Estimation of Anthropogenic CH₄ and CO₂ Emissions in Taiyuan-Jinzhou Region: One of the World's Largest Emission Hotspots

Cheng Hu¹,², Wei Xiao³,⁴, Timothy J. Griffis⁴, Qitao Xiao⁴, Shumin Wang⁵, Yuzhong Zhang⁷,⁸, Weifeng Wang¹, Lingyun Zhu⁶, Xin Chen⁴, Xueying Yu⁴, and Xuhui Lee¹⁰

¹College of Biology and the Environment, Joint Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing, China, ²Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science & Technology, Nanjing, China, ³Center on Atmospheric Environment, International Joint Laboratory on Climate and Environment Change (ILCEC), Nanjing University of Information Science & Technology, Nanjing, China, ⁴Department of Soil, Water, and Climate, University of Minnesota-Twin Cities, Minneapolis, MN, USA, ⁵Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China, ⁶Shanxi Province Institute of Meteorological Sciences, Taiyuan, China, ⁷Key Laboratory of Coastal Environment and Resources of Zhejiang Province, School of Engineering, Westlake University, Hangzhou, China, ⁸Institute of Advanced Technology, Westlake Institute for Advanced Study, Hangzhou, China, ⁹Guang’an Vocational & Technical College, Guangan, China, ¹⁰School of the Environment, Yale University, New Haven, CT, USA

Abstract  Coal mining ranks as the largest anthropogenic CH₄ source in China with emission factors (EFs) varying up to 30-fold among inventories when applied to different provinces. The lack of independent evaluation of coal mining CH₄ EFs in China is one of the main uncertainties in estimating national total CH₄ emissions. Shanxi province, which supplies 25% of the national coal production, is the largest coal mining CH₄ emission region in China and even among the world’s largest coal production regions. This area is also a significant anthropogenic CO₂ source because of high-density power and industrial activities. Given the large uncertainties in CH₄ and CO₂ inventories from provincial to city scales, questions remain whether state-of-the-art inventories have accurately estimated these emission hotspots. Here, we evaluate CH₄ and CO₂ emissions from one of the world’s largest coal production regions near Taiyuan City, the capital of Shanxi province, China. CH₄ and CO₂ concentrations were measured from March 2018 to February 2019 from a 30-m tower. These data were used within an inverse modeling framework to simulate both CH₄ and CO₂ concentrations and to evaluate EFs for this region. Results show generally good agreement between observed and simulated CH₄ concentrations. However, the CO₂ simulations were much lower compared to the observations. Given the minor role of NEE-induced CO₂ enhancements, we believe that the large difference is attributed to the underestimation of anthropogenic CO₂ emissions. In general, the derived posteriori anthropogenic CH₄ emissions were 85.2±(18.1)% of a priori emissions, where fugitive CH₄ from coal mining accounted for ~92.7% of total anthropogenic emissions. The derived coal mining EF was 23.2±(4.9) m³ CH₄/ton coal, close to the default value of high CH₄-content coal, but twofold the province average that were reported by previous observation-based studies in Shanxi province, indicating large spatial inhomogeneity in the coal mining CH₄ EF. The posteriori CO₂ emissions were 1.6-fold of the a priori emissions, highlighting underestimation of CO₂ emissions in industrial cities and some potential large emission sources that are missing from state-of-the-art inventories. Finally, we also emphasize the use of satellite observations and denser tower-based networks are essential in resolving the spatial inhomogeneity of greenhouse gas emissions.

Plain Language Summary  The understanding of anthropogenic CH₄ and CO₂ emissions is basis for climate mitigation especially for global top emitting countries, but the largest issue before addressing above question is that many previous studies have found considerable bias of greenhouse gas emission for almost all inventories from city to regional scales. These facts hindered the government to make and evaluate corresponding mitigation policies. Here, to quantify CH₄ and CO₂ emissions at one of global largest CH₄ and CO₂ hotspot in China, we conducted 1 year tower-based atmospheric CH₄ and CO₂ concentration measurements and used atmospheric inversion method to constrain and evaluate their emissions, we found coal mining CH₄ emission factor has less bias but CO₂ emissions were underestimated by 1.6-fold, highlighting underestimation of CO₂ emissions in industrial cities and some potential large emission sources that are missing
from state-of-the-art inventories. Our findings indicate more work is needed for urban government to fully understand their greenhouse gases emissions.

1. Introduction

Methane (CH$_4$) and carbon dioxide (CO$_2$) are the main anthropogenic greenhouse gases contributing to global warming. Together they accounted for $\sim$81% of direct anthropogenic radiative forcing (Seto et al., 2014). Hence, the implementation of emission reduction policies for both CH$_4$ and CO$_2$, especially in the top emitting countries, is needed to mitigate future climate change. China is the largest emitter for anthropogenic CH$_4$ and CO$_2$, with annual emissions reaching 0.065 Gt for CH$_4$ and 12 Gt for CO$_2$ in 2020. The largest anthropogenic sources for CH$_4$ were fugitive emissions from coal mining which accounted for 31% of national total emissions, and power industry accounted for 41% in anthropogenic CO$_2$ emissions (Crippa et al., 2021). In 2021, China has made a commitment to reduce CH$_4$ emissions under the Glasgow Agreement. Further, an ambitious commitment was also made in 2020 to reach a CO$_2$ emission peak by year 2030. To achieve these emission reductions, a national plan has been developed using a hierarchal approach from the national, regional, and local scales and includes different units such as provinces, cities, and specific facilities. Accurate quantification of emission changes is the basis to evaluate whether these commitments are effectively fulfilled. However, large uncertainties in estimating CH$_4$ and CO$_2$ emissions from local to regional scales may hinder these ambitious emission reduction goals (Han, Zeng, Oda, et al., 2020; Lin et al., 2021; Seto et al., 2014).

Activity data and emission factor (EF)-based “bottom-up” inventories are treated as baselines for local government to evaluate emission reduction rates. Many recent studies have reported large uncertainties in almost all state-of-the-art CH$_4$ and CO$_2$ inventories for China, where relative difference of these inventories can be $>200\%$ at city scale, especially for industrial cities, with averages of $\sim$57%–162% at provincial scales among different sectors (Han, Zeng, Oda, et al., 2020; Han, Zeng, Zhang, et al., 2020; Lin et al., 2021). Uncertainties in CH$_4$ emissions are comparable with CO$_2$ for some specific sectors such as waste treatment, with a relative difference larger than 2 times for China (Lin et al., 2021). Further, the underlying anthropogenic CH$_4$ emissions and its changes were still under debate with different conclusions including largely increased at the rate of 1.1 Tg/year (Miller et al., 2019), slightly increased by 0.1–0.4 Tg/year (Lu et al., 2021; Saunois et al., 2020; Sheng et al., 2021), and slightly decrease (Gao et al., 2020; Liu et al., 2021; Sheng et al., 2019).

One of the main uncertainties in estimating CH$_4$ emission is caused by underground coal mining. Such emission is mainly from ventilation shafts, where atmospheric air was pumped into the mines to keep lower CH$_4$ concentration (<0.5%) and to avoid dangerous conditions in the workplace. Hence, emissions can be affected by burial depths, different CH$_4$-content coal, mining methods/technologies, and the utilization rate. The available CH$_4$ EFs from underground coal mining for China were mainly based on nationwide measurements prior to the year of 2006 for $\sim$800 underground mines, and are highly spatially heterogeneous (Sheng et al., 2019; Zhu et al., 2017). The province-based EFs considerably varied by 30-fold from smaller than 1 m$^3$ CH$_4$/ton coal in Beijing to greater than 30 m$^3$ CH$_4$/ton coal as Chongqing. Further, observations also found that EFs largely varied within the same province by sevenfold (Sheng et al., 2019). While these field measurements provided important EFs, information for specific regions remain very limited. Further, improved coal mining technologies can cause significant bias in EFs when estimating recent national CH$_4$ emissions (Lin et al., 2021; Zhu et al., 2017). Coal mining CH$_4$ EFs used for China have shown large discrepancies among different inventories, i.e., the EDGAR v4.2 was found to overestimate EFs in China by a factor of 2 for the coal mining because it applied the European averaged EFs (Saunois et al., 2016). These disagreements indicate that a robust approach is needed to monitor anthropogenic CH$_4$ emissions, especially from the coal mining. It can help to fill the knowledge gap in understanding their spatiotemporal evolutions, and provide independent constraint on regional CH$_4$ emissions.

Urban areas are also emission hotspots for both CH$_4$ and CO$_2$ and account for 21% and 70% of global anthropogenic emissions, respectively (Marcotullio et al., 2013; Seto et al., 2014). Such contributions are thought to be much larger in Northern China considering there are more anthropogenic emission sources concentrated in industrial cities and coal production areas. The above facts make these regions to be the most important target region to conducting emission mitigation policies for China. Hence, understanding the main sources and its underlying driving factors in these industrial cities is urgently needed for mitigation policy making and application. Among all provinces in China, Shanxi is a representative province which not only provide 25% ($8.93 \times 10^8$ tons) of
national total coal production in 2018, even larger than the world's second largest coal production country India (7.45 × 10^4 tons). Taiyuan City and Shanxi province are also among the most severely polluted regions on earth because of the high density of heavy industry, coal combustion, and anthropogenic CO_2 emissions (Mu et al., 2019; Wei et al., 2018). A recent bottom-up study found five of the top 10 largest Chinese CH_4 emission cities were located in Shanxi province (Wang et al., 2021). These combined factors indicate this region is one of the largest hotspots in China and even the world for both anthropogenic CH_4 and CO_2 emissions.

Tower-based atmospheric inversions provide a useful tool to independently evaluate local emissions and reduce corresponding inventory uncertainty (Gurney et al., 2021; Le et al., 2021; Sargent et al., 2018; Turner et al., 2020). To our best knowledge, no study has used tower-based concentration and atmospheric inverse modeling to evaluate coal mining EFs in China. Here, we constrain anthropogenic CH_4 and CO_2 emissions and provide assessment of inventory uncertainty for CH_4 emission leakage from coal mining in one of the largest anthropogenic emission regions of China. This study is based on 1-year continuous atmospheric CH_4 and CO_2 concentration observations from March 2018 to February 2019 using a 30-m tower in Taiyuan City, the capital of Shanxi province. The hourly concentration observations and three different inversion approaches were used to provide a robust constraint on anthropogenic CH_4 and CO_2 emissions, especially for coal mining CH_4 EFs.

2. Materials and Methods

2.1. Atmospheric CH_4 and CO_2 Concentration and Supporting Materials

Atmospheric CH_4 and CO_2 concentration measurements were conducted on a tower at the Xiaodian meteorological station (112°33′E, 37°44′N). The Xiaodian station is located between Jinzhong and Taiyuan City (Figure 1a). There are no obvious industrial sources within a 10-km radius of the observation site. The air sampling inlet was mounted at a 30-m height on the tower. The CH_4 and CO_2 concentrations were measured near continuously using a wavelength scan cavity ring-down spectroscopy-based analyzer (model G2301, Picarro Inc., Sunnyvale, CA). The measurement system was calibrated every 2 hr with standard gases from the Atmospheric Composition Observation and Service Center of the China Meteorological Administration that was traceable to NOAA/GML (NOAA Global Monitoring Laboratory) standards. The measurement cycle consisted of air sampling for 115 min followed by standard gas calibration for a period of 5 min. The details of the calibration system are described in Fang et al. (2014). In general, 97.0% of hourly CH_4 observations and 96.9% of the CO_2 observations were available for the whole year observation period.

Daily column-averaged dry air CH_4 mixing ratio (XCH_4) from TROPOMI measurements between May 2018 and April 2019 were also used to illustrate the spatial distribution of CH_4 concentration within Shanxi province. The TROPOMI retrievals are from the polar-orbiting Sentinel-5 Precursor satellite with spatial resolution of 7 km × 7 km (Zhang et al., 2020). Observations from the Xiaodian meteorological station were used to evaluate simulated meteorological fields from the Weather Research and Forecasting (WRF version 4.2.2) model including air temperature at 2 m (T_2m), relative humidity (RH), downward solar radiation (S↓), wind speed (WS), and wind direction (WD) at 10 m.

The observation site is located between Taiyuan and Jinzhong Cities as shown in Figure 1a. The main urban areas in both cities are closely connected and are likely to have a strong influence on the CH_4 and CO_2 observations. Here, we have named the core urban area in both cities as the Taiyuan-Jinzhou Metropolitan (TJM) area. Based on a recent bottom-up estimate of CH_4 emissions for all Chinese cities in the year 2015, Taiyuan ranked as the highest and Jinzhong as the ninth highest CH_4 emission city in China (Wang et al., 2021). Hence, the TJM area is likely the highest anthropogenic CH_4 emission urban area in China. The population of the TJM area was 8.7 million in 2020 and raw coal production was 12.18 × 10^8 tons in the year 2018. Shanxi province is surrounded by four provinces including Inner Mongolia, Hebei, Henan, and Shaanxi (Figure 1b). The raw coal production in these five provinces accounted for 74% of the national total production for China in 2018. It ranked as the world's greatest hotspot for CH_4 emissions from coal production and also for total anthropogenic CH_4 emissions. The coal productions within Shanxi province were obtained from Shanxi Statistical Yearbook 2018–2019 (Shanxi Provincial Bureau of Statistics, 2018–2019).

2.2. Simulation of CH_4 and CO_2 Concentration

In general, atmospheric CH_4 (and CO_2) at the observation site is the sum of background air (i.e., when an air mass enters the study domain) and the enhancement contributed by regional sources/sinks within the study domain.
The enhancements are simulated by coupling regional CH$_4$ (and CO$_2$) flux information with dynamic concentration footprint estimates that considers the atmospheric transport process within the study domain. The simplified equations for CH$_4$ (and CO$_2$) concentration simulations can be expressed as

$$\text{CH}_4_{\text{sim}} = \text{CH}_4_{\text{bg}} + \sum_{i=1}^{168} f_i \times (E_{\text{CH}_4_{\text{anthro}}})_i + \sum_{i=1}^{168} f_i \times (E_{\text{CH}_4_{\text{wetland}}})_i$$

(1)

$$\text{CO}_2_{\text{sim}} = \text{CO}_2_{\text{bg}} + \sum_{i=1}^{168} f_i \times (E_{\text{CO}_2_{\text{bio}}})_i + \sum_{i=1}^{168} f_i \times (E_{\text{CO}_2_{\text{anthro}}})_i$$

(2)

where CH$_4_{\text{sim}}$, CH$_4_{\text{bg}}$, $f_i$, $E_{\text{CH}_4_{\text{anthro}}}$, and $E_{\text{CH}_4_{\text{wetland}}}$ represent the simulated atmospheric CH$_4$ concentration, background, footprint, enhancement contributed by anthropogenic and wetland flux (Equation 1), respectively. The number of 168 hr indicates air particles are released and tracked for 7 days. The same as CH$_4$ for CO$_2$ in Equation 2.

Figure 1. (a) Land use map for Taiyuan-Jinzhuang Metropolitan (TJM) area and the figure for tall tower, (b) domain setups in WRF-STILT model, and meteorological fields comparisons for (c) 2-m air temperature, (d) wind speed, (e) wind direction, (f) relative humidity, (i) downward shortwave radiation. Note D-STILT: STILT model domain, D1: outer nested domain in WRF model, D2: inner nested domain in WRF model, red symbol “×” and “Δ” represent tall tower location and WLG observation site, respectively, other two sites of Beijing (BJ) and Xianghe (XH) are also labeled.
Following our previous work (Hu et al., 2019, 2021), the CH$_4$ and CO$_2$ background values were derived from weekly sample observations at the Waliguan site (hereafter WLG, 36°17′N, 100°54′E; https://www.esrl.noaa.gov/gmd/dv/data/ last accessed: 2 January 2022). The site location is shown in Figure 1b. Here, the CCGCRV curve fitting algorithm (see Thoning et al. (1989) as recommended by NOAA, https://www.esrl.noaa.gov/gmd/ccgg/ml/crvfit/crvfit.html) was used to calculate the hourly background values from the weekly observations

\[ y = a_0 + a_1t + a_2t^2 + \sum_{n=1}^{4} [c_n \times \sin(2\pi t) + \varphi_n \times \cos(2\pi t)] \]

where \( y \) represents the CH$_4$ and CO$_2$ observations, \( t \) is time, \( a_0, a_1, a_2, c_1, c_2, \varphi_1, \varphi_2, c_3, c_4, \varphi_3, \varphi_4 \) are regression parameters. Here, we note that only one site has been selected to characterize the influence of background CH$_4$ and CO$_2$ concentrations for the atmospheric inversions. The associated uncertainties for this approach are discussed below.

### 2.2.1. Atmospheric Transport Model

The Stochastic Time-Inverted Lagrangian Transport model (hereafter STILT model) was used to simulate the hourly concentration footprint (Fasoli et al., 2018; Lin et al., 2003), which is estimated by releasing air particles at the receptor and tracking their movement backward in time and represents the sensitivity of the upstream source area to observations at the tower site (or receptor). Specifically, the footprint was calculated from the integration of released air particles and calculating their residence time in each grid box within the planetary boundary layer. Using the STILT model, we released 500 particles for each hour and tracked each particle location back in time for 7 days or until the particle left the rectangle domain (hereafter D-STILT, Figure 1b), which is centered at Shanxi province and covers an area with length around 2,000 km. The simulated hourly concentration footprint contains 7 days of atmospheric transport information, and represents the dynamic changed 168 footprints after 500 particles were released at the observation site. The WRF version 4.2.2 model was used to simulate high spatial-temporal resolution of meteorological fields to drive the STILT model. Two nested domains and a two-way feedback option were applied in WRF, with a spatial resolution of 27 km (outer nested domain, hereafter D1, Figure 1b) and 9 km (inner domain, hereafter D2), respectively. The outer nested domain covers eastern and central China, and the inner domain covers northern China, which was the largest anthropogenic CH$_4$ and CO$_2$ hotspot in China. The physical schemes used in the WRF model are the same as those used in our previous studies (Hu et al., 2019, 2021).

### 2.2.2. The a Priori CH$_4$ and CO$_2$ Flux

The a priori CH$_4$ and CO$_2$ fluxes are needed to simulate a priori enhancements that contain both anthropogenic and natural sources. The natural CH$_4$ source is mainly from wetlands, and biological CO$_2$ flux was emissions from biomass burning and net ecosystem exchange (NEE). The Emission Database for Global Atmospheric Research (EDGAR) provides anthropogenic CH$_4$ and CO$_2$ emissions at monthly spatial resolution with 0.1°. We used the most recent v6.0 inventory for the year 2018 (Crippa et al., 2019; Janssens-Maenhout et al., 2020). Wetland CH$_4$ flux is from the WetCHARTs ensemble mean with 0.5° spatial resolution using monthly averages (Bloom et al., 2017). The EDGAR CH$_4$ inventory contains 22 different categories. The main sources are fugitive from coal mining (PRO_coal), waste water handling (WWT), and agricultural soils (AGS, mainly rice paddies). Considering that wetland CH$_4$ flux in WetCHARTs includes rice paddies (main source as AGS) as one wetland type, we exclude AGS in EDGAR and assume that WetCHART contains all of the anthropogenic and natural wetland CH$_4$ emissions. Biomass burning CO$_2$ flux and NEE are from the Carbon Tracker assimilation system with 3-hourly averages and 1° spatial resolution. The EDGAR anthropogenic CO$_2$ emissions contain 28 different categories. The main sources include power industry (ENE), combustion for manufacturing (IND), nonmetallic minerals production (NMM), and energy for buildings (RCO). Temporal profiles of the main categories were applied to CO$_2$ to derive hourly varied emissions (Hu et al., 2022). No temporal profiles were applied to CH$_4$ emissions considering that coal mining accounted for >90% of CH$_4$ emissions in the study region and does not have diurnal variations.

The enhancements contributed by each category and the source area are tracked as separate tagged tracers in the WRF-STILT framework. To quantify CH$_4$ and CO$_2$ enhancement contributions from different source regions, enhancement from all 11 cities in Shanxi province and surrounding four provinces as Inner Mongolia, Hebei, Henan, and Shaanxi (Figure 1b) will be simulated separately. To provide robust evaluations of bias in a priori
inventories, we also used more available inventories to compare with EDGAR v6.0 and posteriori emissions. They included Two different EDGAR versions (v4.0 and v4.3.2) and another independent coal mining CH$_4$ inventory (0.1° × 0.1°, Scarpelli et al., 2020), and four other different anthropogenic CO$_2$ emissions (ODIAC version 2020b (1 km × 1 km, Oda et al., 2018), PKU-CO$_2$ (0.1° × 0.1°, Liu et al., 2015), Carbon Tracker 2019B (1° × 1°, Peters et al., 2007), MEIC (0.25° × 0.25° Li et al., 2017)) will also be compared with EDGAR v6.0. Note only the year 2014 is available for PKU-CO$_2$ inventory, and year 2018 is available for other inventories.

2.2.3. Representation Error Analysis

Representation error is defined as the model setup mismatch versus the true field conditions for both horizontal and vertical resolutions (Agusti-Panareda et al., 2019; Wang et al., 2022). It also contains the observed representation error that is defined as the mismatch between point measurements and grid-cell-averaged values in models, which is usually caused by coarse spatial resolution of emissions and driving meteorological files (Gerbig et al., 2003). For example, aggregation errors are caused when aggregating heterogeneous fluxes from finer spatial resolutions into a coarser resolution. It can be expressed as the simulated enhancement difference in emissions resulting from using different spatial resolutions (Turner & Jacob, 2015; Zhao et al., 2009). The representation error can have a large influence when emissions are heterogeneously distributed, where the nearby higher (or lower) emissions in finer resolution inventory are smoothed to the regional average and can lead to a corresponding underestimation (or overestimation) of simulated enhancements (Wang et al., 2022). Theoretically, a higher spatial resolution flux map and corresponding footprint can better represent the inhomogeneity of point sources of anthropogenic emissions. Aggregation errors occur in urban areas, especially during the nocturnal period (0:00–06:00, local time) for stable boundary layer stratification. Under these conditions, atmospheric CH$_4$ (and CO$_2$) enhancements can be more easily influenced by local emissions. Here, we will analyze the aggregation errors and quantify its influence on atmospheric concentration and emission inversions in Section 3.3.

2.3. Inverse Modeling: Scale Factor Bayesian Inversion (SFBI) Method

We applied an SFBI method to interpret the tower-based CH$_4$ (and CO$_2$) observations in terms of quantitative constraints on their emissions. This Lagrangian inverse approach is an efficient way to evaluate bottom-up inventories, based on atmospheric CH$_4$ and CO$_2$ concentration observations and simulation of these concentrations by using these a priori inventories (Kim et al., 2013; Miller et al., 2008). The relationship between observed and simulated tower CH$_4$ and CO$_2$ concentrations can be expressed as follows:

\[ y = K \Gamma + \epsilon \]  \hspace{1cm} (4)

where \( y \) is vector of observed enhancement (concentration minus background) for target gases, \( K \) corresponds to simulated enhancements from different categories or tagged area or both, \( \Gamma \) consists the a posteriori scaling factors (hereafter SFs) for each source components in \( K \), \( \epsilon \) is the observing system error.

The optimal way to get a posteriori SFs is to minimize the cost function \( J(\Gamma) \), which represents the error weighted mismatch between observed and simulated CH$_4$ (and CO$_2$) enhancements and mismatch between the a priori and posteriori SFs, the cost function \( J(\Gamma) \) is expressed as

\[ J(\Gamma) = \frac{1}{2} \left[ (y - K \Gamma)^T S_e^{-1}(y - K \Gamma) + (\Gamma - \Gamma_0)^T S_a^{-1}(\Gamma - \Gamma_0) \right] \]  \hspace{1cm} (5)

where \( S_e \) and \( S_a \) are the observational and a priori error covariance matrices, and \( S_e \) consists of measurement and model errors. Each element of \( \Gamma_0 \) is equal to 1. Therefore, the solution for minimizing this cost function and obtaining the posteriori SFs is to solve \( \nabla J(\Gamma) = 0 \), more details can be referred in Kim et al. (2013) and Hu et al. (2019, 2022). The “averaging kernel” will be used to quantify the sensitivity of the retrieved emissions to their true values (Kim et al., 2013), with value of 1 representing the highest contribution and sensitivity, and value of 0 of the lowest.

2.4. Constructions of Error Covariance in Bayesian Inversion Approach

We applied the same approach as previous studies (Chen et al., 2018; Griffis et al., 2017; Hu et al., 2013, 2019, 2022; Kim et al., 2013), which did not consider the off-diagonal elements of the error variance-covariance matrix. The
error covariance and state vector metrics in SFBI approach for observations \((S_o)\) and a priori values \((S_a)\) are specified as follows. The measurement errors lie in the accuracy of calibration strategy and background, where enhancements \((\gamma)\) are evaluated instead of concentration observations. Because our measurement system has high precision and was calibrated every 2 hr, here measurement errors mainly came from \(\text{CH}_4\) and \(\text{CO}_2\) background. Based on comparisons of enhancements and background concentrations for \(\text{CO}_2\) and \(\text{CH}_4\) with different background sites (Section S5 and Table S1 in Supporting Information S1), a 10% relative uncertainty is assigned to measurement error. The model error contains two primary sources as representatives of released particles numbers \((S_{\text{particle}})\) in STILT model, and uncertainties in meteorological field simulations \((S_m)\) especially from wind speed, wind direction, and boundary layer height (hereafter PBLH). Following previous work, an uncertainty value of 13% was assigned to \(S_{\text{particle}}\) \((\text{Chen et al., 2018; Hu et al., 2022; Kim et al., 2013})\), and 20% was assigned to \(S_{\text{met}}\) based on comparisons with midday PBLH with Carbon Tracker simulations and evaluations of wind speed and directions \((\text{Gourdji et al., 2018; Kim et al., 2013})\).

Note our WRF-STILT model setup can simulate enhancement contributions from different categories and regions as displayed below, but considering the limited ability of SFBI approach in resolving SFs for both categories and regions, we only derive posteriori SFs for the main emission categories following previous studies \((\text{Griffis et al., 2017; Hu et al., 2019, 2022})\). For the uncertainty in the a priori flux, previous studies have compared different inventories and found the anthropogenic \(\text{CH}_4\) differences were evident between 50% and 200% \((\text{Han, Zeng, Oda, et al., 2020; Han, Zeng, Zhang, et al., 2020; Lin et al., 2021})\), and \(\text{CO}_2\) emission differences in Chinese industrial cities can reach to 200%. Because coal mining is the largest contributor to \(\text{CH}_4\) in TJM area, we applied a 3-element (coal mining, reset anthropogenic sources, and wetland) state vector in for constraining \(\text{CH}_4\) emissions. Considering the potential uncertainty of city scale coal mining \(\text{CH}_4\) emissions can reach to 200%, and also to analyze whether the derived posteriori SFs for coal mining \(\text{CH}_4\) emissions were sensitive to a priori uncertainty, we applied three uncertainties combination cases, where the uncertainty were 50% (Case 1), 100% (Case 2), 200% (Case 3) for coal mining, respectively. And all three cases assigned 200% for the rest anthropogenic categories and wetland which together only accounted for <10% of regional \(\text{CH}_4\) total emissions.

To constrain anthropogenic \(\text{CO}_2\) emissions, we first applied 4-element (ENE, NMM, others, biological sources) state vector and 3-element (ENE, others, biological sources) state vector. But our preliminary test results in TJM area indicated there should be missing of large \(\text{CO}_2\) hotspots that have not been considered in EDGAR v6.0 inventory. Besides, results from above setups also indicated the spatial distributions of different anthropogenic \(\text{CO}_2\) categories are not accurate. All these factors lead to the condition that there is not enough information from the observations to resolve many elements and the exact locations of missing \(\text{CO}_2\) sources, and anthropogenic \(\text{CO}_2\) categories cannot be resolved by single factor Bayesian inversion approach in this study. Here, we finally choose to use a 2-element (anthropogenic sources, biological sources) state vector following the study by Hu et al. (2013) and Kim et al. (2013), which does not require the partition of all anthropogenic emissions into different sources. We found that the SFs were not sensitive to a priori uncertainty and finally choose 200% for anthropogenic \(\text{CO}_2\) emissions and 100% for biological flux based on the concentration comparisons in Section 3.3.

3. Results and Discussions

3.1. Observed \(\text{CH}_4\) and \(\text{CO}_2\) Concentration Variations

We first evaluated the WRF model simulated meteorological fields (including \(T_{2\text{m}},\) \(\text{RH},\) \(S_z,\) \(\text{WS,}\) and \(\text{WD}\) at 10 m) with observations at the Xiaodian station (Figure 1c). The comparisons showed relatively good WRF model performance (see Table S2 in Supporting Information S1) and indicate that the WRF simulated meteorological fields can be used to drive the STILT atmospheric transport for \(\text{CH}_4\) and \(\text{CO}_2\). Figure 2a shows hourly concentration observations of both \(\text{CH}_4\) and \(\text{CO}_2\) and reveals relatively strong seasonal patterns with \(\text{CH}_4\) and \(\text{CO}_2\) reaching maximum values from November to January. This pattern is likely driven by a combination of higher emissions, lower PBLH variations, and more stable boundary layer stratification during this period. Here, the local home heating period general occurred from 1 November 2018 to 31 March 2019, indicating higher \(\text{CO}_2\) emissions from coal and natural gas combustion. The seasonal concentration averages (enhancements are shown in parentheses) for \(\text{CH}_4\) were 2466.8 ppb (555.0 ppb), 2562.2 ppb (624.1 ppb), 2616.7 ppb (692.9 ppb), 2701.6 ppb (777.8 ppb) in spring, summer, autumn, and winter, respectively; and were 436.9 ppm (22.6 ppm), 431.4 ppm (26.1 ppm), 453.5 ppm (44.6 ppm), 465.1 ppm (51.8 ppm) for \(\text{CO}_2\) during the same periods. The results demonstrate that \(\text{CO}_2\) enhancements in autumn and winter were >2 times that in spring. Further, these data also indicate that the
seasonality in CO₂ flux was larger than for CH₄ when considering that both gases were subject to similar atmospheric transport processes and PBLH variations.

Note the enhancement ratio between two different gases is considered as their emission ratios (Tohjima et al., 2020; Turnbull et al., 2011, 2015; Vardag et al., 2015), we further displayed the relationship between observed midday CH₄ and CO₂ enhancement in Figure 2b, they show strong linear relationship \( R = 0.78, P < 0.001 \) with slope of 12.2 (±0.7) (ppb/ppm). The linear slope indicates the emission ratio between CH₄ and CO₂ was around 0.0122, this slope will be compared with model simulations to examine bias in the a priori emissions in Section 3.6.

To evaluate whether the CH₄ and CO₂ concentration was comparable with other cities, we further compared the annual averages of CH₄ and diurnal averages of CO₂ with some typical urban area in China (Beijing, Nanjing, Xianghe) and other countries including three cities (London, Baltimore/Washington, Los Angeles) for CH₄ and four cities for CO₂ (Salt Lake City, Los Angeles, Boston, and Indianapolis, Figure S2 and Section S2 in Supporting Information S1). The comparisons illustrate midday CO₂ concentration at our observation site is one of the

Figure 2. (a) Time series of hourly CO₂ and CH₄ observations, and (b) scatter plots of observed CH₄ and CO₂ enhancement, each point is midday averages (12:00–18:00, local time), color bar indicates the corresponding months for each point.
highest and only slightly lower than Beijing. Given the lack of observed diurnal CH₄ variations reported in previous studies, we compared our results to annual averages. In these previous studies, all CH₄ observations were calibrated frequently with standard gases that can be traceable to NOAA/GML (NOAA Global Monitoring Laboratory) standards. We believe their absolute accuracy is better than 1 ppbv, which can be the largest systematic bias. Considering the magnitude of CH₄ enhancements at our site is about 2–3 times larger than other major cities, the systematic bias could not explain the large enhancement difference between our study and above-mentioned cities. These comparisons indicate strong CO₂ and CH₄ emissions from the TJM area compared to other major cities.

3.2. The a Priori Flux and Simulated Source Footprint

3.2.1. The a Priori CH₄ and CO₂ Flux

The a priori anthropogenic CH₄ emissions from EDGAR v6.0 and WetCHART wetland emissions are displayed in Figures 3a and 3c. The main anthropogenic sources are located in the central and southern Shanxi province. Wetland emissions were relatively low when compared with the anthropogenic emissions. We also show the satellite-based CH₄ column concentration data in Figure 3b. These data show strong coherency ($R = 0.36$, $P < 0.001$, Figure S6 in Supporting Information S1) with the EDGAR v6.0 CH₄ inventory in Shanxi. The spatial distribution of the high CH₄ column concentration data (>1,900 ppb) is also consistent with the location of coal mines (Figure S3 in Supporting Information S1), and indicates that coal mining CH₄ dominated the regional atmospheric CH₄ distributions. The a priori anthropogenic CO₂ emissions from EDGAR v6.0 and ODIAC (other inventories not shown), and NEE are displayed in Figures 3d–3f, respectively. The annual averaged NEE was negligible when compared with the anthropogenic emissions. However, we should also note the hourly NEE may also be comparable with anthropogenic emissions for some specific days in summer which will be displayed below. Further, CO₂ emissions from within the TJM area represent a significant hotspot (i.e., values are in the upper 10% percentile).

Considering the EDGAR inventories have been widely used as a priori emission in many previous studies, with the highest spatial resolution in China for both CH₄ and CO₂ emissions, and the spatial distributions of satellite-based CH₄ concentrations also showed high consistence with EDGAR emissions in Shanxi province, indicating the general good spatial distributions of anthropogenic CH₄ especially for coal mining. Hence, only EDGAR v6.0 inventory was used in inversion for both CO₂ and CH₄ emissions in this study.

The EDGAR v6.0 anthropogenic CH₄ and CO₂ emissions from the top five emission sources for both TJM area and Shanxi province are displayed in Table 1. We found that CH₄ emissions from coal mining accounted for >90% of the total anthropogenic CH₄ emissions in both the TJM area and Shanxi province. The second most important source was waste water, which accounted for <4% of emissions. These a priori emissions proportions indicate that the CH₄ enhancements at our observation site are dominated by coal mining CH₄ emissions. We also calculated the coal mining CH₄ emissions in Shanxi province using a different inventory available for the year 2016 by ScarPELLi et al., (2020), where the emission was $7.4 \times 10^6$ tons and 88% of $8.4 \times 10^6$ tons for EDGAR v6.0 in the same year. Previous studies have found large discrepancies in the coal mining CH₄ emissions and its mean EF for Shanxi province, with observed mean EF values ranging from 5.6 to 12.7 m³ CH₄/ton coal and largely varied from 3.3 to 22.1 m³ CH₄/ton coal for different coal quality types within Shanxi (Table S3 in Supporting Information S1). We derived the EFs by using three different EDGAR versions (v4.0, v4.3.2, v6.0) in the year 2010 (the most recently released versions for all three EDGAR versions). The EFs decreased from 16 m³ CH₄/ton coal for v4.0 to 12 m³ CH₄/ton coal for v6.0 in the same year 2010, note most of EFs applied by EDGAR inventories are IPCC recommended default values, but EGDRA will update them for the regions with localized EFs (Solazzo et al., 2021). Hence, we assume that such discrepancies will lead to large uncertainty in the estimated CH₄ emissions by bottom-up methods.

The EDGAR v6.0 CO₂ emissions from the power industry accounted for ~60% of the total anthropogenic CO₂ emissions in TJM area and 50% for Shanxi province, indicating large emissions from coal burning-based power stations. The second and third largest sources were cement production and combustion for manufacturing. We also calculated the annual anthropogenic CO₂ emissions from EDGAR v6.0, ODIAC, MEIC, and PUK-CO₂ inventories, which were $5.1 \times 10^8$ tons, $3.2 \times 10^8$ tons, $4.9 \times 10^8$ tons, and $3.9 \times 10^8$ tons for Shanxi province, respectively. The anthropogenic CO₂ emissions in TJM area were $9.3 \times 10^7$ tons, $6.2 \times 10^7$ tons, and $7.8 \times 10^7$ tons from EDGAR v6.0, ODIAC, and PUK-CO₂. Here, the MEIC and Carbon Tracker emissions are not shown because of its coarse spatial resolution at the city scale. As evaluated by Han, Zeng, Oda, et al. (2020) and Han, Zeng, Zhang, et al. (2020) for the anthropogenic CO₂ emissions in Shanxi province based on seven different inventories,
emissions varied from ∼3.0 × 10^8 tons to ∼7.5 × 10^8 tons in year 2012 by 2.5 times, which also indicated relatively large uncertainty in Shanxi province and even larger uncertainty at the city scale.

3.2.2. Tower-Based Concentration Footprint and Source Regions

The WRF-STILT simulated footprint represents potential source area variations for each hourly observation. The climatological source footprint was obtained by averaging the hourly values over the whole year (Figures 3g and 3h).

Figure 3. (a) EDGAR v6.0 CH₄ emissions, units: log₁₀(mol m⁻² s⁻¹), (b) column-averaged CH₄ concentration from Sentinel satellite retrievals, units: ppb, (c) wetland CH₄ emissions, units: log₁₀(mol m⁻² s⁻¹), (d) EDGAR v6.0 CO₂ emissions, units: log₁₀(mol m⁻² s⁻¹), (e) ODIAC CO₂ emissions, units: log₁₀(mol m⁻² s⁻¹), and (f) annual averaged net ecosystem exchange (NEE), units: mol m⁻² s⁻¹, (g) annual average of hourly footprint for CO₂/CH₄ observation site, (h) enlarged figure of footprint.
Table 1  
The a Priori Anthropogenic CO$_2$ and CH$_4$ Emissions for TJM Area and Shanxi Province in Year 2018 From EDGAR v6.0 Inventory

<table>
<thead>
<tr>
<th></th>
<th>CH$_4$ ($\times10^3$ tons)</th>
<th>CO$_2$ ($\times10^6$ tons)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PRO_COAL       WWT          SWD_LDF</td>
<td>RCO</td>
</tr>
<tr>
<td>Taiyuan</td>
<td>642.1          27.8          19.6</td>
<td>9.1</td>
</tr>
<tr>
<td>Proportion</td>
<td>90.4%          3.9%          2.8%</td>
<td>1.3%</td>
</tr>
<tr>
<td>Jinzhong</td>
<td>1584.8         17.6          9.2</td>
<td>6.5</td>
</tr>
<tr>
<td>Proportion</td>
<td>95.5%          1.1%          0.6%</td>
<td>0.4%</td>
</tr>
<tr>
<td>Shanxi</td>
<td>8325.7         260.3         142.0</td>
<td>92.6</td>
</tr>
<tr>
<td>Proportion</td>
<td>91.9%          2.9%          1.6%</td>
<td>1.0%</td>
</tr>
</tbody>
</table>

Figure 4. Monthly mean enhancement contributions from Taiyuan, Jinzhong, Shanxi province, and other regions for (a) CH$_4$ and (b) CO$_2$.

and 3h, enlarged on the right). The most intense source footprint area (i.e., values greater than 10$^{-3}$ ppm m$^2$ s/μmol (displayed in yellow)) includes the TJM area. Further, the tower-based observation site is located in the middle of TJM area, indicating that our CH$_4$ and CO$_2$ observations are representative of the anthropogenic signals originating from within the study domain (Figures 1a and 1b).

To quantify the source area representation of observed CH$_4$ and CO$_2$ concentration, we calculated how many hours were needed to accumulate different proportions of total enhancements, where simulated CO$_2$ enhancements were used because they reflect the hourly variation of emissions. The temporal dynamics of the cumulative anthropogenic CO$_2$ enhancements are shown in Figure S4 in Supporting Information for all four seasons and are discussed in details in Section S3 in Supporting Information S1. There are key differences in the time of day and time of year required to accumulate the same enhancement proportion. The average time to accumulate a 50% of total enhancement was 3 hr in near midnight and 6–12 hr near midday for all seasons. We further used annual wind speed average and the accumulated time to calculate the source area by atmospheric transport, where observed wind speed of 2.6 m/s in the midnight and 3.5 m/s in the midday were used. Results indicate that the 50% enhancement source area was with radius of 28 km in midnight and 76–152 km in midday. The midday radius was comparable with the size of TJM area, where the length and width varied by 100–200 km, but the 28-km radius for midnight only represents local emission sources.

To further quantify the enhancement contributions from different source regions (i.e., TJM area, Shanxi province, and other provinces), enhancements were tracked as separate tagged tracers within the WRF-STILT framework. The daily variations are displayed in Figure S5 in Supporting Information S1 and the monthly mean values are shown in Figures 4a and 4b. We found that the CH$_4$ enhancement (CO$_2$ shown in parentheses) proportions had a strong seasonal pattern, which increased from the lowest value 74.6% (59.6%) in summer to a maximum of 93.1% (82.3%) in winter for Shanxi province's contributions and from 38.2% (41.0%) to 62.6% (64.3%) for the TJM area. The annual mean CH$_4$ enhancement proportions were 85.0%, 1.3%, 1.7%, 2.6%, and 1.9% for Shanxi, Inner Mongolia, Hebei, Henan, and Shanxi province, respectively. The CO$_2$ enhancement proportions were 71.6%, 5.2%, 4.5%, 4.2%, and 3.7% for corresponding provinces. The TJM area accounted for 51.8% (54.5%) of the CH$_4$ and CO$_2$ enhancements. Here, Taiyuan accounted for 36.8% (35.3%) and Jinzhong accounted for 15.0% (19.2%).

3.3. Representation Error Analysis

Before constraining the anthropogenic CH$_4$ and CO$_2$ emissions, we assessed the representation errors in the model simulation and their potential influence on derived posteriori SFs. Instead of using all-day observations in the
atmospheric inversion study, previous studies have found that atmospheric transport simulations are likely more reliable for greenhouse gas concentration during midday compared to early morning or nocturnal periods. For example, the evaluation of five different Eulerian atmospheric transport models with multiple observation sites indicated that the model performed better in the daytime because of well-mixed atmospheric conditions than under stable nighttime conditions (Geels et al., 2007). In Section S4 and Figure S7 in Supporting Information S1, we quantified and discussed the representation error caused by both spatial resolution (aggregation error) and vertical resolution between midday and midnight. We first calculated the relative difference of simulated midday anthropogenic CO₂ enhancement between using 0.1° and 1° EDGAR v6.0 CO₂ emissions, which were 7.9%, 4.4%, 19.1%, and 20.2% for spring, summer, autumn, and winter, respectively. Similarly, the differences were 0.8%, 26.9%, 11.5%, and 19.0% for simulated anthropogenic CH₄ enhancements. The relative difference of simulated midnight anthropogenic CO₂ enhancements between using 0.1° and 1° EDGAR v6.0 emissions were 1.5–2 times and were up to 6 times for CH₄ enhancement simulations. The midday relative difference was much smaller than using midnight data and indicate that the midday data can substantially minimize the aggregation error. Note although much finer spatial resolution of emissions (i.e., 1-km spatial resolution of ODIAC inventory) can better represent spatial distributions of CO₂ emissions, it still needs the same spatial resolution of WRF-STILT model simulated footprints, which is >100 times of computation resources than using 0.1°. Here, we only compared the aggregation errors by using 0.1° and 1° to keep the balance between deriving conclusions and computation cost.

The relatively large differences at midnight are likely caused by: (a) there were strong point sources within a 1° radius (~100 km) of the observation site (Figure 1). Aggregation from 0.1° to 1° emissions will act to smooth higher emissions that is much further/closer with observation site to regional averages; (b) lower PBLH and stable stratification at midnight will assign larger footprint weight to emissions sources that are closer to the tower; (c) the STILT model assumes that all emissions are surface sources. However, there can be strong point sources (i.e., stack height and subsequent plume rise height for coal burning power station CO₂ and coal mining CH₄ emissions), which will lead to overestimation of simulated concentrations when the PBLH is below the emission height. But the emission height is less important when atmospheric mixing is strong during midday and the PBLH is considerably higher than the stack height (Brunner et al., 2019). A recent study comparing emissions with or without using emission height in WRF-STILT found that their differences were negligible near midday (Maier et al., 2022). Based on these analyses, we conclude that CH₄ and CO₂ observations in the midday have smaller representation errors and will be used here for comparison with observed tower concentrations to constrain the emissions.

3.4. Comparisons Between Observed and Simulated Midday CH₄ and CO₂ Concentrations

Time series of observed and simulated midday CH₄ and CO₂ concentrations are shown in Figures 5a and 5b. In general, the annual midday averages for observations and simulations were 2340.9 and 2301.5 ppb for CH₄, 436.7 and 422.7 ppm for CO₂. The statistical analyses of R (correlation coefficient), RMSE (root mean square error), and MB (mean bias) were 0.41, 384.7, and 39.5 ppb for CH₄ and were 0.69, 22.4, and 14.0 ppm for CO₂. The lower R for CH₄ than CO₂ and the more scattered enhancements found in Figures 5c and 5d indicate the spatial heterogeneity of CH₄ emissions is not as well resolved for CO₂ in the EDGAR (0.1° spatial resolution) inventory. More detailed and updated information are needed to further improve its spatial resolution. Previous studies also found large bias in the numbers and locations of coal mining for present inventories in China (Lin et al., 2021; Zhu et al., 2017). The simulated CO₂ concentrations (and enhancements) were much smaller than the observations as displayed in Figure 5b. CH₄ did not show such overestimate/underestimate. Regression slopes of observed and simulated enhancements were 1.12 ± (0.11) for CH₄ and 0.55 ± (0.04) for CO₂ (Figures 5c and 5d), also supporting that the simulated CO₂ enhancement was underestimated. Such large discrepancy for CO₂ could be caused by the following factors: (a) The PBLH was overestimated in the WRF-STILT model and caused the underestimation of CO₂ simulations; (b) an underestimation in CO₂ emissions in the a priori EDGAR inventory or an overestimation of CO₂ uptake by plants; and (c) a combination of both factors.

To help address these concerns, CO₂ simulations from the Carbon Tracker model was added to Figure 5b and PBLH simulations using two models were also compared in Figure S8 in Supporting Information S1. Here, we used Carbon Tracker because it is widely used to simulate global CO₂ distributions. Further, it is an Eulerian atmospheric transport model, which takes a different approach than the STILT model. Hence, the comparison of two different models with observations can provide us with a robust evaluation of potential bias in the a priori emissions and performance of model simulations. These comparisons indicate that: (a) Both WRF-STILT and
Figure 5. Time series of observed and simulated midday (a) CH$_4$ concentrations, (b) and CO$_2$ concentrations, note CO$_2$ simulations from Carbon Tracker model is also displayed, and each dot represent daily averages; (c) scatter plots between observed and simulated CH$_4$ and (d) CO$_2$ using data excluded statistical outliers (>98th percentile in simulations), uncertainty values indicate 95% confidence limits; and simulated midday enhancement components for (e) CH$_4$ and (f) CO$_2$ with anthropogenic and biological enhancements displayed.
Carbon Tracker simulations show similar daily variations with CO₂ observations, indicating these models can well simulate weather patterns and the atmospheric transport process. (b) The intermodel comparisons only show slightly higher CO₂ simulation in WRF-STILT than Carbon Tracker. This is likely because each Carbon Tracker grid cell represents averages of 2° × 3° area, such that the aggregation error will smooth out the higher emissions within the Taiyuan-Jinzhong urban area (with area of <0.5° × 0.5° in Figure 1a). The grid-cell resolution in WRF-STILT is 0.1° × 0.1° and can better resolve the spatial distribution of anthropogenic emissions. Further, because only midday averages are used, the aggregation error is minimized and cannot lead to such large underestimation. (c) The PBLH in both models showed high consistency, although the WRF simulated PBLH was lower than Carbon Tracker for winter. However, this would be expected to result in higher CO₂ concentrations and cannot explain the underestimation of CO₂ value. (d) The comparisons between observed and simulated CO₂ enhancement illustrated underestimation of simulated midday CO₂ enhancement occurred in most study period of four reasons for both models, even though the source contribution from TJM area changed from ~10% to ~100% (Figure 4 and Figure S7 in Supporting Information S1, discussed above), these analysis indicate such bias exist when air flows from different directions and represent common bias in a priori emissions at regional scale.

We also displayed the enhancement components from each category for both CH₄ and CO₂ (Figures 5e and 5f). These results show that anthropogenic sources dominated the seasonal variations compared to wetland source of CH₄ and NEE for CO₂. Simulated annual a priori CH₄ enhancement averages were 374.6 ppb for anthropogenic sources and 1.6 ppb for wetlands. Simulated annual a priori CO₂ enhancement averages were 12.7 ppm for anthropogenic sources and ~0.7 ppm for NEE. Based on the above analysis, we concluded that the underestimation of simulated CO₂ concentration throughout the year cannot be explained by PBLH biases, but is more likely caused by the underestimation of anthropogenic CO₂ emissions in the a priori inventory. In the next section, we will use the SFBI method and atmospheric CH₄ and CO₂ observations to constrain their emissions and derive the posterior emissions.

### 3.5. Constraint of Anthropogenic CH₄ and CO₂ Emissions

The time series of hourly wind direction Figure 1e indicate large hour-to-hour variability. Because our site is located near the center of the TJM area, the observations represent contributions from different sources with varied wind directions. Further, as shown in Figure 5a, the underestimation of CO₂ exists for the majority of the study period, indicating the underestimation of CO₂ emissions for the region. The simulated a priori CH₄ concentrations also showed relatively good agreement from March to October, and an overestimation from November to February (Figure 5b). Based on the above analyses, we have decided to constrain CH₄ and CO₂ emissions at the seasonal scale.

In the SFBI method, hourly CH₄ and CO₂ concentration observations at midday were used to constrain anthropogenic sources in each season. As discussed in Section 2, three cases of uncertainty combinations will be used to provide robust evaluation of emissions inversion. Here, the derived seasonal posteriori SFs of three cases for coal mining, other anthropogenic categories, and wetlands are displayed in Table S4 in Supporting Information S1. The high averaging kernel (>0.8) for coal mining and low averaging kernel (<0.1) for wetland indicates atmospheric CH₄ observations were more sensitive to coal mining emissions than wetland. To provide a robust constraint on the SFs, we averaged the SFs from all three cases to estimate the final posteriori SFs. The uncertainties were represented by the range of SF estimates. The posteriori SFs were 0.94±0.06, 1.08±0.08, 0.80±0.23, and 0.80±0.31 for coal mining in spring, summer, autumn, and winter, respectively; and were 1.86±0.07, 1.44±0.09, 1.02±0.15, and 0.96±0.08 for the other categories.

The posteriori results only show slight seasonality for anthropogenic CH₄ emissions, because this area is dominated by fugitive CH₄ from coal mining. Further, home heating in winter only played as small role in the total regional CH₄ emissions. After resimulating CH₄ concentrations by applying these posteriori CH₄ emissions, the regression slopes for CH₄ concentrations decreased from 1.12±0.11 to 0.92±0.08, and the y-intercept changed from −354.3 to 100.4 ppb (Figure 6a). The time series comparisons between a priori and posteriori CH₄ concentrations also showed that the overestimation of a priori emissions from November to February was significantly improved following the inverse optimization (Figure S9a in Supporting Information S1). Further, the RMSE decreased from 384.5 to 339.8 ppb. After applying these SFs, the derived annual posteriori anthropogenic CH₄ emissions for the TJM area was 21.9±4.4 × 10⁵ tons, which was 7.6% lower than the a priori emissions of 23.7 × 10⁵ tons. The posteriori CH₄ emission from coal mining was 20.2±4.3 × 10⁵ tons, 9.4% lower.
than the a priori emission of $22.3 \times 10^5$ tons. We then divided emission by the coal production of $1.22 \times 10^8$ tons in 2018 for this region, and the posteriori coal mining EF was $23.2(\pm 4.9)$ m$^3$ CH$_4$/ton coal. Here, the coal production was obtained from local government, which conducted provincial policies as “consolidation to large coal mines” and “phase out of small coal mines” to increase coal production efficiency (Zhang et al., 2022), but some small coal mines have been frequently reported still under operation by local news. It indicates the real coal production may be slightly higher than public data and EF can be slightly smaller than $23.2(\pm 4.9)$ m$^3$ CH$_4$/ton coal.

The derived seasonal posteriori SFs for anthropogenic CO$_2$ emissions were $1.70 \pm 0.02$, $1.60 \pm 0.01$, $1.59 \pm 0.05$, and $1.56 \pm 0.09$ in spring, summer, autumn, and winter, respectively. And SFs for NEE were $0.89 \pm 0.07$, $0.43 \pm 0.02$, $-0.40 \pm 0.31$, and $1.96 \pm 0.87$ in corresponding seasons. The posteriori SF of NEE in fall indicated the sign of a priori Carbon Tracker NEE in our study region is wrong, and also with large bias in other seasons.

The SFs for anthropogenic emissions indicate slight seasonal variations, where emissions in winter were the lowest. We hypothesize that the reduced industrial emissions from factories during the traditional Chinese spring festival resulted in decreased fossil fuel consumption. To evaluate whether these SFs improved the CO$_2$ emission estimates, we resimulated CO$_2$ concentration by using the posteriori anthropogenic/biological emissions and compared them with the CO$_2$ observations. These results indicate the posteriori CO$_2$ concentration simulations largely improved when compared with using the a priori emissions (Figure 6b). The RMSE decreased from 22.4 to 10.6 ppm, MB decreased from 14.0 to 5.6 ppm. The SFs calibrated the underestimation of simulated CO$_2$ concentrations. Here, the slope improved from 0.59 (±0.04) to 0.82(±0.04), the $y$-intercept decreased, and the scatter plots show less scatter (Figure 6b). The time series comparisons between a priori and posteriori CO$_2$ concentrations also showed that the underestimation of a priori emissions was calibrated reasonably well across the whole year following the inverse optimization (Figure S9b in Supporting Information S1). Based on these derived SFs for anthropogenic emissions, the posteriori annual anthropogenic CO$_2$ emissions for TJM area were $14.9(\pm 0.4) \times 10^7$ tons, which was 60.2% higher than the a priori EDGAR v6.0 inventory. Considering the a priori anthropogenic CO$_2$ emissions in TJM area were $9.3 \times 10^7$ tons, $6.2 \times 10^7$ tons, and $7.8 \times 10^7$ tons from EDGAR v6.0, ODIAC, and PKU-CO$_2$ inventories, our results show the underestimation was common among all available inventories, and indicate some large emission sources and hotspots are missing in recent inventories for this industrial region. Furthermore, to quantify the potential uncertainty of background concentration on emission inversions, we also compared the CO$_2$ and CH$_4$ concentration at WLG background site with other three backgrounds in different directions. The results indicated the uncertainty was within 5% at the annual averages for both CO$_2$ and CH$_4$ emissions, as discussed in details in Section S5, Table S1, and Figures S10–S12 in Supporting Information S1.
3.6. Further Evaluation of the a Priori Emissions

The comparisons of seasonally averaged diurnal averages between observations and simulations are displayed in Figure 7. The model simulations in midday illustrated similar diurnal patterns but with different amplitude for CO₂. And simulations for CH₄ showed overall consistency with the observations. The derived ratios between observed and simulated midday CO₂ enhancements were 1.88(±0.40), 1.81(±0.46), 2.26(±0.34), and 2.01(±0.31) in spring, summer, autumn, and winter, respectively, also indicating large underestimation of anthropogenic CO₂ emissions as found by SFBI method. Further, the scatter plots between CO₂ and CH₄ concentration (and enhancement) for both observations and simulations are shown in Figure 8. Each observation represents different source areas and the linear relationship illustrates the general emission ratio between CO₂ and CH₄. We found that the regression slopes were 0.085(±0.05) for observed CO₂ and CH₄ concentration ratio, and 0.082(±0.05) for their enhancement ratio. For the simulations, the slopes were 0.036(±0.02) for concentration and 0.032(±0.02) for the enhancement ratio. Both are close to the emission slope of 0.028 (±0.02) derived from the EDGAR a priori inventory (Figure 8b). These large discrepancies between observed and simulated CO₂ and CH₄ slopes also indicate large bias for the a priori anthropogenic CO₂ emissions.

3.7. Comparisons With Previous Studies

The comparison between the a priori EDGAR v6.0 and posteriori emissions is displayed in Figure 9 for TJM area. It illustrates that the posteriori emissions for coal mining and total anthropogenic CH₄ emissions were 92.4% and 90.6% of a priori EDGAR v6.0 inventory. Overall, there are only a few available atmospheric inversion studies for constraining coal mining CH₄ emissions in China. Miller et al. (2019) and Chen et al. (2022) used satellite observations for China, which included Shanxi province. Lu et al. (2021) and Zhang et al. (2022) used both satellite and tower-based observations to constrain CH₄ emissions. The spatial resolutions for these studies were 2.0° × 2.5° latitude-longitude, 0.25° × 0.3125°, 4° × 5°, and 0.5° × 0.625°, respectively. Zhang et al. (2022) used seven atmospheric CH₄ sites in China (site locations were displayed in Zhang et al. (2022)), and even the closest atmospheric CH₄ tower site (SDZ site) was ~500 km (5°) from the TJM area. The concentration footprint of a tall tower based on our WRF-STILT model simulations indicates that the 50% enhancement source area was within about 100 km. Hence, we believe that the observation site used in our study is the most sensitive to both CH₄ and CO₂ emissions within the TJM area.

Among these previous studies, only Zhang et al. (2022) provided the a priori and posteriori CH₄ emissions from different source categories. We examined the ratio between the posteriori and a priori coal + gas + oil...
CH$_4$ emissions for Shanxi province and TJM area. We believe their results for Shanxi province were mainly constrained by the satellite observations (not the tower observations i.e., when considering the footprint of the tower and sensitivity of a priori emissions). We found that the derived ratio was $\sim$1 with slight spatial variations (Figure S13 in Supporting Information). This supports relatively good agreement at the regional scale for CH$_4$ emissions. Overall, our study in the central Shanxi is the first one that combined tower-based CH$_4$ concentration observations to constrain CH$_4$ emissions from coal mining in China with a much finer spatial resolution of 10 km and our results showed high coal mining EF at TJM area and no significant bias in anthropogenic total and coal mining CH$_4$ emissions.

To evaluate whether the derived coal mining EFs for these areas was comparable with other regions in China, we have derived the a priori EF for Shanxi and all of China and compared them with previous studies (Table 2 and Table S1 in Supporting Information S1). The posteriori EF for the TJM area was 23.2(±4.9) m$^3$ CH$_4$/ton coal, and was close to the high CH$_4$-content coal EF (22.1 m$^3$ CH$_4$/ton coal), but 7 times greater than the low CH$_4$-content coal EF (3.3 m$^3$ CH$_4$/ton coal) for Shanxi province as reported by Sheng et al. (2019). This posteriori EF was $\sim$2 times the Shanxi average and 3 times the national average when compared with previous “bottom-up” observations (Lin et al., 2021; Liu et al., 2021; Sheng et al., 2019), indicating large spatial inhomogeneity in coal mining EFs in China’s largest coal production province.

**Figure 8.** (a) Comparisons of observed and simulated CH$_4$-CO$_2$ enhancements, (b) emission ratio between CO$_2$ and CH$_4$ with different radius of source area in EDGAR v6.0 inventory, (c) comparisons of observed and simulated CO$_2$:CH$_4$ concentrations, and (d) CO$_2$:CH$_4$ enhancement.
Some previous studies that conducted local observations in China reported EFs in Shanxi were 6.5–7.0 m$^3$ CH$_4$/ton coal (Gao et al., 2020; Zhu et al., 2017), which is about half the EF in the EDGAR v6.0 inventory. Although the comparisons in the TJM area suggested overall good agreement at the local scale in the EDGAR v6.0 inventory, our single observation site cannot represent coal mining EFs for the entire Shanxi province because of its high spatial inhomogeneity. Besides, as mentioned before that the available CH$_4$ EFs from underground coal mining for China (including Shanxi province) were mainly based on nationwide measurements prior to 2006 from ~800 underground mines, and were highly spatially heterogeneous (Sheng et al., 2019; Zhu et al., 2017). Besides, the observations in a few mines may not be suitable to represent averages at regional or local scales as found in this study. Furthermore, these previous EFs cannot reflect EFs changes by coal mining technology and government policy (i.e., “consolidation to large coal mines” and “phase out of small coal mines”) related EFs changes in recent years (Zhang et al., 2022).

We highlight the area with low CH$_4$-content coal should also be evaluated to better understand the spatial inhomogeneity of coal mining CH$_4$ EFs. A denser atmospheric CH$_4$ observation network in Shanxi province would be need to better resolve its spatial inhomogeneity for coal mining CH$_4$ EFs. A new CH$_4$ observing network is under construction and will be implemented in our future work. Note although TROPOMI observations have not been used in CH$_4$ emission inversion, we also emphasize the use of satellite observations are essential in resolving the spatial inhomogeneity of greenhouse gas emissions especially for hotspots for CH$_4$ and CO$_2$.

As mentioned above, the total anthropogenic and coal mining CH$_4$ emissions in China are still subject to considerable debate, not only between “top-down” atmospheric inversions and “bottom-up” inventories, but also among different inventories (i.e., EDGAR v5.0, PKU-CH$_4$ v2, IEA). The inversion study by Miller et al. (2019) stated that total anthropogenic CH$_4$ emissions increased after 2000 at rate of 1.0–1.2 × 10$^6$ tons/year and kept increasing from 2010 to 2015 at a similar rate. This increase was mainly caused by coal mining emissions. However, the comparisons between six inventories illustrated the opposite decreasing trends of coal mining emissions from 2010 to 2016. The coal mining CH$_4$ emissions then reversed to increasing trends that were mainly caused by increased demand for coal instead of slightly decreasing EFs. And coal production decreased from highest 4.0 × 10$^8$ tons in 2012 to the lowest 3.4 × 10$^8$ tons in 2016 by 15%, and national mean EF only decreased from 9 m$^3$ CH$_4$/ton coal to 8 m$^3$ CH$_4$/ton coal from 2010 to 2020 (Liu et al., 2021; Peng et al., 2016; Sheng et al., 2019). Our findings suggest that the estimation of coal mining CH$_4$ emissions at the regional scale, using only a single SF and without careful treatment of inhomogeneity, can potentially bring large biases.

Contrary to CH$_4$, comparison for CO$_2$ shows that state-of-the-art inventories (EDGAR v6.0, ODIAC, PKU-CO$_2$) considerably underestimated anthropogenic emissions by 61%. For the SFBI method, the linear regression slope between observations and posteriori midday concentration was 0.82 ± 0.06. This indicates that the underestimation in the a priori inventory has not been fully calibrated. Therefore, the true emissions are expected to be slightly higher than the SFBI result. Furthermore, the comparisons also illustrated the underestimate commonly exists in all present inventories and that some large emission sources are missing in the inventories for this heavily industrialized urban area.

There are a few previous studies conducted at the city scale or regional scale (i.e., in Nanjing and Nanchang), which showed less than a 20% bias in inventories (Hu et al., 2021, 2022). The extremely large SFS for the TJM area reveal large underestimation in high-density industrialized cities. A previous study using a “bottom-up” approach that compared eight inventories in Beijing-Tianjin-Hebei region found the relative difference was twofold in almost all industrial cities. Further, emissions from the national database inventories (i.e., EDGAR, ODIAC) were 53%–75% lower than

---

**Table 2** Comparisons of Coal Mining EF for Between a Priori and Posteriori Results, Notes the a Priori Results was Derived From EDGAR v6.0 Inventory

<table>
<thead>
<tr>
<th></th>
<th>China (a priori)</th>
<th>Shanxi (a priori)</th>
<th>Taiyuan-Zhangzhu (a priori)</th>
<th>Taiyuan-Zhangzhu (posteriori)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH$_4$ emissions (×10$^5$ tons)</td>
<td>201.1</td>
<td>84.5</td>
<td>23.1</td>
<td>20.2(±4.3)</td>
</tr>
<tr>
<td>coal production (×10$^5$ tons)</td>
<td>35.46</td>
<td>8.93</td>
<td>1.22</td>
<td>1.22</td>
</tr>
<tr>
<td>EF (m$^3$ CH$_4$/ton coal)</td>
<td>7.9</td>
<td>13.3</td>
<td>26.2</td>
<td>23.2(±4.9)</td>
</tr>
</tbody>
</table>
provincial-level data estimated emissions (i.e., CH₄, CO₂, CH₄) in the high-emitting industrial cities (Han, Zeng, Zhang, et al., 2020). Our findings provide independent support from the perspective of a “top-down” approach for this large underestimation in a heavily industrialize city in Northern China, and suggest atmospheric CH₄ and CO₂ concentration observations are urgently needed to monitor the true emissions from industrial cities.

4. Conclusions

To evaluate the potential CO₂ and CH₄ emission biases at the largest coal mining CH₄ emission region in China and one of the largest CO₂ emission regions, the WRF-STILT model was coupled with the most recent EDGAR v6.0 inventory to simulate the a priori concentration and then applied SFBI method to evaluate and constrain their emissions for TJM area. We found that (a) the simulated a priori CH₄ concentrations showed much better agreement with observations than CO₂ concentrations, where the regressions slope was 1.12 ± (0.01) for CH₄ but 0.59 ± (0.04) for CO₂, indicating the a priori CO₂ emission was largely underestimated. (b) The derived postaieri anthropogenic CH₄ emissions were 92.5% of the a priori emissions, where CH₄ fugitive from coal milling accounting for 92.7% of total anthropogenic emissions. The derived CH₄ EF from coal mining was 23.2 (±4.9) m³ CH₄/ton coal, close to the default value of high CH₄-content coal but twofold greater than the province average, indicating large spatial inhomogeneity in coal mining CH₄ EFs for China’s largest coal production area. (c) Finally, the postaieri CO₂ emissions were nearly 1.6-fold of the a priori emissions, illustrating some large hotspots were missing in this industrial city.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The posterior CH₄ emissions (Zhang et al., 2022) that were compared with our study can be obtained from https://doi.org/10.57760/sciencedb.02269, the EDGAR CH₄ and CO₂ emission data set (Crippa et al., 2019) can be obtained from https://edgar.jrc.ec.europa.eu/, the ODIAC CO₂ emission inventory (Oda et al., 2018) can be obtained from http://db.cger.nies.go.jp/dataset/ODIAC_DL_odiac2018.html, the MEIC inventory (Li et al., 2017) can be obtained from http://meicmodel.org/?page_id=1917, PKU inventory for CO₂ (Liu et al., 2015) can be obtained from http://inventory.pku.edu.cn/download/download.html, the Carbon Tracker 2019B emissions (Peters et al., 2007) can be obtained from https://gml.noaa.gov/aftp/products/carbontracker/CT2019B. The STILT model (Lin et al., 2003) can be downloaded from http://www.stilt-model.org/.

Acknowledgments

Wei Xiao was supported by the National Key R&D Program of China (Grant 2020YFA0607501), Cheng Hu was supported by the Natural Science Foundation of Jiangsu Province (Grant BK2020802) and National Science Foundation of China (Grant 42105117). Shumin Wang is supported by the Shang-dianzi Background Station Open Research Project (Grant SDZ2020617) and Project on Meteorological Situation in Shannxi Province (Grant SXKMSDW2023620).

References


References From the Supporting Information


