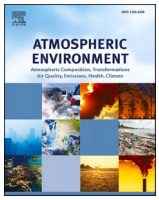
Atmospheric Environment 331 (2024) 120604

Contents lists available at ScienceDirect

Atmospheric Environment

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



MEIAT-CMAQ: A modular emission inventory allocation tool for Community Multiscale Air Quality Model

Haofan Wang [a,b,c], Jiaxin Qiu [d], Yiming Liu [a,b,c,*], Qi Fan [a,b,c], Xiao Lu [a,b,c], Yang Zhang [e], Kai Wu [f], Ao Shen [a,b,c], Yifei Xu [a,b,c], Yinbao Jin [a,b,c], Yuqi Zhu [a,b,c], Jiayin Sun [a,b,c], Haolin Wang [a,b,c]

[a] School of Atmospheric Sciences, Sun Yat-sen University, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, 519082, China
[b] Guangdong Provincial Field Observation and Research Station for Climate Environment and Air Quality Change in the Pearl River Estuary, Guangzhou, 510275, China
[c] Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, Sun Yat-sen University, Zhuhai, China
[d] Shenzhen Key Laboratory of Ecological Remediation and Carbon Sequestration, Institute of Environment and Ecology, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, 518055, China
[e] College of Resources and Environment, Chengdu University of Information Technology, Chengdu, 610225, China
[f] Department of Civil and Environmental Engineering, University of California, Irvine, CA, 92697, USA

HIGHLIGHTS

- Advanced spatial allocation in MEIAT-CMAQ enhances transportation emissions accuracy.
- MEIAT-CMAQ improves air quality model precision for O_3, NO_2, CO, and $PM_{2.5}$.
- The vertical allocation in emission inventories can enhance the accuracy of simulations.

ARTICLE INFO

Keywords:
MEIAT-CMAQ
Emission
Air quality
Allocation
Model-ready inventory

ABSTRACT

The Modular Emission Inventory Allocation Tool for Community Multiscale Air Quality Model (MEIAT-CMAQ) refines emission inventories by providing detailed spatial (horizontal and vertical), temporal, and species allocations, enhancing the accuracy of CMAQ performance. Its efficient algorithm and modular design offer flexibility for managing both gridded and tabulated inventories, widely used in various sectors. In addition, the shapefiles with specific shapes supported by MEIAT-CMAQ can address the allocation challenges in transportation emissions. The evaluation of MEIAT-CMAQ, using model-ready inventories before (BASE scenario), and after allocation without (EXPR scenario) or with (EXPR-V scenario) vertical allocation, demonstrates significant improvements in the mean bias (MB) of gaseous pollutants (O_3, NO_2, CO). In both the EXPR and EXPR-V scenarios, the MB for O_3 exhibits notable enhancements, with respective improvements of 5.7% and 26.9%. For NO_2, corresponding MB improvements are even more pronounced, reaching 27.6% and 61.7% in the EXPR and EXPR-V scenarios, respectively. Likewise, enhancements are observed in the MBs of CO, demonstrating increases of 8.4% and 45.2% in the EXPR and EXPR-V scenarios, respectively. Moreover, with regard to spatial accuracy, the incorporation of the MEIAT-CMAQ model yields significant improvements. Specifically, in the EXPR scenarios, spatial accuracy for O_3 and NO_2 demonstrates respective enhancements of 13.5% and 9.5%. Furthermore, the inclusion of vertical allocation leads to additional enhancements in CO, NO_2, and $PM_{2.5}$, resulting in improvements of 17.6%, 16.6%, and 23.2%, respectively. MEIAT-CMAQ provides an efficient method for transforming coarse-resolution emission inventories into high-resolution files directly useable in the model, offering enhanced flexibility for users to select any period for generating model-ready emission files. This capability provides substantial technical support for automating processes within business departments and significantly improves the performance of high-resolution modeling and forecasting.

* Corresponding author. School of Atmospheric Sciences, Sun Yat-sen University, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, 519082, China.
E-mail address: liuym88@mail.sysu.edu.cn (Y. Liu).

https://doi.org/10.1016/j.atmosenv.2024.120604
Received 9 March 2024; Received in revised form 16 May 2024; Accepted 20 May 2024
Available online 22 May 2024

1. Introduction

High-resolution emission inventories of air pollutants are essential for formulating effective atmospheric environmental mitigation strategies (Lu et al., 2023; Qiu et al., 2023; Wang et al., 2022), improving ambient air quality forecasts (Wu et al., 2014), and promoting air pollution research (Wang et al., 2021; Yang et al., 2020). Determining pollutant concentrations necessitates employing an air quality model (AQM) (Liu and Wang, 2020a, 2020b), such as the Weather Research and Forecast (WRF)–Community Multiscale Air Quality (CMAQ) model. For an effective analysis of regional air quality using regional numerical models like WRF-CMAQ, emission inventories with a 1–5 km resolution are required. For instance, the Multi-resolution Emission Inventory for China (MEIC, http://meicmodel.org.cn, last accessed: May 7, 2023), which adopts a 'bottom-up' methodology, offers a resolution of 0.25° (Li et al., 2017; Zheng et al., 2018), while satellite-derived NO_x emissions, developed via a 'top-down' approach, have a resolution of 9 km (Yang et al., 2021).

Moreover, due to limitations in observational methods and economic constraints, emission inventory production typically focuses on monthly or annual scales for conventional species such as NO_x, NH_3, SO_2, VOC, $PM_{2.5}$, and PM_{10} (McDuffie et al., 2020; Veldeman et al., 2014). Yet, AQM demands hourly scale inventory emissions with more detailed species descriptions based on chemical mechanisms, such as the Carbon-Bond VI (CB06) (Luecken et al., 2019) and the chemical mechanism of Statewide Air Pollution Research Center (SAPRC07) (Hutzell et al., 2012; Xie et al., 2013). This discrepancy necessitates significant effort in adapting original emission inventories for AQM use. Consequently, there is a need for a tool to facilitate spatial, temporal, and species allocation in emission inventory processing. In detail, the spatial allocation usually maps original emission inventories to the AQM domain based on spatial surrogates (Zheng et al., 2021; Zhou et al., 2017), and hourly emissions can be determined through temporal allocation, which involves analyzing temporal profiles such as monthly production statistics, heating degree days, and variations in traffic activity (Wang et al., 2021). Speciated emissions for specific mechanisms, such as CB06 or SAPRC07, are generated through chemical speciation. This process utilizes speciation profiles to determine the composition of organic gases and particulate matter in different sectors (Huang et al., 2015).

Spatial allocation is a critical component in the downscaling of original emission inventories, given that temporal and species allocations are typically reliant on established profiles. This component, spatial allocation, involves employing a spatial surrogate to distribute a proportion of total national and regional emissions across a specified grid, with values ranging from 0 to 1, as noted by Eyth and Habisak (2003). To reflect the true intensity of activities, proxies such as population density, land use, and road maps have been crafted as spatial surrogates in diverse sectors including residential, agriculture, and transportation (Lin et al., 2022). Commonly, the "nearest" approach is employed for downscaling, which entails selecting the emission value closest to the desired grid, according to Lin et al. (2022). Nevertheless, this technique may exacerbate emission discrepancies due to a lack of consideration for potential projection variances between the AQM and emission inventories, which may arise from alterations within the research domain. Recognizing these challenges, some studies have been motivated to devise more accurate tools and methodologies. For example, the Traffic Emission Modelling and Mapping Suite (TEMMS) introduces an 'intersect' method to mitigate this issue, it is limited to managing only the transportation emissions (Namdeo et al., 2002). Conversely, the tool for Hourly Specification of core inventory of air annual emissions (THOSCANE) is capable of processing emissions from various sectors; however, it falls short in generating model-ready emission inventories for AQM (Monforti and Pederzoli, 2005). Recently, the advancement of the Inventory Spatial Allocation Tool (ISATv2.0) has shown promising results in facilitating the processing of emission inventories for WRF-AQM by allowing the definition of WRF nested domains and the generation of model-ready emission inventories (Wang et al., 2023). Despite this, the "sub-grid" method in ISATv2.0 supports only spatial surrogates in a grid format, which may not be fully suitable for specific transportation sectors.

In this study, the Modular Emission Inventory Allocation Tool for CMAQ (MEIAT-CMAQ) was developed to downscale original emission inventories from various formats to CMAQ model-ready emission inventories in a user-friendly and modular way. MEIAT-CMAQ uses the "intersect" method to complete spatial allocation with ArcGIS ports, which allows for parallel processing on computers and improves program efficiency. Furthermore, MEIAT-CMAQ includes a component to enable users to allocate emissions vertically for original emission inventories that only cover one layer. This enhancement enables a more accurate representation of upper atmospheric emissions, including those from aviation and a portion of power emissions. Additionally, this tool benefits from modularity, which allows it to be flexible to process not only original emission inventories in grid but also in tabulation. Moreover, the original emission inventories in a tabular format that are organized in terms of administrative districts for total emissions usually come from business units using a "bottom-up" approach. However, the tools that have been developed in the past are typically only applicable to gridded original emission inventories, which is a major constraint for business development. The architecture of MEIAT-CMAQ has been revamped to enhance efficiency and accuracy in assigning original emission inventories. As a result, it can now adapt to various formats of original emission inventories, making it an ideal tool for CMAQ users in both research and business developments.

This paper provides a detailed overview of the algorithmic principles underlying the MEIAT-CMAQ and demonstrates its applicability through simulations with the WRF-CMAQ model. Apart from the introduction in Section 1, we offer an in-depth introduction to the workflow and algorithmic principles underlying MEIAT-CMAQ in Section 2. Section 3 shows the evaluation of various emission inventories from different allocation methods, and Section 4 presents our conclusions. Additionally, a step-by-step manual for MEIAT-CMAQ has been furnished in the appendix.

2. Model structure

MEIAT-CMAQ has been specifically designed for the Windows operating system and is developed using Python 3 and GIS (Geographic Information System) interfaces. To optimize processing efficiency, the tool is configured to utilize 50% of the CPU (Central Processing Unit) by default, enabling parallel processing of the entire allocation process. Fig. 1 provides an overview of MEIAT-CMAQ, which consists of three main parts that transform the original emissions inventories into model-ready emission inventories.

Part 1 serves as a pre-processing module that includes components enabling the conversion of annual-scale original emission inventories into monthly-scale emission inventories based on monthly profiles. Additionally, the pre-processing module features scripts for processing commonly used emission inventories into GeoTIFF format, as is required by MEIAT-CMAQ. This format, which straightforwardly describes spatial data, must adhere to the World Geodetic System 1984 (WGS-1984) projection with longitude and latitude dimensions (https://gisgeography.com/wgs84-world-geodetic-system/, last access: May 22, 2024). Emission values within these GeoTIFF files should be expressed in Million grams (Mg)/year or Mg/month. Notably, most open-source emission inventories, such as the Emissions Database for Global Atmospheric Research (EDGAR) and the Community Emissions Data System (CEDS), already use WGS-1984 projections, facilitating easy conversion to the required MEIAT-CMAQ format. Although units may vary across inventories, they can all be standardized to Mg/year or Mg/month with minimal effort.

Part 2 comprises two scripts which encompass spatial allocation,

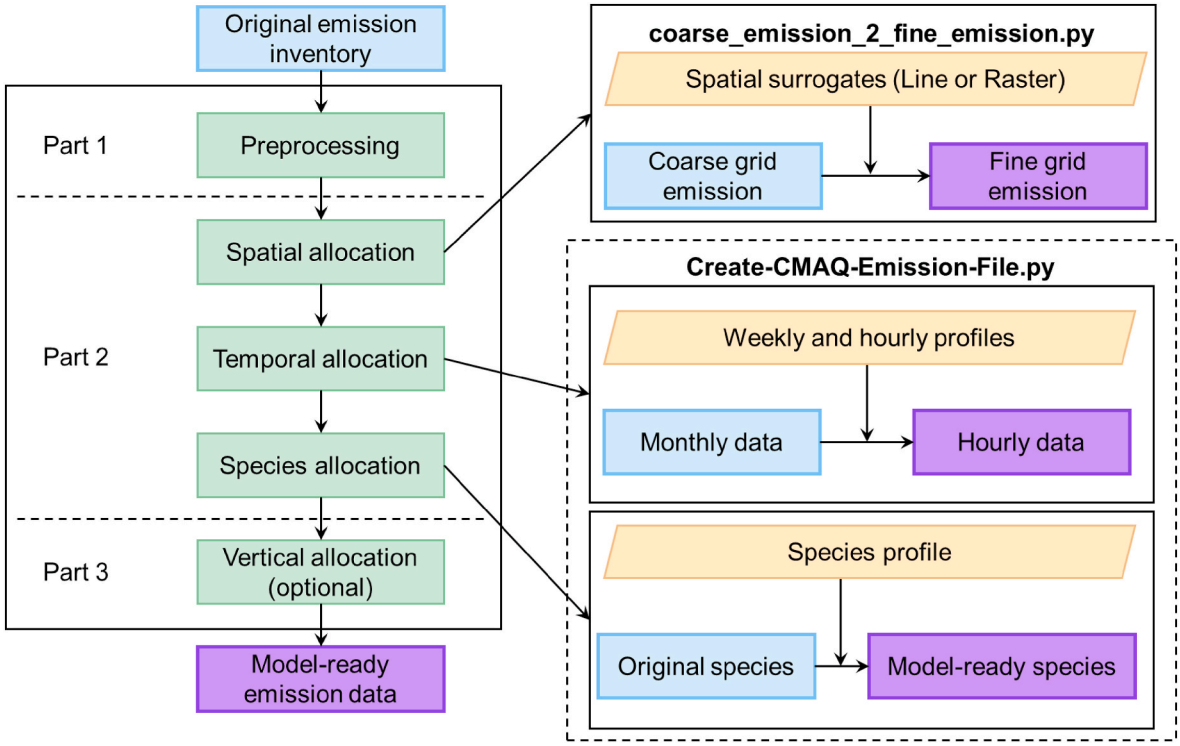


Fig. 1. Operational workflow for MEIAT-CMAQ.

temporal allocation, and species allocation. The first script, 'coarse_emission_2_fine_emission.py,' performs spatial allocation by converting coarse-resolution emission inventories (i.e., the resolution of original emission inventories) to fine-resolution emission inventories (i.e., the resolution of the AQM domain) using spatial surrogates. The second script, 'Create-CMAQ-Emission-File.py', handles temporal and species allocation, converting the assigned results from the first script into model-ready emission inventories. The temporal allocation process supports only monthly scale emission inventories and assumes that there are four weeks in a month with emissions remaining constant across each week and the module can be modified to account for weekly and diurnal variations in weekly and hourly profiles, facilitating allocation of monthly emissions to hourly emissions.

Part 3 is an optional component that handles vertical allocation. This part is omitted for emissions occurring solely near the surface, while it is essential for sectors where emissions may occur at aloft.

The assignment of spatial, temporal, and species components constitutes the fundamental technological foundation of the MEIAT-CMAQ tool. This section is devoted to an in-depth examination of these three critical areas. Meanwhile, the approach for vertical allocation will be comprehensively elucidated in Section 2.3.

2.1. Spatial allocation

Spatial allocation constitutes a critical component of MEIAT-CMAQ, with temporal and species allocation relying on its output for subsequent processing. Fig. 2 displays a conceptual representation of spatial allocation, which aims to transform emissions from the original emission inventory resolution (coarse resolution) to the CMAQ domain resolution (fine resolution) using user-specified spatial surrogates that react to the real intensity of the activity map of the specified sectors. For instance, in Fig. 2, a coarse emission grid with a value of 40 Mg is converted into a 3 × 3 grid configuration (i.e., the spatial allocation process). Spatial surrogates are then employed to represent the intensity of activity within the coarse grid at a finer resolution, ultimately enabling the calculation of fine grid emissions (Eq. (1)).

$$E_{fine} = E_{coarse} \times \frac{ss_{fine}}{ss_{coarse}} \quad (1)$$

In Eq. (1), E_{fine} is the fine grid (i.e., CMAQ simulation grid) emission in Mg, E_{coarse} denotes the coarse grid (i.e., original emission inventory grid or administrative area) emission in Mg, ss_{fine} corresponds to the value of the spatial surrogate grid associated with the fine grid, and ss_{coarse} is the sum value of the spatial surrogate grid within the coarse grid range. Thus, the sum value of ss_{fine}/ss_{coarse} within the same coarse grid equals 1, allowing the coarse grid to be considered as a computational cell.

Fig. 3 elucidates the process of spatial distribution within MEIAT-CMAQ, exhibiting the constituents of this allocation, which encompass modules such as "create fine grid", "fine grid information", "calculate allocation factor", "calculate coarse grid emission", and "calculating fine grid emissions". The GRIDDESC file used in the "create fine grid" module, an output from the Meteorology-Chemistry Interface Processor (MCIP) which can provide the CMAQ domain projection and grid configuration, is employed in the fabrication of a shapefile delineating the simulation grid, which will be used for calculating ss_{fine} in Eq. (1). Each fine grid should be associated with a distinct coarse grid because of the requirement of MEIAT-CMAQ to ensure that the sum value of ss_{fine}/ss_{coarse} within the same coarse grid equals 1; thus, the "fine grid information" module in MEIAT-CMAQ is developed for addressing this problem. Furthermore, spatial surrogates in raster or line shapefile format are used in the calculation of allocation factors in the "calculate allocation factor" module and the previous tools cannot support spatial surrogates in line shapefile format (Namdeo et al., 2002; Wang et al., 2023). The original emission inventories with coarse resolution are segregated and quantified by a coarse grid in shapefile format via the

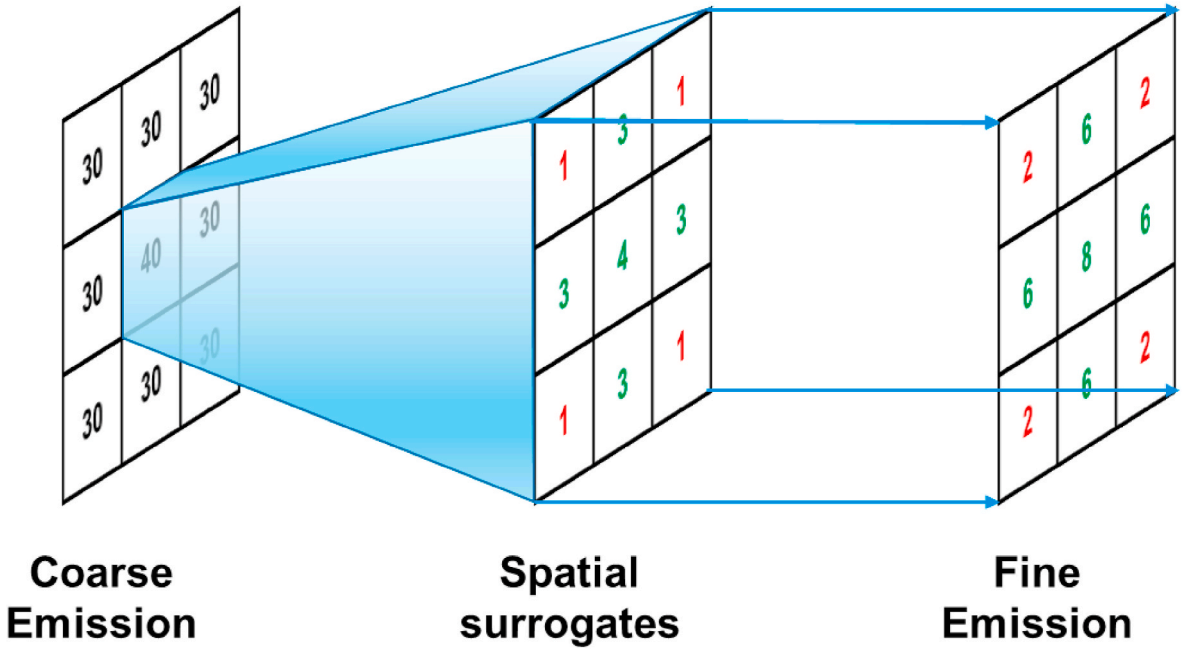


Fig. 2. Conceptual representation of spatial allocation.

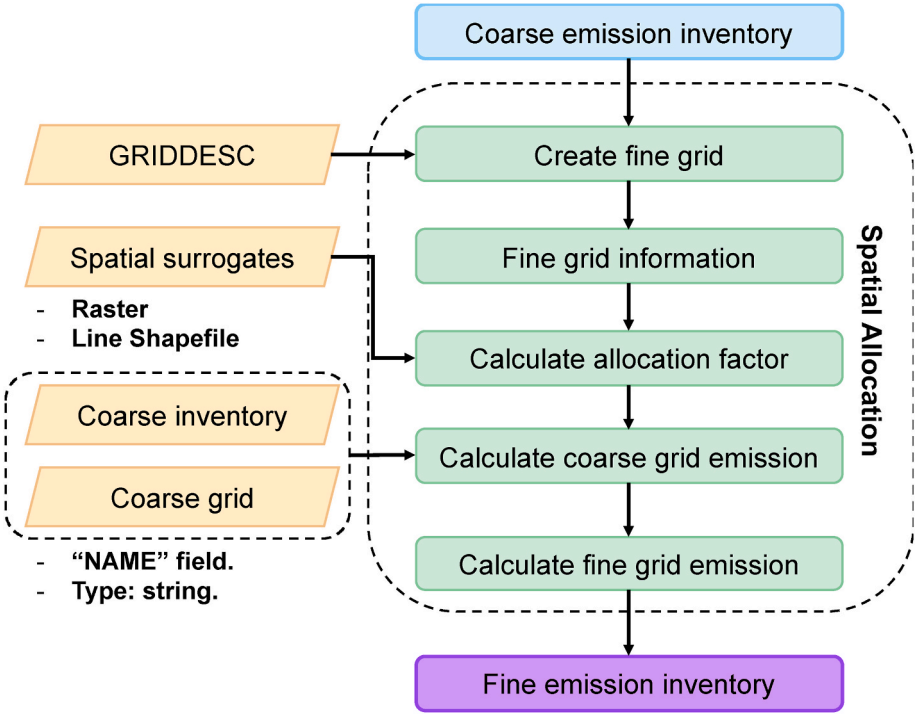


Fig. 3. Operational workflow for spatial allocation in MEIAT-CMAQ.

"calculate coarse grid emission" module, and the final refined grid emission is eventually calculated by the allocation factor and coarse grid emission via the "calculating fine grid emissions" module. We will provide a comprehensive explanation of the calculation and framing of each module in the following subsections.

2.1.1. *Create fine grid*

The "create fine grid" module is designed to transform the CMAQ simulation domain into a shapefile based on the GRIDDESC file with Lambert Conformal Conic (LCC) projection, a projection utilized in regional numerical models, including WRF and CMAQ. This shapefile fundamentally comprises multiple small rectangles, with each representing a grid within the CMAQ simulation domain. To locate each small rectangle, four points are required, which can be computed using Eqs. (2)–(5). The points correspond to the left bottom position $(x_{\min}, y_{\min})$, the right bottom position $(x_{\max}, y_{\min})$, the left upper position $(x_{\min}, y_{\max})$, and the right upper position $(x_{\max}, y_{\max})$. Additionally, x_{org} and y_{org} represent the coordinates in the horizontal and vertical directions, respectively, while dx and dy denote the resolutions in these respective directions. Additionally, the column and row numbers are represented

by i and j (start counting from zero), respectively. The aforementioned projection parameter and total number of columns and rows in the grid can be located in the GRIDDESC file.

$$x_{min} = x_{org} + dx \times i \tag{2}$$

$$x_{max} = x_{org} + dx \times (i+1) \tag{3}$$

$$y_{min} = y_{org} + dy \times j \tag{4}$$

$$y_{max} = y_{org} + dy \times (j+1) \tag{5}$$

2.1.2. Fine grid information

Fig. 4 (a) depicts that gridded emission inventories often face challenges in precisely aligning with multiple fine grids due to variations in coordinate systems. Additionally, the borough boundaries of some tabulated emission inventories might exhibit irregular shapes, which cannot seamlessly overlap the fine grids. These circumstances can result in the smaller grid being divided into one or more sub-grids along the boundary of the coarse grid. Consequently, reasonably addressing these split sub-grids becomes the primary challenge for MEIAT-CMAQ.

Fig. 4 (b) presents a conceptual illustration of the "fine grid information" module, demonstrating how the coarse grid subdivides the fine grid. As shown in Fig. 4 (b), the fine grid is partitioned into four sub-grids (F1–F4), each belonging to distinct coarse grids (C1–C4). The primary objective of the "fine grid information" module is to identify the sub-grid with the largest area and determine the corresponding coarse grid as the current attribution for the fine grid. Compared to other emission inventory process tools, this approach from MEIAT-CMAQ addresses the issue of insufficient accuracy in pinpointing grid information and accelerates computation through the implementation of parallel computing techniques.

2.1.3. Calculate allocation factor

The allocation factor is a critical parameter for spatial allocation, and the "calculate allocation factor" module employs high-resolution spatial surrogates to represent the intensity of activity in each sector. Notably, this module supports both raster and line shapefile formats as spatial surrogates, with the inclusion of line shapefiles marking a significant breakthrough in transportation sector allocation. Furthermore, the "calculate allocation factor" module integrates the output file containing fine grid information about coarse grid labels from the "fine grid information" module to compute ss_{coarse} and ss_{fine}.

Spatial surrogates in raster format exhibit distinct boundaries, which can divide coarse or fine grids. Consequently, the "calculate allocation factor" module employs the "nearest" method for these values. Eq. (6) illustrates the calculation of the allocation factor, where v_{fine} represents the sum value within the fine grid and v_{coarse} denotes the sum value within the coarse grid.

$$A_{fine} = \frac{V_{fine}}{V_{coarse}} \tag{6}$$

Line shapefile format spatial surrogates commonly represent roads, aircraft flight paths, and shipping channels, acting as essential indicators of activity intensity within the transportation sector. Nevertheless, existing inventory assignment tools typically handle roads by converting them into raster format to serve as spatial surrogates. However, due to the significant computational costs and format limitations associated with spatial surrogates, these road-based raster often lack sufficient resolution, resulting in distortions in the final emissions within the fine grid.

Fig. 5 displays the fine grid NO_x emissions for the transportation sector in January, allocated from the 2017 Multiscale Emissions Inventory of China (MEIC) developed by Tsinghua University, using both a 3 km resolution road raster and line shapefile spatial surrogates. The 3 km road raster in Fig. 5 (a) is converted from the four roads with various levels (i.e. motorway, primary road, secondary road, and residential road) from Open Street Map (https://www.openstreetmap.org/#map=6/54.910/-3.432, last accessed: January 11, 2024) depicted in Fig. 5 (b), with values in the raster representing the sum of the four levels road lengths, each multiplied by their respective coefficients indicating activity intensity. Fig. 5 (c) reveals that the fine grid emissions allocated by raster spatial surrogates exhibit pronounced striping patterns due to the low resolution of the road raster. When the resolution of spatial surrogates is close to the fine grids, there is competition in the fine grids because the spatial surrogates do not have enough values for each fine grid, leading to some grids not having values. In contrast, Fig. 5 (d) demonstrates that employing line shapefile spatial surrogates effectively addresses this issue.

Calculating the allocation factor for line shapefile spatial surrogates is similar to that for raster spatial surrogates. The length in the coarse grid is treated as V_{coarse} and the length in the fine grid as V_{fine}. The allocation factor (A_{fine}) is computed following Eq. (6). As most spatial surrogates for the transportation sector are provided in the form of one or more-line shapefiles, converting these line shapefiles into raster format with a specific factor is highly computationally intensive. MEIAT-CMAQ enables users to input an unlimited number of shapefile

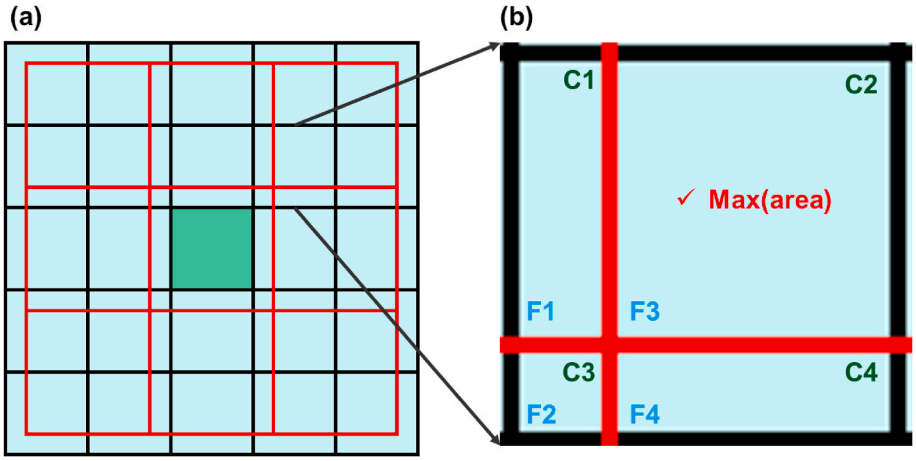


Fig. 4. Conceptual illustration of the "Fine Grid Information" module. In panel (a), red grids represent coarse grids, while black grids denote fine grids. C1–C4 serve as labels for each coarse grid, and F1–F4 designate labels for each sub-grid originating from the fine grid.

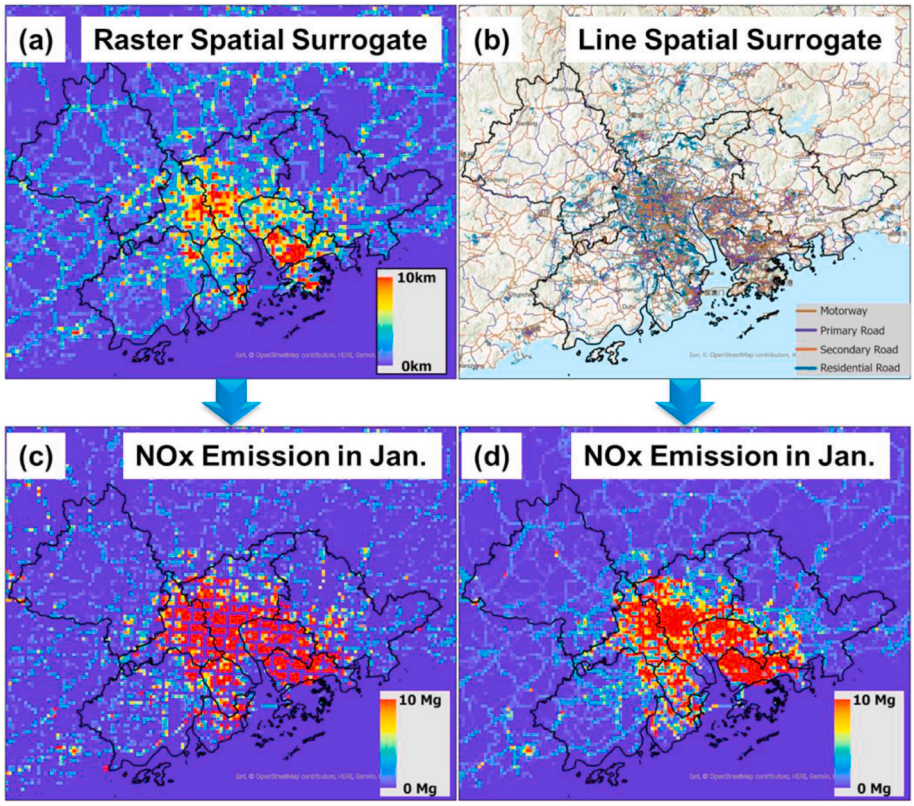


Fig. 5. (a) Road spatial surrogate in raster format, (b) road spatial surrogate in line shapefile format, (c) NO_x fine grid emissions for January using MEIAT-CMAQ with spatial surrogate from panel (a), and (d) same as (c) but from panel (b).

files and easily specify coefficients for each line shapefile, offering a crucial function for allocating emission inventories in the transportation sector.

2.1.4. Calculate coarse and fine grid emissions

Based on the boundaries of each coarse grid, the sum value within the grid represents the coarse grid emission. Thus, the "calculate coarse grid emission" module functions as an independent processor, requiring only the input of the coarse grid in shapefile format and the original coarse grid emission. Additionally, users can create a fishnet shapefile as the coarse grid for gridded emission inventories, ensuring a complete overlap of the boundaries, or they can bypass this step to directly generate the output file of the "calculate coarse grid emission" module for tabulated emission inventories representing countries or cities within each coarse grid. The "calculate fine grid emission" module combines the coarse emission grid table, grid information table, and allocation factor table to compute fine grid emissions.

2.2. Temporal and species allocation

The "CMAQ emission file" module incorporates both temporal and species allocation, conforming to the stringent stipulations of CMAQ model-ready emission inventories. These specifications dictate that the unit for gaseous species must be reported in "mol/s" and particulate species in "g/s." While annual emission inventories are more accessible than their monthly counterparts, the MEIAT-CMAQ model solely accommodates monthly-scale emission inventories across all modules due to the higher temporal resolution in monthly-scale emission inventories.

To address this limitation, MEIAT-CMAQ introduces an efficient alternative procedure, "year2month.py", which effectively converts annual emission inventories into monthly emission inventories. This method harnesses monthly-scale temporal profiles, enabling the use of transformed data within the MEIAT-CMAQ framework. The "year2month.py" script is governed by a monthly temporal profile ("monthly.csv") found in the "temporal" folder.

Alongside the monthly-scale temporal profile, the "CMAQ emission file" model also employs weekly and hourly-scale temporal profiles arranged by emission sectors, resulting in a highly organized and efficient system. Thus, the temporal allocation process aims to assign monthly emissions inventories to hourly emission inventories. Nonetheless, species allocation is also necessary to fulfil the requirements of CMAQ emission files. Therefore, the complete calculation for both temporal and species allocation adheres to Eq. (7).

$$E = E_{m,sector} \times F_{w,sector} \times F_{h,sector} \begin{cases} \times 10^6 \times \frac{1}{3600} \times S_f \ (U_o = Mmol \ \& \ U_t = mol/s) \\ \times 10^6 \times \frac{1}{3600} \times S_f \ (U_o = Mg \ \& \ U_t = g/s) \\ \times 10^6 \times \frac{1}{3600 \times MW} \times S_f \ (U_o = Mg \ \& \ U_t = mol/s) \end{cases} \quad (7)$$

In Eq. (7), E represents the final emission value within the CMAQ model-ready emission inventories, while $E_{m,sector}$ denotes the monthly emission for each sector. $F_{w,sector}$ and $F_{h,sector}$ correspond to the weekly and hourly temporal allocation factors for each sector, as provided by "weekly.csv" and "hourly.csv" files, respectively. Furthermore, S_f, MW, U_o, and U_t symbolize the split factor for the same species, molecular weight, unit of original emission inventories, and unit of CMAQ model-ready emission inventories, respectively. These parameters can be defined in the species files located within the "species" folder.

2.3. Vertical allocation

The input for the "vertical allocation" module is derived from the output file of the "CMAQ emission file" module, which has undergone

spatial, temporal, and species allocation. However, current model-ready emission inventories primarily characterize near-surface sources and are not yet suitable for high-altitude emissions (Zhao et al., 2022; Zheng et al., 2017, 2023). Consequently, accurately characterizing emissions at high altitudes, such as stack emissions from the industrial and power sectors, proves challenging. The "vertical allocation" module can utilize user-provided vertical profiles to complete the vertical allocation according to Eq. (8).

$$E_{layer} = F_{layer} \times E_0 \quad (8)$$

In Eq. (8), E_{layer} denotes the emission value for each respective layer, whereas E_0 signifies the emission value of the initial layer. F_{layer} represents the vertical allocation factor provided by the vertical profile.

3. Model evaluation

3.1. Methods of evaluation

The WRFv4.4.2 (Skamarock et al., 2019) and CMAQv5.4 (https://zenodo.org/record/7218076, last accessed: June 3, 2023) models were compiled and executed on a server within a Linux environment. The WRFv4.4.2 model simulated meteorological conditions, with initial and boundary conditions obtained from the NCEP 1° × 1° Final (FNL) reanalysis dataset (http://dss.ucar.edu/datasets/ds083.2/, last access: May 22, 2024). Fig. 6 illustrates the adoption of three nested domains with horizontal resolutions of 27, 9, and 3 km, respectively. The outermost domain encompasses China and its surrounding countries, while the innermost domain concentrates on the Pearl River Delta (PRD).

To assess the influence of the model-ready emission inventories allocated with the utilization of spatial surrogates (post-assignment emission inventories) compared to the model-ready emission inventories allocated without the utilization of spatial surrogates (pre-assignment inventories) on the performance of CMAQ simulation, a comprehensive study was devised. This study incorporates two distinct scenarios to examine their impact and the first scenario, named BASE, employs pre-assignment emission inventories. Conversely, the second scenario, known as EXPR, utilizes post-assignment emission inventories to quantify the influence of MEIAT-CMAQ on CMAQ performance. Furthermore, an additional scenario, EXPR-V, was established to evaluate the impact of vertical allocation, which encompasses both the assignment processing in the EXPR scenario and the vertical allocation, providing insights into the effects of vertical allocation on the overall evaluation. Table 1 shows the details of the all-scenario configuration and presents that all scenarios employ the identical species (Tables S3–S7) (Shi et al., 2015; Yuan et al., 2010; Zeng et al., 2021) and temporal profiles (Tables S8–S9) (Cai et al., 2018; Zhang et al., 2018), which allows for the attribution of pollutant concentration disparities across scenarios to the distinct spatial allocation algorithms employed by various emissions.

Table 1
Scenario configuration.

Scenario Names	Spatial Surrogates	Vertical Profile
BASE	NO	NO
EXPR	YES	NO
EXPR-V	YES	YES

In the BASE scenario, the original emission inventories with coarse resolution (0.25°) are averaged into the simulation grid with fine resolution (3 km), resulting in the model-ready emissions inventories with the simulation grid resolution. Conversely, the EXPR scenario utilizes population-gridded data in 2017, which can reflect the residential activity intensity, sourced from the LandScan Population Data Explorer (https://landscan.ornl.gov/, last accessed: May 7, 2023), to allocate residential sector emissions. Road data in line shapefile format, obtained from Open Street Map (https://www.openstreetmap.org/, last accessed: Jan 4, 2024), which can describe the activity intensity of transportation sources, is for the allocation of transportation sector emissions. In addition, the China Land Cover Dataset (CLCD) was used for the allocation of agriculture, industry, and power sectors (Yang and Huang, 2021). Additionally, the model-ready emission inventories for the EXPR-V scenario are based on the model-ready emission inventories in the EXPR scenario, but with vertical allocation via the vertical profile (Table S10) (Terrenoire et al., 2015).

All simulations were conducted for four months, specifically, January, April, July, and October in 2017, with each simulation period having ten days of spin-up. Additionally, all scenarios utilized identical physical parameterizations as detailed in Table S1 and Text S1, and all scenarios use the same original emission inventory (MEIC) and simulation periods; thus, the differences in surface pollutant concentrations from various simulations are due to the different emission allocation

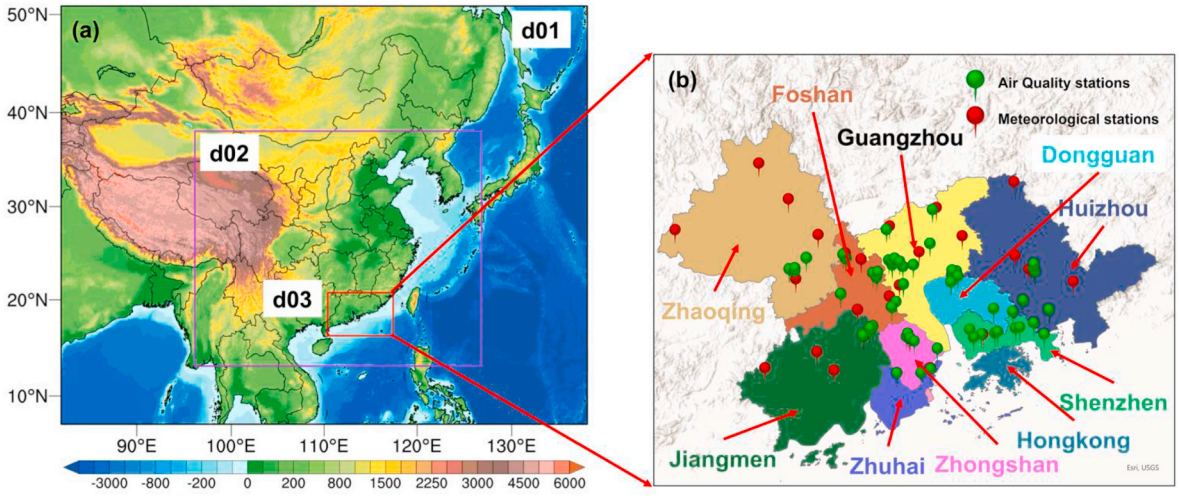


Fig. 6. Depiction of the simulation domains and observation station locations. (a) Illustrates the spatial extent of the simulation domains, and (b) highlights the positions of the observation stations within the simulation domains.

methods. In addition, this study produces the surface pollutant concentrations via WRF-CMAQ and compared to hourly ambient levels of air pollutants observation data from the China National Environmental Monitoring Centre (CNEMC), which include 56 monitoring stations and the locations of these stations are depicted in Fig. 6. The collected data undergo stringent quality control procedures and are subsequently utilized to evaluate the performance of the WRF-CMAQ model. Before evaluating the surface air pollutant concentration, we employ a range of metrics, such as the correlation coefficient (R, Eq. (9)), mean bias (MB, Eq. (10)), and normalized mean bias (NMB, Eq. (11)) for the WRF and CMAQ evaluation. Text S2 and Table S2 illustrate the WRF simulation meteorological fields can well capture the real meteorological processes; thus, this simulation can be used in the CMAQ model.

$$R = \frac{1}{\mathrm{m}} \times \Sigma_{s=1}^{\mathrm{m}} \frac{\sum_{i=1}^{n} \left(M_{s,i} - \overline{M}_{s,i}\right)\left(o_{s,i} - \overline{O}_{s,i}\right)}{\sqrt{\sum_{i=1}^{n} \left(M_{s,i} - \overline{M}_{s,i}\right)^2} \sqrt{\sum_{i=1}^{n} \left(O_{s,i} - \overline{O}_{s,i}\right)^2}} \tag{9}$$

$$MB = \frac{1}{\mathrm{m}} \times \Sigma_{s=1}^{\mathrm{m}} \frac{\sum_{i=1}^{n} \left(M_{s,i} - O_{s,i}\right)}{\mathrm{n}} \tag{10}$$

$$NMB = \frac{1}{\mathrm{m}} \times \Sigma_{s=1}^{\mathrm{m}} \frac{\sum_{i=1}^{n} \left(M_{s,i} - O_{s,i}\right)}{\sum_{i=1}^{n} O_{s,i}} \tag{11}$$

where s and i represent the modeled (M) or observed (O) values for each site and each hour, respectively. $\overline{M}_i$ and $\overline{O}$ are the average of the simulated and observed data, respectively. n and m are the number of total hours in the analysis periods and air quality sites shown in Fig. 6 (B), respectively.

Furthermore, to evaluate the performance of the spatial distribution of pollutants, we first calculate the MB of simulated pollutant concentrations against the observations at each site (Eq. (12)), and then calculate the standard deviation (σ) of them (Eq. (13)). The calculation of σ is shown as follows.

$$MB_s = \frac{\sum_{i=1}^{n} \left(M_{s,i} - O_{s,i}\right)}{n} \tag{12}$$

$$\sigma = \sqrt{\frac{\sum_{s=1}^{m} \left(MB_s - \overline{MB}\right)^2}{n}} \tag{13}$$

where $\overline{MB}$ is the average value of the MBs of all sites and the smaller the standard deviation, the better the performance of the simulated spatial distribution.

3.2. Comparison of the pre-assessment and post-assessment emissions

MEIAT-CMAQ supports raster and line shapefile format spatial surrogates to allocate original emission inventories, and the land use cover in raster format for agriculture and road line in shapefile format for transportation can be used for reflecting the allocating performance. Therefore, Fig. 7 illustrates the pre-assessment emissions and post-assessment emissions in the agriculture and transportation sectors. Specifically, (a)-(d) and (i)-(l) represent the pre-assessment emissions in

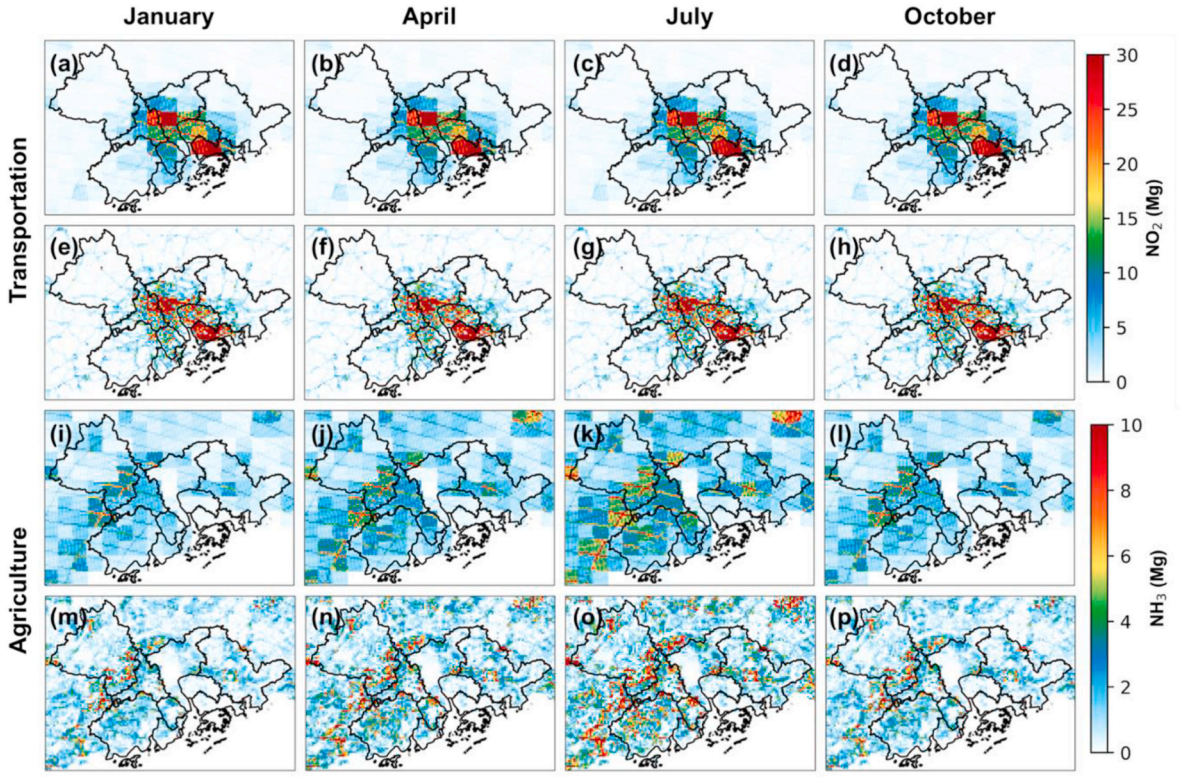


Fig. 7. The emissions after spatial allocation. (a)–(d) and (i)–(l) are the emissions of the BASE scenario in the transportation and agriculture sectors, respectively. (e)–(h) and (m)–(p) are the emissions of EXPR scenario in transportation and agriculture, respectively. The pollutants of the transportation and agriculture sectors are NO_2 and NH_3.

the transportation and agriculture sectors, respectively, for the BASE scenario. On the other hand, (e)-(h) and (m)-(p) depict the post-assessment emissions in the transportation and agriculture sectors, respectively, for the EXPR scenario. Notably, the transportation sector is associated with the pollutant NO_2 while NH_3 is linked to the agriculture sector due to the NO_2 and NH_3 are the major pollutant in the transportation and agriculture sectors, respectively.

It is worth noting that the pre-assessment emission inventories in the BASE scenarios exhibit distinct grid boundaries for both the transportation and agricultural sectors because the original emission inventories resolution is 0.25° (~25 km) while the resolution of the simulation domain is 3 km, and the grid boundaries in the pre-assessment emission inventories are from the original emission inventories. Thus, the pre-assessment emission inventories in the BASE scenarios cannot capture the detailed spatial patterns due to there are no spatial surrogates to reflect the activity intensity. In the case of transportation sources, the pre-assessment emission inventories in the BASE scenario provide only a general overview of high emission values at the junction of Guangzhou and Foshan, as well as in the Shenzhen area. Conversely, the post-assessment emission inventories of transportation sources offer a more detailed portrayal of the spatial distribution along roads, thus providing a finer view of the emission patterns. Similarly, for agricultural source inventories assigned using grid data, the post-assessment emission inventories offer a more detailed representation of the relationship between agricultural source emissions and agricultural land compared to the pre-assessment emission inventories.

Additionally, the model-ready emission inventories in the EXPR-V scenario are based on the post-assessment emission inventories in the EXPR scenario, but with vertical allocation via the vertical profile (Terrenoire et al., 2015), which suggests only the emission inventories from power and industry sources have the aloft emissions. Fig. 8 depicts the vertical distribution of power and industrial source emissions and the vertical profile shows that the emissions from the power sector are mainly concentrated at a pressure level range of 985-980 hPa, whereas the emissions from the industry sector are mainly concentrated at the 990 hPa pressure level. It is crucial to consider that industrial emissions are predominantly released at higher altitudes, a factor that is often overlooked in certain studies (Wang et al., 2022; Wu et al., 2023).

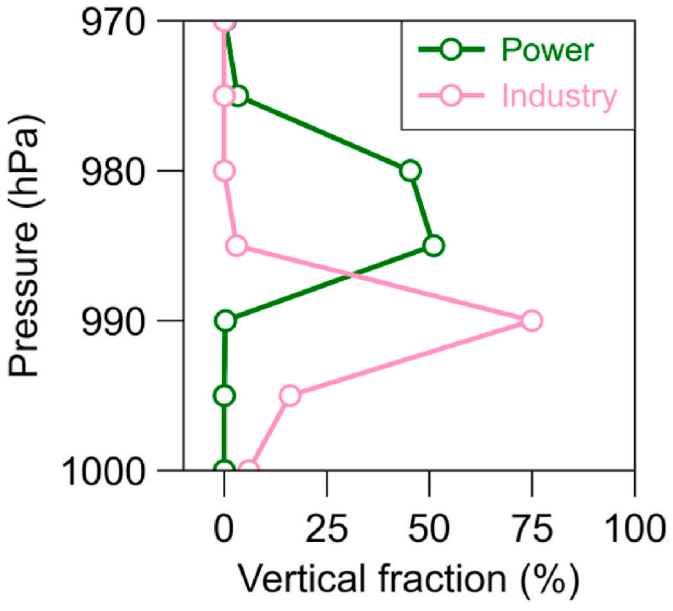


Fig. 8. The vertical profile for power and industry.

Table 2
The summarized metrics for evaluation. The 'Sim.' And 'Obs.' represent the mean value of simulation and observation, respectively and the unit in the first column is only for Sim., Obs., and MB.

	Scenario	Sim.	Obs.	MB	NMB	R
O_3 (μg/m^3)	BASE	50.9	56.1	−5.2	−10.4%	0.721
	EXPR	51.2	56.1	−4.9	−9.6%	0.716
	EXPR-V	52.3	56.1	−3.8	−7.6%	0.732
NO_2 (μg/m^3)	BASE	42.3	37.6	4.7	17.6%	0.546
	EXPR	41.0	37.6	3.4	13.3%	0.540
	EXPR-V	39.4	37.6	1.8	8.8%	0.535
CO (μg/m^3)	BASE	470.6	411.2	59.4	12.7%	0.770
	EXPR	465.6	411.2	54.4	11.3%	0.779
	EXPR-V	443.7	411.2	32.5	6.0%	0.794
$PM_{2.5}$ (μg/m^3)	BASE	26.5	33.9	−7.4	−21.4%	0.786
	EXPR	24.1	33.9	−9.8	−28.5%	0.789
	EXPR-V	22.5	33.9	−11.3	−33.2%	0.791

3.3. Various emissions performance in CMAQ

To assess the performance of various model-ready emission inventories within the WRF-CMAQ framework, this study calculated the average concentrations of key pollutants for selected months—January, April, July, and October. With no significant differences found across these periods (Fig. S1), the average concentrations from these four months were used for further evaluation analysis. These calculated averages were then validated through comparison with the observational data presented in Section 3.1. The results of these comparisons, including the relevant metrics for each pollutant, are concisely summarized in Table 2. This table reveals a slight underestimation of O_3 and $PM_{2.5}$ concentrations, with NMBs ranging from −10.4% to −7.6% and −33.2% to −21.4%, respectively. Conversely, NO_2 and CO concentrations are somewhat overestimated, with NMBs ranging from 8.8% to 17.6% and 6.0%–12.7%, respectively. Additionally, Rs for all assessed pollutants, except for NO_2, exceed 0.7. Specifically, the Rs for O_3 are between 0.716 and 0.732, for CO between 0.770 and 0.794, and for $PM_{2.5}$ between 0.786 and 0.791, while NO_2 presents lower Rs, ranging

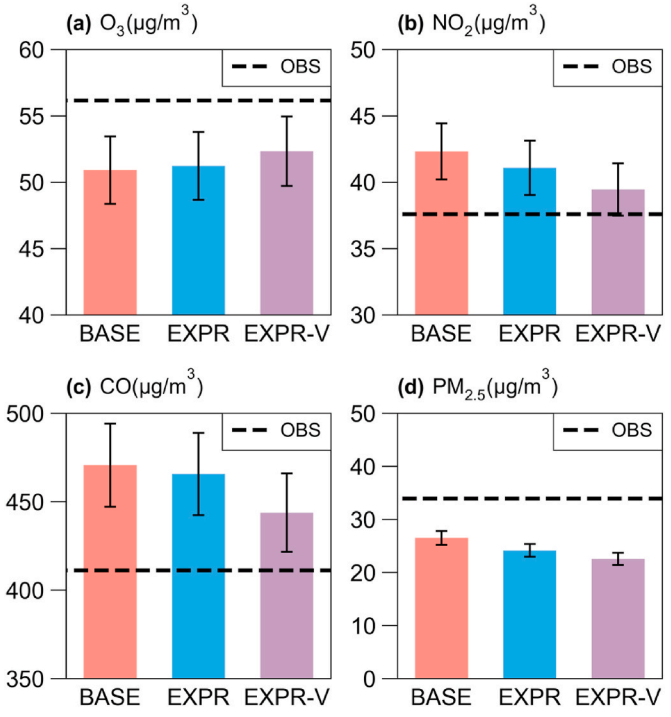


Fig. 9. Simulation average values for (a) O_3, (b) NO_2, (c) CO, and (d) $PM_{2.5}$ in various scenarios with 5% error range and the black line is the observation average values for corresponding pollutants.

from 0.535 to 0.546. Despite these variances, the overall modeling indices for all pollutants demonstrate a comparatively satisfactory level of accuracy and are indicative of the capability of the MEIAT-CMAQ tool to create reasonable emission inventories to simulate atmospheric pollutants.

Fig. 9 illustrates the bias—the discrepancy between simulated and observed values—concerning the gas species O_3, NO_2, and CO, noting a marked amelioration, which suggests that the adoption of post-assessment emissions leads to reduced MBs for O_3, NO_2, and CO, with CO experiencing the most notable reduction. In the EXPR and EXPR-V scenarios, the MB of O_3 decreased from −5.2 μg/m^3 to −4.9 μg/m^3 and -3.8 μg/m^3, respectively, reflecting improvements of 5.7% and 26.9%. For NO_2, the MB declined from 4.7 μg/m^3 to 3.4 μg/m^3 and 1.8 μg/m^3, showing enhancements of 27.6% and 61.7%, respectively. Similarly, the MB of CO dropped from 59.4 μg/m^3 in the BASE scenario to 54.4 μg/m^3 (an 8.4% improvement) and 32.5 μg/m^3 (a 45.2% improvement) in the EXPR and EXPR-V scenarios, respectively. Given that CO is an inert gas frequently utilized as a tracer in studies, its concentration is primarily influenced by regional meteorological conditions. Consequently, the enhanced simulation performance of CO that uses post-assessment emissions underscores the capability of the MEIAT-CMAQ tool to accurately capture the characteristics of actual emissions. Moreover, the marked improvement in MB of CO in the EXPR-V scenario underscores the critical role of vertical allocation in processing emission inventories.

However, it is important to note an increase in the bias for $PM_{2.5}$, which may be attributed to an overestimation of wind speeds and a systematic underestimation of the original emission inventories. In addition, the higher dry deposition rate of the accumulation mode aerosols in CMAQv5.4 can also lead a lower modeling $PM_{2.5}$ concentrations (Pleim et al., 2022). This is evident from the rise in R for $PM_{2.5}$ following the transition from pre-assessment to post-assessment emissions, indicating a more accurate representation of $PM_{2.5}$ emissions by the MEIAT-CMAQ tool.

To evaluate the spatial distribution of pollutants, Fig. 10 presents the standard deviation of the MBs at each site (Eq. (12) and Eq. (13)), reflecting the spatial performance for each scenario. The results indicate a significant enhancement in spatial performance across all pollutants when using post-assessment emission inventories, both with and without vertical allocation. Compared to the BASE scenario, the EXPR-V scenario demonstrates a greater overall improvement in spatial simulation, attributable to its incorporation of vertical allocation, which more accurately represents emission characteristics. Notably, the spatial performance of O_3 and NO_2 has significantly improved, with reductions in standard deviation by 1.2 μg/m^3 (an improvement of 13.5%) and 1.0 μg/m^3 (an improvement of 9.5%) respectively, when compared to the BASE scenario. Additionally, for CO, NO_2, and $PM_{2.5}$, the spatial improvement effect is further enhanced when using post-assessment emission inventories with vertical allocation, resulting in reductions in standard deviation by 1.9 μg/m^3 (an improvement 17.6%), 20.8 μg/m^3 (an improvement 16.6%), and 1.7 μg/m^3 (an improvement 23.2%) respectively.

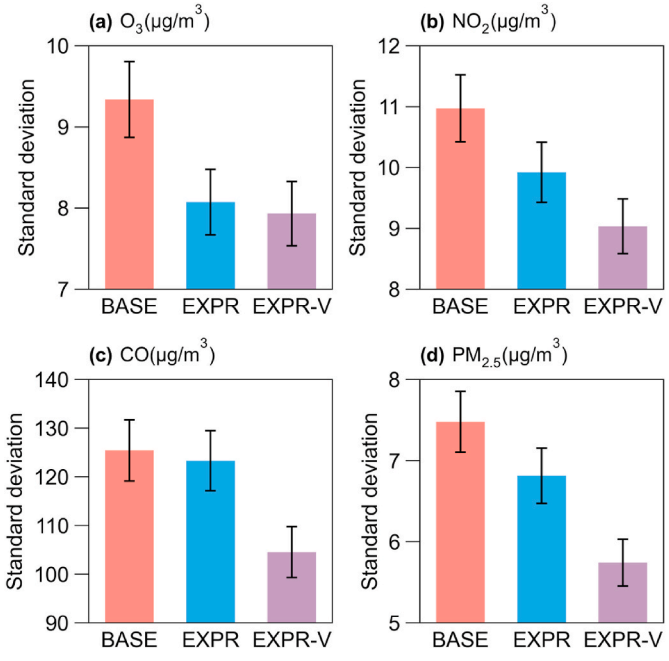


Fig. 10. The standard deviation of MBs with 5% error range in each site for each scenario. (a)–(d) are the results of O_3, NO_2, CO, and $PM_{2.5}$, respectively.

The evaluation outcomes from the WRF-CMAQ model demonstrate that applying post-assessment emission inventories, with and without vertical allocation, effectively reduces simulation biases for gaseous pollutants (O_3, NO_2, and CO) and more accurately captures $PM_{2.5}$ trends. Moreover, these post-assessment emission inventories significantly enhance the spatial performance of the CMAQ model.

4. Conclusion

We have developed a user-friendly modular tool called MEIAT-CMAQ for preparing emission files specifically designed for the CMAQ model. MEIAT-CMAQ encompasses spatial, species, and temporal allocations for the original emission inventory. While we primarily use gridded emission inventories as input to generate model-ready emission files in this paper, the modular structure of MEIAT-CMAQ allows for flexibility in handling various formats, including tabulated emission inventories, which are commonly used for local emissions in a statistical format, which feature is not conducted in previous integrated tools such as ISATv2.0 (Wang et al., 2023).

Furthermore, we have included pre-packaged spatial surrogates in the program to facilitate the spatial allocation process. Additionally, we have prepared temporal allocation profiles for the five main sectors (residential, transportation, power, industry, and agriculture), species allocation profiles for the CB06 mechanism and the SAPRC07 mechanism, as well as vertical allocation files for the industry and power sectors. These pre-packaged files are readily available for users to directly utilize. However, if these files do not fully meet the specific requirements of users, they can easily customize them by utilizing the pre-packaged files as templates and making necessary modifications as per their needs.

In specific terms of emission allocation methods, TEMMS employs the “intersect” method to allocate transportation source emissions, offering greater refinement (Namdeo et al., 2002). However, the main drawbacks of “intersect” are its computationally intensive nature and slow processing speed, along with limitations specific to the source sector (transportation). ISAT addresses these issues through the “sub--grid” method, which speeds up calculations but at the cost of reduced accuracy (Wang et al., 2023). In contrast, MEIAT-CMAQ incorporates an advanced GIS interface and utilizes the “intersect” method, not only retaining the accuracy benefits but also enhancing computational efficiency. Moreover, MEIAT-CMAQ is capable of allocating emissions not just to the transportation sector but to other pollution sources as well. Additionally, MEIAT-CMAQ has enhanced its functionality with support for vertical allocation, which is not included in other tools such as TEMMS, THOSCANE (Monforti and Pederzoli, 2005), and ISAT, allowing for a more realistic representation of emissions from key sectors including industry and power plants, and after completing all allocation operations, users can directly export model-ready emissions with the assistance of MEIAT-CMAQ.

Additionally, we conducted a comparison between the pre-assessment emissions and the post-assessment emissions to evaluate their impact on the performance of the CMAQ model. The results show

that the application of post-assessment emission inventories, both with and without vertical allocation, ameliorates the simulation biases for gaseous pollutants (O_3, NO_2, and CO) and more accurately captures the $PM_{2.5}$ trends, and these post-assessment emission inventories significantly enhance the spatial performance of the CMAQ model. Consequently, it is evident that the model-ready emission inventories, downscaled through MEIAT-CMAQ, substantially enhance the efficacy of the CMAQ model. These findings suggest that employing high-resolution emission inventories through MEIAT-CMAQ can enhance the accuracy and scientific validity of the model, especially when conducting higher-resolution regional air quality simulations.

MEIAT-CMAQ offers an efficient approach for assigning inventories, simplifying the process of transforming coarse-resolution emission inventories into high-resolution emission files that can be directly inputted into the model. Furthermore, MEIAT-CMAQ supports user flexibility by allowing the selection of any desired period for generating the model-ready emission file, providing valuable technical support for automating operations within the business department.

Code and data availability

The source code of MEIAT-CMAQv1.0 and input data used to produce the results used in this paper are archived on Zenodo at https://zenodo.org/record/8001532.

CRediT authorship contribution statement

Haofan Wang: Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Jiaxin Qiu:** Methodology, Conceptualization. **Yiming Liu:** Writing – review & editing, Supervision. **Qi Fan:** Writing – review & editing, Supervision. **Xiao Lu:** Writing – review & editing, Supervision. **Yang Zhang:** Methodology, Conceptualization. **Kai Wu:** Formal analysis. **Ao Shen:** Formal analysis. **Yifei Xu:** Formal analysis. **Yinbao Jin:** Formal analysis. **Yuqi Zhu:** Formal analysis. **Jiayin Sun:** Formal analysis. **Haolin Wang:** Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have shared the link to my data and code in the manuscript.

Acknowledgements

This work was supported by the Guangdong Major Project of Basic and Applied Basic Research (Grant No. 2020B0301030004), Science and Technology Program of Guangdong Province (Science and Technology Innovation Platform Category) (No. 2019B121201002), Key-Area Research and Development Program of Guangdong Province (No. 2020B1111360003), and the National Natural Science Foundation of China (No. 42075181, 42105097). This work was also supported by the high-performance grid-computing platform of Sun Yat-sen University.

Appendix A. MEIAT-CMAQ step-by-step guide

Taking the nested domains in Fig. 6 with a spatial resolution of 3 km as an example, and the MEIC in 2017 as the original emission inventories.

Step 1: namelist.input configuration

(a) namelist.input attribution introduction

Table A1
the namelist.input attribution.

Attribution	Description	Process
Global		
griddesc_file	The path of GRIDDESC file.	A
griddesc_name	The grid name of domain.	A
big_grid_file	The path of coarse grid shapefile.	S
geotiff_dir	The path of original emission directory (GeoTIFF format).	S
inventory_label	The prefix in the original emission inventory file name.	S
inventory_year	The year identifier in the original emission inventory file name.	S
sectors	The sectors identifier in the original emission inventory file name.	A
allocator	The file name of spatial surrogates in the “allocator” directory.	S
allocator_type	The type of spatial surrogates (raster or line). If this option is line, the allocator will read the “line” part.	S
inventory_mechanism	The chemical mechanism of original emission inventories.	Sp. & T
target_mechanism	The chemical mechanism of model-ready emission file.	Sp. & T
start_date	The start date of model-ready emission file.	Sp. & T
end_date	The end date of model-ready emission file.	Sp. & T
cores	Number of parallel cores.	A
Line		
line_files	The file name of line shapefile in the “allocator” directory.	S
line_factors	The weight of each road mentioned in line_files.	S
Control		
create_grid	The control of ‘Create fine grid’ module.	S
grid_info	The control of ‘Fine grid information’ module.	S
create_factor	The control of ‘Calculate allocation factor’ module.	S
coarse_emission	The control of ‘Calculate coarse grid emission’ module.	S
create_source	The control of ‘Calculate fine grid emission’ module.	S

A = All processes; S = Spatial allocation process; Sp. = Species allocation process; T = Temporal allocation process.

(b) namelist.input configuration in this example

Table A2
Contents of case namelist.input

```
&global
  griddesc_file = "input/GRIDDESC.PRD274x181"
  griddesc_name = "PRD274x181"
  big_grid_file = "shapefile/MEIC-0P25.shp"
  geotiff_dir = "input/MEIC"
  inventory_label = "MEIC"
  inventory_year = "2017"
  sectors = 'transportation', 'residential', 'power', 'agriculture', 'industry'
  allocator = 'line', 'landscan-global-2017_nodata.tif', 'power.tif', 'agriculture.tif', 'industry.tif',
  allocator_type = "line", "raster", "raster", "raster", "raster"
  inventory_mechanism = "MEIC-CB05"
  target_mechanism = "CB06"
  start_date = "2017-01-01"
  end_date = "2017-01-02"
  cores = 4
/

&line
  line_files = "motorway.shp", "primary.shp", "residential.shp", "secondary.shp"
  line_factors = 0.435798, 0.326848, 0.081712, 0.155642,
/

&control
  create_grid = 1,
  grid_info = 1,
  create_factor = 1,
  coarse_emission = 1,
  create_source = 1,
/
```

Step 2: spatial allocation

Enter the command “python coarse_emission_2_fine_emission.py” into the terminal.

Step 3: create model-ready emission file (species and temporal allocation)

(a) Species allocation profile introduction

Table A3 displays a portion of the transportation source species allocation profile with the “pollutant” column representing the species name identifier in the original emission inventory file name, while the “emission_species” column denotes species names recognized by CMAQ that will serve as variable names in the model-ready emission file. For the same species, the sum of “split_factor” should be theoretically equal to 1, which represents the weight of splitting the same species into species that can be identified by CMAQ. Typically, original emission inventories are provided in units of “Mg” or “Mmol”, which, even if not directly available, can be easily converted to “Mg” or “Mmol”. Hence, the “inv_unit” column offers the user a choice between the two options. However, for use in CMAQ, emission files must be in “mol/s" (gaseous species) or “g/s" (solid species) units. Therefore, only these units are supported in the “emi_unit” column. The “divisor” column denotes the molecular mass, which serves as the unit conversion factor.

Table A3
Portions of the transportation species profile

pollutant	emission_species	split_factor	divisor	inv_unit	emi_unit
SO2	SO2	1	64	Mg	mol/s
SO2	SULF	0	98	Mg	mol/s
NH3	NH3	1	17	Mg	mol/s
CO	CO	1	28	Mg	mol/s
PMC	PMC	1	1	Mg	g/s
NOx	NO	0.9	30	Mg	mol/s
NOx	NO2	0.092	46	Mg	mol/s
NOx	HONO	0.008	47	Mg	mol/s
OC	POC	1	1	Mg	g/s

(b) temporal allocation profile introduction

It is important to highlight that when allocating monthly-scale emission inventories to individual weeks, the model assumes four weeks per month, with equal total emissions distributed across each week. Subsequently, the module apportions emissions for each day based on the weekly temporal

profile. In this profile, values ranging from 0 to 6 in the "weekly" column (Table A4) correspond to Sunday through Saturday, respectively. Similarly, in the hourly temporal file, values from 0 to 23 in the "hourly" column (Table A5) signify 0 to 23 UTC, respectively. Except for the first column, the column names of the other columns are the emission sectors, which will keep consistent with the sector identification in the original emission inventory file name.

Table A4
The content of weekly temporal profile.

weekly	power	industry	residential	transportation	agriculture
0	0.13	0.078	0.143	0.108	0.143
1	0.147	0.162	0.143	0.155	0.143
2	0.147	0.162	0.143	0.155	0.143
3	0.147	0.162	0.143	0.155	0.143
4	0.147	0.162	0.143	0.155	0.143
5	0.147	0.162	0.143	0.155	0.143
6	0.135	0.112	0.143	0.117	0.143

Table A5
The content of hourly temporal profile.

hourly	power	industry	residential	transportation	agriculture
0	0.032	0.026	0.038	0.017	0.026
1	0.03	0.007	0.038	0.013	0.019
2	0.029	0.007	0.03	0.014	0.019
3	0.028	0.007	0.045	0.015	0.018
4	0.029	0.007	0.045	0.016	0.019
5	0.032	0.007	0.038	0.016	0.021
6	0.035	0.007	0.03	0.029	0.029
7	0.04	0.029	0.03	0.056	0.033
8	0.0433	0.045	0.038	0.0599	0.0473
9	0.0457	0.068	0.038	0.059	0.0576
10	0.0479	0.068	0.03	0.0594	0.07
11	0.0495	0.068	0.045	0.0501	0.0885
12	0.0495	0.068	0.045	0.0501	0.0885
13	0.0497	0.068	0.038	0.0588	0.0823
14	0.0501	0.068	0.03	0.06	0.0803
15	0.05	0.068	0.03	0.062	0.07
16	0.0497	0.068	0.038	0.0594	0.0597
17	0.0489	0.066	0.075	0.0574	0.0453
18	0.0477	0.063	0.075	0.0557	0.0309
19	0.0473	0.037	0.075	0.049	0.0268
20	0.0466	0.037	0.075	0.0454	0.0226
21	0.044	0.037	0.054	0.0417	0.0206
22	0.0397	0.037	0.018	0.0308	0.0206
23	0.0352	0.037	0.018	0.0216	0.0206

(c) complete species and temporal allocation

Enter the command " python Create-CMAQ-Emission-File.py" into the terminal.

Step 4: vertical allocation

(a) Vertical allocation profile introduction

Table A6 presents the vertical allocation profile, listing "vglytop" as sigma coordinates. These values can be obtained by querying the METCRO3D file, an MCIP output where any layers not involved are considered as unemitted for this layer. The "fraction" column denotes the relevant distribution factor

Table A6
vertical allocation profile

Vglvltop	fraction
1.000	0.06
0.995	0.16
0.990	0.75
0.985	0.03

(b) Complete vertical allocation

Enter the command " python vertical_allocation.py" into the terminal.

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosenv.2024.120604.

References

Cai, S., Li, Q., Wang, S., Chen, J., Ding, D., Zhao, B., Yang, D., Hao, J., 2018. Pollutant emissions from residential combustion and reduction strategies estimated via a village-based emission inventory in Beijing. Environ. Pollut. 238, 230–237. https://doi.org/10.1016/j.envpol.2018.03.036.

Eyth, A.M., Habisak, K., 2003. The MIMS spatial allocator: a tool for generating emission surrogates without a geographic information system. In: Proceedings, 12th International Emission Inventory Conference. San Diego.

Huang, C., Wang, H.L., Li, L., Wang, Q., Lu, Q., de Gouw, J.A., Zhou, M., Jing, S.A., Lu, J., Chen, C.H., 2015. VOC species and emission inventory from vehicles and their SOA formation potentials estimation in Shanghai, China. Atmos. Chem. Phys. 15, 11081–11096. https://doi.org/10.5194/acp-15-11081-2015.

Hutzell, W.T., Luecken, D.J., Appel, K.W., Carter, W.P.L., 2012. Interpreting predictions from the SAPRC07 mechanism based on regional and continental simulations. Atmos. Environ. 46, 417–429. https://doi.org/10.1016/j.atmosenv.2011.09.030.

Li, M., Liu, H., Geng, G., Hong, C., Liu, F., Song, Y., Tong, D., Zheng, B., Cui, H., Man, H., Zhang, Q., He, K., 2017. Anthropogenic emission inventories in China: a review. Natl. Sci. Rev. 4, 834–866. https://doi.org/10.1093/nsr/nwx150.

Lin, P., Gao, J., Xu, Y., Schauer, J.J., Wang, J., He, W., Nie, L., 2022. Enhanced commercial cooking inventories from the city scale through normalized emission factor dataset and big data. Environ. Pollut. 315, 120320 https://doi.org/10.1016/j.envpol.2022.120320.

Liu, Y., Wang, T., 2020a. Worsening urban ozone pollution in China from 2013 to 2017 – Part 1: the complex and varying roles of meteorology. Atmos. Chem. Phys. 20, 6305–6321. https://doi.org/10.5194/acp-20-6305-2020.

Liu, Y., Wang, T., 2020b. Worsening urban ozone pollution in China from 2013 to 2017 – Part 2: the effects of emission changes and implications for multi-pollutant control. Atmos. Chem. Phys. 20, 6323–6337. https://doi.org/10.5194/acp-20-6323-2020.

Lu, Y., Yang, X., Wang, H., Jiang, M., Wen, X., Zhang, X., Meng, L., 2023. Exploring the effects of land use and land cover changes on meteorology and air quality over Sichuan Basin, southwestern China. Front. Ecol. Evol. 11, 1131389 https://doi.org/10.3389/fevo.2023.1131389.

Luecken, D.J., Yarwood, G., Hutzell, W.T., 2019. Multipollutant modeling of ozone, reactive nitrogen and HAPs across the continental US with CMAQ-CB6. Atmos. Environ. 201, 62–72. https://doi.org/10.1016/j.atmosenv.2018.11.060.

McDuffie, E.E., Smith, S.J., O'Rourke, P., Tibrewal, K., Venkataraman, C., Marais, E.A., Zheng, B., Crippa, M., Brauer, M., Martin, R.V., 2020. A global anthropogenic emission inventory of atmospheric pollutants from sector- and fuel-specific sources (1970–2017) (preprint). Antroposhere - Energy and Emissions. https://doi.org/10.5194/essd-2020-103.

Monforti, F., Pederzoli, A., 2005. THOSCANE: a tool to detail CORINAIR emission inventories. Environ. Model. Software 20, 505–508. https://doi.org/10.1016/j.envsoft.2004.07.001.

Namdeo, A., Mitchell, G., Dixon, R., 2002. TEMMS: an integrated package for modelling and mapping urban traffic emissions and air quality. Environ. Model. Software 17, 177–188. https://doi.org/10.1016/S1364-8152(01)00063-9.

Pleim, J.E., Ran, L., Saylor, R.D., Willison, J., Binkowski, F.S., 2022. A new aerosol dry deposition model for air quality and climate modeling. J. Adv. Model. Earth Syst. 14, e2022MS003050.

Qiu, J., Fang, C., Tian, N., Wang, H., Wang, J., 2023. Impacts of land use and land cover changes on local meteorology and PM2.5 concentrations in Changchun, Northeast China. Atmos. Res. 289, 106759 https://doi.org/10.1016/j.atmosres.2023.106759.

Shi, J., Deng, H., Bai, Z., Kong, S., Wang, X., Hao, J., Han, X., Ning, P., 2015. Emission and profile characteristic of volatile organic compounds emitted from coke production, iron smelt, heating station and power plant in Liaoning Province, China. Sci. Total Environ. 515–516, 101–108. https://doi.org/10.1016/j.scitotenv.2015.02.034.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Liu, Z., Berner, J., Wang, W., Powers, J.G., Duda, M.G., Barker, D.M., Huang, X.-Y., 2019. A Description of the Advanced Research WRF Model Version 4. UCAR/NCAR. https://doi.org/10.5065/1DFH-6P97.

Terrenoire, E., Bessagnet, B., Rouïl, L., Tognet, F., Pirovano, G., Létinois, L., Beauchamp, M., Colette, A., Thunis, P., Amann, M., Menut, L., 2015. High-resolution air quality simulation over Europe with the chemistry transport model CHIMERE. Geosci. Model Dev. (GMD) 8, 21–42. https://doi.org/10.5194/gmd-8-21-2015.

Veldeman, N., Op 't Eyndt, T., Mensink, C., 2014. Developing air pollutant emission inventories for use in atmospheric transport models. In: Qatar Foundation Annual Research Conference Proceedings Volume 2014 Issue 1. Presented at the Qatar Foundation Annual Research Conference Proceedings, Hamad Bin Khalifa. University Press (HBKU Press), Qatar National Convention Center (QNCC), Doha, Qatar. https://doi.org/10.5339/qfarc.2014.EEPP0079.

Wang, H., Liu, Z., Wu, K., Qiu, J., Zhang, Y., Ye, B., He, M., 2022. Impact of urbanization on meteorology and air quality in chengdu, a basin city of southwestern China. Front. Ecol. Evol. 10, 845801 https://doi.org/10.3389/fevo.2022.845801.

Wang, H., Liu, Z., Zhang, Y., Yu, Z., Chen, C., 2021. Impact of different urban canopy models on air quality simulation in Chengdu, southwestern China. Atmos. Environ. 267, 118775 https://doi.org/10.1016/j.atmosenv.2021.118775.

Wang, K., Gao, C., Wu, K., Liu, K., Wang, H., Dan, M., Ji, X., Tong, Q., 2023. ISAT v2.0: an integrated tool for nested-domain configurations and model-ready emission inventories for WRF-AQM. Geosci. Model Dev. (GMD) 16, 1961–1973. https://doi.org/10.5194/gmd-16-1961-2023.

Wu, K., Zhu, S., Mac Kinnon, M., Samuelsen, S., 2023. Unexpected deterioration of O3 pollution in the South Coast Air Basin of California: the role of meteorology and emissions. Environ. Pollut. 330, 121728 https://doi.org/10.1016/j.envpol.2023.121728.

Wu, Q.Z., Xu, W.S., Shi, A.J., Li, Y.T., Zhao, X.J., Wang, Z.F., Li, J.X., Wang, L.N., 2014. Air quality forecast of PM_{10} in Beijing with community multi-scale air quality modeling (CMAQ) system: emission and improvement. Geosci. Model Dev. (GMD) 7, 2243–2259. https://doi.org/10.5194/gmd-7-2243-2014.

Xie, Y., Paulot, F., Carter, W.P.L., Nolte, C.G., Luecken, D.J., Hutzell, W.T., Wennberg, P. O., Cohen, R.C., Pinder, R.W., 2013. Understanding the impact of recent advances in isoprene photooxidation on simulations of regional air quality. Atmos. Chem. Phys. 13, 8439–8455. https://doi.org/10.5194/acp-13-8439-2013.

Yang, J., Huang, X., 2021. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. Earth Syst. Sci. Data 13, 3907–3925. https://doi.org/10.5194/essd-13-3907-2021.

Yang, X., Wu, K., Wang, H., Liu, Y., Gu, S., Lu, Y., Zhang, X., Hu, Y., Ou, Y., Wang, S., Wang, Z., 2020. Summertime ozone pollution in Sichuan Basin, China: meteorological conditions, sources and process analysis. Atmos. Environ. 226, 117392 https://doi.org/10.1016/j.atmosenv.2020.117392.

Yang, Y., Zhao, Y., Zhang, L., Zhang, J., Huang, X., Zhao, X., Zhang, Y., Xi, M., Lu, Y., 2021. Improvement of the satellite-derived NOx emissions on air quality modeling and its effect on ozone and secondary inorganic aerosol formation in the Yangtze River Delta, China. Atmos. Chem. Phys. 21, 1191–1209. https://doi.org/10.5194/acp-21-1191-2021.

Yuan, B., Shao, M., Lu, S., Wang, B., 2010. Source profiles of volatile organic compounds associated with solvent use in Beijing, China. Atmos. Environ. 44, 1919–1926. https://doi.org/10.1016/j.atmosenv.2010.02.014.

Zeng, X., Kong, S., Zhang, Q., Ren, H., Liu, J., Feng, Y., Yan, Q., Qin, S., Zheng, S., Yao, L., Fan, Z., Zhang, Y., Liu, X., Yan, Y., Zhu, K., Ding, F., Liu, W., Liu, D., Qi, S., Fu, P., 2021. Source profiles and emission factors of organic and inorganic species in fine particles emitted from the ultra-low emission power plant and typical industries. Sci. Total Environ. 789, 147966 https://doi.org/10.1016/j.scitotenv.2021.147966.

Zhang, L., Chen, Y., Zhao, Y., Henze, D.K., Zhu, L., Song, Y., Paulot, F., Liu, X., Pan, Y., Lin, Y., Huang, B., 2018. Agricultural ammonia emissions in China: reconciling bottom-up and top-down estimates. Atmos. Chem. Phys. 18, 339–355. https://doi.org/10.5194/acp-18-339-2018.

Zhao, Y., Xi, M., Zhang, Q., Dong, Z., Ma, M., Zhou, K., Xu, W., Xing, J., Zheng, B., Wen, Z., Liu, X., Nielsen, C.P., Liu, Y., Pan, Y., Zhang, L., 2022. Decline in bulk deposition of air pollutants in China lags behind reductions in emissions. Nat. Geosci. 15, 190–195. https://doi.org/10.1038/s41561-022-00899-1.

Zheng, B., Cheng, J., Geng, G., Wang, X., Li, M., Shi, Q., Qi, J., Lei, Y., Zhang, Q., He, K., 2021. Mapping anthropogenic emissions in China at 1 km spatial resolution and its application in air quality modeling. Sci. Bull. 66, 612–620. https://doi.org/10.1016/j.scib.2020.12.008.

Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., Li, H., Li, X., Peng, L., Qi, J., Yan, L., Zhang, Y., Zhao, H., Zheng, Y., He, K., Zhang, Q., 2018. Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions. Atmos. Chem. Phys. 18, 14095–14111. https://doi.org/10.5194/acp-18-14095-2018.

Zheng, H., Chang, X., Wang, S., Li, S., Zhao, B., Dong, Z., Ding, D., Jiang, Y., Huang, G., Huang, C., An, J., Zhou, M., Qiao, L., Xing, J., 2023. Sources of organic aerosol in China from 2005 to 2019: a modeling analysis. Environ. Sci. Technol. 57, 5957–5966. https://doi.org/10.1021/acs.est.2c08315.

Zheng, Y., Xue, T., Zhang, Q., Geng, G., Tong, D., Li, X., He, K., 2017. Air quality improvements and health benefits from China's clean air action since 2013. Environ. Res. Lett. 12, 114020 https://doi.org/10.1088/1748-9326/aa8a32.

Zhou, Y., Zhao, Y., Mao, P., Zhang, Q., Zhang, J., Qiu, L., Yang, Y., 2017. Development of a high-resolution emission inventory and its evaluation and application through air quality modeling for Jiangsu Province, China. Atmos. Chem. Phys. 17, 211–233. https://doi.org/10.5194/acp-17-211-2017.