Atmospheric Environment 327 (2024) 120490

Contents lists available at ScienceDirect

Atmospheric Environment

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

Hybrid IFDMB/4D-Var inverse modeling to constrain the spatiotemporal distribution of CO and NO_2 emissions using the CMAQ adjoint model

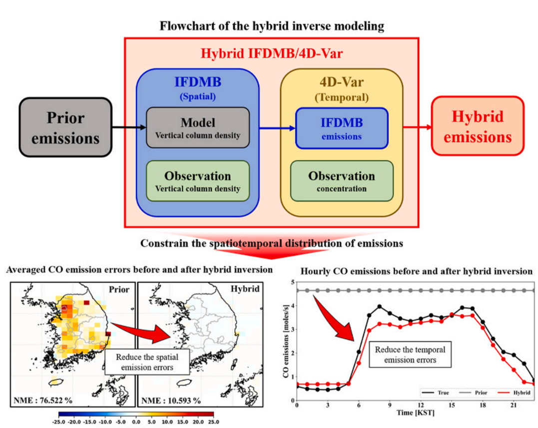
Jeonghyeok Moon [a], Yunsoo Choi [b], Wonbae Jeon [c,d,*], Hyun Cheol Kim [e,f], Arman Pouyaei [g], Jia Jung [h,i], Shuai Pan [j], Soontae Kim [k], Cheol-Hee Kim [c,d], Juseon Bak [l], Jung-Woo Yoo [l], Jaehyeong Park [a], Dongjin Kim [a,d]

[a] *Division of Earth Environmental System, Pusan National University, Busan, 46241, South Korea*
[b] *Department of Earth and Atmospheric Sciences, University of Houston, Houston, TX, 77204, USA*
[c] *Department of Atmospheric Sciences, Pusan National University, Busan, 46241, South Korea*
[d] *BK21 School of Earth and Environmental System, Pusan National University, Busan, 46241, South Korea*
[e] *Air Resources Laboratory, National Oceanic and Atmospheric Administration, College Park, MD, 20740, USA*
[f] *Cooperative Institute for Satellite Earth System Studies, University of Maryland, College Park, MD, 20742, USA*
[g] *Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, 08540, USA*
[h] *Biospheric Sciences Branch, NASA Ames Research Center, Moffett Field, CA, 94035, USA*
[i] *Bay Area Environmental Research Institute, Moffett Field, CA, 94035, USA*
[j] *Emergency Management College, Nanjing University of Information Science and Technology (NUIST), Nanjing, Jiangsu, 210044, China*
[k] *Department of Environmental and Safety Engineering, Ajou University, Suwon, 16499, South Korea*
[l] *Institute of Environmental Studies, Pusan National University, Busan, 46241, South Korea*

HIGHLIGHTS

- Constraints of CO and NO_2 emissions via a hybrid approach that combines the IFDMB and 4D-Var inversions was conducted.
- The IFDMB can constrain the spatial distribution of emissions, and the 4D-Var can estimate their temporal variations.
- The hybrid inversion showed the best performance in constraining the spatio-temporal distribution of emissions.
- The inversion errors of NO_2 were larger than those of CO due to the larger nonlinearity between emissions and concentrations.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Inverse modeling
IFDMB

ABSTRACT

We performed a hybrid approach that combines the iterative finite difference mass balance (IFDMB) and four-dimensional variational data assimilation (4D-Var) methods to effectively constrain the spatiotemporal

* Corresponding author. Department of Atmospheric Sciences, Pusan National University, Busan, 46241, South Korea.
E-mail address: wbjeon@pusan.ac.kr (W. Jeon).

https://doi.org/10.1016/j.atmosenv.2024.120490
Received 13 November 2023; Received in revised form 26 March 2024; Accepted 28 March 2024
Available online 31 March 2024

4D-var
Adjoint model
CMAQ
South Korea

distribution of emissions. To quantitatively compare the performance of inverse modeling in constraining CO and NO_2 emissions in South Korea spatiotemporally, we conducted a model-based twin experiment for three inverse modeling methods: IFDMB, 4D-Var, and hybrid inversions. We performed numerical modeling using the Community Multi-scale Air Quality (CMAQ) and its adjoints to calculate the values required for inverse modeling. As a result, the IFDMB inversion can effectively constrain the average spatial distribution of emissions. Meanwhile, the 4D-Var inversion can help estimate temporal variations in emissions, but it is not effective in regions with large prior emission errors. The hybrid inversion showed the best performance in constraining the spatiotemporal distribution of emissions because it combined the strengths of the two aforementioned methods. Furthermore, to compare the performance of the inverse modeling of pollutants with different chemical properties, we conducted additional inverse modeling for highly reactive NO_2. After the application of inverse modeling, the emission errors of NO_2 (18.339%) were larger than those of CO (10.593%). This difference in inverse modeling errors was due to the greater nonlinear relationship between emissions and concentrations in the inverse modeling process for NO_2, which is more reactive compared to CO. In this study, the ideal modeling tests were performed to quantitatively assess the performance of inverse modeling. In future studies, we expect to apply the hybrid inversion approach used in this study to inverse modeling using actual observations.

1. Introduction

The accuracy of the air quality simulation results depends on several factors, including meteorological fields, chemical reaction mechanisms, and emissions (Guo et al., 2014; Miao et al., 2017, 2018; Zhang et al., 2015). Emission inventories provide crucial input data for air quality modeling to investigate the causes of air pollution (Cheng et al., 2021). The emission inventories were estimated using two methods: bottom-up and top-down. The bottom-up approach estimates the annual or monthly total emissions from direct observations of all identifiable sources based on statistical emission factors derived from human activities (Crippa et al., 2018; Hu et al., 2023; Li et al., 2017; McQuilling and Adams, 2015; Woo et al., 2012). In contrast, the top-down approach estimates the optimized emissions for air quality modeling by minimizing the differences between observations and models (Kwon et al., 2021; Mun et al., 2023; Wang et al., 2012).

The Emissions Database for Global Atmospheric Research (EDGAR), a prominent bottom-up emissions inventory, calculates the spatiotemporal distribution of emissions using annual totals and temporal profiles at a resolution of $0.1^\circ \times 0.1^\circ$ for the entire global domain (Crippa et al., 2020). Temporal profiles are provided based on country, substance, and emission sectors, allowing for the identification of general temporal trends in regional emissions (Crippa et al., 2018). However, it is challenging to represent temporal variations in emissions caused by unidentified sources (Gu et al., 2016; Ma and van Aardenne, 2004; Martin et al., 2003). We need to allocate the total emissions spatiotemporally to hourly grids for air quality modeling (Hu et al., 2023; Woo et al., 2012), which can increase uncertainty owing to the resolution differences between the two datasets. Therefore, to reduce the uncertainties associated with the bottom-up emission inventory, it is necessary to employ a top-down inverse modeling approach to constrain emissions considering the spatiotemporal characteristics of the model.

Inverse modeling is an effective method for constraining emissions by considering the processes in the model (e.g., transport, deposition, and chemistry) of pollutants based on observational data (Souri et al., 2016). Representative inverse modeling methods include the ensemble Kalman filter (EnKF) method (Chen et al., 2019; Dai et al., 2021; Jia et al., 2022; Peng et al., 2017), the mass balance method (Cooper et al., 2017; Lamsal et al., 2011; Li et al., 2019; Martin et al., 2003; Mun et al., 2023), and the four-dimensional variational data assimilation (4D-Var) method (Chen et al., 2021; Henze et al., 2009; Kurokawa et al., 2009; Li et al., 2019; Wang et al., 2012). Mass balance-based methods, including Basic Mass Balance (BMB), Finite Difference Mass Balance (FDMB), and Iterative Finite Difference Mass Balance (IFDMB), are the most computationally efficient approach for constraining emissions, even though they are vulnerable to smearing errors caused by pollutant transport (Choi et al., 2022; Li et al., 2019). Among others, the IFDMB method is designed to reduce smearing errors through iterative calculations (Cooper et al., 2017; Mun et al., 2023). On the contrary, the 4D-Var inversion approach explicitly calculates source influences, which can effectively minimize smearing errors (Li et al., 2019). However, it requires a complex adjoint model demanding a higher computational cost compared to the mass balance inversion (Cooper et al., 2017). Several recent studies have reported a hybrid inversion approach combining mass balance and 4D-Var inversion methods, resulting in more accurate emission estimations at relatively low computational costs (Chen et al., 2021; Qu et al., 2017, 2019).

The aforementioned inverse modeling methods primarily focus on estimating the average spatial distribution of emissions over a specific period, and there is a dearth of research on estimating the hourly emissions. Meanwhile, the 4D-Var inversion can constrain the initial conditions of the forecast model using an adjoint model (Chai et al., 2007; Park et al., 2016). In this study, to constrain hourly emissions, the 4D-Var inversion was performed for each hour instead of at the initial time. We performed a hybrid inversion method by combining the IFDMB and the 4D-Var to effectively constrain the spatiotemporal distribution of emissions. Before dealing with the actual observational data requiring substantial calculation on observational and background errors, we conducted a model-based twin experiment to quantitatively evaluate the performance of the inverse modeling methods in constraining the spatiotemporal distribution of emissions. We performed inverse modeling of CO and NO_2 to compare the effects of inversion on pollutants with different chemical properties. Thus, we quantitatively assessed the spatiotemporal constraint errors of emissions using three inverse modeling methods (IFDMB, 4D-Var, and hybrid inversion) for two different pollutants through a synthetic data inversion experiment.

2. Data and methodology

2.1. CMAQ forward model

We used the Community Multi-scale Air Quality (CMAQ) model (v5.0) (Byun and Schere, 2006), a photochemical model developed by the U.S. Environmental Protection Agency (EPA), for the air quality simulation. Meteorological input fields for CMAQ modeling were simulated using Weather Research and Forecasting (WRF, v3.8.1, Skamarock et al., 2008). The initial input data for WRF modeling were ERA5 reanalysis data ($0.25^\circ \times 0.25^\circ$) (Hersbach et al., 2020) provided by the European Center for Medium-Range Weather Forecasts (ECMWF), and grid nudging was applied to optimize the meteorological field (Jeon et al., 2015). The WRF meteorological field was converted into meteorological input data for CMAQ modeling using the Meteorology Chemistry Interface Processor (MCIP, v4.2).

The CMAQ domains include the upper domain (D1) covering East Asia and the lower domain (D2) covering South Korea (Fig. 1). Inverse modeling was performed for a subdomain (D2), and the grid resolution was set to 27 km, which showed the most effective inverse modeling results in Mun et al. (2023). For the anthropogenic emission input data used in the CMAQ modeling, we utilized the EDGAR v6.1 emissions data (Crippa et al., 2022) for the East Asian region, and the Clean Air Policy Support System (CAPSS) 2019 for South Korea. Biogenic emissions were estimated using the Model of Emissions of Gases and Aerosols from Nature (MEGAN) version 2.1. We compared the results of the IFDMB inversion using different modeling periods (1 day, 5 days, 10 days, 14 days) to determine the optimal modeling period for inverse modeling. The results indicated that a 14-day period resulted in the lowest emission errors (Fig. S1). Thus, the simulations were conducted from 2022.06.25.00 UTC to 2022.07.15.00 UTC, including a 5-day spin-up time, during the Satellite Integrated Joint Monitoring of Air Quality (SIJAQ) campaign in 2022. For the inverse modeling, we used model data from 2022. 07. 01. 00 UTC to 2022. 07. 14. 23 UTC (two weeks). The detailed configurations applied to the meteorological and air quality models are presented in Tables 1 and 2, respectively.

Table 1
Detailed configurations of the Weather Research and Forecasting (WRF) simulation.

WRF	Domain 01(D1)	Domain 02(D2)
Horizontal Resolution	27 km	27 km
Vertical layers	35 Layers	
Microphysics option	WSM 3-class simple ice scheme	
Radiation option	RRTM scheme(long-wave) Dudhia scheme(short wave)	
Surface layer option	Revised MM5 Monin-Obukhov scheme (Jimenez)	
Land-surface option	Unified Noah land-surface model	
PBL option	YSU scheme	
Cumulus option	Kain-Fritsch (new Eta) scheme	
Initial data	ERA5 (0.25° × 0.25°)	
Simulation period	2022. 06. 25. 00 UTC – 2022. 07. 15. 00 UTC	

Table 2
Detailed configurations of the Community Multi-scale Air Quality (CMAQ) simulation.

CMAQ	Domain 01(D1)	Domain 02(D2)
Horizontal Resolution	27 km	27 km
Vertical layers	15 Layers	
Biogenic emission	MEGANv2.1	
Anthropogenic emission	EDGARv6.1	CAPSS 2019
Chemical mechanism	cb5cl_ae5_aq	
Simulation period	2022. 06. 25. 00 UTC – 2022. 07. 15. 00 UTC	

2.2. Inverse modeling methods

2.2.1. IFDMB inversion

Mass balance inversion is performed by assuming a linear relationship between surface emissions and column density (Cooper et al., 2017), and it is a method of constraining surface emissions by comparing the column density of observational data and model results (Martin et al., 2003). The IFDMB method reduces emission errors through iterative calculations (Cooper et al., 2017; Mun et al., 2023). Thus, we used the IFDMB method to constrain the spatial distribution of the emissions. The IFDMB inversion involved the following four steps: (1) CMAQ modeling was conducted using prior emissions (E_a) and emissions perturbed by 10%. (2) The scaling factor (β), which represents the change in the column concentration ($\frac{\Delta\Omega_a}{\Omega_a}$) relative to the emission variation ($\frac{\Delta E_a}{E_a}$), was calculated based on the numerical simulation results (Eq. (1)). (3) The updated emissions (E_t) were calculated using Eq. (2). (4) The processes of steps (1)–(3) are repeated until the Normalized Mean Difference (NMD) between the updated emissions and the previous emissions is less than 1% (Eq. (3)). We performed the IFDMB inversion only for grids with non-zero prior emissions because it is impossible to calculate β for grids with zero values in the prior emissions. Theoretically, the constraints of temporal emission variations can be achieved by implementing hourly IFDMB, but this method requires satellite data that are available hourly. Unfortunately, it is difficult to obtain hourly satellite observations because polar-orbiting satellites (e.g., Ozone Monitoring Instrument (OMI), Tropospheric Monitoring Instrument (TROPOMI)) provide observations around South Korea only once or twice a day, and geostationary satellites (e.g., Geostationary Environmental Monitoring Spectrometer (GEMS)) only eight times a day. Furthermore, to mitigate the smearing effects induced by the transport of pollutants, it is imperative to use the average column concentration over the entire study period. Consequently, in this study, we opted to employ the average column concentration over the study period (2 weeks) to constrain the average spatial distribution of emissions. It is worth noting that this approach comes with the inherent limitation of not addressing the temporal variability in emissions.

$$\beta = \frac{\Delta\Omega_a/\Omega_a}{\Delta E_a/E_a} \quad \text{(Eq. 1)}$$

$$E_t = E_a\left(1 + \frac{\Omega_o - \Omega_a}{\beta\Omega_a}\right) \quad \text{(Eq. 2)}$$

$$NMD = \frac{\sum|E - E_t|}{\sum E_t} \times 100\ [\%] \quad \text{(Eq. 3)}$$

2.2.2. 4D-Var inversion and the CMAQ adjoint model

While the IFDMB method proves effective in constraining the average spatial distribution of emissions, it is limited in its ability to correct for temporal variations. This limitation arises from using average column concentrations in the IFDMB inverse modeling process conducted in this study. In this section, we introduce the 4D-Var inversion

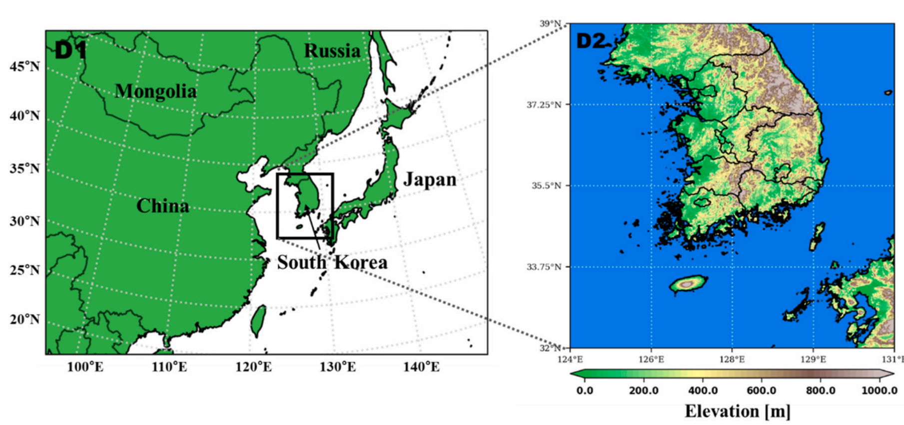


Fig. 1. The simulation domains of the Weather Research and Forecasting (WRF) and Community Multi-scale Air Quality (CMAQ) models.

method, which allows for an effective constraint on the temporal variation of emissions.

The 4D-Var assimilation aims to find a control variable that minimizes the difference between the model results and observations within a time window. The cost function ($J(X_t)$) of the 4D-Var is defined as

$$J(X_t) = (\alpha_t - 1)^T B^{-1} (\alpha_t - 1) + \frac{1}{2} \sum_{i=0}^{t_f} (H(X_t) - y_i)^T R^{-1} (H(X_t) - y_i) \quad \text{(Eq. 4)}$$

where the emissions at time t (X_t) are the control variables, the emission scaling factor (α_t) is the ratio of the control variable to the prior emissions, H is the observation operator of the forward model, y_i is the observation vector, and R and B are the error covariance matrices for the observation and prior emissions, respectively. We utilized the column concentrations for the observation. The two error covariance matrices, R and B, were assumed to be diagonal matrices, assuming independence of the data in each grid. To compare the results of the 4D-Var inverse modeling for CO and NO_2, this study controlled for uncertainties in the error estimation. Since each pollutant has different concentration and emission levels, scaling of observation and emission errors was necessary for comparison under the same conditions. Therefore, we set the observation error as 1% of the observed value and the prior emissions error as 100% of the prior emissions value. We calculated the cost function for model emissions only for grids with non-zero prior emissions because α_t cannot be mathematically computed for grids with zero prior emissions.

To minimize the cost function of the 4D-Var, we used the L-BFGS-B quasi-Newton optimization algorithm (Zhu et al., 1994). The minimization algorithm, L-BFGS-B, requires the values of the cost function and its gradient. We calculated the cost function using the forward model and its gradient with the adjoint model (Zhao et al., 2020), which enabled explicit quantification of source influences by tracking processes, including transport (Choi et al., 2022; Li et al., 2019). In other words, the adjoint model can be employed to quantify the influence of the initial emissions on the concentrations within an assimilation time window. In several inverse modeling studies, the 4D-Var has been employed to assimilate the initial conditions to enhance forecast accuracy within the time window (Chai et al., 2007; Park et al., 2016; Peng and Xie, 2006). To constrain the hourly emissions, an approach different from the traditional 4D-Var assimilation of the initial conditions is required. The concentration at a specific time is influenced by the pollutants emitted up to the previous time. If we constrain emissions from 1 h ago based on the difference in current concentrations between observations and model results, it can lead to overestimation or underestimation of emissions due to the mixing of the influence of pollutants emitted prior to that time. To prevent the temporal mixing of emissions' impacts on concentration, this study updated the prior emissions on an hourly basis. Therefore, in this study, we set the time window to 1 h and performed the 4D-Var inversion every hour. In this approach, the 4D-Var inversion was conducted under the assumption that the difference between the observed and modeled concentrations within the 1-h window was solely attributed to emissions from the previous hour. We controlled for other conditions that could contribute to differences between observations and models, excluding only emissions. This study assumed these ideal conditions to evaluate the possibility of constraining the temporal variation of emissions using the 4D-Var inversion. However, for the application of this approach to real-world scenarios, it is essential to consider other error sources such as meteorological conditions, observation distribution, and chemical mechanisms.

Fig. 2 presents a flowchart of the hourly 4D-Var inversion methodology proposed in this study. First, forward and adjoint modeling was performed with a simulation time of 1 h. Subsequently, we solved the 4D-Var cost function using the modeled and observed concentrations. The process was iterated until the cost function converged or the change in emissions was less than 1%. In this case, α_t is limited within the range of 0.25–4 to ensure realistic emissions correction. The procedure was

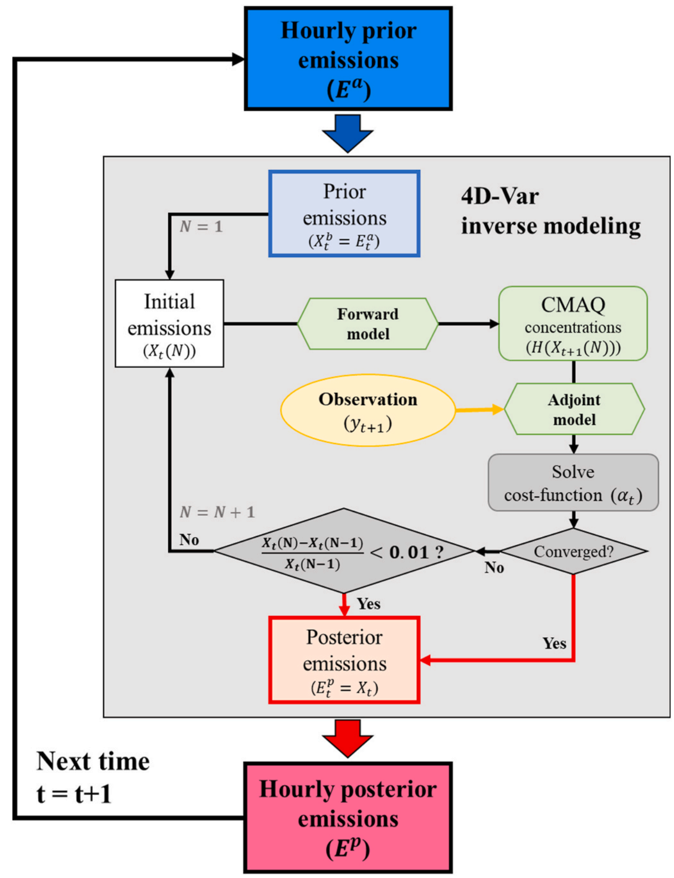


Fig. 2. Flowchart of the hourly four-dimensional variational data assimilation (4D-Var) inversion.

repeated every hour to constrain hourly emissions. Finally, we update the hourly prior emissions (E^a) to hourly posterior emissions (E^p).

2.2.3. Hybrid inversion

In previous studies, a hybrid approach that combined mass balance and 4D-Var methods was utilized to efficiently correct the spatial distribution of emissions (Chen et al., 2021; Qu et al., 2017). However, this approach primarily focused on constraining the average spatial distribution of emissions and did not address the constraints on hourly emissions. Unlike the previous approach, the 4D-Var inversion method proposed in this study can constrain hourly emissions. Therefore, we propose a novel hybrid inversion approach that combines the IFDMB method, which effectively constrains the average spatial distribution of emissions, with the 4D-Var method, which estimates the temporal variations in emissions. The hybrid inversion method was performed in the following two steps independently. First, the IFDMB inversion was employed to spatially refine prior emissions and mitigate the quantitative bias in the average spatial distribution of emissions. Subsequently, the 4D-Var inversion was applied to correct emissions constrained by the IFDMB inversion on an hourly basis, thereby rectifying temporal variations in emissions.

2.3. Model-based twin experiment design

While we constrained emissions using observational data, it remains practically unfeasible to directly verify the accuracy of emissions before and after the constraints due to the elusive nature of true emissions (Mun et al., 2023). Therefore, this study seeks to quantitatively assess the effectiveness of inverse modeling methods (IFDMB, 4D-Var, and hybrid) through a model-based twin experiment before applying actual observational data. The twin experiment is an ideal modeling test to assess the capability of the inverse modeling system to reproduce the true answer using a true solution for the simulation (Kim et al., 2020). In this study, the model-based twin experiment was executed in three sequential steps. Firstly, two emission inventories were prepared, with one designated as “true" emissions and the other as “prior" emissions. The “prior" emissions were based on the EDGARv6.1 emissions, while the “true" emissions were derived from the CAPSS 2019 emissions. We configured the experiments by setting the “prior” emissions only for surface grids to the EDGARv6.1 emissions to compare the three inverse modeling methods under the same conditions. This is because the IFDMB method constrained only the surface emissions by assuming a linear relationship between surface emissions and column concentrations. To assess the impact of inverse modeling on the temporal emission distribution, the “true" emissions were allocated according to their temporal profiles, while the “prior" emissions were uniformly distributed over time. Secondly, we operated on the assumption that model results, generated using the “true" emissions, accurately represent gridded data available continuously without any missing pixels. These model-generated data were treated as observations in the inverse modeling process. Finally, a comparative analysis was conducted between the emissions pre- and post-inverse modeling with the “true" emissions to quantitatively evaluate the efficacy of the inverse modeling application.

3. Results and discussion

3.1. IFDMB inversion

We constrained the average CO emissions over the study period using the IFDMB inversion. Before conducting the inverse modeling, we compared the average spatial distribution of the “prior” and “true” emissions (Fig. 3). Overall, the “prior” emissions were overestimated compared to the “true” emissions, especially in Seoul Metropolitan Area (SMA), Inje, and Pohang of South Korea.

To evaluate the spatial correction effect of the IFDMB inversion, we analyzed the spatial difference between the “true" emissions and the emissions before and after inversion (Fig. 4). After applying the IFDMB inversion, the bias in the “prior” emissions were substantially reduced in most grid cells. However, there was a negative bias in the SMA, an upwind area of heavy emissions, and a positive bias in the northeast region and downwind areas of low emissions. These biases can be attributed to the smearing errors caused by the transport of pollutants by westerly winds. Additionally, we examined the Normalized Mean Error (NME) of the emissions before and after the IFDMB inversion to quantitatively compare the average spatial emission errors. The results showed a significant reduction in NME from 76.522% in the “prior" emissions to 12.207% after IFDMB inversion. Thus, the IFDMB inversion effectively constrained the spatial distribution of the emissions.

To compare the temporal variations in emissions before and after the IFDMB inversion, we calculated the average emissions for each time across all grid cells (Fig. 5). In this study, the IFDMB inversion was performed using average column concentrations, hence, the IFDMB emissions were constrained by the same ratio for all hours. As a result, the IFDMB method significantly reduced the spatial bias in the “prior” emissions, but it had a limitation in constraining the temporal variation of emissions.

3.2. 4D-Var inversion

To compensate for the limitations of the IFDMB inversion, this study utilized the 4D-Var inversion method based on the adjoint model. We evaluated the spatial correction effect of the 4D-Var inversion by presenting the spatial distribution of emission errors, averaged over the study period, in Fig. 6. The 4D-Var inversion showed smaller spatial errors in emissions than the IFDMB inversion for most grid cells. Owing to the explicit calculation of source effects using the adjoint model, the 4D-Var inversion was more effective in constraining emissions, reducing susceptibility to smearing effects. However, the method was less effective in correcting emissions for specific grids (SMA, Inje, and Pohang) where significant errors were observed in the “prior” emissions. Unlike the IFDMB inversion, which directly constrains emissions, the 4D-Var inversion focused on constraining the emission scaling factor, limiting its effectiveness for grids with substantial emission errors. The NME for the 4D-Var inversion was 28.161%, higher than the 12.207% observed with the IFDMB inversion, mainly due to grids with large emission errors. Therefore, in regions where model and actual emission distributions differ significantly, the 4D-Var inversion may not be less effective. In other words, for estimating emission sources that cannot be identified through a bottom-up emission inventory, the IFDMB inversion may be more effective than the 4D-Var inversion in spatially averaging emissions.

We examined the temporal variations in the emissions from the 4D-Var inversion (Fig. 5). Unlike the IFDMB inversion, the 4D-Var inversion could estimate the temporal variation in emissions since it constrained the emissions on an hourly basis. However, significant emission biases were observed almost every hour. This occurred in regions such as SMA, Inje, and Pohang, where the average spatial emission errors of the 4D-Var inversion were considerable, leading to substantial errors each hour (Fig. S2).

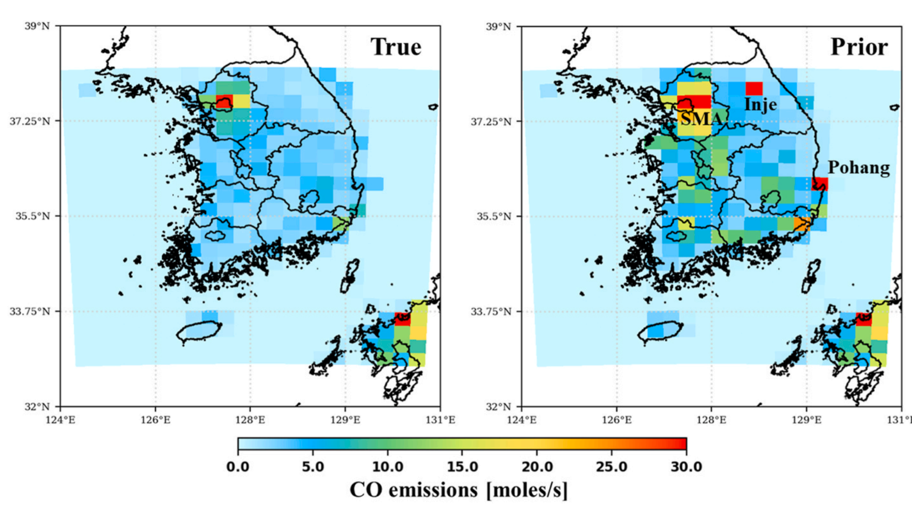


Fig. 3. Distributions of “true” (left) and “prior” (right) CO emissions.

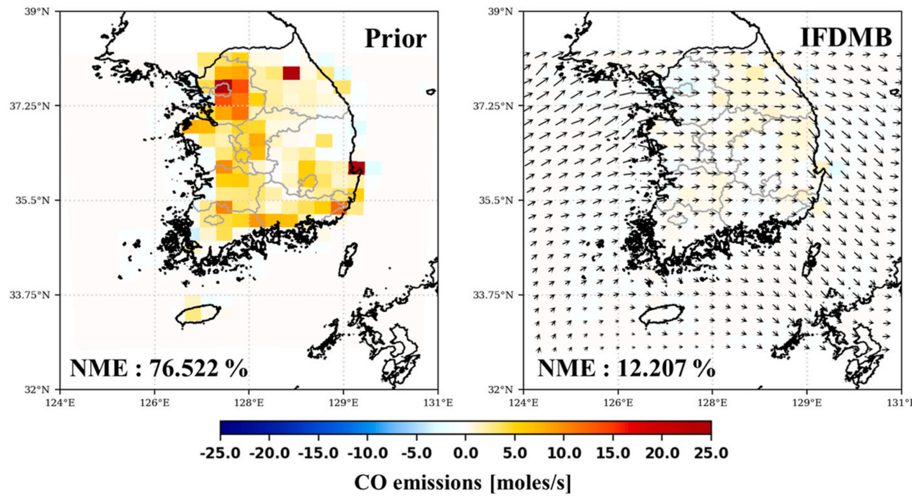


Fig. 4. Spatial distributions of CO emission errors before (left) and after (right) iterative finite difference mass balance (IFDMB) inversion. The errors are calculated as the difference from "true" emissions.

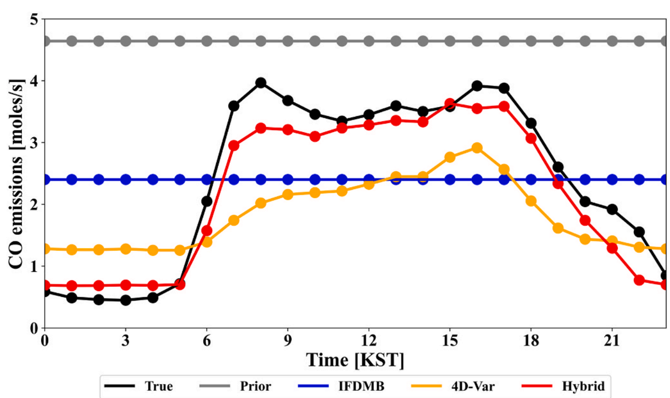


Fig. 5. Diurnal variations of CO emissions averaged over study periods for each experiment: "true" emissions (True), "prior" emissions (Prior), constrained emissions using the iterative finite difference mass balance inversion (IFDMB), constrained emissions using the four-dimensional variational data assimilation inversion (4D-Var), and constrained emissions using the hybrid inversion (Hybrid).

In summary, the 4D-Var inversion enables the estimation of hourly emission profiles by constraining the emissions at each hour. However, the method was not effective in regions with significant discrepancies between the model and the actual emissions. Therefore, we concluded to propose a hybrid inversion method that combines both the IFDMB and 4D-Var methodologies to compensate for these limitations.

3.3. *Hybrid inversion*

To evaluate the influence of the hybrid inversion on the spatial distribution of emissions, we examined the spatial distribution of emission errors (Fig. 7). Among the three inversion methods, the hybrid inversion method showed the least NME of 10.593%, suggesting that it was the most effective in constraining the average spatial distribution of emissions. Such efficacy is because the IFDMB inversion was first conducted over the study period, effectively compensating for the shortcomings of the 4D-Var inversion, which led to large emission discrepancies in specific areas.

We could ascertain the diurnal variations in emissions using the hybrid inversion, which incorporated the 4D-Var inversion (Fig. 5). Overall, the diurnal variations of the emissions constrained by the hybrid inversion showed the closest resemblance to those of the "true"

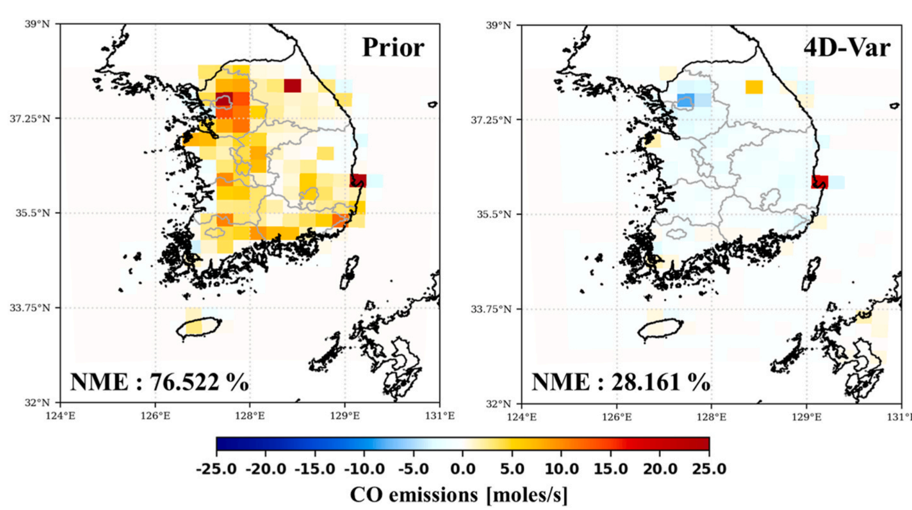


Fig. 6. Spatial distributions of CO emission errors before (left) and after (right) four-dimensional variational data assimilation (4D-Var) inversion. The errors are calculated as the difference from "true" emissions.

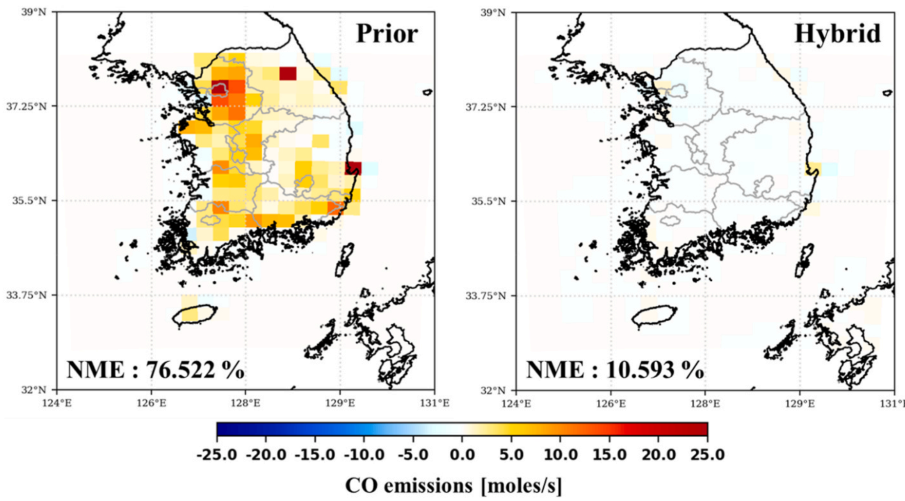


Fig. 7. Spatial distributions of CO emission errors before (left) and after (right) hybrid inversion. The errors are calculated as the difference from "true" emissions.

emissions. Additionally, to quantitatively compare the emission errors for each experiment, the RMSE time series for CO emissions were depicted in Fig. S3. The RMSE was calculated based on the difference between each experiment's emissions and the "true" emissions. The RMSE of emissions using the hybrid inversion was the smallest for almost every hour. Meanwhile, the RMSE of the constrained emissions using the 4D-Var inversion was higher than that of the IFDMB inversion for some hours. This is attributed to the ineffectiveness of the 4D-Var inversion in specific grids with significant errors observed in the "prior" emissions. Consequently, the hybrid inversion, which performed the 4D-Var inversion after enhancing the consistency of the emission spatial distribution through the IFDMB inversion, provided the most accurate constraint on the spatiotemporal distribution of emissions. In this study, we adopted a hybrid inversion approach, aiming to

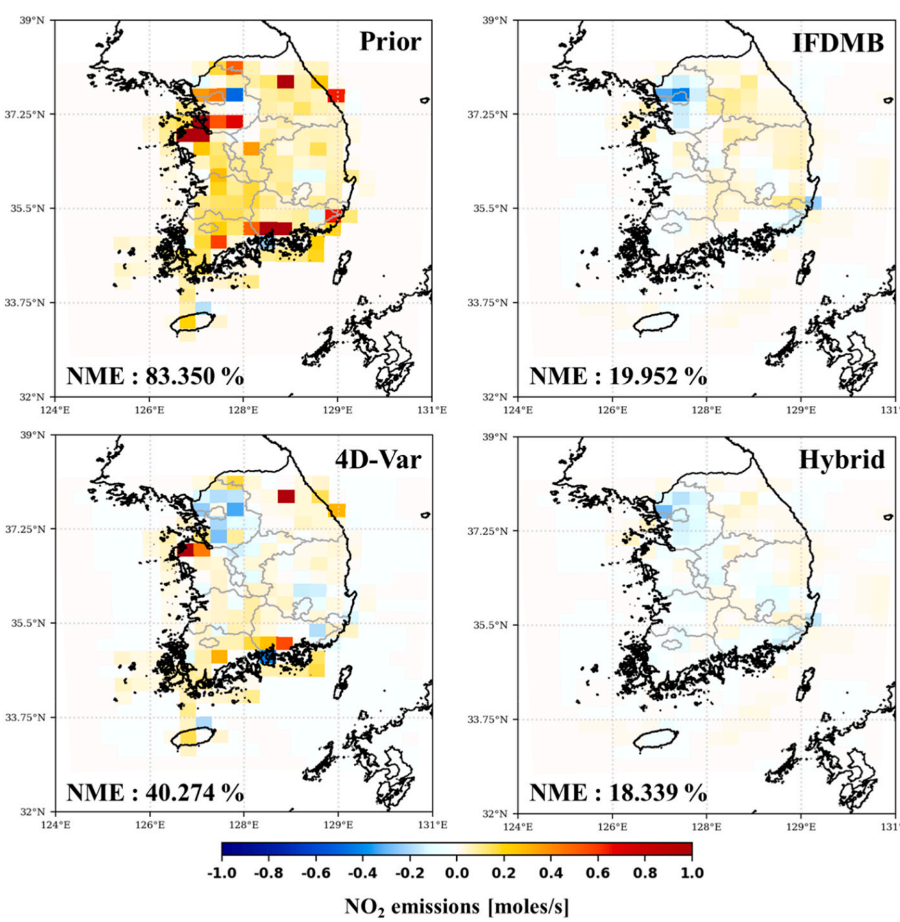


Fig. 8. Spatial distributions of NO_2 emission errors for each experiment: "prior" emissions (Prior), constrained emissions using the iterative finite difference mass balance inversion (IFDMB), constrained emissions using the four-dimensional variational data assimilation inversion (4D-Var), and constrained emissions using the hybrid inversion (Hybrid). The errors are calculated as the difference from "true" emissions.

effectively constrain the spatiotemporal distribution of emissions by using the strengths of both inverse modeling methods and compensating for their limitations.

3.4. Inverse modeling of NO_2 emissions

We introduced the IFDMB, 4D-Var, and hybrid inversion methods to provide spatiotemporal constraints on minor reactive CO emissions. In this section, we investigate the impact of employing these aforementioned inverse modeling methods on the spatiotemporal distribution of NO_2 emissions. NO_2 is highly reactive and acts as a significant precursor to secondary pollutants. Throughout the inverse modeling process, we exclusively used NO_2 concentrations to constrain its emissions.

We assessed the average spatial distribution of NO_2 emission errors to gauge the spatial correction effectiveness of applying the inverse modeling methods (Fig. 8). The NME for the "prior" emissions, initially at 83.350%, was significantly dropped to 19.952% following the IFDMB inversion. These findings affirm that the IFDMB inversion is effective in constraining the spatial distribution of NO_2 emissions. However, echoing the outcomes from the CO inverse modeling, smearing errors were observed in a few regions. In contrast, the NME of the 4D-Var inversion was 40.274%, which surpassed the IFDMB result. A high NME occurred in regions with significant discrepancies in the "prior" emissions. This result underscores that the 4D-Var inversion was not effective in constraining NO_2 emissions in these regions. The hybrid inversion, which combines the IFDMB and 4D-Var inversions, showed the smallest emission error of 18.339%. This result mirrors the results from CO inverse modeling and can be ascribed to the IFDMB inversion counterbalancing the inaccuracies inherent in the 4D-Var inversion. Thus, the hybrid inversion approach emerged as the most effective in constraining the average spatial distribution of NO_2 emissions.

Next, we analyzed the average hourly NO_2 emission errors to gauge the temporal correction efficacy of the inverse modeling methods (Fig. 9). While the IFDMB inversion method constrained NO_2 emissions to the same ratio for all hours, both the 4D-Var and hybrid inversions proved capable of constraining emissions on an hourly basis. Among these, the temporal distribution of emissions from the hybrid inversion emissions closely mirrored that of the "true" NO_2 emissions. For a quantitative comparison of the errors in NO_2 emissions for each experiment, the RMSE time series were shown in Fig. S4. Similar to the CO results, the emissions using the hybrid inversion had the smallest RMSE for nearly every hour. Thus, the hybrid inversion emerged as the most precise method for rectifying temporal discrepancies in NO_2 emissions.

Finally, we compared the results of the hybrid inversion for CO and NO_2, taking into account their distinct chemical properties. Assessing the impacts of inverse modeling on the average spatial distribution of emissions, the emission errors displayed comparable reductions for both CO (65.929%) and NO_2 (65.011%). Fig. 10 illustrates the temporal variations in emissions for both pollutants under the hybrid inversion. Given the quantitative disparities between the CO and NO_2 emissions, the hourly emission error was normalized to ensure an equitable comparison. The Normalized Emission Error (NEE) was estimated by dividing the hourly bias by the sum of the biases for all hours (Eq. (5)). As a result, the bias of NO_2 was larger than that of CO in most hours. Unlike CO, which is less reactive, NO_2 plays a role in generating various secondary pollutants, including O_3, NO_3, and N_2O_5. The larger bias in NO_2 emissions compared to CO emissions is attributed to the fact that this study did not consider the effects of these chemical reactions. In detail, in the NO_2 inverse modeling process of this study, NO_2 emissions were constrained to optimize only for NO_2 concentrations, and the concentrations of other pollutants involved in the production and loss of NO_2 were not considered. Notably, the hourly bias of NO_2 emissions was primarily evident during rush hour and early morning. Inverse modeling refines emissions by leveraging the relationship between concentration and emissions. Owing to the relatively linear relationship between the emissions and CO concentration, inverse modeling using only the CO concentration is effective. Conversely, for NO_2, variations in emissions impact not just the NO_2 concentration but also the concentrations of secondary pollutants such as O_3 during daytime and NO_3 and N_2O_5 at night. Because of this nonlinear relationship between NO_2 emissions and concentrations, the inverse modeling results for NO_2 were not as good as those for CO. This disparity became especially evident during peak rush hours and dawn when secondary pollutant production was more active. In future explorations, we anticipate enhanced constraints that consider observational data related to secondary pollutants, stemming from NO_2 chemical reactions, within the inverse modeling framework.

$$NEE = \frac{E_t^{hybrid} - E_t^{true}}{\sum_{i=0}^{23}\left(E_i^{hybrid} - E_i^{true}\right)} \quad \text{(Eq. 5)}$$

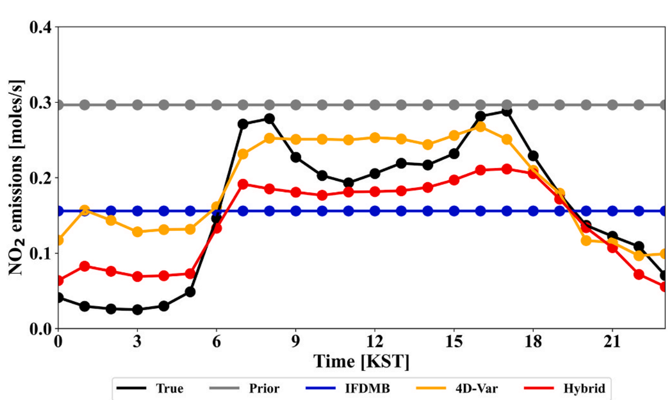


Fig. 9. Diurnal variations of NO_2 emissions averaged over study periods for each experiment: "true" emissions (True), "prior" emissions (Prior), constrained emissions using the iterative finite difference mass balance inversion (IFDMB), constrained emissions using the four-dimensional variational data assimilation inversion (4D-Var), and constrained emissions using the hybrid inversion (Hybrid).

4. Summary and conclusion

To effectively constrain the spatiotemporal distribution of emissions, this study implemented IFDMB, 4D-Var, and hybrid inversions, the latter combining both methods. Before utilizing actual observations, we quantitatively gauge the impact of inverse modeling by performing a model-based twin experiment. Inverse modeling was performed for less reactive CO and more reactive NO_2. The inverse modeling spanned a two-week period from July 1 to July 14, 2022, during the SIJAQ 2022 campaign.

The IFDMB inversion constrained the average spatial distribution of emissions by comparing the modeled and observed column densities

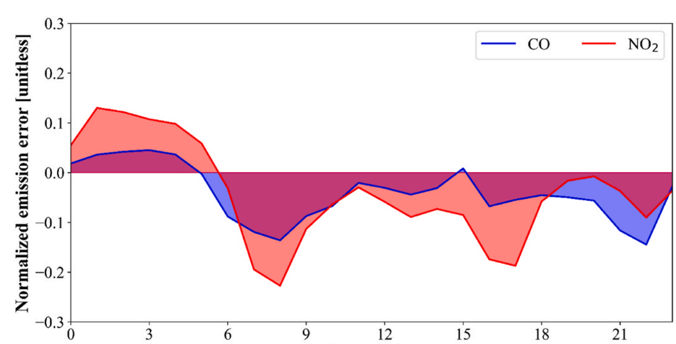


Fig. 10. Diurnal variations of CO and NO_2 emission errors. The errors are calculated by dividing the hourly bias by the sum of the bias for all hours.

over the study period. Although emission errors saw notable reductions across most regions, smearing errors occurred in the source and receptor regions due to the transportation of pollutants. Despite the smearing errors in the source and receptor regions, the spatial emission errors for CO and NO_2 declined by an average of 64.315% and 63.398%, respectively, throughout the study period. These reductions highlight the efficacy of the IFDMB inversion in spatially constraining emissions. In this study, a limitation of the IFDMB inversion, however, stems from its method of comparing averaged observations with the model, making it unsuitable for constraining emissions on an hourly scale. To counteract this shortcoming of the IFDMB, we employed the 4D-Var inversion, rooted in the adjoint model, to constrain the temporal distribution of emissions. This enabled us to constrain emissions hourly and generate a temporal emission profile via the 4D-Var inversion. However, for certain regions characterized by substantial errors in the "prior" emissions, the 4D-Var inversion proved inadequate. Due to these regions, the 4D-Var inversion exhibited higher average spatial emission errors compared to the IFDMB inversion.

Therefore, this study introduces a hybrid inversion approach, blending the strengths of the IFDMB inversion effective for constraining the average spatial distribution of emissions and the 4D-Var inversion, which constrains hourly emissions. The hybrid inversion approach first reduced the average emission bias through the IFDMB inversion and subsequently refined the hourly emissions via the 4D-Var inversion. As a result, the hybrid inversion results in the smallest spatiotemporal error among the aforementioned inverse modeling methods, enabling the most effective constraints on the spatiotemporal distribution of emissions.

Furthermore, the hybrid inversion was conducted for both CO and NO_2 to discern the differences in the effects of inverse modeling on these two pollutants, which have distinct chemical properties. First, we compared the impact on the average spatial distribution of emissions, which was closely matched at 65.929% for CO and 65.011% for NO_2. While the CO emission errors remained under 70% for most of the hours, the errors for NO_2 surged to 195%. This disparity arises from the unique chemical reaction properties of each pollutant. Because CO is less reactive, the results of the inverse modeling for CO were effective, using only the CO concentrations. Conversely, due to NO_2's heightened reactivity, only NO_2 concentrations appear inadequate for constraining its emissions. Hence, it is imperative to incorporate concentrations of other pollutants associated with NO_2, such as O_3, NO_3, and N_2O_5.

This study quantitatively assesses the impact of inverse modeling on spatiotemporal constraints by using a model-based twin experiment. The results of this study were based on the assumption that differences in concentration between observations and models come from solely emissions, controlling for complex error sources such as meteorological conditions, observation distribution, and chemical mechanisms. However, in the real world, inverse modeling errors can arise from the aforementioned conditions. Therefore, future inverse modeling studies utilizing actual observational data should consider realistic inverse modeling error sources.

CRediT authorship contribution statement

Jeonghyeok Moon: Writing – original draft, Methodology, Formal analysis, Conceptualization. **Yunsoo Choi:** Writing – review & editing, Supervision. **Wonbae Jeon:** Writing – review & editing, Supervision, Conceptualization. **Hyun Cheol Kim:** Investigation, Conceptualization. **Arman Pouyaei:** Writing – review & editing, Methodology. **Jia Jung:** Writing – review & editing. **Shuai Pan:** Writing – review & editing. **Soontae Kim:** Writing – review & editing. **Cheol-Hee Kim:** Project administration, Funding acquisition. **Juseon Bak:** Formal analysis. **Jung-Woo Yoo:** Writing – review & editing. **Jaehyeong Park:** Investigation, Formal analysis. **Dongjin Kim:** Validation, Resources.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgments

This research was partially supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2020R1A6A1A03044834) and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2023R1A2C1002608). H.C.K was also supported by NOAA grant NA19NES4320002 (CISESS).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosenv.2024.120490.

References

Byun, D., Schere, K.L., 2006. Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system. Appl. Mech. Rev. 59, 51–77. https://doi.org/10.1115/1.2128636.

Chai, T., Carmichael, G.R., Tang, Y., Sandu, A., Hardesty, M., Pilewskie, P., Whitlow, S., Browell, E.V., Avery, M.A., Nédélec, P., Merrill, J.T., Thompson, A.M., Williams, E., 2007. Four-dimensional data assimilation experiments with international consortium for atmospheric research on transport and transformation ozone measurements. J. Geophys. Res. 112, D12S15 https://doi.org/10.1029/2006JD007763.

Chen, D., Liu, Z., Ban, J., Chen, M., 2019. The 2015 and 2016 wintertime air pollution in China: SO_2 emission changes derived from a WRF-Chem/EnKF coupled data assimilation system. Atmos. Chem. Phys. 19 (13), 8619–8650. https://doi.org/10.5194/acp-19-8619-2019.

Cheng, X., Hao, Z., Zang, Z., Liu, Z., Xu, X., Wang, S., Liu, Y., Hu, Y., Ma, X., 2021. A new inverse modeling approach for emission sources based on the DDM-3D and 3DVAR techniques: an application to air quality forecasts in the Beijing-Tianjin-Hebei region. Atmos. Chem. Phys. 21 (18), 13747–13761. https://doi.org/10.5194/acp-21-13747-2021.

Chen, Y., Shen, H., Kaiser, J., Hu, Y., Capps, S.L., Zhao, S., Hakami, A., Shih, J.-S., Pavur, G.K., Turner, M.D., Henze, D.K., Resler, J., Nenes, A., Napelenok, S.L., Bash, J.O., Fahey, K.M., Carmichael, G.R., Chai, T., Clarisse, L., Coheur, P.-F., van Damme, M., Russell, A.G., 2021. High-resolution hybrid inversion of IASI ammonia columns to constrain US ammonia emissions using the CMAQ adjoint model. Atmos. Chem. Phys. 21, 2067–2082. https://doi.org/10.5194/acp-21-2067-2021.

Choi, J., Henze, D.K., Cao, H., Nowlan, C.R., González Abad, G., Kwon, H., Lee, H., Oak, Y.J., Park, R.J., Bates, K.H., Maasakkers, J.D., Wisthaler, A., Weinheimer, A.J., 2022. An inversion framework for optimizing non-methane VOC emissions using remote sensing and airborne observations in northeast Asia during the KORUS-AQ field campaign. J. Geophys. Res. Atmos. 127, e2021JD035844 https://doi.org/10.1029/2021jd035844.

Cooper, M., Martin, R.V., Padmanabhan, A., Henze, D.K., 2017. Comparing mass balance and adjoint methods for inverse modeling of nitrogen dioxide columns for global nitrogen oxide emissions. J. Geophys. Res. Atmos. 122, 4718–4734. https://doi.org/10.1002/2016JD025985.

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., van Aardenne, J.A., Monni, S., Doering, U., Olivier, J.G.J., Pagliari, V., Janssens-Maenhout, G., 2018. Gridded emissions of air pollutants for the period 1970-2012 within EDGAR v4.3.2. Earth Syst. Sci. Data 10 (4), 1987–2013. https://doi.org/10.5194/essd-10-1987-2018.

Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Monforti-Ferrario, F., Banja, M., Pagani, F., Solazzo, E., 2022. EDGAR v6.1 Global Air Pollutant Emissions. European Commission, JRC129555.

Crippa, M., Solazzo, E., Huang, G., Guizzardi, D., Koffi, E., Muntean, M., Schieberle, C., Friedrich, R., Janssens-Maenhout, G., 2020. High resolution temporal profiles in the Emissions Database for Global Atmospheric Research. Sci. Data. 7 (1), 121. https://doi.org/10.1038/s41597-020-0462-2.

Dai, T., Cheng, Y., Goto, D., Li, Y., Tang, X., Shi, G., Nakajima, T., 2021. Revealing the sulfur dioxide emission reductions in China by assimilating surface observations in WRF-Chem. Atmos. Chem. Phys. 21 (6), 4357–4379. https://doi.org/10.5194/acp-21-4357-2021.

Gu, D., Wang, Y., Yin, R., Zhang, Y., Smeltzer, C., 2016. Inverse modelling of NO_x emissions over eastern China: uncertainties due to chemical non-linearity. Atmos. Meas. Tech. 9 (10), 5193–5201. https://doi.org/10.5194/amt-9-5193-2016.

Guo, S., Hu, M., Zamora, M.L., Peng, J., Shang, D., Zheng, J., Du, Z., Wu, Z., Shao, M., Zeng, L., Molina, M.J., Zhang, R., 2014. Elucidating severe urban haze formation in China. Proc. Natl. Acad. Sci. USA 111, 17373–17378. https://doi.org/10.1073/pnas.1419604111.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G. D., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146, 1999–2049. https://doi.org/10.1002/qj.3803.

Henze, D.K., Seinfeld, J.H., Shindell, D.T., 2009. Inverse modeling and mapping US air quality influences of inorganic $PM_{2.5}$ precursor emissions using the adjoint of GEOS-Chem. Atmos. Chem. Phys. 9, 5877–5903. https://doi.org/10.5194/acp-9-5877-2009.

Hu, Y., Li, Y., Ma, X., Liang, Y., You, W., Pan, X., Zang, Z., 2023. The optimization of SO_2 emissions by the 4DVAR and EnKF methods and its application in WRF-Chem. Sci. Total Environ. 888 https://doi.org/10.1016/j.scitotenv.2023.163796.

Jeon, W., Choi, Y., Lee, H.W., Lee, S.H., Yoo, J.W., Park, J., Lee, H.J., 2015. A quantitative analysis of grid nudging effect on each process of $PM_{2.5}$ production in the Korean Peninsula. Atmos. Environ. 122, 763–774. https://doi.org/10.1016/j.atmosenv.2015.10.050.

Jia, G., Huang, Z., Tang, X., Ou, J., Lu, M., Xu, Y., Zhong, Z., Sha, Q., Wu, H., Zheng, C., Deng, T., Chen, D., He, M., Zheng, J., 2022. A meteorologically adjusted ensemble Kalman filter approach for inversing daily emissions: a case study in the Pearl River Delta, China. J. Environ. Sci. 114 (4), 233–248. https://doi.org/10.1016/j.jes.2021.08.048.

Kim, H.C., Chai, T., Stein, A., Kondragunta, S., 2020. Inverse modeling of fire emissions constrained by smoke plume transport using HYSPLIT dispersion model and geostationary satellite observations. Atmos. Chem. Phys. 20 (17), 10259–10277. https://doi.org/10.5194/acp-20-10259-2020.

Kurokawa, J. ichi, Yumimoto, K., Uno, I., Ohara, T., 2009. Adjoint inverse modeling of NO_x emissions over eastern China using satellite observations of NO_2 vertical column densities. Atmos. Environ. 43 (11), 1878–1887. https://doi.org/10.1016/j.atmosenv.2008.12.030.

Kwon, H.A., Park, R.J., Oak, Y.J., Nowlan, C.R., Janz, S.J., Kowalewski, M.G., Fried, A., Walega, J., Bates, K.H., Choi, J., Blake, D.R., Wisthaler, A., Woo, J.H., 2021. Top-down estimates of anthropogenic VOC emissions in South Korea using formaldehyde vertical column densities from aircraft during the KORUS-AQ campaign. Elementa-Sci. Anthrop. 9 (1), 00109 https://doi.org/10.1525/elementa.2021.00109.

Lamsal, L.N., Martin, R.v., Padmanabhan, A., van Donkelaar, A., Zhang, Q., Sioris, C.E., Chance, K., Kurosu, T.P., Newchurch, M.J., 2011. Application of satellite observations for timely updates to global anthropogenic NO_x emission inventories. Geophys. Res. Lett. 38, L05810 https://doi.org/10.1029/2010GL046476.

Li, C., Martin, R.V., Shephard, M.W., Cady-Pereira, K., Cooper, M.J., Kaiser, J., Lee, C.J., Zhang, L., Henze, D.K., 2019. Assessing the iterative finite difference mass balance and 4D-Var methods to derive ammonia emissions over North America using synthetic observations. J. Geophys. Res. Atmos. 124, 4222–4236. https://doi.org/10.1029/2018JD030183.

Li, M., Zhang, Q., Kurokawa, J.I., Woo, J.H., He, K., Lu, Z., Ohara, T., Song, Y., Streets, D. G., Carmichael, G.R., Cheng, Y., Hong, C., Huo, H., Jiang, X., Kang, S., Liu, F., Su, H., Zheng, B., 2017. MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP. Atmos. Chem. Phys. 17, 935–963. https://doi.org/10.5194/acp-17-935-2017.

Ma, J., van Aardenne, J.A., 2004. Impact of different emission inventories on simulated tropospheric ozone over China: a regional chemical transport model evaluation. Atmos. Chem. Phys. 4 (4), 877–887. https://doi.org/10.5194/acp-4-877-2004.

Martin, R.v., Jacob, D.J., Chance, K., Kurosu, T.P., Palmer, P.I., Evans, M.J., 2003. Global inventory of nitrogen oxide emissions constrained by space-based observations of NO_2 columns. J. Geophys. Res. Atmos. 108 (17), 4537. https://doi.org/10.1029/2003jd003453.

McQuilling, A.M., Adams, P.J., 2015. Semi-empirical process-based models for ammonia emissions from beef, swine, and poultry operations in the United States. Atmos. Environ. 120, 127–136. https://doi.org/10.1016/j.atmosenv.2015.08.084.

Miao, Y., Guo, J., Liu, S., Liu, H., Li, Z., Zhang, W., Zhai, P., 2017. Classification of summertime synoptic patterns in Beijing and their associations with boundary layer structure affecting aerosol pollution. Atmos. Chem. Phys. 17 (4), 3097–3110. https://doi.org/10.5194/acp-17-3097-2017.

Miao, Y., Guo, J., Liu, S., Zhao, C., Li, X., Zhang, G., Wei, W., Ma, Y., 2018. Impacts of synoptic condition and planetary boundary layer structure on the trans-boundary aerosol transport from Beijing-Tianjin-Hebei region to northeast China. Atmos. Environ. 181, 1–11. https://doi.org/10.1016/j.atmosenv.2018.03.005.

Mun, J., Choi, Y., Jeon, W., Lee, H.W., Kim, C.H., Park, S.Y., Bak, J., Jung, J., Oh, I., Park, J., Kim, D., 2023. Assessing mass balance-based inverse modeling methods via a pseudo-observation test to constrain NO_x emissions over South Korea. Atmos. Environ. 292, 119429 https://doi.org/10.1016/j.atmosenv.2022.119429.

Park, S.Y., Kim, D.H., Lee, S.H., Lee, H.W., 2016. Variational data assimilation for the optimized ozone initial state and the short-time forecasting. Atmos. Chem. Phys. 16 (5), 3631–3649. https://doi.org/10.5194/acp-16-3631-2016.

Peng, S.Q., Xie, L., 2006. Effect of determining initial conditions by four-dimensional variational data assimilation on storm surge forecasting. Ocean Model. 14 (1–2), 1–18. https://doi.org/10.1016/j.ocemod.2006.03.005.

Peng, Z., Liu, Z., Chen, D., Ban, J., 2017. Improving $PM_{2.5}$ forecast over China by the joint adjustment of initial conditions and source emissions with an ensemble Kalman filter. Atmos. Chem. Phys. 17 (7), 4837–4855. https://doi.org/10.5194/acp-17-4837-2017.

Qu, Z., Henze, D.K., Capps, S.L., Wang, Y., Xu, X., Wang, J., Keller, M., 2017. Monthly top-down NO_x emissions for China (2005-2012): a hybrid inversion method and trend analysis. J. Geophys. Res. 122 (8), 4600–4625. https://doi.org/10.1002/2016JD025852.

Qu, Z., Henze, D.K., Theys, N., Wang, J., Wang, W., 2019. Hybrid mass balance/4D-var joint inversion of NO_x and SO_2 emissions in East Asia. J. Geophys. Res. Atmos. 124 (14), 8203–8224. https://doi.org/10.1029/2018JD030240.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.-Y., Wang, W., Powers, J.G., 2008. A Description of the Advanced Research WRF Version 3. National Center for Atmospheric Research, Boulder, CO, USA.

Souri, A.H., Choi, Y., Jeon, W., Li, X., Pan, S., Diao, L., Westenbarger, D.A., 2016. Constraining NO_x emissions using satellite NO_2 measurements during 2013 DISCOVER-AQ Texas campaign. Atmos. Environ. 131 (2), 371–381. https://doi.org/10.1016/j.atmosenv.2016.02.020.

Wang, J., Xu, X., Henze, D.K., Zeng, J., Ji, Q., Tsay, S.C., Huang, J., 2012. Top-down estimate of dust emissions through integration of MODIS and MISR aerosol retrievals with the GEOS-Chem adjoint model. Geophys. Res. Lett. 39 (8), L08802 https://doi.org/10.1029/2012GL051136.

Woo, J.H., Choi, K.C., Kim, H.K., Baek, B.H., Jang, M., Eum, J.H., Song, C.H., Ma, Y. Il, Sunwoo, Y., Chang, L.S., Yoo, S.H., 2012. Development of an anthropogenic emissions processing system for Asia using SMOKE. Atmos. Environ. 58, 5–13. https://doi.org/10.1016/j.atmosenv.2011.10.042.

Zhang, R., Wang, G., Guo, S., Zamora, M.L., Ying, Q., Lin, Y., Wang, W., Hu, M., Wang, Y., 2015. Formation of urban fine particulate matter. Chem. Rev. 115 (10), 3803–3855. https://doi.org/10.1021/acs.chemrev.5b00067.

Zhao, S., Russell, M., Hakami, A., Capps, S., Turner, M., Henze, D., Percell, P., Resler, J., Shen, H., Russell, A., Nenes, A., Pappin, A., Napelenok, S., Bash, J., Fahey, K., Carmichael, G., Stanier, C., Chai, T., 2020. A multiphase CMAQ version 5.0 adjoint. Geosci. Model Dev. (GMD) 2925–2944. https://doi.org/10.5194/gmd-2019-287.

Zhu, C., Byrd, R.H., Lu, P., Nocedal, J., 1994. L-BFGS-B: A Limited Memory FORTRAN Code for Solving Bound Constrained Optimization Problems. Northwestern University. Tech. Rep.