Development of a high-resolution integrated emission inventory of air pollutants for China

Nana Wu¹, Guannan Geng²,³, Ruochong Xu¹, Shigan Liu¹, Xiaodong Liu², Qinren Shi², Ying Zhou⁴, Yu Zhao⁵, Huan Liu²,³, Yu Song⁶, Junyu Zheng⁷, Qiang Zhang¹, and Kebin He²,⁸

¹Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China
²State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China
³State Environmental Protection Key Laboratory of Sources and Control of Air Pollution Complex, Beijing 100084, China
⁴Key Laboratory of Beijing on Regional Air Pollution Control, Faculty of Environment and Life, Beijing University of Technology, Beijing 100124, China
⁵State Key Laboratory of Pollution Control and Resource Reuse and School of the Environment, Nanjing University, 163 Xianlin Rd., Nanjing, Jiangsu 210023, China
⁶State Key Joint Laboratory of Environmental Simulation and Pollution Control, College of Environmental Sciences and Engineering, Peking University, Beijing 100871, China
⁷Sustainable Energy and Environmental Thrust, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511458, China
⁸Institute for Carbon Neutrality, Tsinghua University, Beijing 100084, China

Correspondence: Guannan Geng (guannangeng@tsinghua.edu.cn)

Received: 7 January 2024 – Discussion started: 26 January 2024
Revised: 15 April 2024 – Accepted: 19 April 2024 – Published: 21 June 2024

Abstract. Constructing a highly resolved comprehensive emission dataset for China is challenging due to limited availability of refined information for parameters in a unified bottom-up framework. Here, by developing an integrated modeling framework, we harmonized multi-source heterogeneous data, including several up-to-date emission inventories at national and regional scales and for key species and sources in China to generate a 0.1° resolution inventory for 2017. By source mapping, species mapping, temporal disaggregation, spatial allocation, and spatial–temporal coupling, different emission inventories are normalized in terms of source categories, chemical species, and spatiotemporal resolutions. This achieves the coupling of multi-scale, high-resolution emission inventories with the Multi-resolution Emission Inventory for China (MEIC), forming the high-resolution INTegrated emission inventory of Air pollutants for China (INTAC). We find that INTAC provides more accurate representations for emission magnitudes and spatiotemporal patterns. In 2017, China’s emissions of sulfur dioxide (SO₂), nitrous oxides (NOₓ), carbon monoxide (CO), non-methane volatile organic compounds (NMVOCs), ammonia (NH₃), PM_{10} and PM_{2.5} (particulate matter), black carbon (BC), and organic carbon (OC) were 12.3, 24.5, 141.0, 27.9, 9.2, 11.1, 8.4, 1.3, and 2.2 Tg, respectively. The proportion of point source emissions for SO₂, PM_{10}, NOₓ, and PM_{2.5} increases from 7%–19% in MEIC to 48%–66% in INTAC, resulting in improved spatial accuracy, especially mitigating overestimations in densely populated areas. Compared with MEIC, INTAC reduces mean biases in simulated concentrations of major air pollutants by 2–14 μg m⁻³ across 74 cities, compared against ground observations. The enhanced model performance by INTAC is particularly evident at finer-grid resolutions. Our new dataset is accessible at http://meicmodel.org.cn/intac (last access: 15 April 2024) and https://doi.org/10.5281/zenodo.10459198 (Wu et al., 2024), and it will provide a solid data foundation for fine-scale atmospheric research and air-quality improvement.
1 Introduction

In recent years China has achieved remarkable progress in improving air quality and public health through the active implementation of clean-air policies (Liu et al., 2020; Xiao et al., 2022; Zhang and Geng, 2019; Q. Zhang et al., 2019). To further unlock the potential of targeted clean-air actions, there is an urgent need for an accurate and detailed depiction of emissions encompassing their magnitudes and spatial–temporal patterns. Developing a reliable, highly resolved emission inventory for China is also crucial for studies of atmospheric chemistry and climate change (Cheng et al., 2021a; Geng et al., 2021; Q. Zhang et al., 2019).

The construction of high-resolution emission inventories for China poses significant challenges due to the diversity and complexity of emission sources and technology distributions. Additionally, the limited availability of localized measurements for emission factors (EFs) and source profiles, along with the exact locations of the emission facilities, further compounds the difficulties (Li et al., 2017a). The widely used bottom-up approach involves the establishment of a unified framework that encompasses source categories, chemical speciation processes, spatial–temporal allocation profiles, and emission estimation methods (An et al., 2021; Huang et al., 2021). However, achieving both wide coverage and high accuracy in compiling an emission inventory for China through this approach remains a formidable task for individual research institutions.

Comprehensive national-scale emission inventories developed using the unified framework typically provide extensive coverage of space, species, and sectors (Li et al., 2017a; S. Li et al., 2023) but tend to exhibit limitations in spatial accuracy (Wu et al., 2021; Zhao et al., 2015; Zheng et al., 2021; Zhou et al., 2017b). Previous studies have indicated that the spatial allocation in large-scale emission inventories relies on spatial proxies (e.g., population and road networks), rather than latitude–longitude coordinates of emission sources, due to the unavailability of extensive spatial information (Li et al., 2017b; Zhang et al., 2009). The assumption of a linear correlation between emissions and spatial proxies might lead to an overestimation of emissions in urban areas especially at resolutions finer than 0.25° (Wu et al., 2021; Zheng et al., 2021, 2017). Biases introduced by the proxy-based method are found to be propagated as the grid size diminishes, resulting in uncertainties for chemical transport models (CTMs) (Zheng et al., 2021, 2017).

Emission inventories focused on a specific region (An et al., 2021; Huang et al., 2021; Liu et al., 2018), sector (Chen et al., 2016; Deng et al., 2020; Zhou et al., 2017a), or key species (Huang et al., 2012b; Li et al., 2021; Wang et al., 2023) under the aforementioned unified framework demonstrate enhanced accuracy but fail to achieve comprehensive coverage. These inventories assimilate substantial detailed foundational data from various statistical datasets, on-site measurements, or surveys to represent real-world emission magnitudes, including energy consumption, removal efficiencies, and localized speciation profile (An et al., 2021; Huang et al., 2021; Liu et al., 2018). Innovative data, such as measurements from continuous emission monitoring systems (Bo et al., 2021; Tang et al., 2023; Wu et al., 2022), or methodologies like process-based models (Kang et al., 2016; Zhao et al., 2020), are implemented to enable a more accurate characterization of complex emission dynamics. Facility-level geographical location is incorporated to optimize the representation of spatial patterns (F. Liu et al., 2015; Wang et al., 2019; Wu et al., 2023). The reliability of these local-scale, sector-specific, or species-specified inventories has been validated against satellite and ground-based measurements (F. Liu et al., 2016; Zhang et al., 2021; Zheng et al., 2019).

The other strategy for developing bottom-up emission inventories is commonly known as the integrated method. This method consolidates multiple emission datasets for specific regions, species, or sectors into a unified product ensuring extensive representation (Li et al., 2017b). Taking advantage of existing inventories derived from localized data and advanced methods, the integrated method facilitates the efficient generation of highly resolved emission inventories at large scales. However, the heterogeneity of different emission datasets presents challenges for the fusion, which are manifested in diverse data formats, sector categories, species, and spatial–temporal resolution. In recent years, there has been growing interest in adopting the integrated approach to enrich inventories with local insights, particularly at the global (Crippa et al., 2023; Janssens-Maenhout et al., 2015) and Asian scales (Kurokawa et al., 2013; M. Li et al., 2024; Li et al., 2017b; Zhang et al., 2009). Research on establishing integrated inventories for China is constrained due to the inherent complexity and challenging accessibility of the data. These efforts are concentrated in specific regions such as the Yangtze River Delta (YRD) (An et al., 2021).

In this work, with the support of several research institutions, we use an emission integration model to construct a high-resolution integrated emission inventory at a spatial resolution of 0.1° for China in 2017, denoted as INTAC (INTEGRated emission inventory of Air pollutants for China). The challenges associated with coupling multi-source heterogeneous data are addressed through the implementation of an inventory integration framework. Then, leveraging the strengths of inventories enriched with local knowledge, we compile a comprehensive, highly resolved emission product to enhance the accurate representation of emissions from crucial regions, sectors, and species. Finally, the improved accuracy of emission magnitude and spatial distribution is evaluated using atmospheric chemistry models.
Figure 1. Methodology framework of INTAC development.

2 Methodology and data

Figure 1 schematically illustrates the integration process of INTAC. We collect seven emission inventories: the Multi-resolution Emission Inventory for China (MEIC) developed by Tsinghua University (Li et al., 2017a; Zheng et al., 2018), the industrial point source emission inventory for China by the MEIC team (Zheng et al., 2021; Zheng et al., 2017), the YRD air pollutant emission inventory led by Nanjing University (An et al., 2021; Zhou et al., 2017b), the Pearl River Delta (PRD) emission inventory by Jinan University (Huang et al., 2021; Sha et al., 2021), the open biomass burning emission inventory in China by Peking University (Huang et al., 2012a; M. Liu et al., 2015; Song et al., 2009; Yin et al., 2019), the shipping emission inventory in East Asia by Tsinghua University (H. Liu et al., 2016; Liu et al., 2019), and the high-resolution ammonia emission inventory in China (PKU-NH$_3$) by Peking University (Huang et al., 2012b; Kang et al., 2016). The details of these inventories and the rationale for choosing them will be described in Sect. 2.1.

An integration model is then established to merge emission inventories which have different sectors, species, spatial–temporal resolutions, and formats (i.e., point, area, and gridded forms). The integration process consists of five steps: source mapping, species mapping, temporal disaggregation, spatial allocation, and spatial–temporal coupling, as detailed in Sect. 2.2. Based on the priority order, multi-source emission inventories are assembled at the standardized species, sector, and grid levels, yielding a standardized data cube. Ultimately, the integrated emission inventory INTAC was created for China, featuring a resolution of 0.1° on a monthly scale and covering nine air pollutants: sulfur dioxide (SO$_2$), nitrogen oxides (NO$_x$), carbon monoxide (CO), non-methane volatile organic compounds (NMVOCs), ammonia (NH$_3$), PM$_{10}$ and PM$_{2.5}$ (particulate matter), black carbon (BC), and organic carbon (OC).

2.1 Components of the integrated emission inventory INTAC

Table 1 lists the essential details about the seven inventories and priority order utilized for integration. Given MEIC’s extensive coverage across species, sectors, and spatial domains, it functions as the default inventory in our integration, supplementing the missing data in other inventories. The remaining six inventories can be categorized into three types in sequence: point-source-based inventory (ranked sixth), regional inventories (ranked fifth and fourth), and process-based inventories (ranked third to first). The point-source-based inventory can directly correct the spatial misallocation of industrial emissions in MEIC at fine resolutions (Zheng et al., 2021, 2017). The regional inventories further enhance local investigations of individual emission sources and simultaneously refine estimation methods for mobile and area sources (Gu et al., 2023; Zhao et al., 2018; Zhou et al., 2017b). Process-based inventories typically adopt advanced methods to improve the characterization for emission processes and parameters specific to particular sectors or species, thereby providing emission totals and distributions that are more in line with measurements (Huang et al., 2012a, 2012b; Kang et al., 2016; H. Liu et al., 2016; Liu et al., 2019; M. Liu et al., 2015; Song et al., 2009; Yin et al., 2019).

2.1.1 MEIC

The integrated inventory INTAC is built upon MEIC, a comprehensive database with extensive coverage across time pe-
Among the seven inventories, MEIC has the lowest priority and is only considered when the other six cannot provide the necessary emissions for a specific city and source.

### 2.1.2 The industrial point source emission inventory for China

The proxy-based method used for spatial allocation in MEIC introduces biases in emission mapping, especially at the kilometer scale (Zheng et al., 2021, 2017). To significantly reduce the uncertainty, we merged an industrial emission inventory with detailed information on ~ 100 000 facilities into INTAC.

Compiled by the MEIC team at Tsinghua University for the year 2013 (Zheng et al., 2021) and updated by them for 2017, this point-based inventory combines three databases investigated under the guidance of the Chinese government, offering a comprehensive overview of industrial facilities. It includes details on the locations, activity rates, production technology, end-of-pipe pollution control devices, and other parameters. It is worth noting that the facility-level activity data were corrected using provincial activity data from MEIC as a total constraint to ensure consistency with national totals from statistics (Zheng et al., 2021). The facility-level, technology-based approach allows for dynamic tracking of emission fluctuations resulting from technological advancements and tightening emission regulations. Crucially, the use of facility geolocations, rather than relying on spatial proxies like urban population, enables the derivation of gridded industrial data at a resolution of ~ 1 km. This approach sig-
The regional emission inventories within the PRD region have demonstrated enhanced reliability compared with previous studies (Huang et al., 2021; Sha et al., 2021; Zheng et al., 2012). The PRD emission inventory developed by Jinan University captures spatial and temporal variations within the PRD region under emission control policies, serving as a foundation for supporting air-quality modeling (Huang et al., 2021; Sha et al., 2021).

The PRD inventory exhibits notable accuracy improvements achieved by means of big-data-driven estimation methods, updated spatial-temporal allocations, and localized NMVOC speciation profiles. Gridded hourly open biomass burning emissions are quantified by fusing the fire radiative power data from three satellites, and hourly shipping emissions are estimated using high-frequency automatic identification system (AIS) records. A total of 31 monthly profiles and 10 spatial proxies are updated to reflect spatial-temporal patterns of emissions influenced by economic growth and energy consumption structural adjustment. Approximately 90% of industrial emissions are disaggregated using exact locations, and novel proxies (e.g., farmland production potential) have been developed for several sectors. The NMVOC speciation is carried out through extensive localized measurements and literature reviews, manifested as a collection of 480 NMVOC source profiles across eight sectors and 380 species. The species relevant to the SAPRC-07 chemical mechanism are collected in this work. Additionally, the inventory encompasses 800 source categories, placing particular emphasis on incorporating new sectors relevant to VOC emissions. Activity rates are improved by utilizing extensive field surveys and data mining efforts involving investigations of production data for 10,000 industrial plants and the gathering of activity-relevant information for 50 million vehicles. Emission factors that reflect the local context are obtained or revised based on source measurements and the latest research findings. These updates help mitigate uncertainties in emission estimates for the PRD region. The PRD inventory is initially collected at a monthly resolution and a spatial resolution of 0.05°, with detailed spatial-temporal allocation proxies as outlined in Huang et al. (2021).

2.1.5 The open biomass burning emission inventory in China

As a significant source of CO₂, BC, OC, and other pollutants, open biomass burning profoundly influences air quality, climate change, and human health (Reisen et al., 2013). A case study in summer 2011 for the YRD region revealed that during a severe haze episode, open biomass burning contributed to 37%, 70%, and 61% of PM$_{2.5}$, OC, and elemental carbon (EC) emissions, respectively (Cheng et al., 2014). To address the absence of this source in MEIC, we integrate a high-resolution open biomass burning emission inventory by

2.1.3 The YRD air pollutant emission inventory

Regional emission inventories within YRD provide a more accurate representation of emissions compared with the national-scale MEIC, as proven by ground and satellite observations (Yang and Zhao, 2019; Zhang et al., 2021; Zhao et al., 2017a, 2018, 2020; Zhou et al., 2017b). This improvement is attributed to the avoidance of outdated or non-localized emission calculation parameters commonly present in large-scale inventories like MEIC. Here, we merge the 2017 YRD air pollutant emission inventory into INTAC to achieve state-of-the-art estimates for rapidly changing emissions over this core area (An et al., 2021; Gu et al., 2023; Zhou et al., 2017b).

Localized field surveys and measurements greatly enhance the reliability of calculation parameters within the YRD inventory. Highly resolved emissions for the power sector are acquired through on-site monitoring with high temporal resolution (Y. Zhang et al., 2019), rather than relying on static and outdated average emission factors. Facility-level information (e.g., the removal efficiencies) obtained from local investigation and a segment-based industrial process method enhances the understanding of both the quantity and spatial patterns of industrial emissions. Considering meteorological factors and land use conditions during agricultural processes results in more accurate seasonal and spatial distributions of NH$_3$ emissions. (Zhao et al., 2020). An investigation of in-use machinery is conducted to capture the seasonal emission patterns from off-road machines (Zhang et al., 2020). Real-world surveys are performed to determine grain straw ratios and household burning proportions, facilitating the quantification of emissions from biomass-fueled stoves. The PM$_{2.5}$ and NMVOC speciation profiles are updated based on multi-instrument sampling and analysis in both current and previous studies (Huang et al., 2018; Zhao et al., 2017a), satisfying the need for simulating PM$_{2.5}$ chemical components and O$_3$. The YRD inventory is collected with a spatial resolution of 0.1° and an annually temporal resolution in this study. Only CB05 volatile organic compound (VOC) species are collected.

2.1.4 The PRD emission inventory

The regional emission inventories within the PRD region significantly avoid misallocating emissions from rural to urban areas at fine grids, as supported by previous studies demonstrating its effectiveness in mitigating simulated biases in air pollutant concentrations within densely populated regions (e.g., Zheng et al., 2021). For temporal variations it employs the same monthly profiles as MEIC, including the production of various industrial goods or gross domestic product (GDP), as outlined in Li et al. (2017b). The NMVOC speciation also aligns with the MEIC model. This inventory takes priority over MEIC, indicating that only few industrial sources not covered in this inventory are substituted with MEIC.
Peking University into INTAC (Huang et al., 2012a; M. Liu et al., 2015; Song et al., 2009; Yin et al., 2019). The inventory applies satellite observations to tackle considerable uncertainties associated with provincial statistical data and overcome the coarse resolution found in previous studies (Ni et al., 2015). The estimation of biomass consumption in the inventory is based on the fire radiative energy (FRE) approach, which depends on the energy emitted by fires. This approach helps reduce the biases introduced by burned-area algorithms, especially for small-scale fires. The inventory utilizes the high-spatial-resolution land cover dataset GlobeLand30 derived from multi-spectral images to classify biomass fuel types. Eventually, daily emissions from forest, grassland, cropland, and shrubland are calculated at a 1 km resolution. The reasonableness is validated by comparing with other datasets such as the fourth version of the Global Fire Emissions Database (GFED). The initially collected inventory lacks model-ready VOC species.

2.1.6 The shipping emission inventory in East Asia

In recent years, maritime trade in the East Asian region has significantly increased (United Nations Conference on Trade and Development (UNCTAD), 2014), resulting in a surge in shipping emissions with substantial impacts on air quality and climate. Previous studies have indicated that East Asian shipping emissions accounted for 16% of the global total in 2013. Shipping emissions made a growing contribution to the rise in annual mean PM$_{2.5}$ concentrations, reaching levels as high as 5.2 µg m$^{-3}$ in 2015 (Lv et al., 2018). To address the omission of this emission source in MEIC, we integrate the shipping emission inventory in East Asia for 2017 into INTAC (H. Liu et al., 2016; Liu et al., 2019).

The inventory introduces an innovative approach based on comprehensive and dynamic ship activity data. A static dataset of approximately 66 000 vessels is compiled as a foundation, using information from Lloyd’s Register and the China Classification Society. This dataset encompasses various ship properties including ship category, hull shape, engine rotational speed, engine capacity, maximum speed capability, and build year. High-quality AIS data are used to capture ship activities, incorporating the Maritime Mobile Service Identification identifier, geographical location, real-time speed, and time-related information. The AIS data are also employed to generate gridded emissions from shipping at a spatial resolution of 0.1°. The inventory enhances our comprehension of regional-level shipping emissions and significantly alleviates biases arising from the misallocation of marine fuels, as observed in global studies (Endresen et al., 2007). The collected shipping inventory provides emissions at an annual resolution for seven species including SO$_2$, NO$_x$, CO, NMVOC, PM$_{2.5}$, BC, and OC.

2.1.7 PKU-NH$_3$

As a prominent alkaline component in the atmosphere, ammonia plays a crucial role in atmospheric chemistry and also terrestrial and aquatic ecosystems through its participation in atmospheric reactions and deposition processes. This study integrates PKU-NH$_3$, a high-resolution ammonia emission inventory for China developed by Peking University. PKU-NH$_3$ is designed to track the evolution of NH$_3$ emissions amid the rapid increase in grain and meat production in China over the past few decades (Huang et al., 2012b; Kang et al., 2016). This inventory offers a better grasp on NH$_3$ emissions in China through the application of a process-based method and more reliable emission factors, in contrast to previous studies (Kurokawa et al., 2013; Li et al., 2017b). Top-down NH$_3$ inversion through satellite observations provides additional validation for the accuracy of PKU-NH$_3$ (Paulot et al., 2014).

Earlier studies of NH$_3$ emissions commonly used fixed EFs, overlooked some ammonia emission sources, and had coarse resolutions (Ohtara et al., 2007; Streets et al., 2003). Unlike previous approaches, PKU-NH$_3$ incorporates dynamic and multifactorial EFs and more comprehensive emissions sources. The determination of emission factors takes into account various parameters related to local conditions and agricultural practices. When estimating NH$_3$ emissions of synthetic fertilizer application, the model considers five types of fertilizers and factors such as soil acidity, ambient temperature, fertilizer application technique and dosage, wind speed, and in situ measurements of NH$_3$ flux. For livestock waste, NH$_3$ emissions are calculated using a mass-flow approach across four phases of manure management, considering variables such as animal rearing types, temperature, and wind speed. In addition, NH$_3$ emissions from other small sources are quantified, including agricultural soil, nitrogen-fixing crops, crop residue compost, excretion of rural populations, open biomass burning, waste disposal, gasoline vehicles, diesel vehicles, and industrial processes. The NH$_3$ emissions are allocated from provinces into 0.1° grids based on spatial proxies such as land cover, rural population, and other relevant indicators. Monthly emission factors shaped by meteorological conditions are used to calculate NH$_3$ emissions from fertilizer application and livestock source at a monthly level.

2.2 The integration of multi-source heterogeneous data

In the integration process, seven heterogeneous inventories are first normalized in terms of emission sources, species, and spatial–temporal resolutions, and then they are integrated following a priority order to produce a standardized, highly resolved data cube.
2.2.1 Source mapping

To merge inventories under a unified emission source classification system, the emission sources in the MEIC model are categorized into 88 standard sectors for mapping (Table S1 in the Supplement). The first-level category comprises 10 subcategories, namely, stationary combustion, industrial process, mobile source, solvent use, agriculture, dust, biomass burning, storage and transportation, waste treatment, and other sources. These subcategories are then further subdivided into 88 second-level sources, which take industrial classification for national economic activities for reference. For example, the industrial process sector encompasses emission sources such as the manufacturing of non-metallic mineral products, manufacturing of chemical fibers, manufacturing of foods, and smelting and pressing of ferrous metals. In the initial step of integration, the sectors in each emission inventory are mapped to the standardized two-level sources.

2.2.2 Species mapping

In the next step, non-methane volatile organic compounds (NMVOCs), particulate matter (PM), and NO₃ in each inventory are converted into model-ready species to support CTMs. The species mapping process is grounded in the inventory are converted into model-ready species to support CTMs. The species mapping process is grounded in the speciation framework to generate model species for the five chemical mechanisms. The seven emission inventories are collected at different temporal resolutions (Table 1) and need to be temporally allocated to a unified monthly scale for integration. Monthly emissions from the open biomass burning inventory are aggregated to monthly scales through summation. For annual inventories (e.g., the YRD inventory), sector-specific monthly profiles derived from the MEIC model are used for disaggregation (Li et al., 2017b). For instance, monthly power generation data from the National Bureau of Statistics describe variations in monthly power emissions. Industrial production or the GDP from the National Bureau of Statistics is employed to account for monthly emission fluctuations related to industrial heating, boilers, cement, iron and steel, and other industrial processes. Monthly emission factors calculated by the International Vehicle Emissions model are applied to on-road vehicles. Considering the insignificant monthly variations in automatic identification system data for marine shipping, the annual shipping emissions are uniformly disaggregated across the months.

2.2.3 Temporal disaggregation

The seven emission inventories are collected at different temporal resolutions (Table 1) and need to be temporally allocated to a unified monthly scale for integration. Monthly emissions from the PKU-NH₃, the PRD inventory, the industrial point source inventory, and MEIC can be directly used for data merge. Daily-level emissions from the open inventory, latitude and longitude coordinates are employed to directly position them within grid locations. Area sources in MEIC are allocated to grids using spatial proxies within the MEIC model (Li et al., 2017b). For instance, industrial sources are assigned to grids based on urban population (Schneider et al., 2009). The road network (Zheng et al., 2014) serves as a proxy for disaggregating emissions of on-road vehicles, while rural population (Schneider et al., 2009) is used as the proxy for fertilizer and livestock sources. It is important to mention that uncertainties may arise at city borders if emissions from adjacent cities come from different inventories during the integration process. To mitigate biases introduced by border issues, all emissions at 0.1° resolution are first uniformly downscaled to 1 km in preparation for the spatial–temporal coupling process, and then they are re-gridded back to 0.1° for the final product.

2.2.4 Spatial allocation

The seven inventories are in different data formats, including point source and gridded formats at varying resolutions, necessitating spatial harmonization for integration. Although the industrial point source inventory and the open biomass burning inventory can accurately pinpoint the specific geographical locations of emission sources, the other five inventories rely on numerous spatial proxies to disaggregate emissions into grids, which inevitably introduce uncertainties at very fine resolutions. Therefore, we re-grid the final product to 0.1° to ensure high-level spatial accuracy. Gridded emissions finer than 0.1° resolution are aggregated to 0.1°, which is performed in the open biomass burning inventory and the PRD inventory. For the industrial point source inventory, latitude and longitude coordinates are employed to directly position them within grid locations. Area sources in MEIC are allocated to grids using spatial proxies within the MEIC model (Li et al., 2017b). For instance, industrial sources are assigned to grids based on urban population (Schneider et al., 2009). The road network (Zheng et al., 2014) serves as a proxy for disaggregating emissions of on-road vehicles, while rural population (Schneider et al., 2009) is used as the proxy for fertilizer and livestock sources. It is important to mention that uncertainties may arise at city borders if emissions from adjacent cities come from different inventories during the integration process. To mitigate biases introduced by border issues, all emissions at 0.1° resolution are first uniformly downscaled to 1 km in preparation for the spatial–temporal coupling process, and then they are re-gridded back to 0.1° for the final product.

2.2.5 Spatial–temporal coupling

Finally, following the procedures outlined in Sects. 2.2.1–2.2.4, all inventories are preprocessed to a standardized format encompassing 88 sectors, various species, a spatial resolution of 1 km, and a monthly temporal resolution. This pre-
The integration is carried out at source-by-source, speciesby-species, and grid-by-grid levels, with the process guided by the priority order of each inventory (Table 1). MEIC serves as the default inventory in our integration, offering extensive spatial and species coverage along with spatial proxies, temporal profiles, and NMVOC speciation methods within the model. The remaining six emission inventories are assigned a predefined priority order. The industrial point source emission inventory for China takes precedence over industrial emissions in MEIC, substituting proxy-based spatial allocation with precise geographical coordinates. This extends the applicability of MEIC from a resolution greater than 0.25° to finer scale (Zheng et al., 2021, 2017). To achieve fine-grained emission characterization in critical areas, the YRD and PRD PM2.5 emission inventories enriched with localized data and advanced methods are incorporated to update emissions in these areas. While MEIC comprehensively estimates emissions for ∼800 source categories in China, there may still be omissions for certain emission sources. The inclusion of inventories for open biomass burning and East Asian shipping helps partially fill this gap. PKU-NH3, generated by a process-based model to provide a comprehensive understanding of China’s NH3 sources, is utilized to replace all NH3 emissions in other inventories. The prioritization is performed city by city. For emissions of a particular species from a specific emission sector, when multiple inventories overlap in city grids, the estimates from the highest-priority inventory are selected as the final emissions. Through this step the integrated inventories are developed based on the configured output settings such as map projection and spatial–temporal attributes.

2.3 Evaluation of the emission inventory using WRF/CMAQ model

We apply Weather Research and Forecasting version 3.9 (WRFv3.9) and Community Multiscale Air Quality version 5.2 (CMAQ5.2) as the air-quality simulation systems. Two nested simulation domains with horizontal resolutions of 36 and 12 km are used (Fig. S1 in the Supplement). The mother domain (172 × 127 cells) covers the entire Chinese mainland and parts of the neighboring countries, and the nested domain (226 × 241 cells) includes the heavily polluted eastern China. Some 4-month (January, April, July, and October) simulations in 2017 are carried out with a 7 d spin-up period preceding each month. The vertical resolution in WRF is set with 45 sigma levels ranging from the surface up to 100 hPa. Subsequently, it is collapsed into 28 layers through the Meteorology–Chemistry Interface Processor (MCIP) before being input into CMAQ.

The configurations of the WRF and CMAQ models in this study follow Cheng et al. (2019). The meteorological initial and boundary conditions for the simulation are provided by the final reanalysis data from the National Centers for Environmental Prediction (NCEP-FNL; https://rda.ucar.edu/datasets/ds083.2/; last access: 15 December 2022). The schemes for shortwave radiation, longwave radiation, land surface processes, boundary layer, cumulus parameterization, and cloud microphysics are selected as the New Goddard scheme (Chou et al., 1998), RRTM scheme (Mlawer et al., 1997), Pleim–Xiu surface layer scheme (Xiu and Pleim, 2001), ACM2 PLB scheme (Pleim, 2007), Kain–Fritsch scheme (Kain, 2004), and WSM6 scheme (Hong and Lim, 2006), respectively. Observational nudging and soil nudging are employed to enhance the meteorological simulation. Regarding the CMAQ model, the chemical mechanisms for gas-phase, aqueous-phase, and aerosol are configured as CB05, the regional acid deposition model (RADM), and AERO6, respectively. Photolysis rates are calculated online using the simulated aerosols and ozone concentrations. Anthropogenic emissions outside China are taken from the MIX inventory (Li et al., 2017b). The integrated inventories INTAC and MEIC are used for comparison within China. Biogenic emissions are calculated using the Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGANv2.1), while dust and lightning emissions are not considered in this study.

The performances of WRF for the meteorological parameters are evaluated against the Integrated Surface Database (ISD) from the National Climatic Data Center (NCDC; ftp://ftp.ncdc.noaa.gov/pub/data/noaa/). Evaluation metrics include the correlation coefficient (R), mean bias (MB), root mean square error (RMSE), normalized mean bias (NMB), and normalized mean error (NME). Table S2 demonstrates good agreement between WRF model results and ground-level observations. Similar configurations also have been validated in previous studies (Cheng et al., 2019, 2021a, b). CMAQ modeling performances are assessed using hourly observed concentrations of air pollutants obtained from the China National Environmental Monitoring Center (http://www.cnemc.cn/, last access: 15 December 2022).

3 Results

3.1 China’s emission characteristics in 2017

We use the integrated emission inventory to analyze air pollutant emissions in China for 2017. Major air pollutant emissions are estimated as follows: 12.3 Tg SO2, 24.5 Tg NOx, 141.0 Tg CO, 27.9 Tg NMVOCs, 9.2 Tg NH3, 11.1 Tg PM10, 8.4 Tg PM2.5, 1.3 Tg BC, and 2.2 Tg OC. The emission data, organized into power, industry, residential, transportation, agriculture, solvent use, shipping, and open biomass burning sectors, are available for download at https://doi.org/10.5281/zenodo.10459198 (Wu et al., 2024) and http://meicmodel.org.cn/intac. The following sections will characterize emissions in detail across sectors, fuel types, and spatial distributions.
Table 2. Anthropogenic emissions (in Gg) of air pollutants by sector in the 2017 INTAC for China. The shipping sector includes inland waterway sources and the marine vessels.

<table>
<thead>
<tr>
<th>Sector</th>
<th>SO$_2$</th>
<th>NO$_x$</th>
<th>CO</th>
<th>NMVOCs</th>
<th>NH$_3$</th>
<th>PM$_{10}$</th>
<th>PM$_{2.5}$</th>
<th>BC</th>
<th>OC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power</td>
<td>1822</td>
<td>3790</td>
<td>4909</td>
<td>152</td>
<td>14</td>
<td>981</td>
<td>568</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Industry</td>
<td>6066</td>
<td>8800</td>
<td>5282</td>
<td>8824</td>
<td>249</td>
<td>5603</td>
<td>3620</td>
<td>308</td>
<td>285</td>
</tr>
<tr>
<td>Residential</td>
<td>2361</td>
<td>861</td>
<td>5589</td>
<td>3676</td>
<td>629</td>
<td>3516</td>
<td>3088</td>
<td>606</td>
<td>1649</td>
</tr>
<tr>
<td>Transportation</td>
<td>341</td>
<td>7751</td>
<td>22597</td>
<td>4123</td>
<td>619</td>
<td>533</td>
<td>493</td>
<td>257</td>
<td>95</td>
</tr>
<tr>
<td>Agriculture</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7609</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Solvent</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9255</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Shipping</td>
<td>1642</td>
<td>3077</td>
<td>391</td>
<td>191</td>
<td>2</td>
<td>73</td>
<td>264</td>
<td>43</td>
<td>49</td>
</tr>
<tr>
<td>Open biomass burning</td>
<td>21</td>
<td>215</td>
<td>4403</td>
<td>1659</td>
<td>76</td>
<td>409</td>
<td>355</td>
<td>35</td>
<td>167</td>
</tr>
<tr>
<td>Total</td>
<td>12253</td>
<td>24494</td>
<td>141023</td>
<td>27881</td>
<td>9198</td>
<td>11117</td>
<td>8388</td>
<td>1255</td>
<td>2245</td>
</tr>
</tbody>
</table>

Figure 2. Sector-specific distributions of emissions in the 2017 INTAC for China. Panels (a)–(c) represent the sectoral contributions for SO$_2$, NO$_x$, NMVOCs, and PM$_{2.5}$, respectively. The figure only displays the top eight contributing sources, while sources excluding these are categorized as “other sources”.

3.1.1 By sector

Table 2 displays emissions specific to power, industry, residential, transportation, agriculture, solvent use, shipping, and open biomass burning sectors in INTAC. For pollutants primarily originating from fuel combustion and industrial processes (e.g., SO$_2$, NO$_x$, CO, PM$_{10}$, and PM$_{2.5}$), the power, industry, and transportation sources collectively contribute significantly to their emissions, ranging from 56% to 83%. Industrial sources take a leading role in various atmospheric pollutants, contributing more than 30% for SO$_2$, NO$_x$, CO, NMVOC, PM$_{10}$, and PM$_{2.5}$ emissions. Due to low combustion efficiency and a lack of emission control measures, residential sources exhibit a high emission factor for products of incomplete combustion, leading to 40% of CO, 48% of BC, and 73% of OC emissions. Solvent sources exclusively produce NMVOC emissions, contributing 33% to the overall emissions. The complexity of VOC emission origins is evident in the diverse range of contributing sources. Agricultural sources dominate NH$_3$ emissions, comprising an 83% share of total emissions. As described in Sect. 2.1.7, PKU-NH$_3$ incorporates a wide variety of NH$_3$ sources, providing a more comprehensive understanding of the sectors contributing to NH$_3$ emissions. Insignificant sources may exert a large influence in specific regions or periods, such as during large wildfires or in cities with heavy traffic. Additionally, the contribution of the supplemented open biomass burning source, especially for OC (7%) and NMVOCs (6%), cannot be overlooked.

Figure 2 consolidates 88 standardized emission sources into 25 categories, allowing for a more detailed analysis of...
sectoral emission patterns compared with Table 2. Owing to substantial coal use in industrial and power sectors, along with sulfur-rich ship fuels, prominent contributors to \( \text{SO}_2 \) emissions include power, shipping, stationary combustion, and manufacture of non-metallic mineral product sources, accounting for 15 %, 13 %, 12 %, and 12 %, respectively, to total \( \text{SO}_2 \) emissions. This indicates that achieving further reductions in \( \text{SO}_2 \) emissions requires the implementation of more energy-efficient end-of-pipe control measures and adoption of low-sulfur fuels. The dominant origins of \( \text{NO}_x \) emissions are from the freight truck, power generation, and shipping sectors, representing 21 %, 15 %, and 13 %, respectively, of the total emissions. Both freight trucks and vessels extensively use compression ignition engines, which are prone to generating \( \text{NO}_x \) emissions under high-temperature and oxygen-rich conditions. Implementing strict vehicle standards is crucial to effectively reduce \( \text{NO}_x \) emissions from exhaust gases. Coatings, other industrial processes, and passenger vehicle sources together account for 51 % of anthropogenic NMVOC emissions. The major contributors to primary \( \text{PM}_{2.5} \) emissions include biomass fuel, the manufacture of non-metallic mineral products, and the smelting and pressing of ferrous metal sources, making up 22 %, 17 %, and 10 %, respectively, of the total emissions. It is noteworthy that the use of biomass fuels (e.g., rice straw and firewood) for cooking or heating in rural areas results in considerable \( \text{PM}_{2.5} \) emissions especially in provinces like Sichuan, Anhui, Shandong, and Heilongjiang.

### 3.1.2 By fuel type

Figure 3 illustrates the proportions of major air pollutant emissions in 2017 for each fuel type. Fossil fuel combustion significantly dominates the emissions of \( \text{PM}_{10} \), \( \text{PM}_{2.5} \), \( \text{CO} \), \( \text{BC} \), \( \text{SO}_2 \), and \( \text{NO}_x \), with proportions ranging from 38 % to 80 %. Coal combustion accounts for 56 % of \( \text{SO}_2 \) emissions, with power, residential activities, and industrial production as the primary emitters. Furthermore, petroleum combustion, mainly from marine vessels, constitutes 20 % of \( \text{SO}_2 \) emissions. For \( \text{NO}_x \) emissions, petroleum combustion contributes 48 % of the total, predominantly arising from freight trucks (5.2 Tg), marine vessels (3.1 Tg), and passenger vehicles (1.0 Tg). Coal combustion processes, such as power and industrial boilers, also result in substantial \( \text{NO}_x \) emissions (31 %). The biomass fuel source causes 53 % of \( \text{OC} \) emissions. Emissions of NMVOCs and \( \text{NH}_3 \) are primarily associated with non-combustion processes.

### 3.1.3 Spatial distribution

We present the gridded emission maps of major air pollutants in Fig. 4. Emissions from anthropogenic sources in China exhibit significant spatial heterogeneity. Due to economic growth and industrial activities, air pollutant emissions are primarily concentrated in the central and eastern regions of China, especially in economically developed urban clusters such as the Beijing–Tianjin–Hebei (BTH) region, the YRD, and the PRD and in regions like Sichuan and Chongqing. These four key areas, as depicted in Fig. S2, collectively account for 25 %, 33 %, 35 %, 37 %, 30 %, 35 %, 33 %, 27 %, and 29 % of the national emissions of \( \text{SO}_2 \), \( \text{NO}_x \), \( \text{CO} \), NMVOCs, \( \text{NH}_3 \), \( \text{PM}_{10} \), \( \text{PM}_{2.5} \), \( \text{BC} \), and \( \text{OC} \), respectively. Moreover, the emission maps at a fine spatial resolution of 0.1° × 0.1° present the local variations in emission patterns, identifying numerous hotspots in small areas and showcasing distinct gradients in emissions. Table 3 shows the provincial-level emissions (except Hong Kong SAR, Macao SAR, and Taiwan), and a map depicting provincial boundaries is displayed in Fig. S2. The emission levels in specific provinces are determined by factors such as resource endowments, industrial structure, energy consumption, and emission control measures. Taking \( \text{SO}_2 \) as an example, the top five provinces are Shanxi, Shandong, Hebei, Guizhou, and Inner Mongolia, collectively accounting for 36 % of the national total \( \text{SO}_2 \) emissions. Guizhou Province, located in the southwest of China, is characterized by high-sulfur coal and a relatively gradual implementation of pollution control measures, which result in elevated \( \text{SO}_2 \) emissions. In the other four provinces, large-scale heavy industries have led to substantial coal consumption and correspondingly higher \( \text{SO}_2 \) emissions. Provinces with a less industry-focused economic structure and lower energy consumption, including Tianjin, Hainan, Qinghai, Beijing, and Tibet, exhibit the low-
Figure 4. Spatial distributions of major air pollutant emissions in the 2017 INTAC for China.

est SO₂ emissions accounting for approximately 2% of the national total.

3.2 Improved accuracy of INTAC for China’s anthropogenic emissions

3.2.1 Comparison of emission magnitudes in INTAC with MEIC across sectors and regions

INTAC improves the representation of anthropogenic air pollutant emissions by incorporating a large number of industrial point sources, integrating high-resolution regional inventories, and supplementing missing emission sources in MEIC. Remarkable differences between INTAC and MEIC are illustrated in Fig. 5 across regions and sectors. Compared with MEIC, INTAC shows higher levels of 16.7%, 11.5%, 10.8%, 11.0%, and 9.1% for SO₂, NOₓ, PM₁₀, PM₂.₅, and OC emissions, respectively. However, it indicates lower levels of 6.3% and 10.6% for NMVOCs and NH₃. CO and BC emissions exhibit good agreement between the two inventories, with differences lower than 3.9%. In comparison with MEIC, the supplementary emission sources in INTAC – specifically open biomass burning and marine shipping – account for the majority of increased emissions, contributing 95%, 89%, and 74% for SO₂, CO, and PM₂.₅, respectively. Additionally, the incorporation of PKU-NH₃ in INTAC leads
Table 3. Anthropogenic emissions (in Gg) of air pollutants by province in the 2017 INTAC for China. Emissions from the shipping emission inventory in East Asia are not included.

<table>
<thead>
<tr>
<th>Province</th>
<th>SO$_2$</th>
<th>NO$_x$</th>
<th>CO</th>
<th>NMVOC</th>
<th>NH$_3$</th>
<th>PM$_{10}$</th>
<th>PM$_{2.5}$</th>
<th>BC</th>
<th>OC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anhui</td>
<td>315</td>
<td>846</td>
<td>5955</td>
<td>1089</td>
<td>341</td>
<td>596</td>
<td>443</td>
<td>49</td>
<td>114</td>
</tr>
<tr>
<td>Beijing</td>
<td>26</td>
<td>231</td>
<td>1394</td>
<td>516</td>
<td>36</td>
<td>61</td>
<td>49</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>Chongqing</td>
<td>396</td>
<td>375</td>
<td>2419</td>
<td>564</td>
<td>150</td>
<td>207</td>
<td>158</td>
<td>22</td>
<td>47</td>
</tr>
<tr>
<td>Fujian</td>
<td>161</td>
<td>530</td>
<td>2344</td>
<td>895</td>
<td>149</td>
<td>202</td>
<td>152</td>
<td>22</td>
<td>49</td>
</tr>
<tr>
<td>Gansu</td>
<td>189</td>
<td>351</td>
<td>2221</td>
<td>358</td>
<td>276</td>
<td>163</td>
<td>126</td>
<td>22</td>
<td>42</td>
</tr>
<tr>
<td>Guangdong</td>
<td>430</td>
<td>1566</td>
<td>6897</td>
<td>1268</td>
<td>351</td>
<td>783</td>
<td>356</td>
<td>67</td>
<td>17</td>
</tr>
<tr>
<td>Guangxi</td>
<td>265</td>
<td>434</td>
<td>3578</td>
<td>808</td>
<td>323</td>
<td>355</td>
<td>275</td>
<td>29</td>
<td>83</td>
</tr>
<tr>
<td>Guizhou</td>
<td>652</td>
<td>355</td>
<td>6629</td>
<td>508</td>
<td>236</td>
<td>459</td>
<td>347</td>
<td>76</td>
<td>125</td>
</tr>
<tr>
<td>Hainan</td>
<td>47</td>
<td>95</td>
<td>584</td>
<td>172</td>
<td>57</td>
<td>46</td>
<td>37</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>Hebei</td>
<td>667</td>
<td>1697</td>
<td>11731</td>
<td>1673</td>
<td>523</td>
<td>708</td>
<td>528</td>
<td>88</td>
<td>125</td>
</tr>
<tr>
<td>Heilongjiang</td>
<td>246</td>
<td>822</td>
<td>7034</td>
<td>1419</td>
<td>379</td>
<td>495</td>
<td>403</td>
<td>65</td>
<td>156</td>
</tr>
<tr>
<td>Henan</td>
<td>367</td>
<td>1256</td>
<td>7962</td>
<td>1500</td>
<td>678</td>
<td>620</td>
<td>459</td>
<td>79</td>
<td>108</td>
</tr>
<tr>
<td>Hubei</td>
<td>513</td>
<td>703</td>
<td>6341</td>
<td>1183</td>
<td>358</td>
<td>455</td>
<td>354</td>
<td>68</td>
<td>118</td>
</tr>
<tr>
<td>Hunan</td>
<td>518</td>
<td>633</td>
<td>6802</td>
<td>953</td>
<td>330</td>
<td>481</td>
<td>363</td>
<td>77</td>
<td>122</td>
</tr>
<tr>
<td>Inner Mongolia</td>
<td>594</td>
<td>1211</td>
<td>5747</td>
<td>831</td>
<td>562</td>
<td>459</td>
<td>340</td>
<td>56</td>
<td>89</td>
</tr>
<tr>
<td>Jiangsu</td>
<td>391</td>
<td>1217</td>
<td>8628</td>
<td>1529</td>
<td>498</td>
<td>667</td>
<td>496</td>
<td>50</td>
<td>105</td>
</tr>
<tr>
<td>Jiangxi</td>
<td>179</td>
<td>449</td>
<td>3676</td>
<td>646</td>
<td>209</td>
<td>273</td>
<td>195</td>
<td>28</td>
<td>52</td>
</tr>
<tr>
<td>Jilin</td>
<td>235</td>
<td>652</td>
<td>3973</td>
<td>847</td>
<td>207</td>
<td>307</td>
<td>237</td>
<td>39</td>
<td>76</td>
</tr>
<tr>
<td>Liaoning</td>
<td>459</td>
<td>1200</td>
<td>5835</td>
<td>1316</td>
<td>268</td>
<td>432</td>
<td>325</td>
<td>54</td>
<td>86</td>
</tr>
<tr>
<td>Ningxia</td>
<td>226</td>
<td>327</td>
<td>766</td>
<td>178</td>
<td>79</td>
<td>91</td>
<td>63</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>Qinghai</td>
<td>43</td>
<td>106</td>
<td>598</td>
<td>129</td>
<td>131</td>
<td>59</td>
<td>45</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Shaanxi</td>
<td>334</td>
<td>549</td>
<td>3781</td>
<td>820</td>
<td>273</td>
<td>294</td>
<td>221</td>
<td>39</td>
<td>68</td>
</tr>
<tr>
<td>Shandong</td>
<td>946</td>
<td>2134</td>
<td>11469</td>
<td>2846</td>
<td>696</td>
<td>897</td>
<td>678</td>
<td>105</td>
<td>150</td>
</tr>
<tr>
<td>Shanghai</td>
<td>114</td>
<td>469</td>
<td>1130</td>
<td>342</td>
<td>29</td>
<td>104</td>
<td>86</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>Shanxi</td>
<td>977</td>
<td>964</td>
<td>6017</td>
<td>756</td>
<td>199</td>
<td>555</td>
<td>415</td>
<td>64</td>
<td>81</td>
</tr>
<tr>
<td>Sichuan</td>
<td>379</td>
<td>777</td>
<td>6362</td>
<td>1478</td>
<td>646</td>
<td>463</td>
<td>371</td>
<td>56</td>
<td>141</td>
</tr>
<tr>
<td>Tianjin</td>
<td>90</td>
<td>333</td>
<td>1434</td>
<td>573</td>
<td>33</td>
<td>81</td>
<td>61</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Xinjiang</td>
<td>257</td>
<td>608</td>
<td>2639</td>
<td>632</td>
<td>516</td>
<td>218</td>
<td>158</td>
<td>23</td>
<td>32</td>
</tr>
<tr>
<td>Xizang</td>
<td>1</td>
<td>52</td>
<td>149</td>
<td>46</td>
<td>149</td>
<td>15</td>
<td>12</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Yunnan</td>
<td>332</td>
<td>435</td>
<td>3823</td>
<td>576</td>
<td>398</td>
<td>302</td>
<td>230</td>
<td>38</td>
<td>75</td>
</tr>
<tr>
<td>Zhejiang</td>
<td>293</td>
<td>670</td>
<td>3009</td>
<td>1342</td>
<td>118</td>
<td>270</td>
<td>195</td>
<td>22</td>
<td>22</td>
</tr>
</tbody>
</table>

to a 21% decrease in NH$_3$ emissions from agricultural sources, while NH$_3$ emissions from residential sources and transportation increase by 99% and 13.1 times, respectively. Such difference in agricultural sources is mainly caused by the estimates of synthetic fertilizer (Kang et al., 2016), particularly concerning the treatment of fertilizer types and corresponding emission factors.

Many discrepancies between MEIC and INTAC arise from the integration of regional emission inventories. As presented in Fig. 5b, notable disparities are observed in the YRD and PRD regions. Estimates for NO$_x$ emissions in the YRD region are approximately 88% of those derived from the MEIC model. This highlights an enhanced precision attributable to reliable assessments of denitrification efficiency in power plants and the measured NO$_x$ emission factors for both power plants and boilers within the integrated YRD inventory, as supported by previous studies (e.g., Zhao et al., 2018). INTAC’s estimates for NMVOC emissions in the YRD region are 26% lower than estimates in MEIC. The overestimation in MEIC mainly results from the uncertainties in solvent use source, particularly coating, printing, and dyeing processes. The integrated YRD emission inventory employs more accurate calculation parameters for NMVOCs such as statistical data from local city yearbooks, industry association reports, and apparent consumption of solvents. Furthermore, the speciation profiles of NMVOCs are localized and corrected based on the literature research and measurements. In the PRD region, the NO$_x$ emissions from INTAC are 41% higher than MEIC estimates, with non-road sources and non-metallic mineral products contributing 45% and 40%, respectively, to this difference. The PRD inventory employs a detailed calculation approach for shipping emissions based on AIS data, in contrast to the simplified approach for inland waterway sources in MEIC. The NO$_x$ emissions from industrial processes of brick and flat-glass manufacturing are not considered in MEIC, a deficiency which is addressed in the integrated PRD inventory. INTAC’s NMVOC emissions are approximately 59% of those from MEIC. The
disparity is particularly notable in industrial and solvent use sources, contributing 49% and 35%, respectively, to the observed difference. In INTAC nearly half of the VOC emission factors for industrial solvent sources are based on local measurements, and a preference for raw-material-based calculations over product-based ones reduces uncertainty in the estimation. For significant VOC-emitting sources like cleaning solvents, MEIC employs an emission factor of 1000 g kg$^{-1}$, whereas the PRD inventory uses 850 g kg$^{-1}$. In the case of oil refineries, the emission factors are 2.76 g kg$^{-1}$ for MEIC and 1.82 g kg$^{-1}$ for the PRD inventory.

3.2.2 Impact of point source contributions

The most accurate method for obtaining emissions at finer-scale grids relies on spatial allocation based on precise geographical coordinates. In MEIC the majority of emission sources are represented as area sources and distributed onto grids using spatial proxies, such as urban population, except for power plants. In contrast, the increased proportion of industrial point source emissions in INTAC significantly constrains the uncertainties associated with spatial proxies. Figure 6 shows the inter-comparisons of the percentage of point, on-road, and area source emissions between INTAC and MEIC. Air pollutants, especially those dominated by industrial combustion sources like SO$_2$, NO$_x$, PM$_{10}$, and PM$_{2.5}$, exhibit a significantly higher proportion of point source emissions within INTAC compared with MEIC. In MEIC the proportion of point source emissions for SO$_2$, PM$_{10}$, NO$_x$, and PM$_{2.5}$ is 17%, 9%, 19%, and 7%, respectively. However, in INTAC these percentages substantially increase to 66%, 54%, 52%, and 48%, respectively, indicating a more accurate representation of spatial patterns. For other species with emissions mainly from area sources (e.g., residential and transportation), there are limited improvements in the proportion of point source emissions in INTAC.

To further assess the impact of point sources, Fig. 7 takes SO$_2$ and the YRD region as an example to compare the spatial emission patterns between INTAC and MEIC. Figure 7c–e reveal that MEIC tends to overestimate emissions in urban centers and underestimate emissions in rural areas compared with INTAC. Amid economic growth and rapid urbanization, MEIC’s use of urban population as a proxy for spatial allocation becomes impractical as many factories relocate from city centers to rural areas. To elucidate the difference between population-based and point-source-based allocation methods.

Figure 5. Inter-comparisons of emission estimates between INTAC and MEIC. Panel (a) shows the difference by sector, and (b) presents the ratio of emissions in INTAC to those in MEIC.
in emission mapping, we present the cumulative percentage of SO\textsubscript{2} emissions in MEIC and INTAC based on descending population orders in Fig. 7f. We use the grid groups where densely populated areas contribute 50\% of SO\textsubscript{2} emissions in MEIC as an example, and we compare them with the cumulative percentage in INTAC across various grid sizes. The results show that at a resolution of 0.05° INTAC only accounts for 17\% of the emissions, while it reaches 48\% as the grid size increases to 1.0°. This suggests that at a fine grid scale MEIC tends to allocate more emissions to densely populated urban areas, while INTAC allocates a larger proportion to suburban and rural areas, aligning better with the real-world emission spatial patterns. This mitigation of bias through INTAC is especially notable at finer resolutions. The close cumulative percentage at 1.0° in the two inventories can be attributed to the fact that urban and suburban areas often fall within the same grid, leading to a decreasing enhancement in the spatial accuracy achieved by INTAC. Figure 7g further presents the correlation between the spatial patterns of SO\textsubscript{2} emissions in INTAC and various spatial proxies. At a resolution of 1.0° the correlation coefficients between emission distributions and factors (i.e., road networks, nighttime lights, total population, urban population, and rural population) fall within the range of 0.55–0.79. Nevertheless, at a resolution of 0.05° the correlation coefficients range from 0.05 to 0.13. This indicates that at higher spatial resolutions INTAC substantially reduces the bias introduced by spatial proxies in MEIC.

3.3 Improvements in air-quality modeling by INTAC

3.3.1 Overall performance in key regions

We conducted simulations using the WRF–CMAQ model driven by INTAC and MEIC separately to evaluate the improvements in modeled air pollutant concentrations. Table 4 evaluates the simulated emissions in 74 major cities (locations depicted in Fig. S2) against in situ observations, with corresponding scatterplots shown in Fig. S3. INTAC demonstrates improved agreement between modeled concentrations and ground-level observations, which benefits from the integrated high-resolution inventories. Compared with MEIC, INTAC leads to a decline in the mean bias of simulated major pollutant concentrations by 2–14 µg m\textsuperscript{-3}, a reduction in the root mean square error by 4–19 µg m\textsuperscript{-3}, and a decrease in the normalized mean error by 4\%–71\%. This finding indicates that INTAC produces a more accurate overall characterization of emissions in China. Furthermore, given that atmospheric pollution monitoring stations are mainly located in urban areas in China, the observed differences suggest that INTAC can mitigate the overestimation of major pollutant concentrations in urban centers. As discussed in Sect. 3.2.2, MEIC overestimates emissions in urban areas and underestimates them in rural and suburban areas, consequently introducing uncertainties into air-quality modeling. The improved accuracy in spatial distributions within INTAC significantly contributes to enhancing the overall accuracy of air pollutant modeling.

Figure 8 further compares the overall simulation performance between INTAC and MEIC in three key regions (BTH, YRD, and PRD), with corresponding scatterplots shown in Figs. S4–S6. Regarding PM\textsubscript{2.5} and its precursors, MEIC shows a considerable mean bias of up to 36 µg m\textsuperscript{-3} and a root mean square error of up to 59 µg m\textsuperscript{-3} in key regions. In contrast, INTAC demonstrates maximum MB values of 15 µg m\textsuperscript{-3} and maximum RMSE values of 40 µg m\textsuperscript{-3}. The correlation coefficients between simulated and observed concentrations of the three air pollutants are generally lower in MEIC compared with those in INTAC. The modeling performance driven by INTAC, particularly for short-lived pollutants, experiences significant improvement due to their strong correlation with spatial distributions of emission sources. Nonetheless, discrepancies between modeled and observed surface concentrations still exist because of uncertainties from meteorological, physical, and chemical processes within chemical transport models. Moreover, emission sources such as residential, transportation, and agriculture are treated in INTAC as non-point sources, and their allocation to grids using spatial proxies can introduce biases in
Figure 7. Spatial pattern analysis of emissions in INTAC using SO₂ emissions as an example. Panels (a) and (b) display the spatial distributions of SO₂ emissions in MEIC and INTAC, respectively. MEIC emissions have been downscaled from 0.25 to 0.1° for comparison. To compare MEIC and INTAC in detail, the YRD region is zoomed in on. Panels (c)–(e) show spatial distributions of SO₂ emissions in MEIC and INTAC, and also their difference. Circles in (e) represent the center of a city. Panel (f) compares cumulative percentage of SO₂ emissions in INTAC with those in MEIC across different spatial resolutions. The gridded SO₂ emissions, ranging in resolution from 0.05 to 1.0°, are cumulated in descending order of population. The percentage annotations in different colors indicate the level of accumulated SO₂ emissions in INTAC at various spatial resolutions when SO₂ emissions in MEIC reach 50% accumulation. Panel (g) shows correlation coefficient between SO₂ emissions in INTAC and multiple spatial proxies at different grid sizes.

3.3.2 Improvements across different spatial resolutions

To provide a more in-depth assessment of improved spatial patterns in INTAC, Fig. 9 categorizes grid cells into different bins based on their urban population and calculates the ratio of simulated pollutant concentrations to ground observations for both INTAC and MEIC in each category. The results demonstrate that as the urban population increases,
Table 4. The discrepancies between simulated SO$_2$, NO$_2$, and PM$_{2.5}$ concentrations, and observed values for 74 major cities at a resolution of 12 km, using MEIC and INTAC as emission inputs. The statistical metrics used for comparison include $R$, MB, and RMSE. The bold font represents the difference in modeling performance between INTAC and MEIC.

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>Inventory</th>
<th>MB (µg m$^{-3}$)</th>
<th>RMSE (µg m$^{-3}$)</th>
<th>NME (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO$_2$</td>
<td>INTAC</td>
<td>11</td>
<td>30</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>MEIC</td>
<td>25</td>
<td>49</td>
<td>163</td>
</tr>
<tr>
<td></td>
<td>Difference</td>
<td>$-$14</td>
<td>$-$19</td>
<td>$-$71</td>
</tr>
<tr>
<td>NO$_2$</td>
<td>INTAC</td>
<td>7</td>
<td>22</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>MEIC</td>
<td>18</td>
<td>31</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Difference</td>
<td>$-$11</td>
<td>$-$9</td>
<td>$-$17</td>
</tr>
<tr>
<td>PM$_{2.5}$</td>
<td>INTAC</td>
<td>6</td>
<td>35</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>MEIC</td>
<td>8</td>
<td>39</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Difference</td>
<td>$-$2</td>
<td>$-$4</td>
<td>$-$4</td>
</tr>
</tbody>
</table>

Figure 8. Comparison of modeling performance across key regions (i.e., BTH, YRD, and PRD) when using MEIC and INTAC, respectively, as emission inputs. The statistical metrics used for comparison include $R$, MB, and RMSE. The regions under comparison comprise BTH, YRD, and PRD.

the enhanced model performance of INTAC over MEIC for SO$_2$, NO$_2$, and PM$_{2.5}$ becomes more evident. Specifically, when the urban population is less than 50,000, both INTAC and MEIC exhibit a median range of simulated-to-observed concentration ratios close to 1. However, as the urban population exceeds 550,000, the average range for MEIC widens to 1.4–5.2, whereas it remains within the range of 0.9–1.0 for INTAC. This indicates a significant improvement in mitigating the overestimation of emissions in densely populated areas by INTAC. The incorporation of the industrial point source emission inventory for China along with the YRD and PRD emission inventories significantly increases point source shares in INTAC, thus producing better spatial representations for real-world emission distributions and smaller simulated deviations.

Model performance differences between MEIC and INTAC are influenced by grid size. Figure 10 presents the comparison between modeled SO$_2$, NO$_2$, and PM$_{2.5}$ concentrations against ground observations for 74 major cities at resolutions of 36 and 12 km. Increasing spatial resolution does not lead to a reduction in simulation errors, especially for MEIC. As the horizontal resolution increases from 36 to 12 km, the mean biases of simulated SO$_2$, NO$_2$, and PM$_{2.5}$ concentrations using MEIC as input show an increase from 37 % to 143 %, 11 % to 46 %, and $-$3 % to 15 %, respectively, when compared with in situ observations. In contrast, the simulation results using INTAC as input exhibit better agreement with ground observations, with mean biases for SO$_2$, NO$_2$, and PM$_{2.5}$ increasing from 23 % to 64 %, $-$0 % to 17 %, and 2 % to 11 %, respectively. This is due to the fact that the deviations in finer grid cells, whether overestimated or underestimated, tend to cancel out at a coarse spatial resolution. The decoupling between emission spatial distributions with proxies at finer grids leads to more noticeable bi-

As in air-quality modeling. Therefore, the findings suggest that the INTAC developed in this study can effectively constrain uncertainties in emissions and the modeling bias, especially at fine spatial scales. The improvement will help tackle emerging challenges in high-resolution air-quality modeling in China.

4 Discussion

Both qualitative and quantitative uncertainty assessments are essential components of a comprehensive inventory for policy or scientific purposes. Approaches such as error propagation and Monte Carlo simulation are commonly used for quantitative uncertainty analysis in China’s emission inventory (Lu et al., 2011; Streets et al., 2003; Zhao et al., 2011, 2017b). However, this study uses an integrated method, rather than a unified framework, to compile the high-resolution emission inventory for China. Collecting only emission quantities from the seven inventories without detailed calculation parameters makes it challenging to assess the overall uncertainties in INTAC here. We have summarized the estimated uncertainty range for components of INTAC in Table 5, where such information is available. Although the uncertainties might be reported for a year other than 2017, they still provide a rough representation of the uncertainty range in major air pollutant emission estimates within INTAC. Species such as SO$_2$ and NO$_x$ exhibit relatively low uncertainties, benefiting from well-established estimates for large-scale combustion sources. The considerable uncertainties observed in BC and OC emissions may be attributed to inaccuracies in the emission factors of the residential sector. Further details regarding the uncertainties in each component inventory can be found in corresponding literature (An et al., 2021; Huang et al., 2021; Kang et al., 2016; H. Liu et al., 2016; Yin et al., 2019; Zhao et al., 2011).

The uncertainties in INTAC also arise from the integration process:

- The emission sectors in all inventories need to be mapped to the 88 standard sectors first. Due to limited foundational information for an aggregated sector’s disaggregation, this process may introduce biases for those who initially provide coarser source categories. For example, if an inventory only offers one aggregated sector for power, which needs to be broken down into four subsectors (i.e., production of power, supply of power, production of industrial heat power, and production of residential heat power), we use the energy consumption for corresponding sectors from the statistical yearbook as a reference basis for this allocation, which is a relatively reliable method despite potential deviations.

- To generate speciated VOC species, sectoral NMVOC emissions in each inventory need to be matched to corresponding source profiles from the MEIC model. Discrepancies in emission source mapping can impact the outcomes, which will be overcome by gathering more detailed sectoral information for each inventory or directly collecting speciated species in future studies.

- INTAC is made publicly available at a monthly scale, given that the majority of its components are gathered on a monthly or annual scale. The temporal disaggregation to finer resolutions for modeling is achieved using empirically selected weighting factors in the MEIC model. However, it is noteworthy that the parameters employed for allocating emissions to daily or hourly scales remain fixed and do not vary over time or region, introducing additional uncertainties. In the future, we plan to incorporate more advanced data or methods (e.g., real-time emission measurements) to enhance temporal accuracy at finer scales, as indicated in the previous work for the power sector (Wu et al., 2022).

- The border issue is inevitable when emissions for the same species in two adjacent cities are derived from
Figure 10. Comparison of modeled air pollutant concentrations and ground observations for 74 cities at 36 and 12 km resolutions using MEIC and INTAC, respectively, as emission inputs. The dashed black line represents the observational mean, and the annotations above the bar charts indicate the mean biases between simulated concentrations and the corresponding observed values.

Table 5. Uncertainties in the inventory components of INTAC (in %) contingent upon the availability of such information.

<table>
<thead>
<tr>
<th>Emission inventory</th>
<th>Reporting year</th>
<th>SO₂</th>
<th>NOₓ</th>
<th>CO</th>
<th>NMVOCs</th>
<th>NH₃</th>
<th>PM₁₀</th>
<th>PM₂.₅</th>
<th>BC</th>
<th>OC</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PKU-NH₃</td>
<td>2012</td>
<td>±4</td>
<td>±4</td>
<td>±5</td>
<td>±4</td>
<td>−26 to 25</td>
<td>±4</td>
<td>±4</td>
<td>±4</td>
<td>±4</td>
<td>Kang et al. (2016)</td>
</tr>
<tr>
<td></td>
<td>2013</td>
<td>±4</td>
<td>±4</td>
<td>±5</td>
<td>±4</td>
<td>−26 to 25</td>
<td>±4</td>
<td>±4</td>
<td>±4</td>
<td>±4</td>
<td>H. Liu et al. (2016)</td>
</tr>
<tr>
<td>Shipping emission inventory for East Asia</td>
<td>2003–2017</td>
<td>±44 to 89</td>
<td>−74 to 84</td>
<td>−65 to 65</td>
<td>−75 to 100</td>
<td>−74 to 81</td>
<td>±4</td>
<td>±4</td>
<td>±4</td>
<td>±4</td>
<td>Yin et al. (2019)</td>
</tr>
<tr>
<td>Open biomass burning emission inventory for China</td>
<td>2017</td>
<td>−17 to 20</td>
<td>−25 to 28</td>
<td>−30 to 39</td>
<td>−34 to 50</td>
<td>−50 to 86</td>
<td>−45 to 60</td>
<td>−43 to 62</td>
<td>−53 to 116</td>
<td>−54 to 160</td>
<td>Huang et al. (2021)</td>
</tr>
<tr>
<td>PRD emission inventory</td>
<td>2017</td>
<td>−29 to 36</td>
<td>−28 to 33</td>
<td>−42 to 75</td>
<td>−44 to 68</td>
<td>−58 to 117</td>
<td>−36 to 62</td>
<td>−30 to 46</td>
<td>−14 to 45</td>
<td>−17 to 54</td>
<td>−25 to 136</td>
</tr>
<tr>
<td>YRD air pollutant emission inventory</td>
<td>2005</td>
<td>−14 to 13</td>
<td>−13 to 37</td>
<td>−13 to 37</td>
<td>−14 to 45</td>
<td>−17 to 54</td>
<td>−25 to 136</td>
<td>−40 to 121</td>
<td>−14 to 45</td>
<td>−17 to 54</td>
<td>−25 to 136</td>
</tr>
</tbody>
</table>

different inventories. Typical examples are the cities located at the boundaries of the YRD and PRD regions. In INTAC we downscale all emissions to 1 km before the spatial–temporal coupling process, thereby mitigating this uncertainty to some extent.

INTAC for 2017 is subject to some limitations: (1) In the integrated method, emission data for the same city and species across different sectors may come from different datasets. Similarly, emission data for different species within the same city and sector may also originate from different datasets. The utilization of species ratios requires careful consideration in these cases. (2) Limited resources present a substantial challenge in gathering emission inventories over extended time series from diverse research institutions within the scope of this study. Consequently, we exclusively present INTAC for 2017, with the possibility of extension to other years in subsequent research.

5 Data availability

Data described in this paper can be accessed via Zenodo at https://doi.org/10.5281/zenodo.10459198 (Wu et al., 2024) and http://meicmodel.org.cn/intac (last access: 15 April 2024).

6 Concluding remarks

Compiling a comprehensive bottom-up emission inventory for China that achieves both extensive coverage and high resolution poses a significant challenge. In this work, we construct a 0.1° resolution integrated inventory for 2017 through the fusion of multi-source emission inventories. An integration model has been developed to effectively couple heterogeneous emission datasets, aimed at generating a standardized data cube with consistent sectors, species, and spatial–temporal resolution. INTAC is created through source mapping, species mapping, temporal disaggregation, spatial allocation, and spatial–temporal coupling. Six representative emission inventories focusing on national and regional scales, as well as key species and sources in China, are merged with MEIC. This integration harnesses the strengths of each inventory, resulting in an improved depiction of emission totals and spatial distribution patterns for China.

We find that the total emissions of SO₂, NOₓ, CO, NMVOCs, NH₃, PM₁₀, PM₂.₅, BC, and OC in INTAC for 2017 are 12.3, 24.5, 141.0, 27.9, 9.2, 11.1, 8.4, 1.3, and
2.2 Tg, respectively. Industrial production serves as the main source of various atmospheric pollutants. Residential sources contribute over 40% to CO, BC, and OC emissions. Apart from agricultural sources, which account for 83% of NH₃ emissions, the contributions from various minor emission sources cannot be overlooked. This study emphasizes the significance of shipping emissions, particularly in contributing to SO₂ (13%) and NOₓ (13%). Fossil fuel combustion dominates the emissions of PM₁₀, PM₂.₅, CO, BC, SO₂, and NOₓ, ranging from 38% to 80%. The enhancement in emission estimates for China in INTAC is demonstrated by the comparison with MEIC. For instance, the incorporation of numerous point sources has notably addressed MEIC’s tendency to overestimate emissions in urban centers, particularly at higher spatial resolutions. In comparison with MEIC, INTAC exhibits a mean bias reduction in simulated concentrations of major pollutants against ground observations across 74 cities, ranging from 2 to 14 µg m⁻³. The improvement in model performance achieved by INTAC is particularly noticeable at finer spatial resolutions.

Our study offers an efficient framework for creating a highly resolved emission inventory on a large scale. This approach integrates advantages from previous studies and holds the potential to support policymakers in making well-informed decisions for improving air quality. In the future, we anticipate the ongoing incorporation of additional emission datasets to offer a more reliable representation of emissions in China over extended time periods.

Supplement. The supplement related to this article has six figures and three tables. The supplement related to this article is available online at: https://doi.org/10.5194/essd-16-2893-2024-supplement.

Author contributions. NW, GG, and QZ designed the study. NW developed the INTAC, conducted chemical transport modeling, and analyzed the emissions. JZ, YS, HL, YuZ, YiZ, and QS provided the emission inventories for the integration. RX helped with the data analysis. SL compiled the chemical transport model. XL contributed to the design of computer programs for the integration model. The manuscript was written by NW and GG, and it was revised and discussed by all coauthors.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Regarding the maps used in this paper, please note that Figs. 1, 4, and 7 contain disputed territories.

Acknowledgements. We thank Zhijiong Huang, Junchi Wang, Mingxu Liu, Wenling Liao, and Chen Gu for their contributions to the handling and transfer of the emission inventories for the integration.

Financial support. This research has been supported by the National Natural Science Foundation of China (grant no. 92044303), the National Key R&D Program of China (grant no. 2022YFC3700605), and the Major Project of High Resolution Earth Observation System (grant no. 30-Y60B01-9003-22/23).

Review statement. This paper was edited by Yuqiang Zhang and reviewed by two anonymous referees.

References


Liu, F., Zhang, Q., Tong, D., Zheng, B., Li, M., Huo, H., and He, K. B.: High-resolution inventory of technologies, activities, and emissions of coal-fired power plants in China


Xiao, Q., Geng, G., Xue, T., Liu, S., Cai, C., He, K., and Zhang, Q.: Tracking PM


