuse $FC_1$ and $FC_2$. Higher values of $W$ indicate that the data fusion results are driven by $FC_1$, which is based on the spatial Kriging interpolation of observational data. On the other hand, lower values of $W$ indicate that the data fusion method weights $FC_2$, the seasonally (i.e., temporally) corrected CMAQ outputs, more. We visualized the averaged weights, $\overline{W}$, during the study period (Fig. 2). As expected, $\overline{W}$ was higher at locations close to monitors. The $PM_{2.5}$ weights near the monitors were higher than the $O_3$ and $NO_2$ weights. This is due to the relatively inferior performance of CMAQ $PM_{2.5}$ simulations, which had an R value of 0.41, compared to 0.73 for $NO_2$ and 0.76 for $O_3$. With the mathematical definition of $W$ in Equation (12), the weights tended to be higher when the correlation between CMAQ and observations was lower.

**Table 1**
Data fusion parameters for $PM_{2.5}$, $O_3$, and $NO_2$.

| | A | $jt_{max}$ (days) | α | β | $R_{coll}$ | r (km) |
|---|---|---|---|---|---|---|
| $PM_{2.5}$ | 0.07 | 200.70 | 0.98 | 1.0 | 0.80 | 686.49 |
| $O_3$ | −0.06 | 260.22 | 1.00 | 1.0 | 0.71 | 1298.9 |
| $NO_2$ | −0.16 | 213.73 | 1.12 | 1.0 | 0.49 | 2337.27 |

### 3.2. Comparisons between CMAQ and Fused-CMAQ

CMAQ and fused-CMAQ had similar spatial distributions for the selected species (Fig. 3). $NO_2$ was concentrated in large cities due to heavy transportation. The western U.S. cities, which have high anthropogenic emissions and favorable meteorological conditions for $O_3$ formation, such as Los Angeles, experience higher $O_3$ levels compared to the eastern U.S. Complex lake and land breezes lead to higher $O_3$ levels (Cleary et al., 2022) in the Great Lakes Region than surrounding regions. $PM_{2.5}$ is higher in the southeastern U.S. and California due to prescribed fire and wildfire smoke, respectively (Jaffe et al., 2020). Depending on the species, concentration differences between averaged fused-CMAQ and CMAQ showed either a positive or negative bias. MDA8-$O_3$ and $PM_{2.5}$ in fused-CMAQ were higher than CMAQ, while $NO_2$ in fused-CMAQ was lower than CMAQ. These results can be explained by the value of the regression slope $\alpha_{year}$ (Table S1) used for CMAQ intensity adjustment (Equation (2)). The intensity-adjusted CMAQ was incorporated in both $FC_1$ (Equation (3)) and $FC_2$ (Equations (4) and (8)), driving the intensity of data fusion results. CMAQ had lower concentrations than fused-CMAQ for MDA8-$O_3$ and $PM_{2.5}$ since a slope less than 1 was employed to adjust fused-CMAQ concentration intensity in most years from 2010 to 2019. A slope greater than 1 was applied for $NO_2$ in most

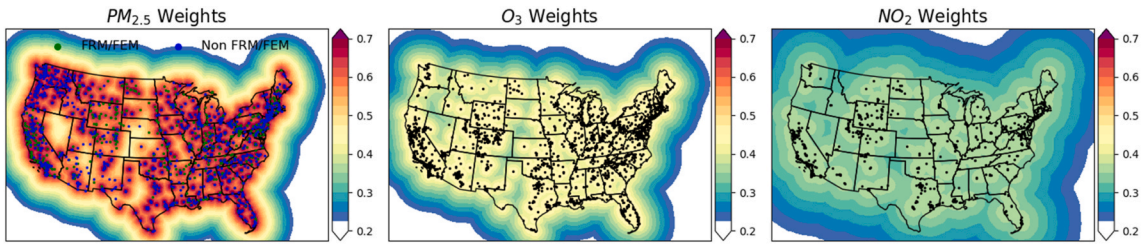



**Fig. 2.** The averaged weights ($\overline{W}$) of $PM_{2.5}$, $O_3$, and $NO_2$ that were used in combining $FC_1$ and $FC_2$ from 2010 to 2019. The dots indicate the locations of monitors utilized in the data fusion.

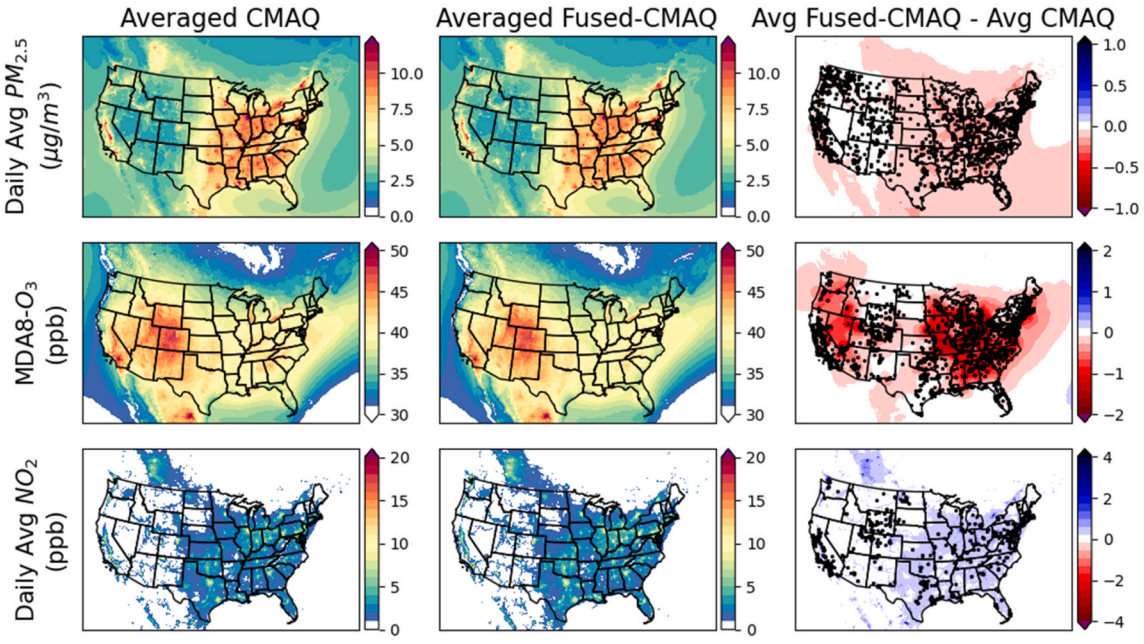



**Fig. 3.** Spatial distribution of daily average $PM_{2.5}$, MDA8-$O_3$, and daily average $NO_2$ from CMAQ (first column), fused-CMAQ (second column). The CMAQ and fused-CMAQ are averaged from 2010 to 2019. The differences between the averaged fused-CMAQ and CMAQ are shown in the third column.

years, causing the fused-CMAQ concentrations to be larger than the CMAQ concentrations. The absolute values of differences between fused-CMAQ and CMAQ were higher in the eastern U.S., where the observational networks were dense, leading to observations dominating the concentrations rather than CMAQ fields.

To assess the performance of the implemented data fusion method, we compared the results of CMAQ and fused-CMAQ with observations. For each monitor, we calculated the Pearson R (correlation coefficient) between CMAQ (or fused-CMAQ) results and observations (Fig. 4). CMAQ displayed higher correlations in the northeastern U.S. for all selected species. However, outside the Northeast, $NO_2$ and $O_3$ simulations had low correlations with observations in states like Wyoming and Colorado. $PM_{2.5}$ had a lower correlation than the other species, especially in the southwestern U.S. The correlations between fused-CMAQ and observations were higher compared to CMAQ. Fused-CMAQ and most observations showed strong correlations ($> 0.9$) for all the selected pollutants. However, $NO_2$ had subpar performance compared to other pollutants. This can be explained by the lower number of monitors for $NO_2$ measurements and the relatively lower $W$ values (i.e., weight of observations) (Fig. 2). The accuracy of Kriging interpolation decreases when the number of observations is small or when the data has limited spatial coverage (CUMSPH). Given the sparse distribution of $NO_2$ monitoring sites, the reduced Kriging interpolation performance could negatively impact the data fusion performance, particularly in areas surrounding the monitors, where the observation interpolation dominated the concentration field. Additionally, the low weights $W$ used for $NO_2$ (Fig. 2) indicated that the data fusion relied more on CMAQ simulations than observations, compared to other pollutants.

We evaluated the performances of CMAQ and fused-CMAQ by applying various statistical metrics suggested by Emery et al. (2017) using observations and corresponding predictions at all monitors (Fig. S2, Table 2). The fused-CMAQ performed better on all the statistical metrics and reached the "goal" performance, one-third of the top performances in past applications of CTMs, as suggested by the study (Emery et al., 2017). The MDA8-$O_3$ performance was the best, followed by the daily average $NO_2$, and then the daily average $PM_{2.5}$. To understand the main factors that impeded the fused $PM_{2.5}$ performance, we visualized the observations and their corresponding simulations in the scatter plots (Fig. 5). Although fused-CMAQ effectively predicted $PM_{2.5}$ lower than 30 $\mu g/m^3$, it did not perform well when observed $PM_{2.5}$ levels were higher. Sometimes, fused-CMAQ highly overestimated $PM_{2.5}$ even though the observed $PM_{2.5}$ was low. This can be partly attributed to biased wind inputs, which lead to uncertainties in concentration simulations, particularly for wildland fire smoke. Biased wind directions can easily make the smoke hit or miss a monitor, leading to high biases.

### 3.3. Data Fused-CMAQ uncertainty

5-fold cross-validation was conducted to understand the fused-CMAQ uncertainty. For each fold, we evaluated the fused-CMAQ uncertainty by comparing the fused predictions with the 20 % observations that were not used in the data fusion process. The overall Pearson R and

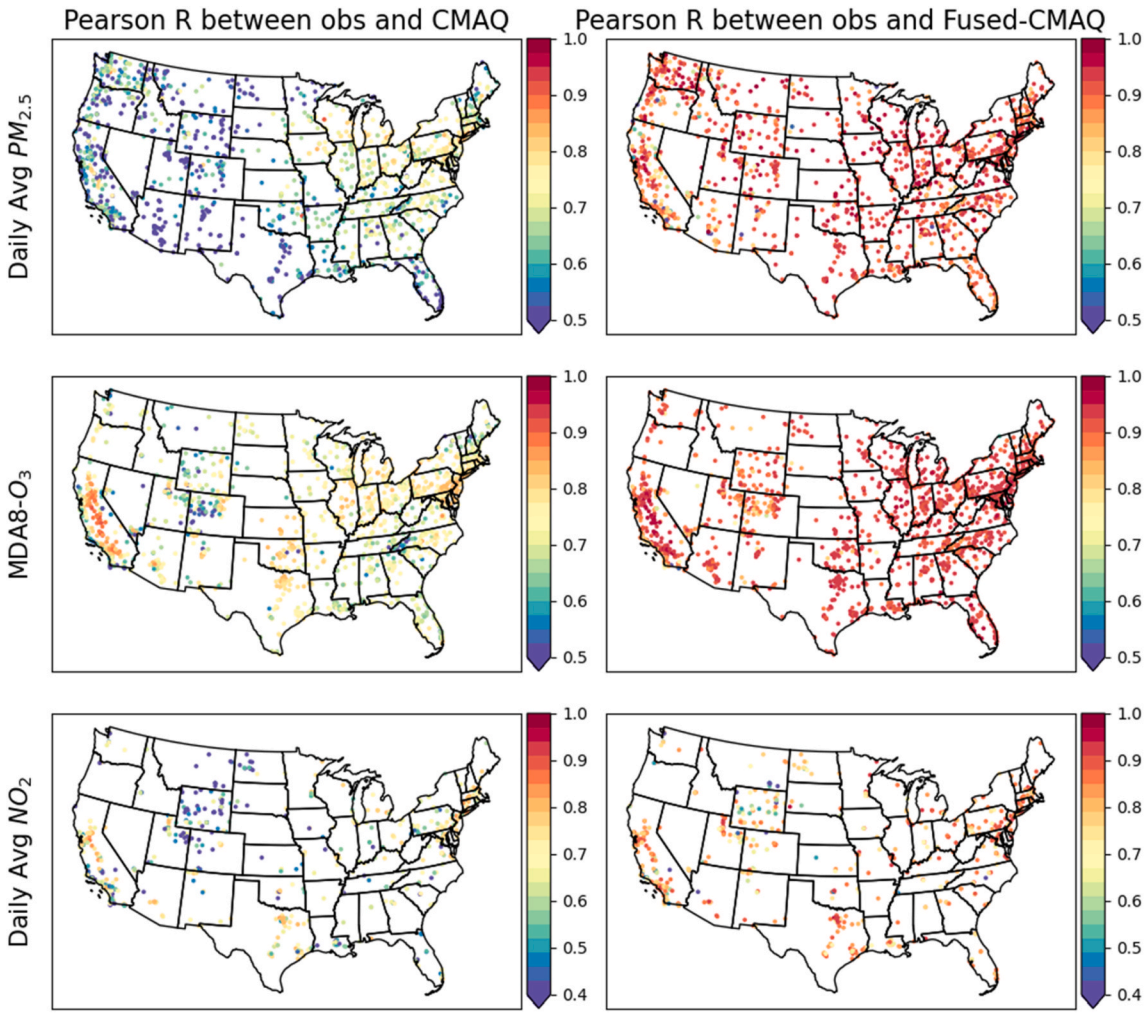



**Fig. 4.** Pearson correlation coefficient (R) values between simulated and observed concentrations at each monitor from 2010 to 2019 for daily average $PM_{2.5}$, MDA8-$O_3$, and daily average $NO_2$.

**Table 2**
Daily averaged $PM_{2.5}$, MDA8-$O_3$, and daily average $NO_2$ performance for CMAQ, fused-CMAQ, and fused-CMAQ under random and site-wise data withholding. The units of MB, ME, RMSE, and CRMSE are μg/$m^3$ for daily averaged $PM_{2.5}$ and ppb for both MDA8-$O_3$ and daily average $NO_2$.

| Species | daily averaged $PM_{2.5}$ | | | | MDA8-$O_3$ | | | | daily average $NO_2$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CMAQ | Fused-CMAQ | Random W/h | Site W/h | CMAQ | Fused-CMAQ | Random W/h | Site W/h | CMAQ | Fused-CMAQ | Random W/h | Site W/h |
| MB | −0.51 | −0.66 | −0.72 | −0.65 | −0.21 | −0.97 | −0.94 | −0.94 | −1.48 | −0.42 | −0.42 | −0.42 |
| ME | 3.27 | 2.23 | 2.41 | 2.57 | 6.53 | 4.38 | 4.73 | 4.75 | 3.58 | 3.06 | 3.22 | 3.22 |
| RMSE | 7.82 | 4.56 | 4.87 | 5.38 | 8.45 | 5.82 | 6.32 | 6.33 | 5.40 | 4.59 | 4.78 | 4.78 |
| CRMSE | 7.80 | 4.51 | 4.82 | 5.34 | 8.45 | 5.73 | 6.25 | 6.26 | 5.19 | 4.57 | 4.76 | 4.76 |
| NMB | −6.16 | −7.90 | −8.56 | −7.86 | −0.52 | −2.35 | −2.27 | −2.26 | −17.38 | −4.97 | −4.96 | −4.96 |
| NME | 39.32 | 26.80 | 28.56 | 30.94 | 15.77 | 10.57 | 11.44 | 11.48 | 41.94 | 35.84 | 37.66 | 37.72 |
| MNB | 15.82 | 9.84 | 7.97 | 13.75 | 5.06 | 1.52 | 2.11 | 2.12 | 34.07 | 42.26 | 52.10 | 52.16 |
| MNE | 51.13 | 36.98 | 37.02 | 43.34 | 18.97 | 12.41 | 13.67 | 13.72 | 80.07 | 73.53 | 84.55 | 84.65 |
| FB | −5.59 | −4.77 | −5.40 | −3.92 | 1.37 | −0.32 | −0.18 | −0.18 | −11.13 | 2.97 | 3.42 | 3.39 |
| FE | 41.88 | 29.21 | 31.29 | 34.08 | 16.82 | 11.42 | 12.36 | 12.40 | 51.46 | 43.48 | 45.94 | 45.96 |
| IOA | 0.60 | 0.83 | 0.81 | 0.77 | 0.86 | 0.93 | 0.92 | 0.92 | 0.83 | 0.89 | 0.88 | 0.88 |
| Pearson R | 0.41 | 0.73 | 0.69 | 0.62 | 0.76 | 0.90 | 0.88 | 0.88 | 0.73 | 0.80 | 0.78 | 0.78 |

RMSE of fused-CMAQ, as calculated with random and site-wise withholding, were greatly improved compared to the original CMAQ performance, although the performance declined compared to the fused-CMAQ with all observation data (Fig. S2, Table 2). For most other statistical metrics, the fused-CMAQ under different data withholding strategies also performed better than CMAQ (Table 2). Additionally, the linear regression relationships of fused data generally had smaller absolute intercepts and slopes closer to 1 compared to CMAQ simulations, revealing that background biases and intensity differences were reduced with the data fusion model (Fig. 5). However, performance declined under site-wise data withholding when compared to random data withholding, particularly for $PM_{2.5}$. This indicates that data fusion is more sensitive to gaps in spatial coverage than to gaps in temporal coverage of the observation data. For instance, $PM_{2.5}$, which had the

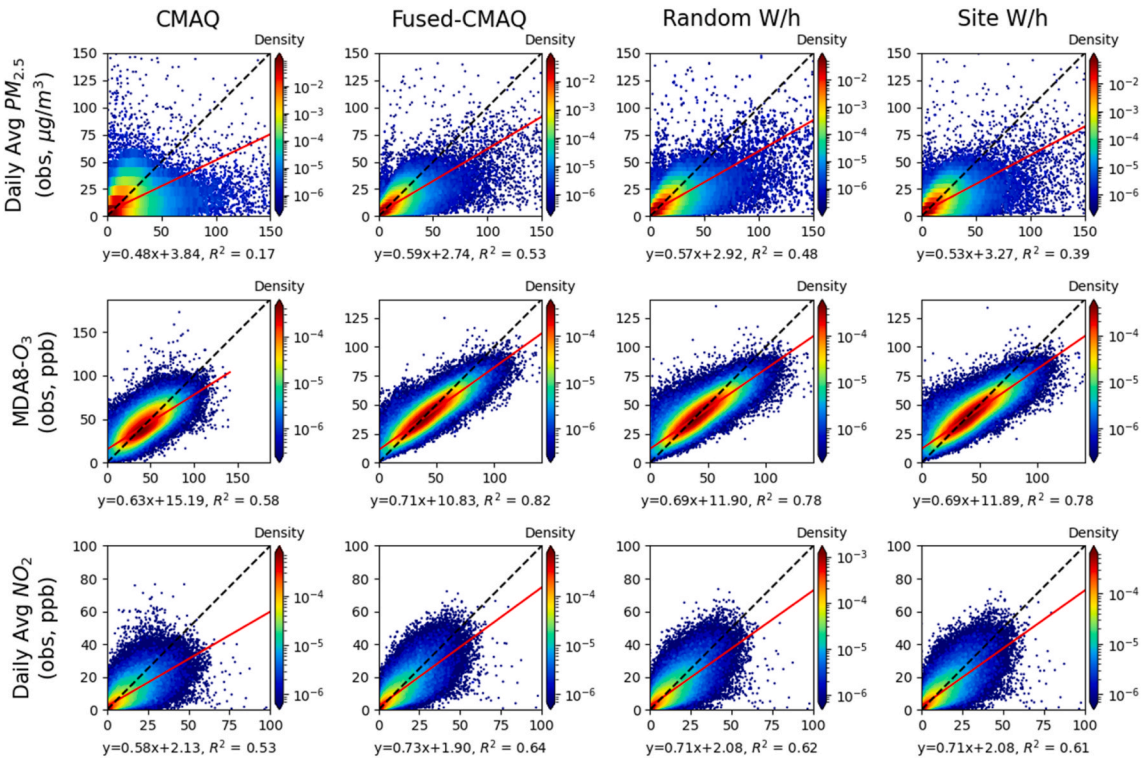



**Fig. 5.** Comparisons between observations and simulations of daily averaged $PM_{2.5}$, MDA8-$O_3$, and daily averaged $NO_2$. The dots show the simulation and the corresponding observations. Each fold's comparisons in the testing dataset (observations not used in data fusion) are indicated for random and site-wise data withholding. The black dashed line is the unity (1:1) slope line. The red line shows the linear regression between simulation and observations. The $R^2$ value and the formula of linear regressions are indicated below the figures.

highest weights (*W*) among all selected species (Fig. 2), relied more on Kriging interpolation of observational data. Since withholding part of the monitors can highly impact spatial Kriging interpolation, the $PM_{2.5}$ data fusion results were less accurate under site-wise data withholding.

Fig. 6 shows the spatial distribution of fused-CMAQ's uncertainty. The uncertainty was estimated by first averaging the data fusion outputs from each cross-validation fold over the study time period. Then, the standard deviation among these averaged 5-fold outputs was calculated for each CMAQ grid cell. Both withholding strategies showed higher uncertainty near the monitors, especially in urban areas where the monitors were densely distributed. However, the relative magnitude of the standard deviation in each grid cell was much lower than the averaged concentration (Fig. 3), indicating the data fusion model was robust. The site-wise data withholding still showed a higher standard deviation than random data withholding. The eastern and western coasts of the U.S. showed higher standard deviations as monitors were more concentrated in these locations. Withholding the observations from these monitors led to high uncertainties in the data fusion outputs near these locations.

### 3.4. Data fusion for other chemical transport models

We applied GF-1 to 2017 GEOS-Chem and WRF-Chem simulations to evaluate the generalizability of the method and our implementation, and whether it is leading to similar performance trends as it did for CMAQ (Tables S2–S5, Figs. S3–S14).

The GEOS-Chem simulated daily average $PM_{2.5}$, MDA8-$O_3$, and daily average $NO_2$ had a Pearson R of 0.37, 0.64, and 0.58, respectively. The fused GEOS-Chem results improved with R values of 0.63, 0.89, and 0.68, respectively. The fused GEOS-Chem results also had a lower RMSE than the GEOS-Chem results. The RMSE for fused GEOS-Chem $PM_{2.5}$, MDA8-$O_3$, and $NO_2$ were 7.07 μg/m$^3$, 6.05 ppb, and 5.77 ppb, which were lower than the GEOS-Chem RMSE of 10.63 μg/m$^3$, 10.75 ppb, and 7.81 ppb. The fused GEOS-Chem under either withholding strategy also showed better performance than the GEOS-Chem simulated fields compared with observations (Table S3). Results indicated that the data fusion method effectively improved the GEOS-Chem model's performance for both methods. Analyses of fused GEOS-Chem performance and uncertainties, similar to those conducted for CMAQ, are available in the Supplementary Information (Tables S2–S3, Figs. S3–S8).

A similar analysis was conducted using WRF-Chem (Tables S4–S5, Figs. S9–S14). The WRF-Chem simulated daily average $PM_{2.5}$, MDA8-$O_3$, and daily average $NO_2$ had Pearson R values of 0.21, 0.63, and 0.62, and RMSE of 16.07 μg/m$^3$, 10.77 ppb, and 8.15 ppb, respectively. The fused-WRF-Chem had much higher Pearson R (0.56, 0.91, 0.70 for daily average $PM_{2.5}$, MDA8-$O_3$, and daily average $NO_2$, respectively) and lower RMSE (7.48 μg/m$^3$, 5.30 ppb, and 5.45 ppb for daily average $PM_{2.5}$, MDA8-$O_3$, and daily average $NO_2$, respectively), indicating a better performance. We also applied the two withholding strategies to demonstrate the robustness of the GF-1 and evaluate the uncertainties of the model (Figs. S12–S13) and found that fused-WRF-Chem under different data withholding strategies still has better performance than WRF-Chem.

### 3.5. Comparison of different CTMs and Fused-CTMs

By using the nearest-grid matching method, we compared the matched grid cell's simulations between GEOS-Chem and CMAQ and between WRF-Chem and CMAQ (Fig. 7), both with and without data fusion. For the simulations without data fusion, CMAQ had a correlation $R^2$ of 0.09, 0.62, and 0.43 with GEOS-Chem for daily average $PM_{2.5}$, MDA8-$O_3$, and daily average $NO_2$ simulations, respectively. The daily average $PM_{2.5}$ and MDA8-$O_3$ simulation correlation between CMAQ and WRF-Chem was relatively low ($R^2$ values are 0.01 and 0.38, respectively), while the daily average $NO_2$ had a higher correlation ($R^2$ = 0.49). For both comparisons, the daily average $PM_{2.5}$ had the lowest

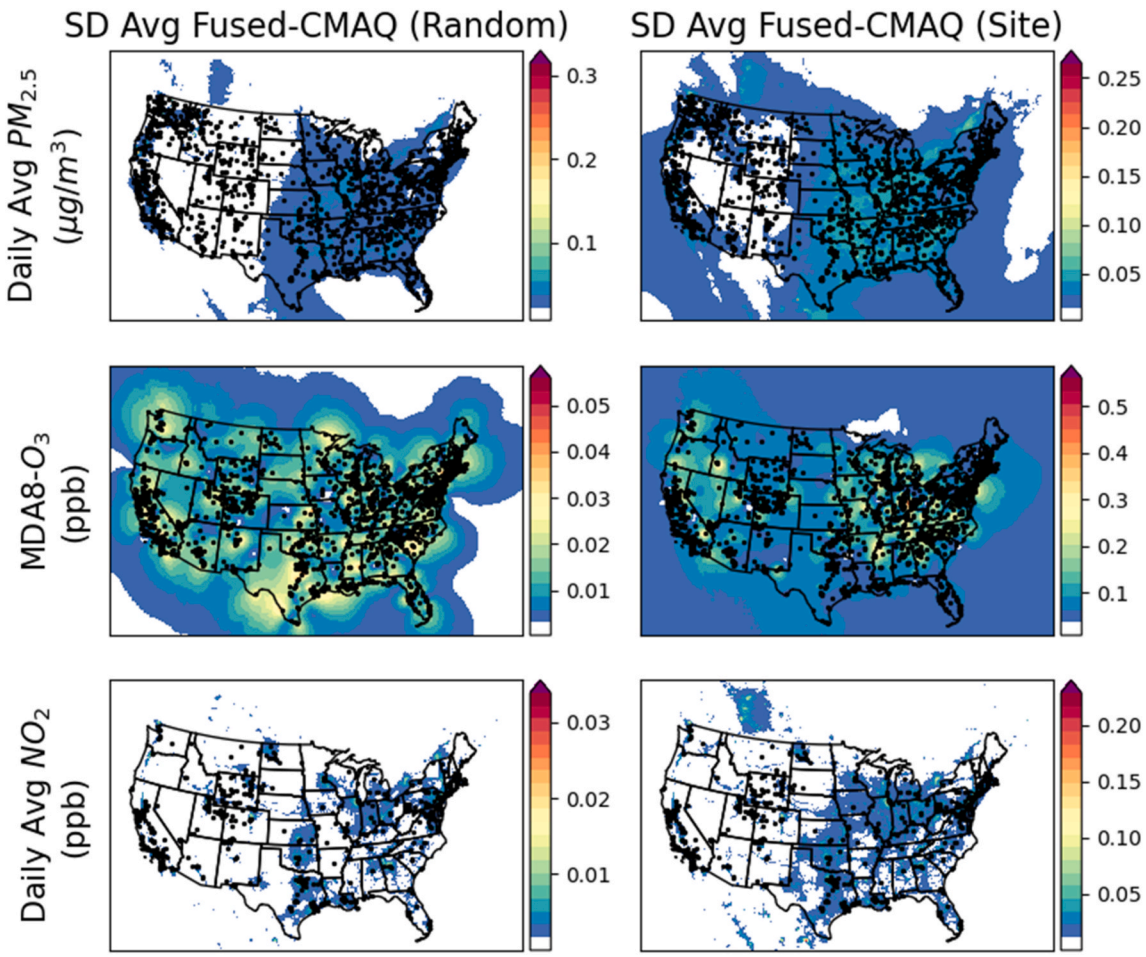



**Fig. 6.** The standard deviation of averaged data fused-CMAQ results under random data withholding and site-wise data withholding strategies. For each grid cell shown in the figures, the value is the standard deviation of 5 withheld data fusion results generated by a 5-fold cross-validation process using the data fusion model.

correlations among the three pollutant species. The $PM_{2.5}$ predictions showed significant differences across different simulations, particularly at high concentration levels. This is likely due to large uncertainties in different fire emission inventories and plume rise models used in these chemical transport models. Correlations between fused-CMAQ and fused-GEOS-Chem and between fused-CMAQ and fused-WRF-Chem improved. The $R^2$ between fused-GEOS-Chem and fused-CMAQ were 0.38, 0.86, and 0.59 for daily average $PM_{2.5}$, MDA8-$O_3$, and daily average $NO_2$. The $R^2$ values between fused-CMAQ and fused-WRF-Chem were 0.36 for daily average $PM_{2.5}$, 0.86 for MDA8-$O_3$, and 0.66 for daily average $NO_2$. As data fusion adjusted CTM outputs closer to observations, the disparities were reduced between the fused CTMs.

The nearest-grid matching method is commonly used when projecting simulation results onto a finer-resolution grid, where each finer grid cell is assigned the concentration value of the nearest coarser grid cell. An alternative comparison approach is to regrid the simulation from different models onto the same grid defined by the coarsest resolution, and such a method is more consistent with the physical meaning of concentrations in chemical transport models, where each grid cell's concentration value represents the spatially averaged concentration over the grid volume. To understand how these two different comparison methods affect comparison results, we regridded the outputs from WRF-Chem, CMAQ, fused WRF-Chem, and fused CMAQ onto the GEOS-Chem grid, which has the coarsest spatial resolution among these models. Specifically, for each GEOS-Chem grid cell, we identified all overlapping grid cells from the other models and averaged their values to obtain the regridded value for that cell (Fig. S15). Overall, the comparison results using the regridding method were generally similar to those obtained from the nearest-grid matching method, as the simulation outputs with data fusion show a higher $R^2$ than those without data fusion. Comparing these two comparison approaches, the regridded approach yielded slightly higher $R^2$ values across the model simulations, particularly for $NO_2$ and $PM_{2.5}$, likely because the averaging process smoothed out the influence of extreme concentration events.

## 4. Discussion

In this study, we implemented a generalized data fusion method for CTMs, the Gen-Friberg method, and applied it to CMAQ, GEOS-Chem, and WRF-Chem models. The fused versions of CMAQ, GEOS-Chem, and WRF-Chem showed improved performance compared to the original simulations, demonstrating the effectiveness of data fusion in reducing biases in chemical transport model outputs. To assess the uncertainty of the data fusion model, we performed 5-fold cross-validation using both random and site-wise data withholding strategies. For the CMAQ model, site-wise data withholding resulted in poorer performance than random withholding, particularly for $PM_{2.5}$, suggesting that fused $PM_{2.5}$ was more sensitive to spatial coverage than temporal coverage of the observation data. For GEOS-Chem and WRF-Chem, the cross-validation results for $NO_2$ and $O_3$ were similar under both withholding strategies, while performance under site-wise data withholding is also poorer than random withholding for $PM_{2.5}$ (Tables S3 and S5). Overall, cross-validation performance for all CTMs in the study was lower than the data fusion with all observations, as each fold was evaluated using observational data not included in that fold's data fusion. However, the cross-validation performance still exceeded the

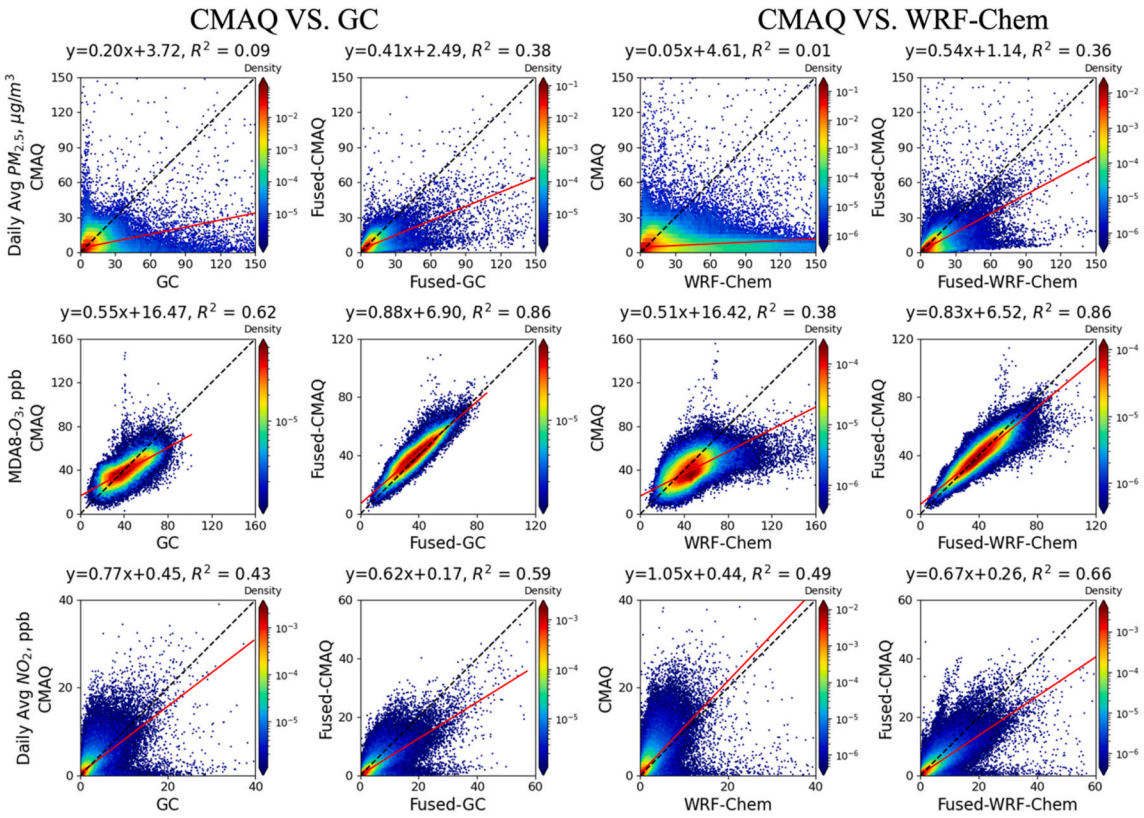



**Fig. 7.** Comparisons between CMAQ and GEOS-Chem (Column 1) or WRF-Chem (Column 3) and between fused-CMAQ and fused-GEOS-Chem (Column 2) or fused-WRF-Chem (Column 4) on daily average $PM_{2.5}$ (first row, unit: μg/m$^3$), MDA8-$O_3$ (second row, unit: ppb), and daily average $NO_2$ (third row, unit: ppb). The black dashed line is the unity (1:1) slope line. The red line shows the linear regression between the compared models. The $R^2$ performance and the formula of linear regressions are indicated at the top of each panel.

original simulations, indicating that the data fusion model was not overfitting and remained robust in predicting spatial and temporal patterns where observations were missing.

### 4.1. Applicability to assessing source impacts and source apportionment

The data fusion method implemented in this study can be used to assess the impacts of specific sources and for source apportionment (SA) of pollutants, although its use is somewhat limited.

Normally, for source impact studies, the concentration impacts ($\Delta conc$) can be calculated from differences of simulations with and without specific sources or estimated from methods such as the direct decoupled method (DDM) (Napelenok et al., 2006), the integrated source apportionment method (ISAM) (Shu et al., 2023; Cohan and Napelenok, 2011), $O_3$ (Dunker et al., 2002) or particulate matter (Yarwood et al., 2007) source apportionment methods (PSAT or OSAT). An "observation-adjusted impact" can be estimated by the following formula (Equation (15)), which was conducted by Huang et al. (2019) for fire-related pollutants:

$$\Delta conc_{adjust} = \frac{conc_{fused}}{conc_{CTM}} \times \Delta conc \quad (15)$$

However, the assumption made by this formula to calculate observation-adjusted impact for specific sources is highly simplified. From the mathematical derivation, the formula implies that the ratio of fused CTM to CTM equals the ratio of fused CTM to CTM without the specific source (Text S4). In reality, the nonlinearity of chemical reactions and the data fusion process can easily break the assumption that the ratio is the same under these two scenarios (with and without the source).

For source apportionment, the results from data fusion can serve as inputs for source apportionment models, such as chemical mass balance (CMB) (Miller et al., 1972) or positive matrix factorization (PMF) (Bhandari et al., 2022). Using fused-CTMs as inputs to CMB or PMF provides the spatial distribution of source contributions, unlike using only observations with source apportionment models, and offers greater computational efficiency compared to the integrated source apportionment method (ISAM) (Shu et al., 2023; Cohan and Napelenok, 2011) or DDM. For instance, Huang et al. used fused-CMAQ as the CMB input and demonstrated the effectiveness of such ensemble models in predicting the spatial distribution of $PM_{2.5}$ source contributions in North Carolina (Huang et al., 2022).

### 4.2. Development of an ensemble data fusion modeling framework

The data fusion model implemented in this study can be applied within an ensemble modeling framework using a boosting (Schapire, 1990; Freund and Schapire, 1997) or bagging method (Breiman, 1996) to further reduce the bias between simulations and observations. Boosting is an ensemble technique to combine different models sequentially. Instead of fusing the simulation with observations in one step, a boosting model combines the predictions of several weak data fusion models to create a strong data fusion model that can additionally reduce the simulation bias. At each step, the following model in the boosting framework used the previous model outputs in the boosting framework as one of the sources of inputs to predict the target concentration (observation). For instance, Senthilkumar et al. (2022) designed a boosting data fusion model by combining a traditional data fusion model with a random forest model to additionally reduce the bias. Bagging (including Bootstrapping and Aggregating steps) is another ensemble method for combining different models or data sources in a parallel way. In the bootstrapping step, several weak models, which will

be used in the ensemble modeling framework, are trained. Instead of using the whole dataset to train models, each model is trained on a subset of data. Then, the predictions from individual models are combined by the aggregating step to produce a final prediction. For example, when CTM, satellite retrieval, and ground measurements are available, GF-1 can be applied to fuse CTM and observations as part of the ensemble framework. Other data fusion methods, such as those proposed in some previous studies (Xiao et al., 2018; Chan et al., 2021), could be included to fuse ground measurements with satellite retrievals. Then, we can easily average these two data fusion models' results to generate the final data fusion field. The resulting averaged concentration field integrates data from CTM, satellite retrievals, and ground measurements, potentially offering more realistic and robust spatiotemporal patterns of pollutant distribution. The user-friendly implementation in this study offers greater flexibility and convenience for researchers to incorporate a variety of data and data fusion methods to develop an ensemble data fusion modeling framework, which can be used to achieve the desired performance.

### 4.3. Limitations

While the generalized data fusion method implemented here effectively reduces biases in concentration fields obtained from CTM simulations, certain limitations may impact how well the procedure captures the real spatiotemporal variations in the concentrations. First, this method is primarily designed for fusing CTM outputs with observations at a daily temporal resolution, as it was originally developed to support air quality exposure assessments used in epidemiological studies, such as estimating mortality, morbidity, or disease-specific relative risks associated with air pollution. Since healthcare claims data for emergency department visits or other hospital records generally are reported daily, the corresponding air pollution concentration fields are also desired at a daily temporal resolution. Accordingly, we used 365.25 days as the period of the trigonometric sinusoidal function in Equation (7) to capture seasonal variations in air pollutants. For applications requiring hourly resolution, one possible solution is to incorporate diurnal variations in Equation (7), for example, by using a variable amplitude with a 24-h period instead of a fixed amplitude. Moreover, although bias is reduced by correcting concentration intensity spatiotemporally and fusing interpolated observations using Kriging interpolation and regression models, poor performance by the underlying CTM is only partially corrected. For instance, meteorological models, which provide input for chemical transport models, often have biases in wind simulations. Inaccurate wind speed and direction can affect pollutant transport, leading to biased concentration fields, especially for extreme pollution events such as fires. Incorrect wind direction can cause smoke to impact a monitor that would not be affected under actual meteorological conditions or vice versa, such that the interpolated field will suffer very large biases. Wind speed also affects smoke concentrations through dilution and influences the timing of smoke arrival via advection. To address concentration bias due to biased wind simulations, both observed and simulated wind data should be incorporated into the data fusion method. Another challenge for our data fusion method is reducing biases in counterfactual scenario simulations. The data fusion model requires both chemical transport model simulations and corresponding observations, so it cannot be applied to counterfactual scenarios where observations are missing. However, counterfactual simulations are critical for health impact assessment or future policy-making. To address this, machine learning or deep learning models that are trained with emission inputs, CTM simulations, and observations could be effective in capturing concentration responses to variations in emissions (Xing et al., 2020; Huang et al., 2021). Additionally, the difference between CTM predictions and observations does not necessarily indicate that the predictions are inaccurate, as CTMs provide average concentrations within a grid cell. However, the data fusion process forces the predictions at monitoring sites to closely align with observed values, which can also influence predicted concentrations in nearby areas. The impact of this alignment is more significant when fusing data into coarse-resolution models.

GF-1 can be applied to various CTMs to correct their biases. We observed that the data fused concentration fields of CTMs exhibited stronger correlations with each other compared to the CTM concentration fields. This is because all fused concentration fields were adjusted toward observations. The higher correlation between different fused-CTM concentrations has the potential to reduce uncertainties when they are applied in epidemiological studies, particularly when examining the correlations between pollutant exposure and health outcomes, morbidity, or mortality. However, notable magnitude differences persisted even after fusing the CTMs' outputs with observations. Among the studied species, daily $PM_{2.5}$ exhibited the largest disparities: fused-CMAQ concentrations were 59 % lower than fused-GEOS-Chem and 46 % lower than fused-WRF-Chem, as indicated by the regression slopes (Fig. 7). Model resolution differences and physics, fire emission inventories, plume-rise schemes, and chemistry mechanisms used in those CTMs could explain these disparities. This raises concerns for health assessment studies that rely on fused concentration fields. Using different CTMs or resolutions, even with identical observational data, can yield varying estimates of mortality, morbidity, and associated economic impacts. This finding underscores the critical need for comprehensive uncertainty analyses in health studies. Such analyses should not only evaluate the fused concentration fields compared with observations but also account for discrepancies stemming from the choice of CTMs. Some previous studies have often overlooked this aspect, which could significantly influence uncertainty estimations. Meanwhile, underlying CTM still plays an important role in affecting the fused concentration intensity and spatiotemporal distributions. Improving the representation of physical and chemical processes within CTMs is essential for further reducing uncertainties in air quality-related health assessments. Future research should focus on harmonizing CTM outputs through advanced fusion techniques and systematically evaluating the implications of model intercomparisons for health studies.

## CRediT authorship contribution statement

**Zongrun Li:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Abiola S. Lawal:** Writing – review & editing, Investigation. **Bingqing Zhang:** Writing – review & editing, Software, Data curation. **Kamal J. Maji:** Writing – review & editing. **Pengfei Liu:** Writing – review & editing, Resources, Data curation. **Yongtao Hu:** Writing – review & editing, Resources, Investigation, Formal analysis, Data curation. **Armistead G. Russell:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization. **M. Talat Odman:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization.

## Code and data availability

The data fusion code is available on GitHub https://github.com/zli867/DataFusion (accessed on August 19th, 2025). EQUATES CMAQ, GEOS-Chem, WRF-Chem, and the AQS observations used in this study are available at: https://zenodo.org/records/16906250 (accessed on August 19th, 2025). The fused CMAQ, fused GEOS-Chem, and fused-WRF-Chem processed by the GF-1 are also available at: https://zenodo.org/records/16906250 (accessed on August 19th, 2025).

## Disclaimer

The contents of this article do not necessarily reflect the views of HEI, or its sponsors, nor do they necessarily reflect the views and policies of EPA or NASA.

## Declaration of competing interest

The authors declare that they have no conflict of interest.

## Acknowledgement

Research described in this article was conducted under Research Agreement #4988-RFA20-1A/21-11 between the Georgia Institute of Technology and the Health Effects Institute (HEI), an organization jointly funded by the United States Environmental Protection Agency (EPA) (Assistance Award No. CR-83998101). The contents of this article do not necessarily reflect the views of HEI or its sponsors, nor do they necessarily reflect the views and policies of the EPA. In addition, this research has been supported by the NASA Health and Air Quality Applied Sciences Team (HAQAST) program (grant no. 80NSSC21K0506), and the US Environmental Protection Agency (EPA; grant no. 84024601). Research cyberinfrastructure resources and services were provided by the Partnership for an Advanced Computing Environment (PACE) at the Georgia Institute of Technology, Atlanta, Georgia, USA.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envsoft.2025.106827.

## References

Archer-Nicholls, S., Lowe, D., Schultz, D.M., McFiggans, G., 2016. Aerosol–radiation–cloud interactions in a regional coupled model: the effects of convective parameterisation and resolution. Atmos. Chem. Phys. 16 (9), 5573–5594.

Bates, J.T., Pennington, A.F., Zhai, X., Friberg, M.D., Metcalf, F., Darrow, L., Strickland, M., Mulholland, J., Russell, A., 2018. Application and evaluation of two model fusion approaches to obtain ambient air pollutant concentrations at a fine spatial resolution (250m) in Atlanta. Environ. Model. Software 109, 182–190.

Berrocal, V.J., Gelfand, A.E., Holland, D.M., 2010. A spatio-temporal downscaler for output from numerical models. J. Agric. Biol. Environ. Stat. 15 (2), 176–197.

Bey, I., Jacob, D.J., Yantosca, R.M., Logan, J.A., Field, B.D., Fiore, A.M., Li, Q., Liu, H.Y., Mickley, L.J., Schultz, M.G., 2001. Global modeling of tropospheric chemistry with assimilated meteorology: model description and evaluation. J. Geophys. Res. Atmos. 106 (D19), 23073–23095.

Bhandari, S., Arub, Z., Habib, G., Apte, J.S., Hildebrandt Ruiz, L., 2022. Source apportionment resolved by time of day for improved deconvolution of primary source contributions to air pollution. Atmos. Meas. Tech. 15 (20), 6051–6074.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24 (2), 123–140.

Byun, D., Schere, K.L., 2006. Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system. Appl. Mech. Rev. 59 (2), 51–77.

Cao, L., Li, S., Sun, L., 2021. Study of different Carbon Bond 6 (CB6) mechanisms by using a concentration sensitivity analysis. Atmos. Chem. Phys. 21 (16), 12687–12714.

Chan, K.L., Khorsandi, E., Liu, S., Baier, F., Valks, P., 2021. Estimation of surface NO2 concentrations over Germany from TROPOMI satellite observations using a machine learning method. Remote Sens. 13 (5), 969.

Cleary, P.A., Dickens, A., McIlquham, M., Sanchez, M., Geib, K., Hedberg, C., Hupy, J., Watson, M.W., Fuoco, M., Olson, E.R., 2022. Impacts of lake breeze meteorology on ozone gradient observations along Lake Michigan shorelines in Wisconsin. Atmos. Environ. 269, 118834.

Cohan, D.S., Napelenok, S.L., 2011. Air quality response modeling for decision support. Atmosphere 2 (3), 407–425.

Columbia University Mailman School of Public Health (CUMSPH). Kriging Interpolation. https://www.publichealth.columbia.edu/research/population-health-methods/kriging-interpolation (accessed March 17, 2025).

Ding, D., Zhu, Y., Jang, C., Lin, C.-J., Wang, S., Fu, J., Gao, J., Deng, S., Xie, J., Qiu, X., 2016. Evaluation of health benefit using BenMAP-CE with an integrated scheme of model and monitor data during Guangzhou Asian Games. J. Environ. Sci. 42, 9–18.

Dodge, M.C., 2000. Chemical oxidant mechanisms for air quality modeling: critical review. Atmos. Environ. 34 (12–14), 2103–2130.

Dunker, A.M., Yarwood, G., Ortmann, J.P., Wilson, G.M., 2002. Comparison of Source Apportionment and Source Sensitivity of Ozone in a Three-Dimensional Air Quality Model. Environmental Science & Technology 36, 2953–2964. https://doi.org/10.1021/es011418f.

EMEP, 2018. European Monitoring and Evaluation Programme.

Emery, C., Liu, Z., Russell, A.G., Odman, M.T., Yarwood, G., Kumar, N., 2017. Recommendations on statistics and benchmarks to assess photochemical model performance. J. Air Waste Manag. Assoc. 67 (5), 582–598.

Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of On-Line learning and an application to boosting. J. Comput. Syst. Sci. 55 (1), 119–139.

Friberg, M.D., Zhai, X., Holmes, H.A., Chang, H.H., Strickland, M.J., Sarnat, S.E., Tolbert, P.E., Russell, A.G., Mulholland, J.A., 2016. Method for fusing observational data and chemical transport model simulations to estimate spatiotemporally resolved ambient air pollution. Environ. Sci. Technol. 50 (7), 3695–3705.

Garcia-Menendez, F., Hu, Y., Odman, M.T., 2013. Simulating smoke transport from wildland fires with a regional-scale air quality model: sensitivity to uncertain wind fields. J. Geophys. Res. Atmos. 118 (12), 6493–6504.

Gilliam, R.C., Hogrefe, C., Godowitch, J.M., Napelenok, S., Mathur, R., Rao, S.T., 2015. Impact of inherent meteorology uncertainty on air quality model predictions. J. Geophys. Res. Atmos. 120 (23), 12,259-12,280.

Grell, G.A., Peckham, S.E., Schmitz, R., McKeen, S.A., Frost, G., Skamarock, W.C., Eder, B., 2005. Fully coupled "online" chemistry within the WRF model. Atmos. Environ. 39 (37), 6957–6975.

Hand, J., Copeland, S., Chow, J., Dillner, A., Hyslop, N., Malm, W., Prenni, A., Raffuse, S., Schichtel, B., Watson, J., 2011. IMPROVE (Interagency Monitoring of Protected Visual Environments): Spatial and Seasonal Patterns and Temporal Variability of Haze and its Constituents in the United States: Report VI.

Hanna, S., Russell, A., Wilkinson, J., Vukovich, J., Hansen, D., 2005. Monte Carlo estimation of uncertainties in BEIS3 emission outputs and their effects on uncertainties in chemical transport model predictions. J. Geophys. Res. Atmos. 110 (D1).

Henneman, L.R., Chang, H.H., Liao, K.-J., Lavoué, D., A Mulholland, J., Russell, A.G., 2017. Accountability assessment of regulatory impacts on ozone and PM 2.5 concentrations using statistical and deterministic pollutant sensitivities. Air Qual. Atmos. Health 10, 695–711.

Huang, Z., Hu, Y., Zheng, J., Zhai, X., Huang, R., 2018. An optimized data fusion method and its application to improve lateral boundary conditions in winter for Pearl River Delta regional PM2. 5 modeling, China. Atmos. Environ. 180, 59–68.

Huang, R., Hu, Y., Russell, A.G., Mulholland, J.A., Odman, M.T., 2019. The impacts of prescribed fire on PM2.5 Air Quality and human health: application to asthma-related emergency room visits in Georgia, USA. Int. J. Environ. Res. Publ. Health 16 (13), 2312.

Huang, L., Liu, S., Yang, Z., Xing, J., Zhang, J., Bian, J., Li, S., Sahu, S.K., Wang, S., Liu, T. Y., 2021. Exploring deep learning for air pollutant emission estimation. Geosci. Model Dev. (GMD) 14 (7), 4641–4654.

Huang, R., Li, Z., Ivey, C.E., Zhai, X., Shi, G., Mulholland, J.A., Devlin, R., Russell, A.G., 2022. Application of an improved gas-constrained source apportionment method using data fused fields: a case study in North Carolina, USA. Atmos. Environ. 276, 119031.

Huijnen, V., Pozzer, A., Arteta, J., Brasseur, G., Bouarar, I., Chabrillat, S., Christophe, Y., Doumbia, T., Flemming, J., Guth, J., 2019. Quantifying uncertainties due to chemistry modelling–evaluation of tropospheric composition simulations in the CAMS model (cycle 43R1). Geosci. Model Dev. (GMD) 12 (4), 1725–1752.

Jaffe, D.A., O'Neill, S.M., Larkin, N.K., Holder, A.L., Peterson, D.L., Halofsky, J.E., Rappold, A.G., 2020. Wildfire and prescribed burning impacts on air quality in the United States. J. Air Waste Manag. Assoc. 70 (6), 583–615.

Jerez, S., Palacios-Peña, L., Gutiérrez, C., Jiménez-Guerrero, P., López-Romero, J.M., Pravia-Sarabia, E., Montávez, J.P., 2021. Sensitivity of surface solar radiation to aerosol–radiation and aerosol–cloud interactions over Europe in WRFv3.6.1 climatic runs with fully interactive aerosols. Geosci. Model Dev. (GMD) 14 (3), 1533–1551.

Kreveld, M., Nievergelt, J., Roos, T., Widmayer, P., 1997. Algorithmic Foundations of Geographic Information Systems. Springer.

Lawal, A.S., Russell, A.G., Kaiser, J., 2022. Assessment of airport-related emissions and their impact on air quality in Atlanta, GA, using CMAQ and TROPOMI. Environ. Sci. Technol. 56 (1), 98–108.

Li, L., Zhou, X., Kalo, M., Piltner, R., 2016. Spatiotemporal interpolation methods for the application of estimating population exposure to fine particulate matter in the contiguous U.S. and a real-time web application. Int. J. Environ. Res. Publ. Health 13 (8), 749.

Li, J., Zhu, Y., Kelly, J.T., Jang, C.J., Wang, S., Hanna, A., Xing, J., Lin, C.-J., Long, S., Yu, L., 2019. Health benefit assessment of PM2.5 reduction in Pearl River Delta region of China using a model-monitor data fusion approach. J. Environ. Manag. 233, 489–498.

Li, Y., Tong, D., Ma, S., Freitas, S.R., Ahmadov, R., Sofiev, M., Zhang, X., Kondragunta, S., Kahn, R., Tang, Y., 2023a. Impacts of estimated plume rise on PM 2.5 exceedance prediction during extreme wildfire events: a comparison of three schemes (Briggs, Freitas, and Sofiev). Atmos. Chem. Phys. 23 (5), 3083–3101.

Li, J., Jang, J.-c., Zhu, Y., Lin, C.-J., Wang, S., Xing, J., Dong, X., Li, J., Zhao, B., Zhang, B., 2023b. Development of a recurrent spatiotemporal deep-learning method coupled with data fusion for correction of hourly ozone forecasts. Environ. Pollut. 335, 122291.

Lin, J., Zhang, A., Chen, W., Lin, M., 2018. Estimates of daily PM2.5 exposure in Beijing using spatio-temporal kriging model. Sustainability 10 (8), 2772.

Lyu, B., Hu, Y., Zhang, W., Du, Y., Luo, B., Sun, X., Sun, Z., Deng, Z., Wang, X., Liu, J., Wang, X., Russell, A.G., 2019. Fusion method combining ground-level observations with chemical transport model predictions using an ensemble deep learning framework: application in China to estimate spatiotemporally-resolved PM2.5 exposure fields in 2014–2017. Environ. Sci. Technol. 53 (13), 7306–7315.

Maji, K.J., Li, Z., Vaidyanathan, A., Hu, Y., Stowell, J.D., Milando, C., Wellenius, G., Kinney, P.L., Russell, A.G., Odman, M.T., 2024a. Estimated impacts of prescribed fires on air quality and premature deaths in Georgia and surrounding areas in the US, 2015–2020. Environ. Sci. Technol. 58 (28), 12343–12355.

Maji, K.J., Ford, B., Li, Z., Hu, Y., Hu, L., Langer, C.E., Hawkinson, C., Paladugu, S., Moraga-McHaley, S., Woods, B., Vansickle, M., Uejio, C.K., Maichak, C., Sablan, O.,

Magzamen, S., Pierce, J.R., Russell, A.G., 2024b. Impact of the 2022 New Mexico, US wildfires on air quality and health. Sci. Total Environ. 946, 174197.
Miller, M., Friedlander, S., Hidy, G., 1972. A chemical element balance for the Pasadena aerosol. J. Colloid Interface Sci. 39 (1), 165–176.
Ministry of Ecology and Environment. China National Environmental Monitoring Centre. http://www.cnemc.cn/en/ (accessed January 6, 2025).
Murphy, B., Yurchak, R., Müller, S., 2024. GeoStat-Framework/PyKrige: V1.7.2 (V1.7.2).
Napelenok, S.L., Cohan, D.S., Hu, Y., Russell, A.G., 2006. Decoupled direct 3D sensitivity analysis for particulate matter (DDM-3D/PM). Atmos. Environ. 40 (32), 6112–6121.
Ng, N.L., Dillner, A.M., Bahreini, R., Russell, A.G., de La Beaujardière, J., Flynn, J., Gentner, D.R., Griffin, R., Hawkins, L.N., Jimenez, J.L., 2022. Atmospheric Science and Chemistry mEasurement NeTwork (ASCENT): a new ground-based high time-resolution air quality monitoring network. AGU Fall Meeting Abstracts, pp. A55R–1360.
Picciotto, S., Huang, S., Lurmann, F., Pavlovic, N., Ying Chang, S., Mukherjee, A., Goin, D.E., Sklar, R., Noth, E., Morello-Frosch, R., Padula, A.M., 2024. Pregnancy exposure to PM2.5 from wildland fire smoke and preterm birth in California. Environ. Int. 186, 108583.
Schapire, R.E., 1990. The strength of weak learnability. Mach. Learn. 5 (2), 197–227.
Senthilkumar, N., Gilfether, M., Metcalf, F., Russell, A.G., Mulholland, J.A., Chang, H.H., 2019. Application of a fusion method for gas and particle air pollutants between observational data and chemical transport model simulations over the contiguous United States for 2005–2014. Int. J. Environ. Res. Publ. Health 16 (18), 3314.
Senthilkumar, N., Gilfether, M., Chang, H.H., Russell, A.G., Mulholland, J., 2022. Using land use variable information and a random forest approach to correct spatial mean bias in fused CMAQ fields for particulate and gas species. Atmos. Environ. 274, 118982.
Shu, Q., Napelenok, S.L., Hutzell, W.T., Baker, K.R., Henderson, B.H., Murphy, B.N., Hogrefe, C., 2023. Comparison of ozone formation attribution techniques in the northeastern United States. Geosci. Model Dev. (GMD) 16 (8), 2303–2322.
Skipper, T.N., Lawal, A.S., Hu, Y., Russell, A.G., 2023. Air quality impacts of electric vehicle adoption in California. Atmos. Environ. 294, 119492.
Tang, B., Stanier, C.O., Carmichael, G.R., Gao, M., 2024. Ozone, nitrogen dioxide, and PM2. 5 estimation from observation-model machine learning fusion over S. Korea: influence of observation density, chemical transport model resolution, and geostationary remotely sensed AOD. Atmos. Environ., 120603
U.S. EPA, AQS. EPA pre-generated AQS data files. https://aqs.epa.gov/aqsweb/airdata/download_files.html (accessed March 28, 2024).
U.S. EPA, CASTNET. Clean Air Status and Trends Network (CASTNET). www.epa.gov/castnet (accessed: January 6, 2025).
U.S. EPA, CMAQ. EPA CMAQ hr2day. https://github.com/USEPA/CMAQ/tree/main/POST/hr2day (accessed September 1, 2024).
U.S. EPA, CSN. EPA Chemical Speciation Network (CSN) - general information. https://www.epa.gov/amtic/chemical-speciation-network-csn-general-information-0 (accessed October 2, 2024).
U.S. EPA, 2021. EQUATESv1.0: Emissions, WRF/MCIP, CMAQv5.3.2 Data — 2002-2019 US_12km and NHEMI_108km. V5 Ed.; UNC Dataverse.
van Donkelaar, A., Martin, R.V., Li, C., Burnett, R.T., 2019. Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. Environ. Sci. Technol. 53 (5), 2595–2611.
Xiao, Q., Chang, H.H., Geng, G., Liu, Y., 2018. An ensemble machine-learning model to predict historical PM2. 5 concentrations in China from satellite data. Environ. Sci. Technol. 52 (22), 13260–13269.
Xing, J., Zheng, S., Ding, D., Kelly, J.T., Wang, S., Li, S., Qin, T., Ma, M., Dong, Z., Jang, C., 2020. Deep learning for prediction of the air quality response to emission changes. Environ. Sci. Technol. 54 (14), 8589–8600.
Xue, T., Zheng, Y., Geng, G., Zheng, B., Jiang, X., Zhang, Q., He, K., 2017. Fusing observational, satellite remote sensing and air quality model simulated data to estimate spatiotemporal variations of PM2. 5 exposure in China. Remote Sens. 9 (3), 221.
Yang, Y., Chen, M., Zhao, X., Chen, D., Fan, S., Guo, J., Ali, S., 2020. Impacts of aerosol–radiation interaction on meteorological forecasts over northern China by offline coupling of the WRF-Chem-simulated aerosol optical depth into WRF: a case study during a heavy pollution event. Atmos. Chem. Phys. 20 (21), 12527–12547.
Yarwood, G., Morris, R.E., Wilson, G.M., 2007. Particulate matter source apportionment technology (PSAT) in the CAMx photochemical grid model. In: Air Pollution Modeling and its Application XVII. Springer, pp. 478–492.
Yu, H., Russell, A., Mulholland, J., Odman, T., Hu, Y., Chang, H.H., Kumar, N., 2018. Cross-comparison and evaluation of air pollution field estimation methods. Atmos. Environ. 179, 49–60.
Yuan, Y., Zhu, Y., Lin, C.-J., Wang, S., Xie, Y., Li, H., Xing, J., Zhao, B., Zhang, M., You, Z., 2023. Impact of commercial cooking on urban PM2.5 and O3 with online data-assisted emission inventory. Sci. Total Environ. 873, 162256.
Zhao, Y., Zhou, Y., Qiu, L., Zhang, J., 2017. Quantifying the uncertainties of China's emission inventory for industrial sources: from national to provincial and city scales. Atmos. Environ. 165, 207–221.