https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

Hyperparameter-Optimized Inversion Modeling Framework for Urban CO_2 Emission Estimation and Uncertainty Evaluation

Wei Xu[1,2], Ge Ren[1,2*], Kailun Du[2], Gege Liu[1,2], Shiqi Zhao[2], Xiaoning Wang[2], Mengjuan Han[2], Hong Lin[1,2*], James Whetstone[3]
[1]Division of Thermophysics Metrology, National Institute of Metrology, Beijing100029, China
[2]Zhengzhou Institute of Metrology, Zhengzhou 450001, China
[3]National Institute of Standards and Technology, Gaithersburg, MD, 20899, USA

*Correspondence to: Ge Ren (reng@nim.ac.cn); Hong Lin (linhong@nim.ac.cn)

Abstract. Urban areas, occupying only 3% of the global land surface yet generating approximately 70% of anthropogenic carbon emissions, are critical targets for climate change mitigation. Accurate emission quantification remains challenging, as most atmospheric inversion studies neglect spatiotemporal correlations in prior fluxes and observation errors, inflating uncertainties in both the spatial distribution and magnitude of greenhouse gases. This study introduces an inversion framework integrating explicit correlation functions, hierarchical Bayesian modeling, and maximum-likelihood estimation, thereby removing reliance on empirical parameterization in error covariance matrices. Application to the urban core of Zhengzhou demonstrated superior performance over conventional approaches. In controlled experiments, the framework achieved superior precision in localizing high-emission sources, reducing root-mean-square error by 21.4% between posterior estimates and assumed true fluxes across multiple emission scenarios. Real-world validation at the two monitoring towers further confirmed the improvements under the EDGAR-based prior, with lower RMSE values (10.31 and 10.05 ppm) and higher correlation coefficients (0.91 and 0.81) than conventional benchmarks. Additionally, the relative uncertainty of posterior emissions declined by approximately 46.4% compared to the traditional method, reflecting the enhanced precision of the approach. Crucially, analysis indicated that the reduction in posterior uncertainty resulted from systematic examination of inter-grid correlations, demonstrating that spatial correlations are essential for rigorous uncertainty quantification.

1 Introduction

Anthropogenic greenhouse gas (GHG) emissions have been a primary driver of global warming since the industrial revolution (IPCC, 2023). Among these, carbon dioxide (CO_2) is considered the most significant anthropogenic GHG due to its chemical stability and long atmospheric residence time (Le Quéré et al., 2018). Although metropolitan regions cover only about 3% of the terrestrial land surface, they generate over two-thirds of anthropogenic carbon emissions globally (Cozzi et al., 2020; Seto et al., 2012). This disproportionate contribution underscores the urgent need for precise quantification of urban emissions as a prerequisite for formulating effective policy interventions and regulatory frameworks.

Two primary strategies are employed to estimate anthropogenic CO_2 fluxes. The traditional "bottom-up" emission estimation method relies on emission inventories derived from activity data and associated

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

emission factors (Andres et al., 2011; IPCC, 2006). The accuracy of these inventories is inherently constrained by uncertainties in energy consumption data, inconsistencies in combustion efficiency measurements, and variability in emission factor estimates (Gurney et al., 2021; Liu et al., 2015; Macknick, 2009). Furthermore, high-resolution, locally validated inventories are available for only a small fraction of municipalities (Lian et al., 2022; Saint-Vincent and Pekney, 2019; Smith et al., 2024). In most cases, reliance on spatial disaggregation of coarse-scale global or national datasets (e.g., EDGAR, ODIAC) is unavoidable (Janssens-Maenhout et al., 2019; Oda et al., 2018), introducing systematic biases through scale-dependent error propagation. As a complementary strategy, the "top-down" atmospheric inversion framework has evolved into a critical operational tool, synergistically integrating real-time atmospheric concentration measurements with atmospheric transport models and prior emission estimates, iteratively refining surface flux distributions through Bayesian optimization (Feng et al., 2016; Lauvaux et al., 2016; Lopez-Coto et al., 2017; Yadav et al., 2021).

The principal sources of uncertainties in atmospheric inversion systems originate from two domains: inaccuracies in prior emission inventories and diverse observation-related errors, including instrumentation limitations, transport modeling discrepancies, spatial representation mismatches, and temporal aggregation errors. In most current implementations, these error components are represented through covariance matrices, specifically, the prior error covariance matrix characterizing emission estimate uncertainties and the observation error covariance matrix encapsulating measurement system imperfections. These mathematically defined covariance structures are fundamental to the inversion optimization framework, determining relative weighting between modeled simulations and observational data during parameter estimation (Chevallier, 2015; Staufer et al., 2016). The prior error covariance matrix governs the spatial propagation of uncertainties in the prior flux field. Standard parameterizations predominantly adopt an exponential decay model to describe spatial covariance, requiring specification of two key parameters: uncertainty magnitude and correlation length. The former is typically expressed as a fixed percentage of the prior flux (e.g., 30%–100%), while the latter is empirically determined from regional characteristics (Lauvaux et al., 2022; Lauvaux et al., 2012; Nassar et al., 2017). The sensitivity of inversion results to these prescribed parameters has been well established by urban-scale studies. Nickless et al. (2019), for example, demonstrated that varying the spatial correlation length in an inversion over Cape Town led to substantial differences in both the magnitude and spatial distribution of posterior emissions, underscoring the dominant influence of this single parameter on the inversion outcome. However, reliance on prescribed exponential decay functions with empirically assigned parameters fails to sufficiently capture the spatial complexity of prior errors in heterogeneous landscapes, particularly in urban systems characterized by multiscale topography and temporally dynamic anthropogenic emissions. The observation error covariance matrix is critical for representing both spatial and temporal correlations in observation errors, with temporal correlations at individual monitoring sites being especially important. Many inversion studies oversimplify this component by assuming a diagonal covariance structure, thereby disregarding temporal error correlations despite clear empirical evidence of diurnal variability in human activity and associated emissions (Koschorreck et al., 2024; Wesloh et al., 2024). For the diagonal elements, parameter values are commonly derived from empirical estimates or sensitivity analyses (Byrne et al., 2022; Hu et al., 2022; Wu et al., 2016). While sensitivity experiments

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

are valuable for isolating the effects of individual parameters in covariance matrix configuration, they fail to account for the complex influence of multiple parameters on inversion performance (G'sell et al., 2014; Guigues, 2011). An alternative approach involves tuning covariance parameters so that the data misfit function conforms to a χ^2 distribution (Lauvaux et al., 2020; Sijikumar et al., 2023). However, such optimization-based parameterizations often yield non-unique solutions and cannot differentiate the relative contributions of prior and observation errors to the total uncertainty budget.

To overcome the limitations inherent in empirically specified covariance matrix parameters and conventional sensitivity analyses, several studies have incorporated maximum-likelihood estimation (MLE) within hierarchical Bayesian frameworks, treating parameters from both prior and observation error covariance matrices as hyperparameters. Michalak et al. (2005) introduced this approach for atmospheric trace gas inversions, demonstrating that covariance parameters can be objectively optimized from available data rather than subjectively prescribed. Subsequently, Gourdji et al. (2010), working within a geostatistical framework, further established that joint optimization of spatial and temporal covariance parameters is necessary to recover accurate posterior uncertainty bounds. Nevertheless, existing implementations remain constrained by critical methodological shortcomings. For instance, Michalak et al. (2005) reduced both covariance matrices to purely diagonal structures, thereby eliminating all spatial and temporal correlations in prior and observation errors. Wu et al. (2013) employed the Balgovind correlation model exclusively for both matrices, retaining an empirically assigned correlation length for the prior error covariance matrix. In both cases, such simplifications were deliberately introduced to lower hyperparameter dimensionality and mitigate computational complexity, facilitating convergence toward an analytical solution of the posterior objective function. However, these constraints preclude accurate representation of error structures in the heterogeneous and dynamically evolving conditions characteristic of urban emission systems.

This study introduces a novel covariance matrix construction framework, known as "optimization method", designed to address the dual challenges in atmospheric inversion: (1) constraints arising from empirical construction of prior and observation error covariance matrices, and (2) inherent deficiencies in conventional MLE approaches for parameterizing these matrices. Our methodology systematically integrates multiple correlation functions, each validated in engineering applications, to define the structures of both covariance matrices. All parameters within these functions are treated as hyperparameters, whose optimal values are jointly estimated through a hierarchical Bayesian model combined with MLE and appropriate numerical optimization algorithms. Pseudo-observation inversion experiments were performed to compare the proposed approach with the conventional empirically driven covariance construction approach, known as the "traditional method". Real-observation inversion experiments were further performed to validate the improvements achieved by the optimization method under operational conditions. Section 2 presents the methodology, including the inversion modeling scheme, pseudo-observation experiment design, and real-observation experimental setup. Section 3 reports and analyses the results from both experimental phases, while Section 4 summarizes the key findings and conclusions.

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

2. Methodology

2.1 Inversion modeling algorithm

This study employed Bayesian inversion, a well-established technique for estimating atmospheric fluxes, to quantify CO_2 emissions. Assuming Gaussian error distributions, we formulated a cost function via Bayes' theorem to simultaneously evaluate concentration discrepancies between modeled concentrations and observations, and flux deviations between prior and unknown fluxes. The posterior flux vector $\mathbf{f}$ and its associated covariance matrix $\mathbf{A}$ were obtained by minimizing the cost function with respect to the unknown fluxes (Enting, 2002; Tarantola, 2005), expressed as:

$$\mathbf{f} = \mathbf{f^b} + \mathbf{BH^T}(\mathbf{R} + \mathbf{HBH^T})^{-1}(\boldsymbol{\mu} - \mathbf{Hf^b}) \quad (1)$$

$$\mathbf{A} = (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1} \quad (2)$$

In Eq. (1), $\boldsymbol{\mu}$ is the observation vector of CO_2 mixing ratios with dimension $d \times 1$, $\mathbf{f^b}$ is the vector of the prior estimates with dimension $n \times 1$, $\mathbf{H}$ is a linear atmospheric transport operator (Jacobian matrix) with dimension $d \times n$, and $\mathbf{R}$ (dimension $d \times d$) and $\mathbf{B}$ (dimension $n \times n$) are the observation flux and prior error covariance matrices, respectively.

Traditionally, the structures of the two error covariance matrices $\mathbf{R}$ and $\mathbf{B}$ are defined by assuming $\mathbf{R}$ to be diagonal, thereby omitting all spatiotemporal correlations in observation errors, while $\mathbf{B}$ is modeled using an exponential decay function. The associated parameters are typically assigned empirically or adjusted so that the data mismatch function conforms to a χ^2 distribution. However, this adjustment procedure does not yield a unique solution and cannot quantify the relative contributions of prior and observation errors. Consequently, both the structural specification of $\mathbf{R}$ and $\mathbf{B}$ and the selection of their parameters remain heavily reliant on empirical and subjective assumptions, which inevitably introduces substantial uncertainty into inversion outcomes. The term "traditional method" is used here to describe this empirically based approach for determining the two error covariance matrices.

In the optimization approach, a hierarchical Bayesian model combined with MLE is employed to derive the corresponding negative log-likelihood function (Michalak et al., 2005; Wu et al., 2013):

$$\mathcal{L}(\boldsymbol{\theta}) = -\ln \mathrm{p}(\boldsymbol{\mu}|\boldsymbol{\theta}) = \frac{1}{2}\ln|\mathbf{R_\theta} + \mathbf{HB_\theta H^T}| + \frac{1}{2}\left(\boldsymbol{\mu} - \mathbf{Hf^b}\right)^{\mathrm{T}}(\mathbf{R_\theta} + \mathbf{HB_\theta H^T})^{-1}\left(\boldsymbol{\mu} - \mathbf{Hf^b}\right) + \mathrm{C} \quad (3)$$

where subscript $\boldsymbol{\theta}$ is the vector of hyperparameters contained in the two error covariance matrices and C represents a constant term irrelevant to the optimization. In Michalak et al. (2005), both $\mathbf{R_\theta}$ and $\mathbf{B_\theta}$ were simplified to diagonal matrices, effectively disregarding all spatiotemporal correlations in observation and prior errors—an assumption at odds with empirical reality. Wu et al. (2013), on the other hand, adopted a single Balgovind correlation function to simultaneously define the structure of both $\mathbf{R_\theta}$ and $\mathbf{B_\theta}$, while retaining one parameter as an empirically fixed value and treating the remaining parameters as hyperparameters. In both cases, reductions in structural complexity and parameter dimensionality were likely motivated by the need to reduce the computational burden associated with MLE, thereby enabling analytical derivation of hyperparameter estimates. However, these simplifications inevitably lead to incomplete representations of error covariance matrices.

To address the constraints of existing MLE approaches in characterizing $\mathbf{R_\theta}$ and $\mathbf{B_\theta}$, this study employed several established correlation functions, combined in various configurations, to construct the full structure of both error covariance matrices. All associated parameters were treated as

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

hyperparameters, significantly increasing the dimensionality and nonlinearity of the negative log-likelihood function defined in Eq. (3), rendering analytical estimation infeasible.

In this framework, the hyperparameters associated with these correlation function combinations were estimated by minimizing the negative log-likelihood function defined in Eq. (3), thereby transforming the hyperparameter estimation problem into a numerical optimization problem. Hyperparameter optimization was performed using the NLopt library within the PyTorch deep learning environment, which provided a range of nonlinear optimization algorithms (Johnson, 2014). A hybrid strategy was applied, combining the Improved Stochastic Ranking Evolution Strategy (ISRES) for global optimization and the Method of Moving Asymptotes (MMA) for local refinement. ISRES, a population-based algorithm, demonstrates superior global search capabilities in high-dimensional, non-convex parameter spaces (Runarsson and Yao, 2000), while MMA, an efficient gradient-based local optimizer, iteratively approximates the original non-convex problem with convex subproblems, guaranteeing convergence to locally optimal solutions (Svanberg, 1987). This two-stage approach, leveraging the respective strengths of evolutionary computation and convex optimization techniques for comprehensive hyperparameter tuning, was designated the "optimization method"—the procedure involving correlation function combinations and hyperparameter optimization in the inversion experiments.

2.1.1 Proposed correlation functions

(1) Exponential correlation function

The exponential (EN) correlation function is a fundamental tool for characterizing spatial or temporal correlations of random variables or physical quantities in a system, particularly when correlations decay exponentially with distance or time interval, expressed as:

$$Cov = \sigma^2 e^{-\frac{r}{l}} \tag{4}$$

where σ^2 represents the variance and l is the correlation length controlling the rate of decay. This correlation function is widely applied in spatial statistics for geostatistical modeling, time-series analysis of meteorological data, and Gaussian process regression within machine learning (Brockwell and Davis, 2002; Cressie, 2015; Williams and Rasmussen, 2006). Given its mathematical simplicity and robust performance, it has become a foundational tool for analyzing correlated data across disciplines.

(2) Exponentially damped cosine (EDC) correlation function

The EDC correlation function describes the behavior of decaying oscillations in certain physical, statistical, or time-series systems, combining exponential decay and cosine oscillations and expressed as:

$$Cov = \sigma^2 \cos(\omega r) e^{-\frac{r}{l}} \tag{5}$$

where parameter l represents correlation length, reflecting the scale of decay and ω is the angular frequency controlling the oscillation frequency of the cosine function. The EDC function is extensively used in signal processing for damped harmonic systems, geophysical data analysis for seismic or atmospheric correlations, financial time-series for cyclical trends with decay, and machine learning as a Gaussian process kernel (Box et al., 2015; Chiles and Delfiner, 2012; Shumway et al., 2000; Williams and Rasmussen, 2006). Due to its ability to capture both periodicity and damping, it is highly versatile for complex correlated systems.

(3) Balgovind (BG) correlation function

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

The BG correlation function is used to describe the spatial correlation of atmospheric turbulence. It is commonly used in geophysics, fluid dynamics, and atmospheric sciences, especially in numerical simulations, to model and analyze the spatial structure of atmospheric turbulence. It is formulated as:

$$Cov = \sigma^2 \left(1 + \frac{r}{l}\right) e^{-\frac{r}{l}} \tag{6}$$

This function is widely used to construct the spatial correlation of error covariance matrices. Its smooth attenuation properties make it effective for modeling the spatial correlations of atmospheric variables such as temperature, humidity, and wind speed. In regional assimilation systems, the BG correlation function is frequently employed to characterize local correlations, thereby improving assimilation accuracy (Balgovind et al., 1983; Bannister, 2008). Beyond atmospheric applications, it is also used in air pollution dispersion modeling, random field simulations, and spatial interpolation (Chiles and Delfiner, 2012).

(4) Squared exponential correlation function

The squared exponential (SE) covariance function is expressed as:

$$Cov = \sigma^2 e^{-\frac{r^2}{2l^2}} \tag{7}$$

The squared distance of the exponent results in slower decay with distance than the usual exponential function, producing a smoother correlation structure. The SE function is widely used in Gaussian process regression and machine learning due to its smoothness and infinite differentiability, making it ideal for modeling continuous and smoothly varying processes. It is commonly used in spatial statistics for geostatistical modeling, time-series forecasting, and Bayesian optimization for hyperparameter tuning (Stein, 1999; Williams and Rasmussen, 2006).

(5) Rational quadratic correlation function

The rational quadratic (RQ) covariance function is expressed as:

$$Cov = \sigma^2 \left(1 + \frac{r^2}{2\alpha l^2}\right)^{-\beta} \tag{8}$$

where parameters $\alpha, \beta > 0$ control the relative influence of the polynomial and exponential terms, respectively. The RQ function combines exponential and polynomial terms and is thus more flexible compared to other common covariance functions. The RQ function is particularly suited to modeling different spatial or temporal stochastic processes, especially when the most suitable form of the covariance function cannot be determined. It is widely used in Gaussian process regression and machine learning, particularly for datasets exhibiting multi-scale behaviors or long-range dependencies. It is applied in spatial statistics, time-series analysis, and environmental modeling, where it effectively captures complex patterns and correlations (Stein, 1999; Williams and Rasmussen, 2006).

2.1.2 Combinations of correlation functions of structural frameworks

Based on the five correlation functions described above, the EN, EDC, and BG functions were used to construct the observation error covariance matrix $\mathbf{R_\theta}$, while the EN, BG, SE, and RQ functions were used to build the prior error covariance matrix $\mathbf{B_\theta}$, yielding the 12 combinations presented in Table 1. The full set of 25 theoretical combinations (5 × 5) were not implemented due to numerical stability issues. In ideal experiment tests, the excluded 13 combinations produced negative values approaching zero for the determinant $\mathbf{R_\theta + HB_\theta H^T}$ in Eq. (3), violating the statistical principles of Gaussian distributions (Gelb,

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

1974). This instability likely arose from numerical errors during multiplication of the positive definite matrix $\mathbf{B_\theta}$ by the Jacobian matrix $\mathbf{H}$ and its transpose $\mathbf{H^T}$. Consequently, only the numerically stable correlation function combinations were retained to establish the structural frameworks of the two error covariance matrices.

Table 1. Combinations of correlation functions used for constructing the observation error covariance matrix R and the prior error covariance matrix B

Combination No.	Correlation function for **R**	Correlation function for **B**
1	EN: $\sigma^2 e^{-\frac{r}{l}}$	EN: $\sigma^2 e^{-\frac{r}{l}}$
2	EN: $\sigma^2 e^{-\frac{r}{l}}$	BG: $\sigma^2\left(1+\frac{r}{l}\right)e^{-\frac{r}{l}}$
3	EN: $\sigma^2 e^{-\frac{r}{l}}$	SE: $e^{-\frac{r^2}{2l^2}}$
4	EN: $\sigma^2 e^{-\frac{r}{l}}$	RQ: $\sigma^2\left(1+\frac{r^2}{2\alpha l^2}\right)^{-\beta}$
5	EDC: $\sigma^2\cos(\omega r)e^{-\frac{r}{l}}$	EN: $\sigma^2 e^{-\frac{r}{l}}$
6	EDC: $\sigma^2\cos(\omega r)e^{-\frac{r}{l}}$	BG: $\sigma^2\left(1+\frac{r}{l}\right)e^{-\frac{r}{l}}$
7	EDC: $\sigma^2\cos(\omega r)e^{-\frac{r}{l}}$	SE: $e^{-\frac{r^2}{2l^2}}$
8	EDC: $\sigma^2\cos(\omega r)e^{-\frac{r}{l}}$	RQ: $\sigma^2\left(1+\frac{r^2}{2\alpha l^2}\right)^{-\beta}$
9	BG: $\sigma^2\left(1+\frac{r}{l}\right)e^{-\frac{r}{l}}$	EN: $\sigma^2 e^{-\frac{r}{l}}$
10	BG: $\sigma^2\left(1+\frac{r}{l}\right)e^{-\frac{r}{l}}$	BG: $\sigma^2\left(1+\frac{r}{l}\right)e^{-\frac{r}{l}}$
11	BG: $\sigma^2\left(1+\frac{r}{l}\right)e^{-\frac{r}{l}}$	SE: $e^{-\frac{r^2}{2l^2}}$
12	BG: $\sigma^2\left(1+\frac{r}{l}\right)e^{-\frac{r}{l}}$	RQ: $\sigma^2\left(1+\frac{r^2}{2\alpha l^2}\right)^{-\beta}$

2.1.3 Settings for hyperparameter optimization

In parametric optimization, defining appropriate upper and lower bounds for the parameters is essential to ensure robust results. The hyperparameters σ^2 and l, representing the variance and correlation length, respectively, are common to all five selected correlation functions. For the uncertainty hyperparameter σ, lower and upper bounds were set to $0.1 < \sigma < 100$, based on the high-emission values in actual prior inventories. For the correlation length l, lower and upper bounds were set to $10 < l < 80$, based on the inversion domain in our study (40 km × 40 km). For the hyperparameters α and β contained in the RQ correlation function, lower and upper bounds were set to $0.01 < \alpha < 20$ and $0.1 < \beta < 5$, based on the application of this function in geophysical exploration and climate modeling (Chiles and Delfiner, 2012; Stein, 1999). For initial parameter values, relatively small values within the upper and lower bounds were selected: 1, 10, 5, 1.5 for σ, l, α, and β, respectively.

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

2.2 Inversion experiment design

2.2.1 Domain and CO_2 monitoring towers

The study area encompasses the main urban area of Zhengzhou, Henan Province, China. Situated along the transitional zone between the middle and lower reaches of the Yellow River, Zhengzhou spans 112°42′–114°14′E and 34°16′–34°58′N, with a total municipal area of 7 567 km^2. For this study, the inversion domain was defined as the central portion of Zhengzhou, covering 34.585°–34.875°N and 113.435°–113.825°E. Within this area, two CO_2 monitoring towers, with heights of 35 m and 50 m, were used for inversion (Fig. 1a, Table A1). The inversion grid resolution was set to 0.01° × 0.01° (latitude × longitude), resulting in 1 600 grid cells of 1 km × 1 km for flux optimization.

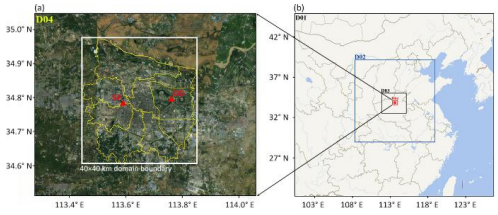

Figure 1: (a) Inversion domain in Zhengzhou (34.585°–34.875°N, 113.435°–113.825°E) and locations of the CO_2 monitoring towers (red triangles). The base map is provided by Google Maps. © Google, 2023. (b) Nested WRF domains (D01–D04) used in this study; the red box indicates the inversion domain in D04.

2.2.2 WRF-STILT model setup

The Weather Research and Forecasting Model (WRF v4.3) was used to drive the Stochastic Time-Inverted Lagrangian Transport (STILT) model, thereby generating the observation operator **H** for both ideal and real-data experiments (Yang et al., 2012). The WRF configuration included 35 vertical levels and four nested domains (Fig. 1b), with a temporal resolution of 1 hour. The horizontal resolutions are configured as 27 km, 9 km, 3 km, and 1 km, respectively. A multiple nesting strategy is adopted to simultaneously capture large-scale circulation patterns and resolve fine-scale local meteorological features, thereby enhancing the spatial resolution over the study area (Madhulatha et al., 2021). To reduce the accumulation of lateral boundary errors caused by multi-level nesting, a nesting ratio of 3:1 is adopted, which minimizes spurious reflections at the nesting boundaries and maintains flux conservation across different grid scales (Warner, 2011). Furthermore, the outermost domain is designed to be sufficiently large, ensuring that the study area is located at least 100 km away from all lateral boundaries, thereby effectively reducing the influence of lateral boundary conditions (Skamarock et al., 2008; Daniels et al., 2016; Imberger et al., 2020).The outermost domain (D01) covers most of China to provide stable large-scale forcing, featuring a 27 km grid resolution with 99 × 91 grid cells in the zonal and meridional directions, respectively. The second domain (D02) encompasses Henan Province and its surrounding areas, configured with a 9 km grid resolution and 120 × 123 grid cells. The third domain (D03) focuses on the north-central region of Henan Province, utilizing a 3 km resolution with 111 × 90 grid cells. The innermost domain (D04) is centered on the main urban area of Zhengzhou and its immediate vicinity, enabling high resolution representation of urban meteorological processes, with a high-resolution 1 km grid with 69 × 57 grid cells. STILT was run in time-reversed mode, driven by WRF-simulated wind

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

fields. Ensembles of 1 000 particles were released from the two monitoring towers and transported backward for 24 hours during the afternoon period (12:00–18:00 local time), when the boundary layer is typically well-developed and more stable. These trajectories yielded hourly footprint data at 1 km spatial resolution across the D04 domain for each measurement period. Approximately 90% of the backward trajectories reached the boundary of the inversion domain within 6 hours, with a mean travel time of less than 4 hours. Based on the monthly averaged footprints derived from the 24-hour trajectories, the D04 domain contributed approximately 40% of the total footprint (Figure A1). Furthermore, most of the D04 domain was identified as the sensitive area for the monitoring sites, defined as the region contributing 50% of the total footprint (Sargent et al. 2018), where emissions can be optimally constrained by the atmospheric inversion model. Detailed physical parameterizations and domain configurations are provided in Ren et al. (2024).

2.2.3 Pseudo-observation inversion experiment

To evaluate the performance of the optimization method under controlled conditions, inversion analyses were conducted using synthetic, or pseudo-observation, datasets. "True" and prior flux fields were prescribed for different scenarios. Forward transport simulations of atmospheric CO_2, driven by the true fluxes, were then performed to generate synthetic concentration data at each tower and time step. The inversion period extended from 1 to 31 October 2023. Under this configuration, the dimensions of the matrices $\mathbf{R}$, $\mathbf{B}$, and $\mathbf{H}$ were 434×434, $1\,600 \times 1\,600$, and $434 \times 1\,600$, respectively. As the experiment was conducted under ideal conditions, background contributions could be neglected. The inversion experiments were then performed using these pseudo-observations to optimize the fluxes relative to the prior estimates.

(1) Prior inventories and true fluxes

For the prior inventory, pixel-level emissions were generated from a normal distribution with a mean of 10 $\mu mol\ m^{-2}\ s^{-1}$ and a variance of $(3\ \mu mol\ m^{-2}\ s^{-1})^2$, reflecting values from real prior inventories in the study region. This approach imposed minimal assumptions on prior emissions. The true flux fields were constructed by introducing localized high-emission sources into the prior flux fields, establishing seven distinct scenarios containing 1, 4, 7, 10, 15, 20, and 30 high-emission points. These high values were based on the ODIAC inventory, with emission strengths generated randomly from normal distributions with means of 500–800 $\mu mol\ m^{-2}\ s^{-1}$ and a variance of $(20\ \mu mol\ m^{-2}\ s^{-1})^2$. Details of prior and true fluxes are provided in the Appendix (see Fig. A2).

(2) Error covariance matrices

The construction of the error covariance matrices for the optimization method is described in Sect.2.1. In the traditional method, both matrices and their parameters were defined empirically. For the observation error covariance matrix $\mathbf{R}$, a diagonal structure was assumed, implying the absence of temporal and spatial correlations in observation errors. Consistent with Bréon et al. (2015), each diagonal element was assigned a variance of $(3\ ppm)^2$ (Staufer et al., 2016; Wu et al., 2016). For the prior error covariance matrix $\mathbf{B}$, an exponential model was used, with the uncertainty for each pixel set to 100% of its net emissions in the prior inventory. Given that all inversions in this study used the same 40 km × 40 km domain, the spatial correlation length was fixed at 10 km.

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

(3) Pseudo-observation perturbations

To further assess the robustness of the optimization method in the presence of observation errors, a range of random perturbations was introduced to the pseudo-observations. Posterior-simulated concentrations derived from both the optimization and traditional methods were compared with the original, unperturbed pseudo-observations. Initially, perturbations with a mean of 0 and a variance of $(0.5\ \mathrm{ppm})^2$ were applied to all seven scenarios. Differences between the posterior-simulated concentrations and pseudo-observations were analyzed for each method. In addition, the scenario with 10 high-emission points was used as a case study for quantitative assessment of the effect of random perturbation magnitude. The variance was incrementally increased from $(0.5\ \mathrm{ppm})^2$ to $(1\ \mathrm{ppm})^2$, $(1.5\ \mathrm{ppm})^2$, $(2.0\ \mathrm{ppm})^2$, $(2.5\ \mathrm{ppm})^2$, and $(3.0\ \mathrm{ppm})^2$, enabling evaluation of the performance of both methods under progressively higher levels of observation error.

2.2.4 Real observation inversion experiment

For the real observation inversion experiment, two towers (SP and DD) were used for flux estimation, with each inversion covering a one-month period. This inversion was conducted from January to March 2023, spanning three consecutive months. In accordance with the restriction of the WRF-STILT model to afternoon-only backward transport, the observation data used for the inversion were selected from the afternoon period (12:00–18:00 local time) for the two towers.

(1) Prior inventories

The EDGAR (Emissions Database for Global Atmospheric Research, v7.0, 2022; https://edgar.jrc.ec.europa.eu; Janssens-Maenhout et al., 2019) inventory was used as the anthropogenic CO_2 prior in the real observation inversion. EDGAR supplies annual emission data, and the original spatial resolution is 0.1° × 0.1° (~10 km × 10 km at the equator). To ensure consistency with the inversion grid resolution, the EDGAR data were resampled to 1 km × 1 km using nearest-neighbor interpolation (Li et al., 2004). The spatial distributions of the prior emissions for the three months are shown in the Appendix (see Fig. A3).

(2) Error covariance matrices

For the real observation inversion experiment, the construction of the two error covariance matrices **R** and **B** under both the optimization and traditional methods followed the same procedures as described for the pseudo-observation inversion experiment in Sect. 2.2.3.

(3) Background mixing ratios

Because flux correction depends on observed concentration enhancements relative to the background level, accurate determination of this background is critical for the real observation inversion experiment. In this study, the background mixing ratio was defined as the minimum value within a 24-hour moving window for each day, following approaches used in previous CO_2 studies (Karion et al., 2021; Nickless et al., 2019; Turnbull et al., 2015). The daily minimum typically occurs during afternoon hours when the boundary layer is fully developed and vertical mixing is strongest, at which point local emission signals are maximally diluted and the observed concentration most closely approximates the regional background. By updating the background on a daily basis, this approach captures synoptic-scale meteorological variations, rather than imposing a single static value across the entire study period. The

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

potential influence of regional transport on the estimated background and the associated uncertainties are discussed in the conclusions.

2.4 Evaluation metrics

(1) Optimal correlation function combination selection

In the ideal experiments, the primary objective was to evaluate the performance of each method in capturing high-emission points in "true" fluxes. The root mean square error (RMSE) between the posterior and true fluxes was used as the selection criterion for the optimal covariance function combination under each true flux scenario, defined as:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(\mathbf{f_i^a}-\mathbf{f_i})^2}{N}} \quad (9)$$

where N denotes the number of grid cells and $\mathbf{f_i^a}$ and $\mathbf{f_i}$ represent the posterior and true fluxes at the ith grid cell, respectively. RMSE provides a quantitative measure of the deviation between posterior and true fluxes. To further verify the robustness of this selection criterion, daily posterior emissions from each covariance function combination were compared with daily emissions derived from the true fluxes. For the real observation experiments, RMSEs between the posterior-simulated concentrations and corresponding observations were used to select the optimal covariance function combination. The optimal combinations, their associated optimized hyperparameters, and the corresponding RMSE values for all inversion experiments conducted with the optimization method are summarized in Table A2.

(2) Relative uncertainty algorithm

In the posterior covariance matrix **A** defined in Eq. (2), diagonal elements represent the variances of the posterior fluxes for each grid cell, while off-diagonal elements denote covariances between different grid cells.

In this study, inversion performance was further assessed by comparing the relative uncertainties of CO_2 emissions derived from the posterior fluxes obtained using the optimization and traditional methods. CO_2 emissions over the inversion domain for a given time period t can be defined as:

$$\mathrm{E} = \mathrm{stf_1} + \mathrm{stf_2} + \cdots + \mathrm{stf_n} = \mathrm{st}(\mathrm{f_1} + \mathrm{f_2} + \cdots + \mathrm{f_n}) \quad (11)$$

where s represents the area size of the inversion domain and $\mathrm{f_1}$ to $\mathrm{f_n}$ represent all elements of the posterior flux **f** corresponding to the posterior fluxes of each grid cell. The synthesis formula of uncertainty is defined as (JCGM, 2008):

$$\mathrm{u_c(y)} = \sqrt{\sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\right)^2 \mathrm{u^2(x_i)} + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}\mathrm{u(x_i, x_j)}} \quad (12)$$

where $\mathrm{u_c(y)}$ is combined standard uncertainty, $\frac{\partial f}{\partial x_i}$ is the sensitivity coefficient, $\mathrm{u(x_i)}$ is the standard uncertainty of the input $\mathrm{x_i}$, and $\mathrm{u(x_i, x_j)}$ is the covariance of $\mathrm{x_i}$ and $\mathrm{x_j}$. For our emission model Eq. (11), $\frac{\partial f}{\partial x_i}$ is equal to st and $\mathrm{u^2(x_i)}$ and $\mathrm{u(x_i, x_j)}$ correspond to the diagonal and off-diagonal elements of matrix **A**, respectively. Finally, based on the relative uncertainty (UR) formula:

$$u_{rel}(y) = \frac{u_c(y)}{|y|} \quad (13)$$

where y denotes the measured quantity, here representing the posterior CO_2 emissions, the relative uncertainty of the emissions can be computed. Notably, the coefficients st cancel out during the

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

calculation, indicating that the relative uncertainty of emissions is independent of both the emission duration and area size.

3 Results and Discussion

This study employed a two-stage validation strategy. First, a theoretical evaluation was conducted by comparing the inversion performance of the optimization method with that of the traditional method using ideal experiment results, thereby establishing the fundamental advantages of the proposed approach. Empirical validation was then carried out through real observation inversion experiments to further demonstrate the superiority of the method.

3.1 Pseudo-observation inversion experiment

In the ideal experiments, where the true fluxes were known, a rigorous comparison was performed between the posterior fluxes from both methods and the true fluxes. To better approximate real-world conditions, controlled perturbations were then applied to the pseudo-observations to simulate observation errors commonly encountered in operational inversions, enabling evaluation of the robustness of each method under error-affected conditions. Finally, a comprehensive uncertainty assessment of posterior emission estimates was performed to directly compare the statistical reliability of the two methods.

3.1.1 Comparisons between true and posterior fluxes

Seven emission scenarios were analyzed, each containing a different number of high-emission points within the true flux distribution (1, 4, 7, 10, 15, 20, and 30) (see Fig. A2 for complete prior inventories). For detailed comparison, Figure 2 presents performance metrics of the optimization method across covariance function combinations, highlighting four representative scenarios with 1, 4, 10, and 20 high-emission points. For the 1-point case, optimal performance was achieved with the BG-BG combination (BG for both matrix **R** and matrix **B**), yielding an RMSE of 0.54×10^{-6} kg CO_2 m^{-2} s^{-1},while in the 4-point case, the EN-EN combination (EN for both matrix **R** and matrix **B**) performed best, with an RMSE of 0.95×10^{-6} kg CO_2 m^{-2} s^{-1}. In the 10-point and 20-point cases, the EN-BG combination and EDC–EN combination yielded the lowest RMSE values, 1.86×10^{-6} and 2.43×10^{-6} kg CO_2 m^{-2} s^{-1}, respectively.

Analysis of daily posterior emissions confirmed that these selected combinations consistently produced relatively small deviations from the true emissions across all scenarios, outperforming most other combinations as illustrated in Figure 2. Compared with the traditional method, the optimization method delivered statistically significant RMSE reductions in every case: 31.65% (0.79×10^{-6} to 0.54×10^{-6} kg CO_2 m^{-2} s^{-1}) for the 1-point case, 37.50%(1.52×10^{-6} to 0.95×10^{-6} kg CO_2 m^{-2} s^{-1}) for the 4-point case, 15.45% (2.20×10^{-6} to 1.86×10^{-6} kg CO_2 m^{-2} s^{-1}) for the 10-point case, and 19.00% (3.00×10^{-6} to 2.43×10^{-6} kg CO_2 m^{-2} s^{-1}) for the 20-point case.

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

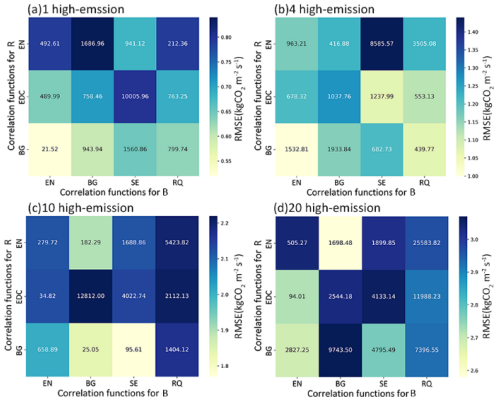

Figure 2: Results of the optimization method for "true" flux scenarios containing 1, 4, 10, and 20 high-emission points. RMSE denotes the root mean square error between posterior fluxes from each error covariance matrix combination and corresponding "true" fluxes. Values in grid cells indicate absolute biases between posterior daily emissions and "true" daily emissions for each combination. Units for daily emissions: tons (t).

Figure 3 shows the spatial distributions of posterior fluxes derived from both the traditional and optimization methods across multiple true flux scenarios. The traditional method exhibited limited spatial resolution, with elevated flux values concentrated primarily in the central inversion domain and poor localization of high-emission points. In contrast, the optimization method achieved markedly improved spatial accuracy, particularly for high-emission sources located near the domain center, as evidenced by the closer alignment between simulated and true flux distributions. Inversion results corresponding to additional high-emission points scenarios are provided in the Appendix (see Fig. A4).

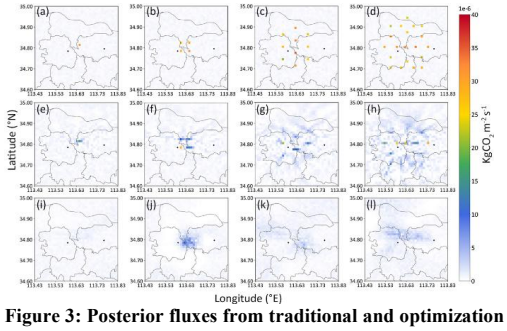

Figure 3: Posterior fluxes from traditional and optimization methods for "true" flux scenarios with 1, 4, 10, and 20 high-emission points. (a–d) "True" fluxes for the four scenarios. (e–h) Posterior fluxes from the optimization method. (i–l) Posterior fluxes from the traditional method.

Figure 4 provides a quantitative comparison of the traditional and optimization methods for estimating posterior fluxes across multiple true flux scenarios. Panel (a) shows that the RMSE between the posterior and true fluxes increased almost linearly with the number of high-emission points for both methods ($R^2 > 0.95$), reflecting the increasing complexity of flux distributions. Across all scenarios, the optimization method consistently outperformed the traditional approach, achieving an average RMSE

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

reduction of 21.40% (1.80×10^{-6} vs 2.29×10^{-6} kg CO_2 m^{-2} s^{-1}). Posterior fluxes from the optimization method also exhibited smaller deviations from the true fluxes than those reported in previous studies (Wu et al., 2018; Turner et al., 2016), highlighting its improved accuracy. Panel (b) indicates that although both methods displayed a similar scaling trend in daily emission discrepancies with increasing high-emission points, their response patterns differed fundamentally. Results from the traditional method showed significant scenario-dependent variability in flux estimation errors, whereas the optimization method exhibited a steady and predictable error growth trend. In all scenarios, the optimization method produced smaller discrepancies. These systematic differences, occurring under identical experimental conditions except for the construction of the error covariance matrix, demonstrate that covariance matrix formulation is a critical determinant of inversion accuracy. The proposed optimization method delivered closer agreement with the true fluxes while maintaining robust performance across increasingly complex emission fields.

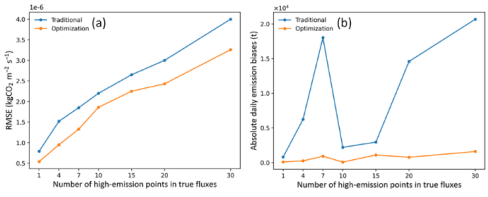

Figure 4: (a) RMSE between posterior fluxes and the "true" fluxes and (b) absolute daily emission biases for the optimization and traditional methods across different ideal flux scenarios.

3.1.2 Performance of the two methods under concentration perturbations

Method robustness was assessed by introducing random perturbations (σ = 0.5 ppm) to the pseudo-observations in the ideal experiments, where background concentration variability and instrumental measurement errors were negligible. Figure 5a presents the RMSE-evaluated differences between posterior-simulated concentrations and pseudo-observations under various true flux scenarios. The central point represents the lowest RMSE values. For both the SP and DD towers, the polygons connecting all scenarios for the optimization method were more closely clustered toward the central point than those from the traditional method, indicating greater consistency and closer agreement between simulations and pseudo-observations. The RMSE values for the optimization method remained consistently low across all emission scenarios, ranging from 0.38 to 1.01 ppm at SP tower and from 0.37 to 1.20 ppm at DD tower, substantially narrower than the ranges obtained using the traditional method (1.66 – 19.99 ppm at SP tower and 1.82 – 37.13 ppm at DD tower). These results demonstrate that the optimization method provides a more stable and accurate representation of concentration responses under different flux perturbations. Detailed comparisons between posterior-simulated concentrations and pseudo-observations under true flux scenarios (Fig. A5) further confirm the closer agreement achieved by the optimization method.

Sensitivity analysis was performed for the 10-point high-emission scenario, with perturbation magnitudes progressively increased from σ = 0.5 ppm to 3 ppm (Fig. 5b). Increasing perturbation magnitude led to growing variability in differences between the posterior-simulated concentrations and

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

observations for both methods. Across all perturbation levels, however, the optimization method consistently maintained smaller concentration differences -- ranging from 0.59 to 2.56 ppm at SP tower and from 0.66 to 2.47 ppm at DD tower -- compared with the traditional method (6.72 – 7.75 ppm at SP tower and 2.42 – 4.16 ppm at DD tower). These results demonstrate that the optimization method consistently achieves higher accuracy and sustained greater robustness under increasing observational uncertainty.

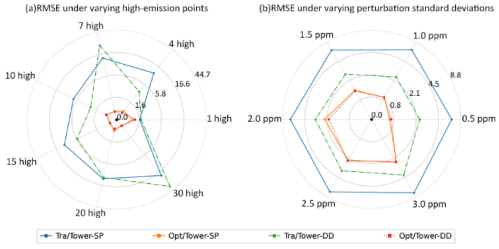

Figure 5: Radar charts showing deviations between posterior-simulated concentrations and pseudo-observations after applying perturbations, evaluated using RMSE (ppm). (a) Results for random perturbations with σ = 0.5 ppm under different ideal flux scenarios. (b) Results for the 10-point high-emission scenario with perturbations ranging from σ = 0.5 ppm to σ = 3.0 ppm. Note that a logarithmic transformation was applied to the RMSE values for improved visualization, while the radial axis is labeled in the original RMSE scale.

3.1.3 Relative uncertainty of posterior emissions for the two methods

The relative uncertainty of the posterior emissions generated by the two distinct inversion approaches across multiple true flux scenarios is shown in Fig. 6a. Across all scenarios, the optimization method consistently produced markedly lower relative uncertainties (0.50% to 1.35%) compared with the traditional method (3.40% to 8.15%), indicating a clear theoretical advantage in producing more reliable posterior estimates. In the traditional method, the relative uncertainty of posterior daily emissions decreased as the number of high-emission points increased. As indicated by Eq. (13), relative uncertainty is a function of the combined uncertainty derived from the posterior error covariance matrix elements in Eq. (12) and posterior flux magnitude. In the traditional framework, parameters including error covariance matrices **B** and **R** and the transport operator **H** in Eq. (2) remain fixed, resulting in identical posterior combined uncertainty across scenarios. However, larger true emissions yield proportionally greater pseudo-observations, increasing the posterior flux magnitude. With the numerator (combined uncertainty) unchanged and the denominator (posterior flux) increasing, the relative uncertainty value decreases, generating a spurious trend of uncertainty reduction with increasing high-emission points. The optimization approach, through rigorous construction of error covariance matrices, avoids this artifactual reduction.

To examine the interplay between individual grid-cell posterior emission uncertainties and inter-grid flux correlations in determining domain-wide combined uncertainty, comparative analyses were performed for 4-, 10-, and 20-point high-emission scenarios. As shown in Fig. 6b–e, the diagonal

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

elements of matrix **A** (individual grid uncertainties) in the optimization method were generally larger than those from the traditional method, with averages of 1.95×10^{-6}, 1.18×10^{-6}, and 1.98×10^{-6} kg CO_2 m^{-2} s^{-1} for the three high-emission scenarios, respectively. In contrast, the traditional method yielded significantly lower average identical individual grid uncertainties (0.27×10^{-6} kg CO_2 m^{-2} s^{-1}), identical across all three scenarios (Fig. 6e), reflecting the invariance of its posterior error covariance matrix **A** to changes in true flux distribution. The optimization method generated sums of off-diagonal elements in matrix **A** (inter-grid correlations) of −3087.29, -1135.80, and -3181.97 $(\text{kg } CO_2 \text{ m}^{-2} \text{ s}^{-1})^2$ across the different emission scenarios, contrasting sharply with the value of 1727.87 $(\text{kg } CO_2 \text{ m}^{-2} \text{ s}^{-1})^2$ obtained for the traditional method (Table A3).

This approximately 200% difference of the summed off-diagonal elements of **A** between the two methods can be attributed to the distinct approaches to constructing the posterior error covariance matrix, $\mathbf{A} = (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}$. In the traditional method, the observation error covariance matrix **R** is assumed to be diagonal, and most elements of the transport operator **H** are inherently close to zero due to the sparse sensitivity of concentration observations to distant grid emissions. As a result, the contribution of $\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}$ relative to $\mathbf{B}^{-1}$ becomes negligible, and **A** effectively approaches $(\mathbf{B}^{-1})^{-1} = \mathbf{B}$. Since **B** is constructed using an exponential-decay correlation function, $\sigma^2 e^{-\frac{r}{l}}$, where the parameter σ is scaled by the magnitude of prior fluxes, all elements of **B** are positive. This leads to a large positive sum of off-diagonal elements in **A**. In contrast, the optimization method employs multiple combinations of correlation functions to construct both **B** and **R**, allowing for more flexible spatial structures in the prior and observation uncertainties. In particular, **R** includes inter-station and inter-grid correlations that can introduce negative or compensating covariance terms. Consequently, the resulting posterior covariance matrix **A** exhibits pronounced negative off-diagonal sums, indicating stronger negative correlations among grids.

These negative correlations are consistent with the spatial characteristics of the true fluxes. As shown in Figure 3, the posterior fluxes derived from the optimization method more accurately capture the location and magnitude of the prescribed high-emission grids, reflecting their closer agreement with the true fluxes. Because the fluxes of these high-value grids are much larger than those of the remaining low-value grids, the resulting disparity induces statistically negative correlations, which account for the negative inter-grid correlations observed in **A**. In contrast, the traditional method yields posterior fluxes characterized by a regionally aggregated high-emission center, where flux values gradually decrease toward surrounding grids (Section 3.1.1). This smooth gradient results in neighboring grids having similar flux magnitudes, thereby increasing inter-grid correlations and generating a positive off-diagonal sum in matrix **A**. Because total uncertainty is a function of both grid-level variances and inter-grid covariances, the higher variances in the optimization method were offset by markedly reduced covariances, resulting in lower overall relative uncertainty in posterior daily emissions. These findings substantiate both the methodological superiority of the optimization framework and the importance of accurately representing spatial correlations in emission inversion systems, particularly for uncertainty quantification.

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

The spatial distributions of posterior uncertainty differed markedly between the optimization and traditional methods (Fig. 6b–e). In the traditional framework, uncertainties were spatially homogeneous across the domain, aligning with the uniform structure of the prior inventory (Fig. A2). As delineated in Eq. (2), the structure of matrices **B** and **R** plays a decisive role in shaping inversion outcomes. Here, **R** was constrained to a diagonal form and **B** adopted an exponential decay structure, a configuration that inherently emphasizes prior errors over observation errors during inversion. Moreover, because the uncertainty parameter σ^2 in the exponential decay model was derived directly from the prior inventory, the resulting grid-level posterior uncertainty field was dominated by prior information, consistent with earlier reports (Lauvaux et al., 2016; Wang et al., 2018). In contrast, the optimization method removed any fixed weighting between prior and observation errors, with parameters **B** and **R** determined through a mathematically rigorous optimization process based on objective numerical criteria (likelihood maximization), ensuring that the posterior uncertainty field reflected a more balanced integration of the prior inventory, observation data, and transport operator. Across all three flux scenarios (Fig. 6b–d), the optimization method produced comparable spatial patterns of posterior uncertainty, with minima near inversion towers and a gradual increase with distance from observation sites, while maintaining relative homogeneity in remote regions. This distribution is consistent with previous studies (Lopez-Coto et al., 2017; Nalini et al., 2022) showing that transport models achieve greater emission characterization accuracy in proximity to monitoring locations, thereby reducing uncertainties locally. Overall, within the controlled conditions of these experiments, the optimization method yielded posterior uncertainty fields that more faithfully represented the underlying atmospheric transport processes and inversion system behavior than those produced by the traditional method.

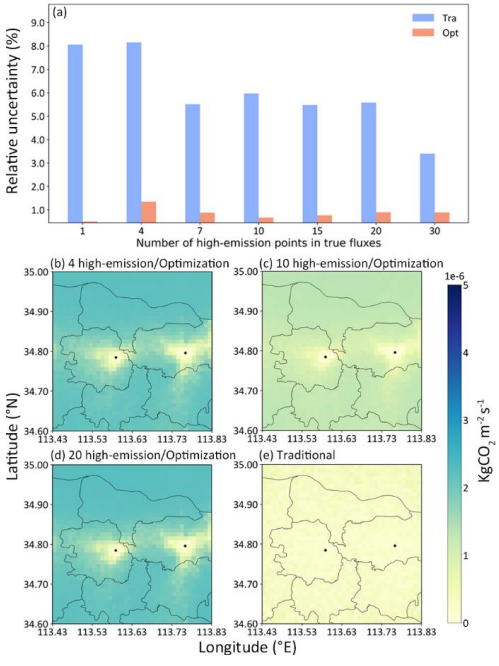

Figure 6: (a) Relative uncertainty of posterior daily emissions obtained using traditional and optimization

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

methods under different ideal flux scenarios. (b–e) Standard uncertainty of posterior fluxes obtained using traditional and optimization methods for different ideal flux scenarios.

3.2 Real observation inversion experiment

The optimization method was further assessed through inversion experiments incorporating CO_2 concentration measurements from monitoring towers collected between January and March 2023, combined with the EDGAR prior inventory. Evaluation focused on quantifying the goodness-of-fit between posterior-simulated concentrations and observations, alongside a comparative assessment of the relative uncertainty of posterior daily emissions obtained from both the optimization and traditional methods.

3.2.1 Comparisons between CO_2 observations and simulated concentrations

Figure 7 presents the correspondence between simulated and observed CO_2 concentrations at the two inversion towers (SP and DD). For both methods, the prior and posterior simulations reproduced the temporal dynamics of the observed concentrations, demonstrating that the transport model effectively captured the influence of local emissions on atmospheric CO_2 levels at the monitoring sites. Nevertheless, the prior simulation exhibited noticeable discrepancies, with RMSE values of 12.57 ppm and 11.66 ppm and correlation coefficients (R) of 0.85 and 0.69 at SP and DD towers, respectively.

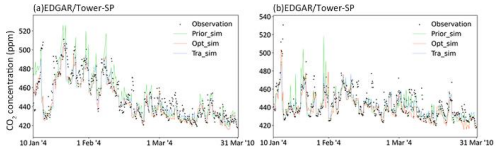

Figure 7. Comparison of hourly observed and simulated CO_2 concentrations at the inversion towers (a) SP and (b) DD under the EDGAR prior inventory from June to August 2023.

Inversion markedly reduced these discrepancies for both approaches (Fig. 8). The traditional method reduced the RMSE to 10.57 ppm at both towers, accompanied by increasing in the correlation coefficients (R) to 0.89 and 0.74, respectively. The optimization method achieved further improvements, yielding RMSE values of 10.31 ppm at SP tower and 10.05 ppm at DD tower, with corresponding R values of 0.91 and 0.81. This improvement is consistent with the results in Sect. 3.1.1, where analogous gains were observed in the ideal experiments with concentration perturbations. Given that both frameworks employ the same transport model and differ solely in their error covariance matrices specification, the improved performance can be attributed to the optimized error structure, which more effectively constrains posterior emissions under the EDGAR prior inventory. Despite these improvements, the absolute RMSE values remain on the order of 10 ppm, indicating that residual discrepancies persist in the real-observation inversions. Such discrepancies are consistent with limitations associated with the coarse spatial and temporal resolution of the EDGAR prior inventory, WRF-STILT simulation errors， as well as uncertainties in the background CO_2 determination based on the 24-hour moving-window minimum. Although the adoption of a four-domain nested configuration enhances the spatial resolution and physical

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

consistency of the WRF model, lateral boundary effects remain difficult to avoid entirely in a limited-area model (Skamarock et al., 2008; Imberger et al., 2020). Inaccurate outer boundary information can propagate inward and accumulate across domains. Moreover, discontinuities in physical processes between different resolution domains may further exacerbate uncertainties near nested boundaries (Daniels et al., 2016), thereby affecting meteorological field simulations and concentration estimates over the study region. Furthermore, air mass trajectories simulated by the STILT model are predominantly oriented from the northeast and west directions (Figure A1). According to the EDGAR inventory, emission contributions from regions along these pathways (e.g., Luoyang, Xinxiang, and Shandong Province) can also interfere with observed concentrations within the inversion domain, leading to additional uncertainties in emission estimates. In addition, while the 24-hour moving-window minimum approach accounts for part of the meteorological variability, it does not fully isolate large-scale or persistent signals, which may influence the simulated concentration enhancements.

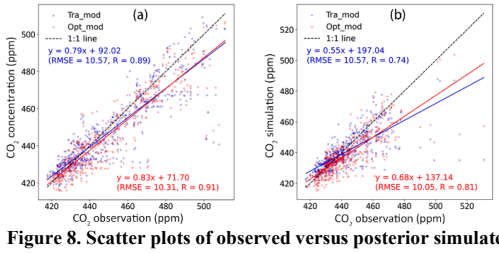

Figure 8. Scatter plots of observed versus posterior simulated CO_2 concentrations at the inversion towers (a) SP and (b) DD under the EDGAR prior inventory.

3.2.2 Relative uncertainty of posterior daily emissions under different prior inventories

Following the approach used in the ideal experiment, the relative uncertainties of the posterior daily emissions were quantified, with the corresponding posterior covariance matrices **A** derived using the two inversion methods under EDGAR prior inventory shown in Fig. 9. The results showed that the optimization method produced a lower relative uncertainty (5.19%) than the traditional method (9.69%) (Fig. 9a). This 46.44% reduction indicates that the optimization framework provides a substantially tighter constraint on posterior emissions. Furthermore, the relative uncertainties achieved with the optimization method were markedly lower than the values reported in related studies (11.54%–25.00%) (Sargent et al., 2018; Lauvaux et al., 2016; Turnbull et al., 2019). Examination of the posterior covariance matrices revealed that the optimization method yielded a larger cumulative grid-level variance (sum of diagonal elements of **A**) than the traditional method yet exhibited a substantially smaller cumulative inter-grid covariance. This combination mirrors the behavior observed in the ideal experiments (Section 3.1.3). The consistency between the ideal and real experiments underscores that spatial correlations among posterior fluxes play a decisive role in determining overall uncertainty in atmospheric inversion analyses.

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

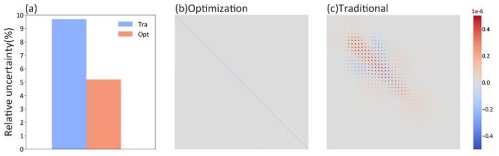

Figure 9. Relative uncertainty of posterior daily emissions (a) and posterior covariance matrices A (dimension 1 600 × 1 600) derived from the optimization (b) and traditional (c) inversion methods under the EDGAR prior inventory.

4 Conclusion

This study introduces an advanced framework for constructing error covariance matrices in atmospheric inversion, integrating correlation-function combinations, a hierarchical Bayesian model, and maximum-likelihood estimation. By explicitly representing coupled spatiotemporal error structures in both prior fluxes and observation errors, the proposed approach circumvents key limitations inherent to traditional inversion methods, which rely on simplified covariance assumptions or sensitivity-based parameter selection. The framework draws on earlier work employing maximum-likelihood estimation to optimize covariance hyperparameters (Michalak et al., 2005; Gourdji et al., 2010; Wu et al., 2013), while simultaneously addressing findings from urban-scale studies that posterior flux estimates are acutely sensitive to prior error covariance specification — particularly the spatial correlation length — despite these parameters being largely subjectively prescribed (Lauvaux et al., 2016; Nickless et al., 2019; Wu et al., 2018; Cohen and Prinn, 2011).

The ideal experiments showed that the optimization method consistently outperformed the traditional approach across diverse emission scenarios, producing more accurate reconstructions of true flux distributions and more precise localization of high-emission sources. For scenarios with multiple high-emission points, the optimization method reduced the RMSE between posterior and true fluxes by an average of 21.4%. When random perturbations (σ = 0.5 ppm) were introduced, the resulting concentration differences remained within 0.37–1.20 ppm, substantially narrower than the 1.66–37.13 ppm range obtained using the traditional method. The relative uncertainty of posterior daily emissions was consistently reduced to 0.50%–1.35%, compared to 3.4%–8.15% for the traditional method. Variance–covariance analysis further showed that this improvement is primarily associated with reduced inter-grid correlations, underscoring the importance of explicitly representing spatial correlation structures rather than relying solely on variance-based metrics. The real-observation inversions confirmed these advantages under the EDGAR prior inventory. At the two inversion towers, the optimization method achieved improved agreement with observed CO_2 concentrations, reducing RMSE to 10.31 ppm and 10.05 ppm and increasing correlation coefficients to 0.91 and 0.81, compared to 10.57 ppm and lower correlations (0.89 and 0.74) from the traditional method. The optimization method also achieved an approximately 46% reduction in the relative uncertainty of posterior daily emissions. This marked improvement in posterior precision carries direct and actionable implications for urban carbon monitoring. First, it strengthens top-down constraints on urban emission inventories, enabling more rigotous independent verification of reported emissions. Second, it improves the detectability of sustained emission trends, allowing earlier and more confident identification of policy-induced emssion

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

reductions. Third, it enhances the identification of high-emission hotspots for targeted mitigation, supporting spatially explicit intervention strategies such as sector-specific retrofits or infrastructure upgrades. Collevtively, these capabilities are essential for evaluating the effectiveness of city-scale emission-control policies and for providing information support for carbon peak and carbon neutrality roadmaps.

Nevertheless, the absolute RMSE values in the real-observation experiments remain on the order of 10 ppm, reflecting uncertainties that lie beyond the scope of covariance optimization alone. Two factors are particularly relevant. First, the prior EDGAR global inventory, while globally consistent, displays insufficient spatial and temporal resolution relative to high-resolution local inventories. Inversion inter-comparison studies that integrate dense observation networks with regionally refined inventories—exemplified by the European CH_4 emissions analysis of Ioannidis et al. (2026)—have yielded substantially superior model–observation agreement. These results imply that prior inventory quality critically constrains the absolute accuracy of posterior concentration fields. Second, the background CO_2 mixing ratio was estimated using a 24-hour moving-window minimum. While this method effectively accounts for meteorological variability, it cannot fully eliminate the influence of persistent upwind emission sources, potentially introducing a conservative bias into the estimated concentration enhancements. Future work will focus on incorporating higher-quality local emission inventories, refining background determination through upwind reference stations or model-derived background fields, and integrating additional observational constraints such as satellite retrievals or denser ground-based networks. Extending the framework to longer-term datasets or multi-species inversions may further improve the detection of emission trends and better constrain temporal variability. Overall, the results demonstrate that explicitly optimized spatial–temporal covariance structures substantially enhance the reliability of atmospheric CO_2 inversions and provide a stronger methodological foundation for urban carbon monitoring and emission verification.

Data availability

The data used in this study are not publicly available due to confidentiality restrictions associated with urban monitoring data. However, the data can be made available from the corresponding author upon reasonable request and with permission from the relevant data providers.

Author contribution

Wei Xu: Conceptualization, Methodology, Software, Writing - original draft. **Ge Ren:** Conceptualization, Writing - review & editing, Validation, Funding acquisition. **Kailun Du:** Writing – review & editing, Software, Investigation. **Xiaoning Wang:** Software, Investigation. **Mengjuan Han:** Writing – review & editing, Investigation. **Hong Lin:** Writing – review & editing, Validation, Supervision. **James Whetstone:** Writing – review & editing, Investigation.

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

Competing interest

The authors declare that they have no conflict of interest.

Acknowledgements

This study was partially supported by the National Key Research and Development Program of China (2023YFE0207200), National Natural Science Foundation of China (41907272), and Fundamental Research Funds for Central Public Welfare Scientific Research Institutes sponsored by the National Institute of Metrology, China (AKYZZ2307 and AKYZD2207-2).

References

Andres, R. J., Gregg, J. S., Losey, L., Marland, G., and Boden, T. A.: Monthly, global emissions of carbon dioxide from fossil fuel consumption, Tellus Series B-Chemical and Physical Meteorology, 63, 309-327, doi: 10.1111/j.1600-0889.2011.00530.x, 2011.

Balgovind, R., Dalcher, A., Ghil, M., and Kalnay, E.: A stochastic-dynamic model for the spatial structure of forcast error statistics, Monthly Weather Review, 111, 701-722, 1983.

Bannister, R. N.: A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics, Quarterly Journal of the Royal Meteorological Society, 134, 1971-1996, doi: 10.1002/qj.340, 2008.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M.: Time series analysis: forecasting and control, John Wiley & Sons, 2015.

Brockwell, P. J. and Davis, R. A.: Introduction to time series and forecasting, Springer, 2002.

Byrne, B., Baker, D. F., Basu, S., Bertolacci, M., Bowman, K. W., Carroll, D., Chatterjee, A., Chevallier, F., Ciais, P., and Cressie, N.: National CO_2 budgets (2015–2020) inferred from atmospheric CO_2 observations in support of the global stocktake, Earth System Science Data Discussions, 2022, 1-59, 2022.

Chevallier, F.: On the statistical optimality of CO_2 atmospheric inversions assimilating CO_2 column retrievals, Atmospheric Chemistry and Physics, 15, 11133-11145, 2015.

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez‐Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morgui, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO_2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, Journal of Geophysical Research: Atmospheres, 115, doi: 10.1029/2010jd013887, 2010.

Chiles, J.-P. and Delfiner, P.: Geostatistics: modeling spatial uncertainty, John Wiley & Sons, 2012.

Cohen, J. B. and Prinn, R. G.: Development of a fast, urban chemistry metamodel for inclusion in global models, Atmospheric Chemistry and Physics, 11, 7629–7656, https://doi.org/10.5194/acp-11-7629-2011, 2011.

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

Cozzi, L., Gould, T., Bouckart, S., Crow, D., Kim, T.-Y., McGlade, C., Olejarnik, P., Wanner, B., and Wetzel, D.: World energy outlook 2020, Energy, 2019, 30, 2020.

Cressie, N.: Statistics for spatial data, John Wiley & Sons, 2015.

Daniels, M. H., Lundquist, K. A., Mirocha, J. D., Wiersema, D. J., and Chow, F. K.: A new vertical grid nesting capability in the Weather Research and Forecasting (WRF) model, Monthly Weather Review, 144, 3725–3747, https://doi.org/10.1175/MWR-D-16-0049.1, 2016.

Enting, I. G.: Inverse problems in atmospheric constituent transport, Cambridge University Press, 2002.

Feng, S., Lauvaux, T., Newman, S., Rao, P., Ahmadov, R., Deng, A., Díaz-Isaac, L. I., Duren, R. M., Fischer, M. L., and Gerbig, C.: Los Angeles megacity: a high-resolution land-atmosphere modelling system for urban CO_2 emissions, Atmospheric Chemistry and Physics, 16, 9019-9045, 2016.

Gourdji, S. M., Hirsch, A. I., Mueller, K. L., Yadav, V., Andrews, A. E., and Michalak, A. M.: Regional-scale geostatistical inverse modeling of North American CO_2 fluxes: a synthetic data study, Atmospheric Chemistry and Physics, 10, 6151–6167, https://doi.org/10.5194/acp-10-6151-2010, 2010.

G'Sell, M. G., Shen-Orr, S. S., and Tibshirani, R.: Sensitivity analysis for inference with partially identifiable covariance matrices, Computational Statistics, 29, 529-546, 2014.

Gelb, A.: Applied optimal estimation, MIT press, 1974.

Guigues, V.: Sensitivity analysis and calibration of the covariance matrix for stable portfolio selection, Computational Optimization and Applications, 48, 553-579, 2011.

Gurney, K. R., Liang, J., Roest, G., Song, Y., Mueller, K., and Lauvaux, T.: Under-reporting of greenhouse gas emissions in US cities, Nature communications, 12, 553, 2021.

Hu, C., Griffis, T. J., Xia, L., Xiao, W., Liu, C., Xiao, Q., Huang, X., Yang, Y., Zhang, L., and Hou, B.: Anthropogenic CO_2 emission reduction during the COVID-19 pandemic in Nanchang City, China, Environmental Pollution, 309, 119767, doi: 10.1016/j.envpol.2022.119767, 2022.

Ioannidis, E., Meesters, A., Steiner, M., Brunner, D., Reum, F., Pison, I., Berchet, A., Thompson, R., Sollum, E., Koch, F.-T., Gerbig, C., Wang, F., Maksyutov, S., Tsuruta, A., Tenkanen, M., Aalto, T., Monteil, G., Lin, H., Ren, G., Scholze, M., and Houweling, S.: An inter-comparison of inverse models for estimating European CH_4 emissions, Earth System Science Data, 18, 167–198, https://doi.org/10.5194/essd-18-167-2026, 2026.

IPCC: Guidelines for national greenhouse gas inventories, Prepared by the National Greenhouse Gas Inventories Programme. Eggleston HS, Buendia L, Miwa K, Ngara T, Tanabe K, editors. Published: IGES, Japan, 2006.

IPCC: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2023.

Imberger, M., Larsén, X. G., Davis, N., and Du, J.: Approaches toward improving the modelling of mid-latitude cyclones entering at the lateral boundary corner in the limited area model WRF, Quarterly Journal of the Royal Meteorological Society, 146, 3225–3244, https://doi.org/10.1002/qj.3836, 2020.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J. G., and Peters, J. A.: EDGAR v4. 3.2 Global Atlas of the three major

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

greenhouse gas emissions for the period 1970-2012, Earth System Science Data, 11, 959-1002, 2019.

Joint Committee for Guides in Metrology (JCGM): Evaluation of measurement data - Guide to the expression of uncertainty in measurement, JCGM 100:2008, https://www.bipm.org/en/publications/guides/gum.html (last access：11 August 2025), 2008.

Johnson, S. G.: The NLopt nonlinear-optimization package, Massachusetts Institute of Technology [code], 2014.

Karion, A., Lopez-Coto, I., Gourdji, S. M., Mueller, K., Ghosh, S., Callahan, W., Stock, M., DiGangi, E., Prinzivalli, S., and Whetstone, J.: Background conditions for an urban greenhouse gas network in the Washington, DC, and Baltimore metropolitan region, Atmospheric chemistry and physics, 21, 6257-6273, 2021.

Koschorreck, M., Kamjunke, N., Koedel, U., Rode, M., Schuetze, C., and Bussmann, I.: Diurnal versus spatial variability of greenhouse gas emissions from an anthropogenically modified lowland river in Germany, Biogeosciences, 21, 1613-1628, 2024.

Lauvaux, T., Giron, C., Mazzolini, M., d'Aspremont, A., Duren, R., Cusworth, D., Shindell, D., and Ciais, P.: Global assessment of oil and gas methane ultra-emitters, Science, 375, 557-561, 2022.

Lauvaux, T., Miles, N. L., Deng, A., Richardson, S. J., Cambaliza, M. O., Davis, K. J., Gaudet, B., Gurney, K. R., Huang, J., and O'Keefe, D.: High-resolution atmospheric inversion of urban CO_2 emissions during the dormant season of the Indianapolis Flux Experiment (INFLUX), Journal of Geophysical Research: Atmospheres, 121, 5213-5236, 2016.

Lauvaux, T., Schuh, A., Uliasz, M., Richardson, S., Miles, N., Andrews, A., Sweeney, C., Diaz, L., Martins, D., and Shepson, P.: Constraining the CO_2 budget of the corn belt: exploring uncertainties from the assumptions in a mesoscale inverse system, Atmospheric Chemistry and Physics, 12, 337-354, 2012.

Lauvaux, T., Gurney, K. R., Miles, N. L., Davis, K. J., Richardson, S. J., Deng, A., Nathan, B. J., Oda, T., Wang, J. A., Hutyra, L., and Turnbull, J.: Policy-relevant assessment of urban CO_2 emissions, Environ Science & Technology, 54, 10237-10245, doi: 10.1021/acs.est.0c00343, 2020.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., and Canadell, J. G.: Global carbon budget 2018, Earth System Science Data, 10, 2141-2194, 2018.

Li, Z., Zhu, C., and Gold, C.: Digital terrain modeling: principles and methodology, CRC press, 2004.

Lian, J., Lauvaux, T., Utard, H., Breon, F. M., Broquet, G., Ramonet, M., Laurent, O., Albarus, I., Cucchi, K., and Ciais, P.: Assessing the effectiveness of an urban CO_2 monitoring network over the Paris region through the COVID-19 lockdown natural experiment, Environ Sci Technol, 56, 2153-2162, doi: 10.1021/acs.est.1c04973, 2022.

Liu, Z., Guan, D., Wei, W., Davis, S. J., Ciais, P., Bai, J., Peng, S., Zhang, Q., Hubacek, K., and Marland, G.: Reduced carbon emission estimates from fossil fuel combustion and cement production in China, Nature, 524, 335-338, 2015.

Lopez-Coto, I., Ghosh, S., Prasad, K., and Whetstone, J.: Tower-based greenhouse gas measurement network design—the national institute of standards and technology north east corridor testbed, Advances in Atmospheric Sciences, 34, 1095-1105, 2017.

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

Macknick, J.: Energy and Carbon Dioxide Emission Data Uncertainties. IIASA Interim Report IR-09-32, International Institute for Applied Systems Analysis, Laxenburg, 2009.

Madhulatha, A., Choi, S.-J., Han, J.-Y., and Hong, S.-Y.: Impact of different nesting methods on the simulation of a severe convective event over South Korea using the Weather Research and Forecasting model, Journal of Geophysical Research: Atmospheres, 126, e2020JD033746, https://doi.org/10.1029/2020JD033746, 2021.

Michalak, A. M., Bruhwiler, L., and Tans, P. P.: A geostatistical approach to surface flux estimation of atmospheric trace gases, Journal of Geophysical Research: Atmospheres, 109, doi: 10.1029/2003jd004422, 2004.

Michalak, A. M., Hirsch, A., Bruhwiler, L., Gurney, K. R., Peters, W., and Tans, P. P.: Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, Journal of Geophysical Research: Atmospheres, 110, 2005.

Nalini, K., Lauvaux, T., Abdallah, C., Lian, J., Ciais, P., Utard, H., Laurent, O., and Ramonet, M.: High-Resolution Lagrangian inverse modeling of CO_2 emissions over the Paris region during the first 2020 lockdown period, Journal of Geophysical Research: Atmospheres, 127, doi: 10.1029/2021jd036032, 2022.

Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B., and Crisp, D.: Quantifying CO_2 emissions from individual power plants from space, Geophysical Research Letters, 44, 10,045-010,053, 2017.

Nickless, A., Rayner, P. J., Scholes, R. J., Engelbrecht, F., and Erni, B.: An atmospheric inversion over the city of Cape Town: sensitivity analyses, Atmospheric Chemistry and Physics, 19, 7789-7816, 2019.

Oda, T., Maksyutov, S., and Andres, R. J.: The Open-source Data Inventory for Anthropogenic CO_2, version 2016 (ODIAC2016): a global monthly fossil fuel CO_2 gridded emissions data product for tracer transport simulations and surface flux inversions, Earth System Science Data, 10, 87-107, 2018.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO_2 inversions, Biogeosciences, 10, 6699-6720, doi: 10.5194/bg-10-6699-2013, 2013.

Ren, G., Du, K., Karion, A., Zhao, S., Lopez-Coto, I., Wang, X., Whetstone, J., and Lin, H.: Site selection and effects of background towers on urban CO_2 estimates: A case study from central downtown Zhengzhou in China, Environmental Research, 263, 120169, 2024.

Runarsson, T. P. and Yao, X.: Stochastic ranking for constrained evolutionary optimization, IEEE Transactions on evolutionary computation, 4, 284-294, 2000.

Saint-Vincent, P. M. and Pekney, N. J.: Beyond-the-meter: Unaccounted sources of methane emissions in the natural gas distribution sector, Environmental science & technology, 54, 39-49, 2019.

Sargent, M., Barrera, Y., Nehrkorn, T., Hutyra, L. R., Gately, C. K., Jones, T., McKain, K., Sweeney, C., Hegarty, J., Hardiman, B., Wang, J. A., and Wofsy, S. C.: Anthropogenic and biogenic CO_2 fluxes in the Boston urban region, Proc Natl Acad Sci U S A, 115, 7491-7496, doi: 10.1073/pnas.1803715115, 2018.

Schulze, E. D., Ciais, P., Luyssaert, S., Schrumpf, M., Janssens, I. A., Thiruchittampalam, B., Theloke,

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

J., Saurat, M., Bringezu, S., Lelieveld, J., Lohila, A., Rebmann, C., Jung, M., Bastviken, D., Abril, G., Grassi, G., Leip, A., Freibauer, A., Kutsch, W., Don, A., Nieschulze, J., Boerner, A., Gash, J. H., and Dolman, A. J.: The European carbon balance. Part 4：integration of carbon and other trace-gas., Global Change Biology, 16, doi: 10.1111/j.1365-2486.2010.02252.x, 2010.

Seto, K. C., Güneralp, B., and Hutyra, L. R.: Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools, Proceedings of the National Academy of Sciences, 109, 16083-16088, 2012.

Shumway, R. H., Stoffer, D. S., and Stoffer, D. S.: Time series analysis and its applications, Springer, 2000.

Sijikumar, S., Raju, A., Valsala, V., Tiwari, Y., Girach, I., Jain, C. D., and Ratnam, M. V.: High-resolution Bayesian inversion of carbon dioxide flux over Peninsular India, Atmospheric Environment, 308, 119868, 2023.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A description of the advanced research WRF version 3, NCAR Technical Note NCAR/TN-475+STR, National Center for Atmospheric Research, https://doi.org/10.5065/D68S4MVH, 2008.

Smith, J. E., Billmire, M., French, N. H., and Domke, G. M.: Application of the wildland fire emissions inventory system to estimate fire emissions on forest lands of the United States, Carbon Balance and Management, 19, 26, 2024.

Staufer, J., Broquet, G., Bréon, F.-M., Puygrenier, V., Chevallier, F., Xueref-Rémy, I., Dieudonné, E., Lopez, M., Schmidt, M., Ramonet, M., Perrussel, O., Lac, C., Wu, L., and Ciais, P.: The first 1-year-long estimate of the Paris region fossil fuel CO_2 emissions based on atmospheric inversion, Atmospheric Chemistry and Physics, 16, 14703-14726, doi: 10.5194/acp-16-14703-2016, 2016.

Stein, M. L.: Interpolation of spatial data: some theory for kriging, Springer Science & Business Media, 1999.

Svanberg, K.: The method of moving asymptotes—a new method for structural optimization, International journal for numerical methods in engineering, 24, 359-373, 1987.

Tarantola, A.: Inverse problem theory and methods for model parameter estimation, SIAM, 2005.

Turnbull, J. C., Sweeney, C., Karion, A., Newberger, T., Lehman, S. J., Tans, P. P., Davis, K. J., Lauvaux, T., Miles, N. L., and Richardson, S. J.: Toward quantification and source sector identification of fossil fuel CO_2 emissions from an urban area: Results from the INFLUX experiment, Journal of Geophysical Research: Atmospheres, 120, 292-312, 2015.

Turnbull, J. C., Karion, A., Davis, K. J., Lauvaux, T., Miles, N. L., Richardson, S. J., Sweeney, C., McKain, K., Lehman, S. J., Gurney, K. R., Patarasuk, R., Liang, J., Shepson, P. B., Heimburger, A., Harvey, R., and Whetstone, J.: Synthesis of urban CO_2 emission estimates from multiple methods from the Indianapolis flux project (INFLUX), Environmental Science & Technology, 53, 287-295, doi: 10.1021/acs.est.8b05552, 2019.

Turner, A. J., Shusterman, A. A., McDonald, B. C., Teige, V., Harley, R. A., and Cohen, R. C.: Network design for quantifying urban CO_2 emissions: assessing trade-offs between precision and network density, Atmospheric Chemistry and Physics, 16, 13465-13475, doi: 10.5194/acp-16-13465-2016,

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

2016.

Warner, T. T.: Quality assurance in atmospheric modeling, Bulletin of the American Meteorological Society, 92, 1601–1610, https://doi.org/10.1175/BAMS-D-10-05004.1, 2011.

Wang, Y., Broquet, G., Ciais, P., Chevallier, F., Vogel, F., Wu, L., Yin, Y., Wang, R., and Tao, S.: Potential of European $^{14}CO_2$ observation network to estimate the fossil fuel CO_2 emissions via atmospheric inversions, Atmospheric Chemistry and Physics, 18, 4229-4250, doi: 10.5194/acp-18-4229-2018, 2018.

Wesloh, D., Keller, K., Feng, S., Lauvaux, T., and Davis, K. J.: Temporal error correlations in a terrestrial carbon cycle model derived by comparison to carbon dioxide eddy covariance flux tower measurements, Journal of Geophysical Research: Biogeosciences, 129, doi: 10.1029/2023jg007526, 2024.

Williams, C. K. and Rasmussen, C. E.: Gaussian processes for machine learning, 3, MIT press Cambridge, MA, 2006.

Wu, K., Lauvaux, T., Davis, K. J., Deng, A., Coto, I. L., Gurney, K. R., and Patarasuk, R.: Joint inverse estimation of fossil fuel and biogenic CO_2 fluxes in an urban environment: An observing system simulation experiment to assess the impact of multiple uncertainties, Elemrnta Science of the Anthropocene, 6, 17, doi: 10.1525/elementa.138, 2018.

Wu, L., Bocquet, M., Chevallier, F., Lauvaux, T., and Davis, K.: Hyperparameter estimation for uncertainty quantification in mesoscale carbon dioxide inversions, Tellus B: Chemical and Physical Meteorology, 65, 20894, 2013.

Wu, L., Broquet, G., Ciais, P., Bellassen, V., Vogel, F., Chevallier, F., Xueref-Remy, I., and Wang, Y.: What would dense atmospheric observation networks bring to the quantification of city CO_2 emissions?, Atmospheric Chemistry and Physics, 16, 7743-7771, 2016.

Yadav, V., Ghosh, S., Mueller, K., Karion, A., Roest, G., Gourdji, S. M., Lopez‐Coto, I., Gurney, K. R., Parazoo, N., and Verhulst, K. R.: The impact of COVID-19 on CO_2 emissions in the Los Angeles and Washington DC/Baltimore metropolitan areas, Geophysical Research Letters, 48, e2021GL092744, 2021.

Yang, B., Zhang, Y., and Qian, Y.: Simulation of urban climate with high-resolution WRF model: A case study in Nanjing, China, Asia-Pacific Journal of Atmospheric Sciences, 48, 227-241, 2012.

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

Appendix

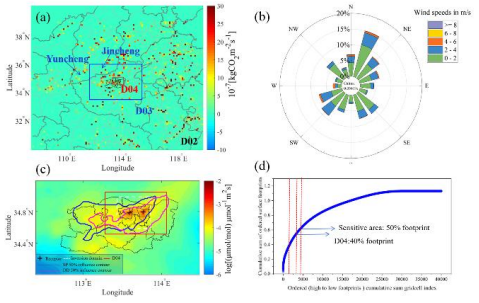

Figure A1: (a) EDGAR mean monthly anthropogenic emissions. The red and blue boxes represent the D04 and D03 boundaries, respectively. (b) Mean wind rose at the meteorological monitoring station for January–March 2023. (c) The average monthly footprint of the urban tower stations at afternoon (1200-1800 Local) for January–March 2023.Surface footprints are provided in log10 (ppm μmol-1m2s). Receptor release point is indicated by the black cross. The magenta (blue) contour encloses 50% of the average footprint initiated at the DD (SP) site (sensitive area). The white and red boxes represent the inversion domain and D04 boundaries, respectively. (d) Cumulative sum of sorted (high to low) average monthly footprint for January–March in 2023. Percentiles selected as points at or below fractional cutoff of summed ordered footprint.

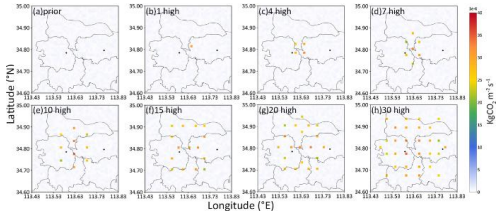

Figure A2: (a) Prior inventory and (b–h) "true" fluxes with varying numbers of high-emission points in the ideal experiment.

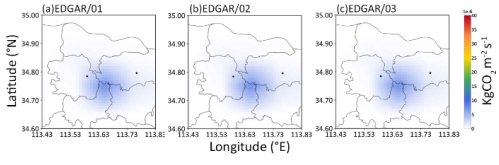

Figure A3: Spatial distributions of the EDGAR prior emissions for (a) January, (b) February, and (c) March used in the real observation experiment.

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

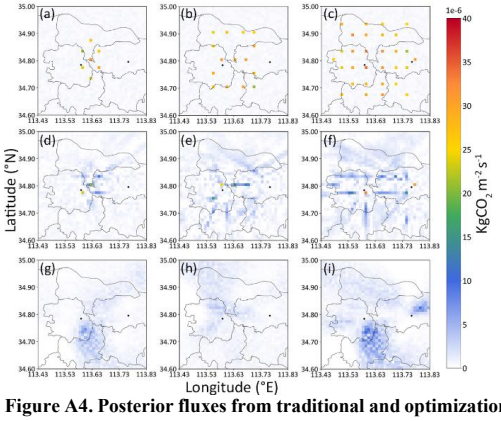

Figure A4. Posterior fluxes from traditional and optimization methods for "true" flux scenarios with 7, 15, 30 high-emission points. (a–c) "True" fluxes for three cases, (d–f) posterior fluxes obtained using optimization method, and (g–i) posterior fluxes derived from traditional method.

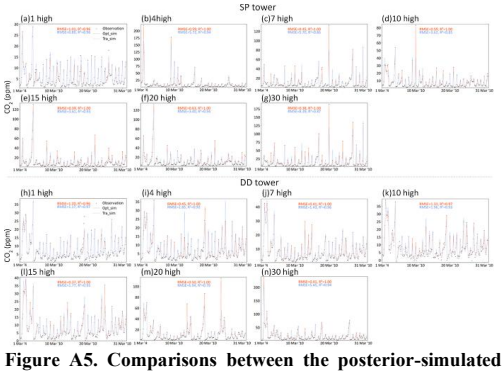

Figure A5. Comparisons between the posterior-simulated concentrations and pseudo-observations after adding perturbations with a standard deviation of 0.5 ppm under different assumed ideal flux scenarios. (a–g) Comparisons for tower SP. (h–n) Comparisons for tower DD.

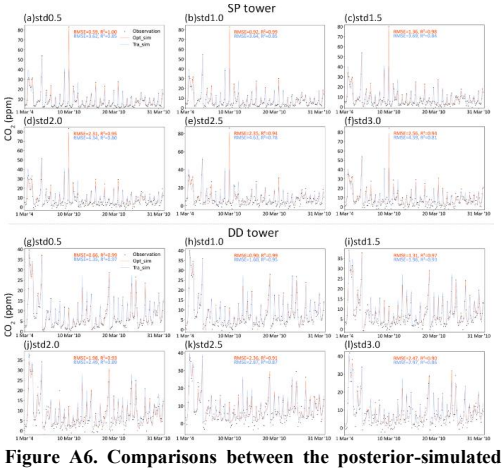

Figure A6. Comparisons between the posterior-simulated concentrations and pseudo-observations after adding perturbations with varying standard deviations under assumed ideal flux scenario with 10 high-

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

emission points. (a–g) Comparisons for tower SP. (h–n) Comparisons for tower DD.

985

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

Table A1. Urban observation towers, including latitude, longitude, and height above ground level (AGL).

Urban tower	Longitude	Latitude	AGL (m)
Sculpture Park Tower (SP)	113.59	34.78	50
Dream of Dragon Community Tower (DD)	113.76	34.80	35

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

Table A2. Optimal combinations of covariance functions, corresponding optimized hyperparameter values, and associated selection metric (RMSE) for ideal and real experiments.

Experiment scenarios	Correlation function(R)	Optimized hyperpara (R)	Correlation function (B)	Optimized hyperpara (B)	RMSE
1-high	$\sigma^2\left(1+\frac{r}{l}\right)e^{-\frac{r}{l}}$	σ = 0.10 l = 44.05	$\sigma^2\left(1+\frac{r}{l}\right)e^{-\frac{r}{l}}$	σ = 25.04 l = 0.10	12.29
4-high	$\sigma^2 e^{-\frac{r}{l}}$	σ = 0.10 l = 23.83	$\sigma^2 e^{-\frac{r}{l}}$	σ = 50.78 l = 0.10	21.69
7-high	$\sigma^2 e^{-\frac{r}{l}}$	σ = 0.10 l = 27.43	$\sigma^2 e^{-\frac{r}{l}}$	σ = 32.06 l = 0.10	30.24
10-high	$\sigma^2 e^{-\frac{r}{l}}$	σ = 0.10 l = 80.00	$\sigma^2\left(1+\frac{r}{l}\right)e^{-\frac{r}{l}}$	σ = 30.87 l = 0.10	42.24
15-high	$\sigma^2\left(1+\frac{r}{l}\right)e^{-\frac{r}{l}}$	σ = 6.98 l = 80.00	$\sigma^2 e^{-\frac{r}{l}}$	σ = 37.11 l = 0.18	51.10
20-high	$\sigma^2\cos(\omega r)e^{-\frac{r}{l}}$	σ = 0.10 l = 70.99	$\sigma^2 e^{-\frac{r}{l}}$	σ = 52.77 l = 0.10	55.21
30-high	$\sigma^2\cos(\omega r)e^{-\frac{r}{l}}$	σ = 0.10 l = 3.36	$\sigma^2 e^{-\frac{r}{l}}$	σ = 51.20 l = 0.10	74.03
EDGAR/01	$\sigma^2\cos(\omega r)e^{-\frac{r}{l}}$	σ = 28.87 l = 30.58	$\sigma^2 e^{-\frac{r^2}{2l^2}}$	σ = 85.39 l = 1.22	13.17
EDGAR/02	$\sigma^2\left(1+\frac{r}{l}\right)e^{-\frac{r}{l}}$	σ = 6.35 l = 0.78	$\sigma^2\left(1+\frac{r}{l}\right)e^{-\frac{r}{l}}$	σ = 49.27 l = 2.70	9.27
EDGAR/03	$\sigma^2\cos(\omega r)e^{-\frac{r}{l}}$	σ = 7.59 l = 4.27	$\sigma^2 e^{-\frac{r}{l}}$	σ = 81.01 l = 12.37	8.38

For ideal experiments, RMSE refers to differences between posterior and true fluxes (μmol m^{-2} s^{-1}); for real experiments, RMSE refers to differences between posterior-simulated concentrations and observations (ppm).

https://doi.org/10.5194/egusphere-2026-2044
Preprint. Discussion started: 23 June 2026

Table A3. Summary of inter-grid covariances, grid-level variances, and corresponding combined posterior uncertainties for optimization and traditional methods under three true flux scenarios.

Experiment scenarios	4-high/ opt	4-high/ tra	10-high/ opt	10-high/ tra	20-high/ opt	20-high/ tra
$\sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\right)^2 u^2(x_i)$	6 860.25	169.46	4 588.95	169.46	6 905.42	169.46
$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}u(x_i,x_j)$	−3 260.00	3 542.51	−2 232.04	3 542.51	−3 269.63	3 542.51
u_c	18.44	85.17	11.17	85.17	19.14	85.17

Inter-grid covariance and grid-level variance sums ($kg\ CO_2\ m^{-2}\ s^{-1})^2$, combined uncertainty $kg\ CO_2\ m^{-2}\ s^{-1}$.