



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

Xinghong Cheng¹, Zilong Hao², Zengliang Zang², Zhiquan Liu³, Xiangde Xu¹, Shuisheng Wang⁴, Yuelin Liu⁵, Yiwen Hu⁶, and Xiaodan Ma⁶

¹State Key Lab of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing 100081, China

²Institute of Meteorology and Oceanography, National University of Defense Technology, Nanjing 211101, China

³National Center for Atmospheric Research, Boulder, CO, USA

⁴GZ Source Clear Tech. Co., Ltd., Guangzhou 510630, China

⁵College of Architecture and Environment, Sichuan University, Chengdu 610065, China

⁶Nanjing University of Information Science and Technology, Nanjing 210044, China

Correspondence: Xinghong Cheng (cxingh@cma.gov.cn) and Zengliang Zang (zzlqxy@163.com)

Received: 6 June 2021 – Discussion started: 14 June 2021

Revised: 3 August 2021 – Accepted: 16 August 2021 – Published: 16 September 2021

Abstract. We develop a new inversion method which is suitable for linear and nonlinear emission source (ES) modeling, based on the three-dimensional decoupled direct (DDM-3D) sensitivity analysis module in the Community Multi-scale Air Quality (CMAQ) model and the three-dimensional variational (3DVAR) data assimilation technique. We established the explicit observation operator matrix between the ES and receptor concentrations and the background error covariance (BEC) matrix of the ES, which can reflect the impacts of uncertainties of the ES on assimilation. Then we constructed the inversion model of the ES by combining the sensitivity analysis with 3DVAR techniques. We performed the simulation experiment using the inversion model for a heavy haze case study in the Beijing–Tianjin–Hebei (BTH) region during 27–30 December 2016. Results show that the spatial distribution of sensitivities of SO₂ and NO_x ESs to their concentrations, as well as the BEC matrix of ES, is reasonable. Using an a posteriori inversed ES, underestimations of SO₂ and NO₂ during the heavy haze period are remarkably improved, especially for NO₂. Spatial distributions of SO₂ and NO₂ concentrations simulated by the constrained ES were more accurate compared with an a priori ES in the BTH region. The temporal variations in regionally averaged SO₂, NO₂, and O₃ modeled concentrations using an a posteriori inversed ES are consistent with in situ observations at 45

stations over the BTH region, and simulation errors decrease significantly. These results are of great significance for studies on the formation mechanism of heavy haze, the reduction of uncertainties of the ES and its dynamic updating, and the provision of accurate “virtual” emission inventories for air-quality forecasts and decision-making services for optimization control of air pollution.

1 Introduction

Since the implementation of the Air Pollution Prevention and Control Action Plan in September 2013, urban air quality in China has improved overall. However, heavy haze frequently occurs over Beijing–Tianjin–Hebei (BTH) and the surrounding region in winter. In recent years, many researchers have studied the formation mechanism of heavy haze in the BTH region (Huang et al., 2014; Cheng et al., 2016; Liu et al., 2016). These studies have shown that rapid conversion from primary gas pollutants to particulates is an internal triggering factor for the “explosive” and “persistent” heavy haze (Wang et al., 2014), and secondary particulate concentrations, such as sulfate and nitrate, account for a significant percentage of PM_{2.5}. Thus, effectively controlling the emissions of precursors of secondary aerosols (such as SO₂ and NO_x) is

important for reducing environmental, economic, and human health problems caused by PM_{2.5} concentrations (Huang et al., 2014).

Emission inventories provide important fundamental data are crucial input data for investigating the causes of air pollution and for atmospheric chemical transport models (ACTMs). Uncertainties in the emission source (ES) are a major factor in determining the simulated and forecast accuracy of ACTMs, and these uncertainties can greatly affect the design of ES control strategies (Tang et al., 2006). The methods for establishing an emission inventory include the bottom-up approach based on human activities, energy consumption statistics, and various emission factors, as well as top-down inversion modeling of the ES based on monitoring data of air pollutants using satellite remote sensing and ground observations. Many studies have established various ES inventories in China using the bottom-up approach (e.g., Bai and Zhou, 1996; Streets et al., 2003; Zhang et al., 2009, 2012; Cao et al., 2011; Zhao et al., 2012, 2015; Li et al., 2017). However, the ES estimated by this method differ greatly due to large uncertainties in the statistical data, emission factors, and spatiotemporal apportionment coefficients (Ma and Van Aardenne, 2004). Moreover, real-time updates of emission inventories are difficult to achieve because of their rapid spatiotemporal variations due to high-speed urbanization and a delay in the release of statistical data of approximately 1–2 years. The top-down approach is a useful supplement to bottom-up estimates, which are subject to uncertainties in emissions factors and activities (Streets et al., 2003). Inverse modeling, in which emissions are optimized to reduce the differences between simulated and observed data, is a powerful method that eliminates the problems of the bottom-up approach.

Over the past decade, many researchers have tried to find an ideal inversion modeling tool that improves the spatiotemporal distribution of ES. With the development of data assimilation technology and ACTMs, constraining the strength of the ES using ACTMs has become one of the main top-down inversion methods (Enting, 2002; Sportisse, 2007). Researchers have primarily constrained the ES of weak active chemical pollutants, such as NO_x, CO, CO₂, SO₂, CH₄, and CHOCHO using the following methods: mass balance (Martin et al., 2003; Wang et al., 2007; Yang et al., 2011), back-trajectory inverse modeling (Manning et al., 2011), adjoint modeling (Liu et al., 2005; Koohkan et al., 2013; L. Zhang et al., 2016; Wang et al., 2018), Bayes estimation theory (Kopacz et al., 2009; Lucas et al., 2017), ensemble Kalman filtering (EnKF; Zhu and Wang, 2006; Zhu et al., 2018; Barbu et al., 2009; Tang et al., 2011, 2016; Miyazaki et al., 2012; Mijling and van der A, 2012; Wang et al., 2016; Peng et al., 2017; Chen et al., 2019; Dai et al., 2021), the four-dimension variational (4DVAR) technique (Elbern et al., 2000, 2007; Gilliland et al., 2006; Napelenok et al., 2008; Henze et al., 2009; Stavrou et al., 2009; Corazza et al., 2011; Jiang et al., 2011), and an adaptive

nudging scheme in the Community Multiscale Air Quality (CMAQ) model (Xu et al., 2008; Cheng et al., 2010). Results show that using an inversion modeling approach to retrieve the spatial distribution of ES can greatly improve air quality simulations and forecasts by the ACTM. Many studies have achieved a certain amount of improvement using the EnKF and 4DVAR methods. The advantage of the EnKF method is that the observation operator is implicit in the assimilation process of ES, and it avoids developing the tangent linear and adjoint models. For example, the fully coupled on-line Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) is used as the forward model to relate the SO₂ emissions to the simulated concentration and efficiently update the emissions based on routine surface SO₂ observations (Dai et al., 2021). However, this method has stricter requirements for error perturbations in ES and the construction of bias-correction models. In addition, the large number of ensemble members in the EnKF method leads to a huge computational cost. Some studies adopted the 4DVAR method to inverse the ES of NO_x and CO based on the Goddard Earth Observing System (GEOS)-Chem adjoint model. However, the GEOS-Chem model is often used to simulate large-scale physical and chemical processes and is rarely utilized in urban air quality forecasts because its spatial resolution is too coarse. This method also has high computational costs due to the gradient calculation of the objective function.

The 3DVAR method is a generalization of optimal interpolation methods. It has the advantages of conveniently adding dynamical constraints and directly assimilating unconventional observation data (Li et al., 2013). 3DVAR is widely used in the assimilation of meteorological and atmospheric chemical data due to its simplicity, ability to use complex observations operators, and low computational cost. However, this method has two requirements: assimilated variables must remain relatively stationary within the assimilation window, and the method must be coordinated between the assimilated initial field and the iterative integration of the model. To apply the 3DVAR method to inverse the ES, it is necessary to construct an inversion model that can satisfy the aforementioned requirements. Firstly, although the ES have monthly, seasonal, and annual variations, the variation of ES is constant within a short period (e.g., for an assimilation window of 1 h). Secondly, the assimilation effect of the ES depends on the quality of observation data and the consistency between observed and simulated values. To ensure consistency between observations and simulations, the sensitivity of the receptor's concentrations with respect to the ES should be accurately calculated (Hu et al., 2009). Using the three-dimensional decoupled direct (DDM-3D) sensitivity analysis method within the CMAQ model, reasonable sensitivity coefficients between the ES and the receptor's concentrations can be calculated. This coefficient matrix is then used in the 3DVAR assimilation process, which ensures consistency between the ES and the modeled results. Thus, the top-down 3DVAR constraint methods for the ES based on the first-

or high-order DDM-3D sensitivity analysis techniques can maintain the coordination between the assimilated field of the ES and the simulated concentration of air pollutants.

The primary methods used to calculate the S–R (source–receptor) relationship include the brute force, the adjoint, and the DDM-3D method. Many studies have shown that these methods can improve ES inventories constructed by bottom-up methods for NO_x (Napelenok et al., 2008), CO (Bergamaschi et al., 2000; Heald et al., 2004), NH_3 (Gilliland et al., 2003), and elemental carbon (EC; Hu et al., 2009). The adjoint method is a backward-sensitivity calculation method, while brute force and DDM-3D are forward-sensitivity calculation methods. For the inverse modeling of pollution sources with a single receptor, the backward-sensitivity method is more suitable, with low computation costs for certain grid sizes in a given time period, but it is not suitable for ES with multiple receptors, which result in high computational costs (Hu et al., 2009; Wang et al., 2013). The forward-sensitivity calculation method is more suitable for inverting the ES based on observed data from satellites or multiple surface stations (Hu et al., 2009). Cohan et al. (2002) introduced the DDM-3D method to the CMAQ model and created the CMAQ-DDM-3D module for low-order sensitivity calculations in early 2010. In 2014, they added a high-order calculation module for particles (High-Order DDM-3D for Particular Matter; HDDM3D/PM) in the newly released version of the CMAQ model. Wang et al. (2013) claim that the sensitivity calculation results using the DDM-3D method are more reasonable than the brute force method. Some studies have used the DDM-3D method (Napelenok et al., 2008; Hu et al., 2009) or a combination of the DDM-3D and a discrete Kalman filter method (Wang et al., 2013) in conjunction with measurements from satellite and ground observations to inverse BC and NO_x ES in the United States. Because inverse modeling of ES based on discrete Kalman filtering is more suitable for linear systems, we use the DDM-3D method to calculate the S–R linear and nonlinear relationship.

The results of inverse modeling are very sensitive to uncertainties in the ESs of NO_x , NH_4 , and inorganic aerosols (L. Zhang et al., 2016). The impact of uncertainties in the ES on the assimilation effects needs to be considered in the top-down inversion model. The top-down 3DVAR inversion methods developed in this study can include the impacts of ES uncertainties by the background error covariance (BEC) matrix of the ES based on multiple sets of ESs. We developed a new inverse modeling approach for the ES that combines the DDM-3D sensitivity analysis method with the 3DVAR assimilation technique and then applied it to a case study during a typical heavy haze episode. This paper is organized as follows: Sect. 2 describes the inversion model and presents results of the sensitivity analysis and the BEC; Sect. 3 provides details of the WRF-CMAQ model and configurations and experiments of the simulation; Sect. 4 presents the results of the control and experiment simulations with a priori

and a constrained posteriori ES, respectively; finally, the discussions and conclusions are provided in Sect. 5.

2 Model and data

We used an offline modeling system that includes two components: the Weather Research and Forecasting (WRF) model (Michalakes et al., 2004) and the CMAQ model (Dennis et al., 1996; hereafter referred to as WRF-CMAQ). This study focuses on the BTH region with 5×5 km grid spacing, 32 vertical layers of varying thickness (between the surface and 50 hPa), and an output interval of 1 h. The WRF-CMAQ simulations are driven by the National Center for Environmental Prediction Final (NCEP FNL) analysis data every 6 h during 27–30 December 2016 and the Multi-resolution Emission Inventory for China (MEIC) data for 2012, with $1^\circ \times 1^\circ$ and $0.25^\circ \times 0.25^\circ$ grid spacing, respectively. The CMAQ model was configured to utilize all layers from the input meteorology. Emissions datasets for CMAQ were generated by the Sparse Matrix Operator Kernel Emissions (SMOKE) model developed by the University of North Carolina (UNC, 2014). Meteorological outputs from the WRF simulations were processed to create model-ready input to CMAQ using the Meteorology-Chemistry Interface Processor (MCIP; Otte and Pleim, 2010). The boundary conditions for chemical trace gases consisted of idealized, northern hemispheric, midlatitude profiles based upon output from the National Oceanic Atmospheric Administration (NOAA) Agronomy Lab Regional Oxidant model. The model simulation started on 27 December 2016. To assess the improved effects of inverse modeling of ES during the heavy haze episode in December 2016, we ran two simulations: a control run with a priori MEIC data for 2012 and an experiment run with constrained a posteriori ES.

Hourly measurements of SO_2 , NO_2 , and O_3 concentrations at 129 stations during 27–30 December 2016 were obtained from the China National Environmental Monitoring Centre. These data are used to validate simulations from the control and experiment runs. The simulation domains and the locations of the 129 stations are shown in Fig. 1.

3 Inverse modeling method

3.1 Constructing the BEC matrix

To construct the BEC matrix for the inversion model, we combined the technique by the National Meteorological Center (NMC) (Parrish and Derber, 1992) with the SMOKE model based on uncertainty analysis of the ES inventories. We created the BEC matrix by four steps, as follows:

1. *Determine the total errors of ES from an a priori bottom-up inventory.* Uncertainty analyses of ESs require detailed information of activities and emission

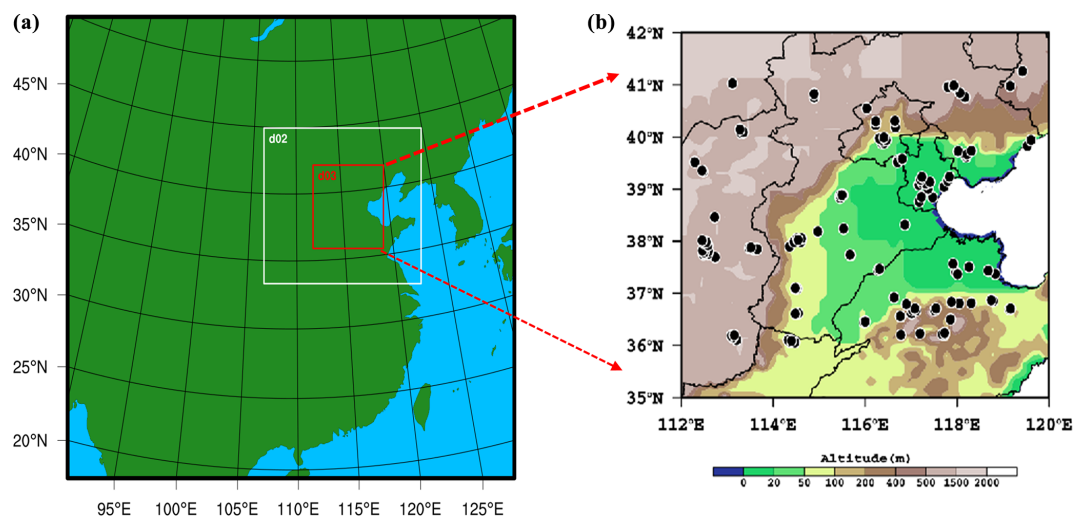


Figure 1. (a) Domain of the WRF-CMAQ model and (b) location of environmental monitoring stations in the innermost domain over the BTH region.

factors from an a priori MEIC emission inventory. The relevant data collected in the China Environmental Yearbook are limited and do not satisfy the requirements of uncertainty analysis. Therefore, we used the available research results relating to SO_2 and NO_x ESs and conducted uncertainty analysis for four types of major sources (industry, power plants, residents, and transportation) based on activities and emission factors from the references (Hong et al., 2017; Zheng et al., 2018; Peng et al., 2019) using the AuvTool software (Frey et al., 2002) and determined the error ranges in total emission rates of SO_2 and NO_x (Table 1). Uncertainties in SO_2 industry and power plant ESs are slightly greater than those for NO_x , while the opposite is true for emissions from the residential and traffic sectors.

2. *Generate multiple sets of inventories using the random perturbation technique.* Based on the aforementioned error ranges in total emission rates, we generated 30 sets of inventories for SO_2 and NO_x with the same resolution as MEIC for each month using a random perturbation method (Kerry et al., 2007). Firstly we obtain the probability distribution of errors of ESs based on uncertainty analysis for four sections, respectively. Then we conduct random perturbation on uncertainties of four sections of ES 30 times according to the probability distributions using the same perturbation coefficients for every perturbation. Lastly we calculate 30 total emission rates using random uncertainties of four sections for every set of inventories, respectively.
3. *Process the 3D gridded ES as input to the CMAQ model.* We used the SMOKE model, national population and road network distribution data in 2016, the temporal apportionment coefficients in the BTH re-

Table 1. Uncertainty of NO_x and SO_2 values used in the SMOKE model and calculation of the BEC matrix.

Categories	NO_x	SO_2
Industry	(−32.4%, 33.0%)	(−37.5%, 38.8%)
Power	(−32.4%, 33.0%)	(−37.5%, 38.8%)
Residential	(−30.0%, 34.0%)	(−15.0%, 16.0%)
Transportation	(−55.4%, 70.3%)	(−17.0%, 20.0%)

gion (Zhang et al., 2007; Wang et al., 2008), and the CB05-ae06-aq chemical species data in the CMAQ model, to process 30 sets of nationwide emission inventories into 3D gridded ES with a grid spacing of 5×5 km. Each grid has 124×130 points, with 12 vertical levels.

4. *Calculate the BEC matrix of each 3D gridded ES.* Finally, the NMC method was used to calculate the BEC matrix of the 3D gridded ES for each month, including horizontal and vertical correlation coefficients and standard deviations. The background error is defined as the difference between 30 sets of 3D gridded ESs generated by the random perturbation method and the 3D gridded background ES directly processed from the original MEIC emission inventory with the SMOKE model, at every hour (24 h diurnal variation of ES for every month).

According to the literature (Liu et al., 2011; Li et al., 2013; Zang et al., 2016), the approximate calculation of the BEC matrix is as follows:

$$\mathbf{B} \approx \frac{1}{2} \langle (\mathbf{e}_t - \mathbf{e}_b)(\mathbf{e}_t - \mathbf{e}_b)^T \rangle, \quad (1)$$

where \mathbf{e}_t is the perturbation field, and \mathbf{e}_b is the background field of an a priori ES. Equation (1) can be written as follows:

$$\mathbf{B} = \mathbf{D}\mathbf{C}^T, \quad (2)$$

where \mathbf{D} is the standard deviation (SD) matrix, and \mathbf{C} is the correlation coefficient matrix. With this factorization, \mathbf{D} and \mathbf{C} can be calculated separately. \mathbf{D} is a diagonal matrix, whose elements are the SD of all state variables in the 3D grids. \mathbf{C} is used to improve the ability of the 3DVAR in representing the impacts of local emissions at one grid on other grids; these impacts vary in the vertical direction, and they are heterogeneous in the horizontal direction.

Figure 2 shows the spatial distribution of averaged emission rates for 30 sets of 3D gridded ESs and the SD of the BEC matrix for SO_2 and NO_x ESs at 08:00 local time in December. SO_2 and NO_x ESs have different spatial distributions in terms of average strength and standard deviation. The NO_x emissions are mainly concentrated in cities and surrounding areas, and they are much greater in Beijing, Tianjin, and Shijiazhuang than other cities. The SO_2 emissions are mainly concentrated in Shijiazhuang, Jinan, the north and east of Shanxi Province, and their surrounding areas. Figure 3 shows variations in the horizontal correlation coefficients by grid distance and the vertical distributions of the SD in \mathbf{B} for SO_2 and NO_x ESs in December 2012. The cross between the correlation curve and the $e^{-1/2}$ line (dashed line) represents the horizontal length scale (L_s), and the L_s of the two species falls between five and six grid distances. Namely, the horizontal scale felt is approximately 25–30 km. The correlation coefficient of SO_2 is slightly larger than that of NO_2 . The difference in the correlation coefficients between SO_2 and NO_x ESs increases with grid distance, and this is related to the regional pollution characteristics of SO_2 . The vertical distributions of the SDs in \mathbf{B} for SO_2 and NO_x ESs vary with height: the SDs of the SO_2 ES are larger on the fourth and eighth model levels than on other levels, while for the NO_x ES, the SD on the first level is the largest, that on the eighth level take the second place, and the SDs on all other levels are smaller.

3.2 Sensitivity analysis

The sensitivity analysis module (DDM-3D) in CMAQ solves a series of equations while simultaneously calculating pollutant concentrations. The local sensitivity of pollutant concentrations with respect to several specified parameters, such as the ES, initial and boundary conditions, and chemical reaction rates, can be calculated by the DDM-3D method. The sensitivity equations about the ES are solved using the governing equations of the model, as follows (Hu et al., 2009):

$$\mathbf{S} = \mathbf{p}_j \frac{\partial \mathbf{C}}{\partial \mathbf{p}_j} = \mathbf{p}_j \frac{\partial \mathbf{C}}{\partial (\boldsymbol{\varepsilon}_j \mathbf{p}_j)} = \frac{\partial \mathbf{C}}{\partial \boldsymbol{\varepsilon}_j}, \quad (3)$$

where \mathbf{S} is the sensitivity of the pollutant j to the parameter \mathbf{p}_j , \mathbf{p}_j is an a priori ES of the pollutant j , \mathbf{C} is the con-

centration of the pollutant j , and $\boldsymbol{\varepsilon}_j$ is the perturbation coefficient of the ES. Theoretically, the DDM-3D method truly captures the sensitivities of pollutant concentrations to ES, and results are more accurate than the brute force method, for the BTH region (Wang et al., 2013). In addition, the results of the DDM-3D method are more accurate and efficient for highly nonlinear pollutants (such as O_3 and $\text{PM}_{2.5}$) and small perturbations.

We used the WRFv3.7.1 and CMAQv5.0.2-DDM-3D models as well as 3D gridded a priori ES from MEIC in 2012 to calculate the sensitivity coefficients of SO_2 and NO_2 concentrations with respect to ESs during the heavy haze episode of 27–30 December 2016. Figure 4 shows the spatial distribution of 96 h averaged sensitivity coefficients for SO_2 and NO_2 concentrations with respect to ESs during 27–30 December 2016. The sensitivity coefficients of SO_2 and NO_2 concentrations all exhibit inhomogeneous distribution. The sensitivity coefficients are higher in Beijing, Shijiazhuang, Baoding, and surrounding regions; i.e., SO_2 and NO_2 concentrations in those areas are greatly affected by the SO_2 and NO_x ESs.

3.3 Observation operators

The relationship between pollutant source and the receptor's concentration is established according to Eq. (3). Next, we create the observation operator matrix between the ES and receptor concentrations as follows:

$$\mathbf{H} = \frac{\partial \mathbf{C}}{\partial \mathbf{E}} = \frac{\partial \mathbf{C}}{\partial (\boldsymbol{\varepsilon}_j \mathbf{E}_0)} = \frac{\mathbf{S}}{\mathbf{E}_0}, \quad (4)$$

where \mathbf{H} is the observation operator matrix, \mathbf{E} denotes an a posteriori ES, which can be written as the product of the perturbation coefficient and an a priori ES during the assimilation window time, and \mathbf{E}_0 denotes an a priori ES. For primary pollutants such as SO_2 and NO_2 , \mathbf{S} is a first-order sensitivity coefficient, and \mathbf{H} is a linear observation operator between the ES and the receptor concentration. For secondary pollutants such as $\text{PM}_{2.5}$ and O_3 , \mathbf{S} is a high-order sensitivity coefficient, and \mathbf{H} is a nonlinear observation operator. In this study, we use the first-order sensitivity coefficient to calculate \mathbf{H} for SO_2 and NO_x ESs.

3.4 Observational error covariance

We firstly performed quality control on the observed SO_2 and NO_2 concentration data. This process involved three steps:

1. Redundant data were removed, and the density of observation data was matched to the model grid. For some grids with more than one observation station, we used the average of those stations.
2. Extrema were controlled; i.e., data exceeding 3 times the SD of observation data were filtered out.

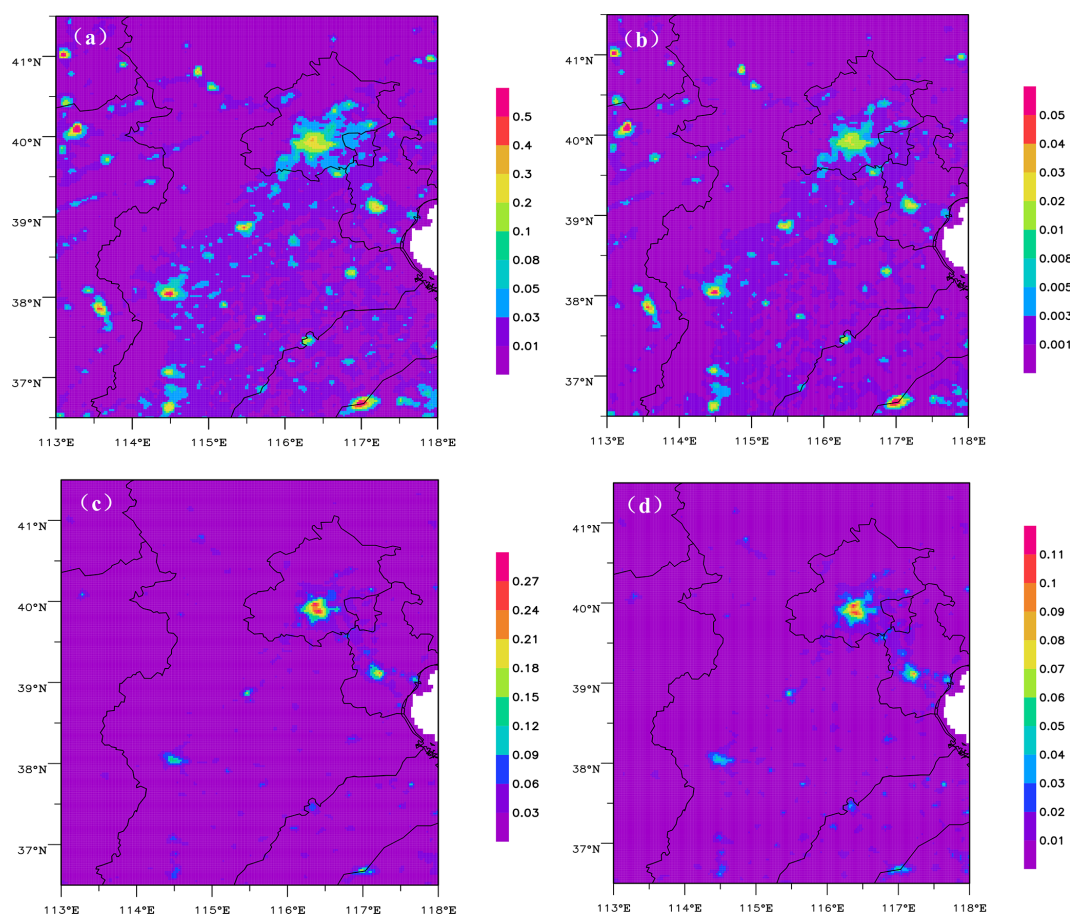


Figure 2. Spatial distributions of (a) averaged emission rates of the SO₂ ES, (b) standard deviation in the BEC of the SO₂ ES, (c) averaged emission rates of the NO_x ES, and (d) standard deviation in the BEC of the NO_x ES at 08:00 local time in December 2012. Unit: mole/s.

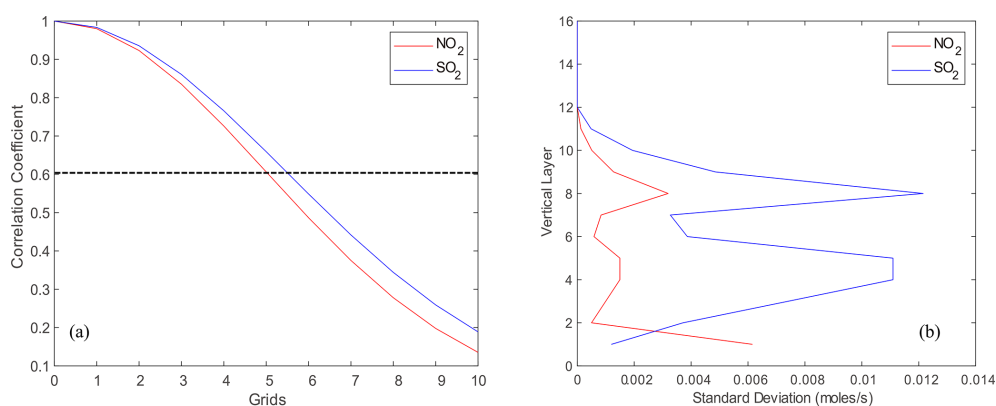


Figure 3. (a) Horizontal correlation coefficients with increasing grid distance and (b) vertical profiles of standard deviations in the BEC of SO₂ and NO₂ ESs in December 2012. The dashed line is the baseline of the horizontal correlation scale.

3. Anomalies were removed; i.e., data that remained constant for 24 consecutive hours, as well as any negative data, were removed.

Data that passed quality control still contained observation or instrument errors. These errors are related to many factors

such as instrument type, calibration design, and environmental conditions. In addition, in the variational assimilation process, representation errors caused by the forward-calculation and variational processes must be considered. Higher resolution models produce smaller representation errors. Repre-

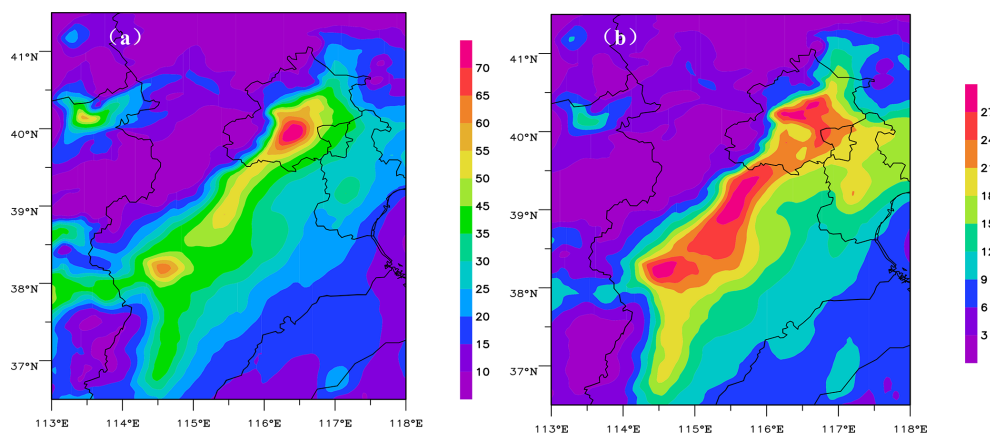


Figure 4. Spatial distributions of 96 h averaged sensitivity coefficients ($\mu\text{g m}^{-3}$) of (a) SO_2 and (b) NO_x concentrations with respect to SO_2 and NO_x ES during 27–30 December 2016.

sensation error, ϵ_r , can be expressed as follows (Pagowski et al., 2010):

$$\epsilon_r = \gamma \epsilon_o \sqrt{\frac{\Delta x}{L_s}}, \quad (5)$$

where γ is the amplification factor, which is used to adjust the instrument error, ϵ_o is related to the SO_2 and NO_2 concentrations, and Δx is grid distance of the model. Note that L_s is usually smaller in urban areas and larger in suburban areas. The amplification parameters of the observing stations in cities, suburbs, and rural areas are 2.5, 5, and 10 km, respectively (Zang et al., 2016). Finally, the total observation error for the SO_2 and NO_2 concentrations, ϵ , is written as

$$\epsilon = \sqrt{\epsilon_o^2 + \epsilon_r^2}. \quad (6)$$

3.5 3DVAR inversion model

We introduce a cost function with respect to the ES in accordance with 3DVAR:

$$J(e) = \frac{1}{2}(e - e_b)^T \mathbf{B}^{-1}(e - e_b) + \frac{1}{2}(\mathbf{H}e - c)^T \mathbf{R}^{-1}(\mathbf{H}e - c), \quad (7)$$

where c is the observation variable, \mathbf{R} is the observation error matrix, and e is the inverting variable of an a posteriori ES. The optimal inversion of the ES for SO_2 and NO_x is obtained using Eq. (7). The 3DVAR solves for the minimum value of $J(e)$ to determine the inverting variable e . This process typically employs a gradient propagation method, with the increment of an ES defined as follows:

$$\delta e = e - e_b. \quad (8)$$

Accordingly, the innovation vector of pollutant concentration is defined as

$$\delta c = c - \mathbf{H}e_b. \quad (9)$$

Therefore, Eq. (7) can be written in gradient form:

$$J(\delta e) = \frac{1}{2}\delta e^T \mathbf{B}^{-1}\delta e + \frac{1}{2}(\mathbf{H}\delta e - \delta c)^T \mathbf{R}^{-1}(\mathbf{H}\delta e - \delta c). \quad (10)$$

After conditionally processing the cost function, a finite-memory quasi-Newton method was used to conduct iterative minimization. The background field was set as the initial iteration values. The maximum number of steps at the end of the iteration and the minimum gradient for convergence were predetermined. The iteration was finished when one of these conditions was met, and the optimal analysis increment, δe , was obtained. Finally, the optimal assimilation analysis field of the ES, $e = \delta e + e_b$, was obtained. The result was a three-dimensional variational inversion model of the ES, using the uncertainty analysis of the ES and sensitivity coefficients between the ES and the receptor's concentrations; the overall framework is shown in Fig. 5.

4 Results and discussion

A typical heavy haze event occurred in the BTH region at the end of December 2016. We applied the 3DVAR inversion model to constrain an hourly a posteriori ES of SO_2 and NO_2 using measurements from 45 and 129 stations, respectively, on 27 December 2016. We validated simulations from the control and experiment run using observational data during 28–30 December 2016.

Figure 6 shows time series of hourly, regional averaged SO_2 and NO_2 simulations from the control run, observations, and sensitivity coefficients at 45 stations in the BTH region during 27–30 December 2016. The trends in modeled concentrations and sensitivity coefficients of SO_2 and NO_2 concentrations with respect to the ES are consistent; therefore the sensitivity coefficients can reasonably reflect the impacts of the ES on concentrations. However, simulated SO_2 and NO_2 concentrations with an underestimated a priori ES

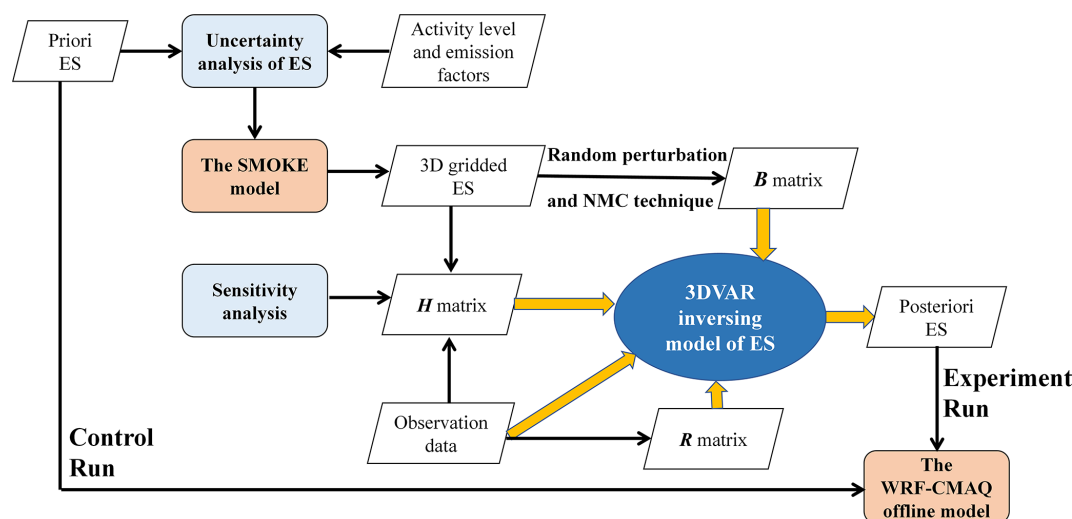


Figure 5. Flowchart of the 3DVAR inversion model of the ES and simulation experiments.

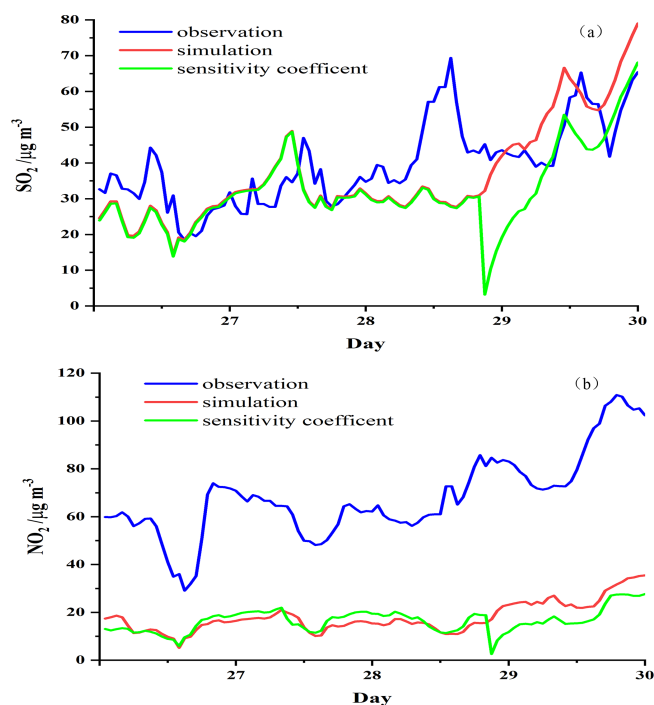


Figure 6. Time series of hourly, regionally averaged (a) SO_2 and (b) NO_2 simulations with an a priori ES, observations, and the first-order sensitivity coefficients between the ES and the receptor's concentration at 45 stations over the BTH region during 27–30 December 2016.

are all significantly lower than observations during the heavy haze period. Thus, it is important to improve a priori ES using the inversion model.

Figures 7 and 8 show the spatial distributions of 24 h averaged emission rates from an a priori and a posteriori ES of SO_2 and NO_2 and their increments on 27 December 2016.

Emission rates of a posteriori ES of SO_2 and NO_2 in the major cities and surrounding areas clearly increase. Compared with a priori ES, the maximum strengths of SO_2 and NO_2 ESs increase by approximately 17 % and 500 %, respectively. Therefore, the strengths of SO_2 and NO_2 in a priori ESs were greatly underestimated, especially for NO_2 .

Using the WRF-CMAQ model and an a posteriori ES, we simulated concentrations of SO_2 , NO_2 , and O_3 in the BTH region during 28–30 December 2016 and validated these simulations with measurements from 45 stations. Figures 9 and 10 show the spatial distributions of 72 h averaged SO_2 and NO_2 concentrations simulated with a priori and a posteriori ESs, increments, and their observations. In general, SO_2 and NO_2 concentrations simulated using an a posteriori ES are closer to observations than an a priori ES, and regional differences in improvements for SO_2 and NO_2 exist. For SO_2 , the improvement is noticeable in the BTH region. However, the simulated concentrations in Beijing with an a posteriori ES are overestimated. This may be related to greater uncertainties in SO_2 sources and the impacts of regional transport from surrounding areas. For NO_2 , simulated differences with a priori and a posteriori ESs are significant in major cities such as Beijing, Tianjin, Shijiazhuang, Baoding, Xingtai, Handan, and Jinan. The simulated concentrations of NO_2 using a posteriori ES are more consistent with measurements, while those with a priori ES are significantly underestimated.

We also investigated temporal variations in regionally averaged SO_2 , NO_2 , and O_3 concentrations simulated using a priori and a posteriori ESs and observations from the 45 stations over the BTH region during 28–30 December 2016 (Fig. 11). In general, simulated SO_2 , NO_2 , and O_3 concentrations using an a posteriori ES are closer to measurements, while the SO_2 and NO_2 concentrations simulated by an a priori ES are significantly lower than observations, and the

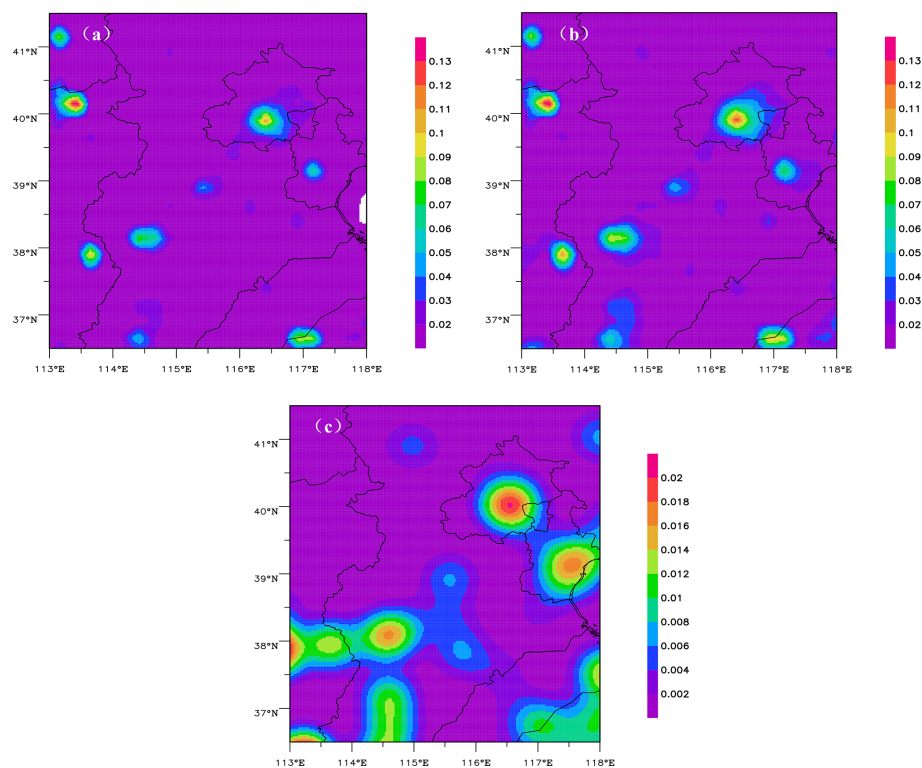


Figure 7. Spatial distributions of 24 h averaged emission rates for SO₂ (mole/s) from (a) a priori and (b) a posteriori ESs and (c) the increment on 27 December 2016.

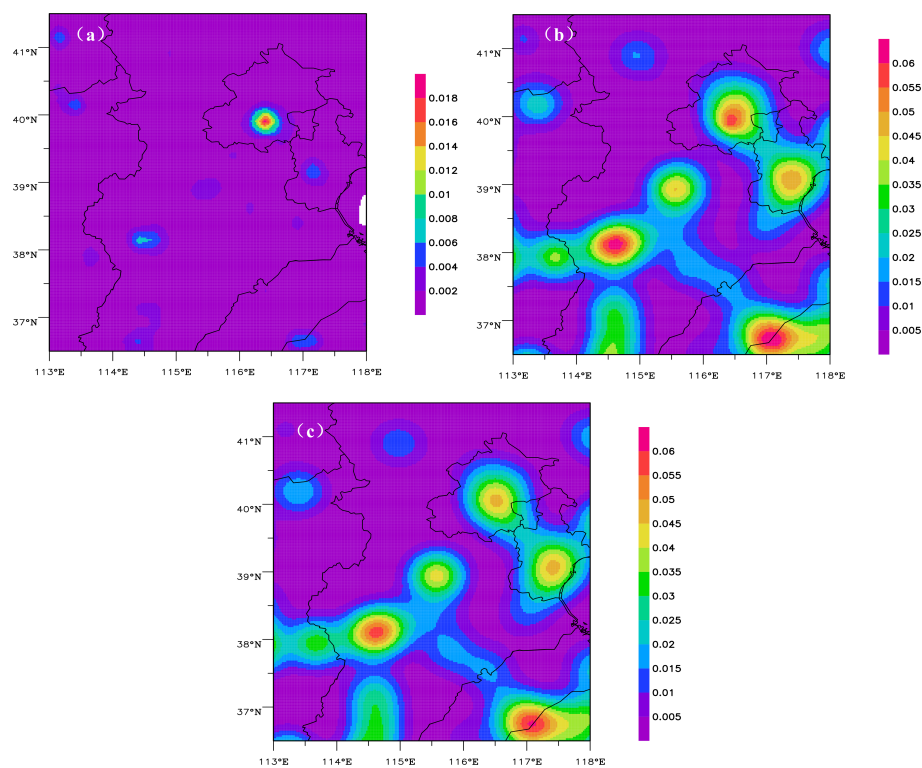


Figure 8. Same as Fig. 7 except for NO₂.

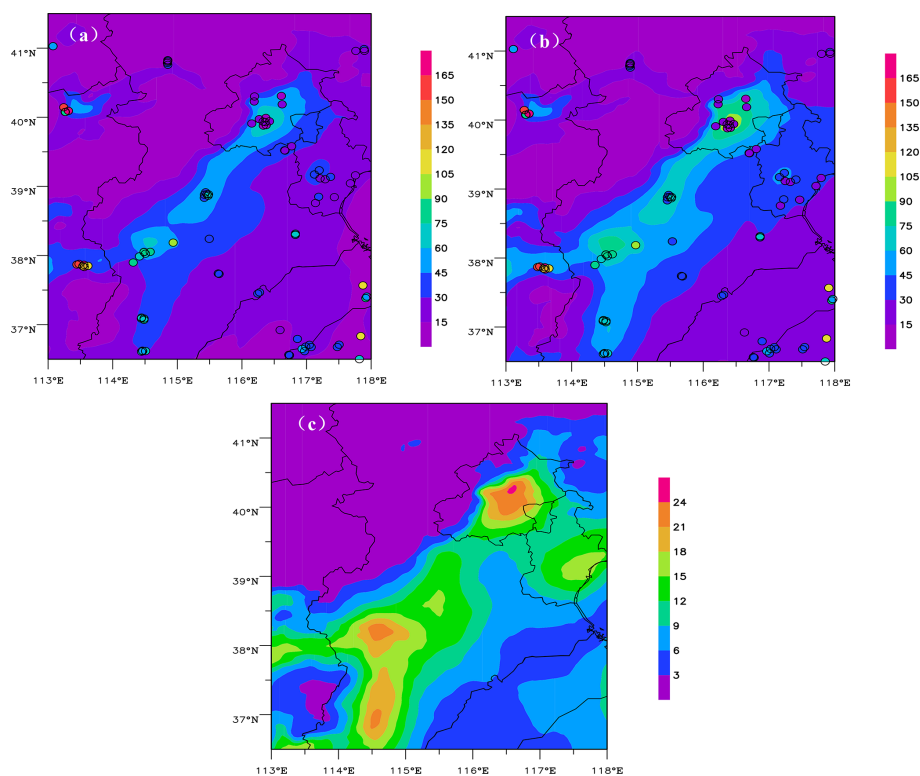


Figure 9. Spatial distribution of 72 h averaged SO_2 concentration simulated with (a) a priori and (b) a posteriori ESs and (c) the increment during 28–30 December 2016. Color solid dots denote the measurements.

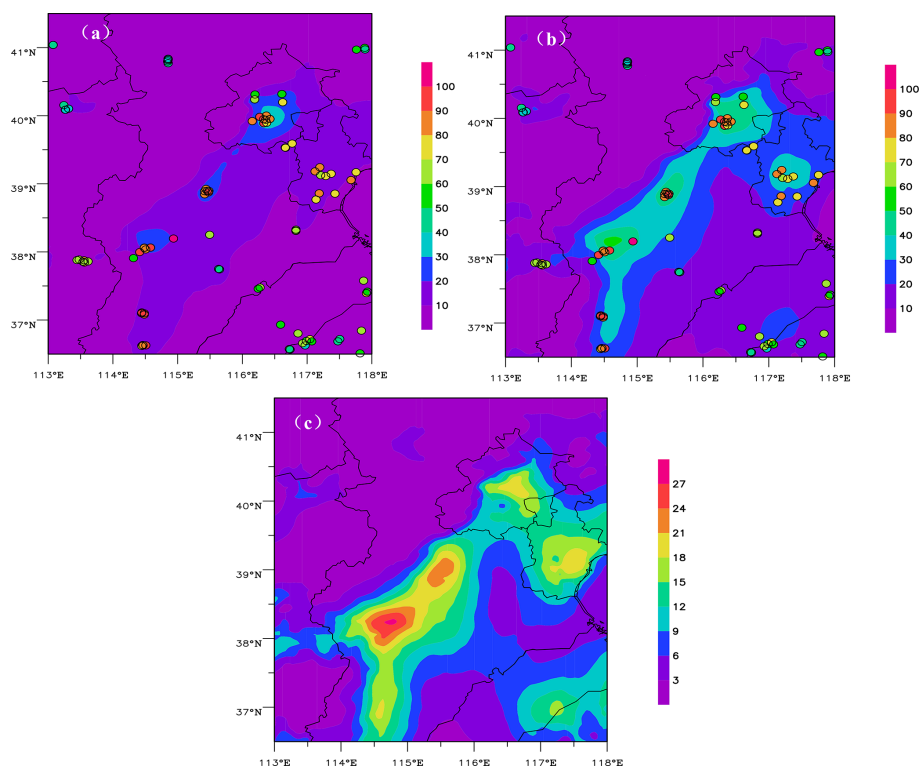


Figure 10. Same as Fig. 9 except for NO_2 .

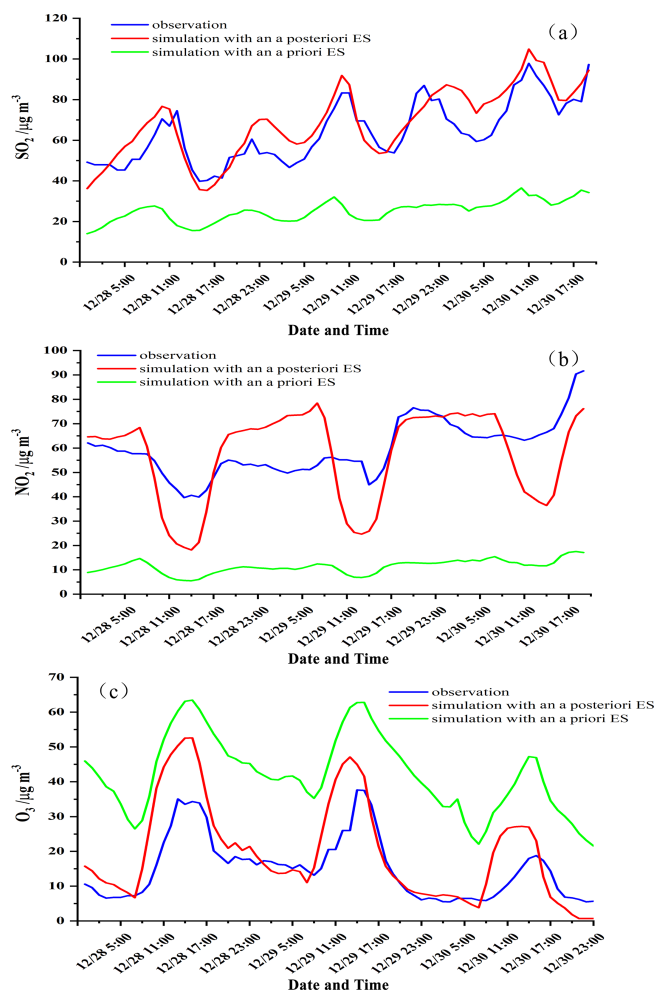


Figure 11. Time series of regional averaged (a) SO_2 , (b) NO_2 , and (c) O_3 concentrations respectively simulated with a priori and a posteriori ESs, and measurements at 45 stations in the BTH region during 28–30 December 2016.

modeled O_3 concentrations are obviously higher than measurements. In addition, the peak of SO_2 simulations with an a posteriori ES is close to measurements, but the peak of NO_2 and valley of O_3 simulations are lower and higher than observations, respectively. This may be related to the absence of inverse modeling of a volatile organic compound (VOC) ES and the uncertainties of sensitivity coefficients calculation. In this study, we used only first-order sensitivity coefficients, but the relationship between the ES of precursors of O_3 such as VOCs and NO_x and their receptor's concentrations is non-linear, and O_3 is generated from both the NO_x and VOC ES. Therefore, higher order sensitivity coefficients are necessary for the inverse modeling of the ES of NO_x and VOCs.

To further assess the simulated accuracy of SO_2 , NO_2 , and O_3 concentrations, we calculated the following statistics (Willmott et al., 2012): correlation coefficient (R), root-mean-squared error (RMSE), mean bias (MB), normalized

Table 2. Statistics for simulated SO_2 , NO_2 , and O_3 from control and experiment runs using a priori and a posteriori inverted ESs at 45 stations in the BTH region during 28–30 December 2016. Bold type indicates better statistical results.

Parameters	Control run			Experiment run		
	SO_2	NO_2	O_3	SO_2	NO_2	O_3
R	0.80	0.82	0.89	0.82	0.52	0.87
RMSE	14.61	8.89	5.02	6.60	8.68	6.31
MB	−40.98	−48.20	26.91	3.23	−2.23	4.70
NMB	−0.61	−0.81	1.78	0.05	−0.04	0.31
IOA	0.42	0.27	0.45	0.89	0.68	0.84

mean bias (NMB), and index of agreement (IOA; see Table 2). Except the result that R values of NO_2 and O_3 decrease and the RMSE of O_3 increases using the constrained ES, other statistics show improvements. In particular, MB and NMB of three pollutants decline significantly, and IOA is closer to 1.0, which means that modeled results of the three pollutants are more consistent with observations. R between SO_2 simulation and observation shows a slight improvement when using an a posteriori ES, whereas R decreases for NO_2 and O_3 , and it may be related with the absence of constraint of a VOC ES.

5 Summary and conclusions

We developed a new inverse approach of ES by combining the sensitivity analysis technique between the ES and the receptor's concentration and the 3DVAR method. Our approach is suitable for solving the linear or nonlinear inversion problems for ESs. It computes quickly and obtains relatively accurate real-time dynamic updates of ESs. First, we used the sensitivity analysis tool in the CMAQ model to construct the explicit observation operator matrix between the ES and the receptor's concentration. Next, we created the BEC matrix for ES based on uncertainty analysis and the NMC statistical method. Finally, we established a three-dimensional variational inverse method of ES based on the observation operator and BEC matrix.

The 3DVAR inversion model was applied to a heavy haze case study in the BTH region during 27–30 December 2016. Results show that the observation operators between SO_2 and NO_2 ESs and their concentrations, as well as spatial distributions of the BEC matrix, are both reasonable. Using the 3DVAR inversion model, a priori SO_2 and NO_2 ESs improved obviously during the heavy haze process, especially for the NO_2 ES. The spatial distributions of SO_2 and NO_2 concentrations simulated using an a posteriori ES are more consistent with measurements than an a priori ES, especially in major cities over the BTH region. Simulation errors of SO_2 , NO_2 , and O_3 concentrations with an a posteriori ES sig-

nificantly decrease, whereas simulations of three pollutants using an a priori ES are underestimated.

Large discrepancies of the simulation and sensitivity coefficient over December 29 may be related with absent calculation of high-order sensitivity coefficient in this case. In the future, we will adopt high-order sensitivity coefficient to improve the constraint effect of SO₂ and NO_x emission sources. In addition, future studies will include the applicability and accuracy of this method for different seasons and regions and different chemical species such as other primary pollutants (e.g., CO) and precursors of secondary pollutants (e.g., PM_{2.5}, PM₁₀, and O₃). An emphasis may be placed on constructing the nonlinear explicit observation operator for precursors of secondary pollutants such as a VOC ES using the high-order sensitivity analysis technique and assessing improvement effects of an a posteriori ES with the 3DVAR inversion method and the CMAQ model.

Data availability. The NCEP-FNL reanalysis data are publicly available at <https://doi.org/10.5065/D6M043C6> (National Center for Atmospheric Research, 2021). The SO₂, NO₂, and O₃ measurements are available at <http://www.cnemc.cn/en/> (last access: 8 September 2021, China National Environmental Monitoring Centre, 2013).

Author contributions. XC and ZZ designed the research. XC and ZH constructed the 3DVAR inversion model, designed model experiments, and performed simulations. XC, ZH, ZZ, YL, YH, and XM contributed to the data processing and analyses. XC and ZH analyzed the results and wrote the paper with inputs from all authors. ZL and XX contributed to the theoretical direction for establishing the inversion model. SW performed the uncertainty analysis of ES.

Competing interests. The contact author has declared that neither they nor their co-authors have any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. Our thanks are due to Yongtao Hu from the Georgia Institute of Technology for assistance with the calculation of the sensitivity. We are grateful to Tsinghua University for providing the emission inventory and the China National Environmental Monitoring Centre for providing surface SO₂, NO₂, and O₃ observation data.

Financial support. This research has been jointly supported by the Fundamental Research Funds for Central Public-interest Scientific Institution from the Chinese Academy of Meteorological Sciences (grant no. 2016Y005), the National Natural Science Foundation of

China (grant no. 91644223), and the National Research Program for Key Issues in Air Pollution Control (grant no. DQGG0104).

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

References

- Bai, N. B. and Zhou, X. J.: Estimation of CO₂, SO₂ and NO_x grid-cell emission sources with a resolution of 1° × 1° in Changes of atmospheric ozone and its impact on climate and environment in China, China Meteorological Press, Beijing, 145–150, 1996.
- Barbu, A. L., Segers, A. J., Schaap, M., Heemink, A. W., and Builtjes, P. J. H.: A multi-component data assimilation experiment directed to sulphur dioxide and sulphate over Europe, *Atmos. Environ.*, 43, 1622–1631, 2009.
- Bergamaschi, P., Hein, R., Heimann, M., and Crutzen, M., P. J.: Inverse modeling of the global CO cycle: 1. Inversion of CO mixing ratios, *J. Geophys. Res.*, 105, 1909–1927, 2000.
- Cao, G. L., Zhang, X. Y., Gong, S. L., Gong, S. L., An, X. Q., and Wang, Y. Q.: Emission inventories of primary particles and pollutant gases for China, *Chin. Sci. Bull.*, 56, 781, <https://doi.org/10.1007/s11434-011-4373-7>, 2011.
- Cheng, X., Xu, X., and Ding, G.: An emission source inversion model based on satellite data and its application in air quality forecasts, *Science China: Earth Sciences*, 53, 752–762, 2010.
- Cheng, Y. F., Zheng, G. J., Wei, C., Mu, Q., Zheng, B., Wang, Z. B., Gao, M., Zhang, Q., He, K. B., Carmichael, G., Pöschl, U., and Su, H.: Reactive nitrogen chemistry in aerosol water as a source of sulfate during haze events in China, *Sci. Adv.*, 2, e1601530, <https://doi.org/10.1126/sciadv.1601530>, 2016.
- Chen, D., Liu, Z., Ban, J., and Chen, M.: The 2015 and 2016 wintertime air pollution in China: SO₂ emission changes derived from a WRF-Chem/EnKF coupled data assimilation system, *Atmos. Chem. Phys.*, 19, 8619–8650, <https://doi.org/10.5194/acp-19-8619-2019>, 2019.
- China National Environmental Monitoring Centre (CNEMC): SO₂, NO₂, and O₃ measurements, CNEMC [data set], available at: <http://www.cnemc.cn/en/> (last access: 8 September 2021), 2013.
- Cohan, D., Hu, Y., Hakami, A., Odman, M. T., and Russell, A.: Implementation of a direct sensitivity method into CMAQ, Models-3 User's Workshop, RTP, North Carolina, 22 October 2002.
- Corazza, M., Bergamaschi, P., Vermeulen, A. T., Aalto, T., Haszpra, L., Meinhardt, F., O'Doherty, S., Thompson, R., Moncrieff, J., Popa, E., Steinbacher, M., Jordan, A., Dlugokencky, E., Brühl, C., Krol, M., and Dentener, F.: Inverse modelling of European N₂O emissions: assimilating observations from different networks, *Atmos. Chem. Phys.*, 11, 2381–2398, <https://doi.org/10.5194/acp-11-2381-2011>, 2011.
- Dai, T., Cheng, Y., Goto, D., Li, Y., Tang, X., Shi, G., and Nakajima, T.: Revealing the sulfur dioxide emission reductions in China by assimilating surface observations in WRF-Chem, *Atmos. Chem. Phys.*, 21, 4357–4379, <https://doi.org/10.5194/acp-21-4357-2021>, 2021.
- Dennis, R., Byun, D., and Novak, J.: The next generation of integrated air quality modeling: EPA's Models-3, *Atmos. Environ.*, 30, 1925–1938, 1996.

- Elbern, H., Schmidt, H., Talagrand, O., and Ebel, A.: 4D-variational data assimilation with an adjoint air quality model for emission analysis, *Environ. Modell. Softw.*, 15, 539–548, 2000.
- Elbern, H., Strunk, A., Schmidt, H., and Talagrand, O.: Emission rate and chemical state estimation by 4-dimensional variational inversion, *Atmos. Chem. Phys.*, 7, 3749–3769, <https://doi.org/10.5194/acp-7-3749-2007>, 2007.
- Enting, I. G.: *Inverse Problems in Atmospheric Constituent Transport*, Cambridge University Press, Cambridge, ISBN 9780521812108, 2002.
- Frey, H. C., Zheng, J., Zhao, Y., Li, S., and Zhu, Y.: Technical Documentation of the AuvTool Software for Analysis of Variability and Uncertainty, North Carolina State University for the Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC, 2002.
- Gilliland, A. B., Dennis, R. L., Roselle, S. J., and Pierce, T. E.: Seasonal NH_3 emission estimates for the eastern United States based on ammonium wet concentrations and an inverse modeling method, *J. Geophys. Res.*, 108, 4477, <https://doi.org/10.1029/2002JD003063>, 2003.
- Gilliland, A. B., Wyatt, A. K., Pinder, R. W., and Dennis, R. L.: Seasonal NH_3 emissions for the continental United States: Inverse model estimation and evaluation, *Atmos. Environ.*, 40, 4986–4998, 2006.
- Heald, C. L., Jacob, D. J., Jones, D. B. A., Palmer, P. I., Logan, J. A., Streets, D. G., Sachse, G. W., Gille, J. C., Hoffman, R. N., and Nehr Korn, T.: Comparative inverse analysis of satellite (MOPITT) and aircraft (TRACE-P) observations to estimate Asian sources of carbon monoxide, *J. Geophys. Res.*, 109, D23306, <https://doi.org/10.1029/2004JD005185>, 2004.
- Henze, D. K., Seinfeld, J. H., and Shindell, D. T.: Inverse modeling and mapping US air quality influences of inorganic $\text{PM}_{2.5}$ precursor emissions using the adjoint of GEOS-Chem, *Atmos. Chem. Phys.*, 9, 5877–5903, <https://doi.org/10.5194/acp-9-5877-2009>, 2009.
- Hong, C., Zhang, Q., He, K., Guan, D., Li, M., Liu, F., and Zheng, B.: Variations of China's emission estimates: response to uncertainties in energy statistics, *Atmos. Chem. Phys.*, 17, 1227–1239, <https://doi.org/10.5194/acp-17-1227-2017>, 2017.
- Hu, Y., Odman, M. T., and Russell, A. G.: Top-down analysis of the elemental carbon emissions inventory in the United States by inverse modeling using Community Multiscale Air Quality model with decoupled direct method (CMAQ-DDM), *J. Geophys. Res.*, 114, D24302, <https://doi.org/10.1029/2009JD011987>, 2009.
- Huang, R. J., Zhang, Y., Bozzetti, C., Ho, K. F., Cao, J. J., Han, Y., Daellenbach, K. R., Slowik, J. G., Platt, S. M., Canonaco, F., Zotter, P., Wolf, R., Pieber, S. M., Brun, E. A., Crippa, M., Ciarelli, G., Piazzalunga, A., Schwikowski, M., Abbasszade, G., Schnelle-Kreis, J., Zimmermann, R., An, Z., Szidat, S., Baltensperger, U., Haddad, I. E., and Prévôt, A. S.: High secondary aerosol contribution to particulate pollution during haze events in China, *Nature*, 514, 218–222, 2014.
- Jiang, Z., Jones, D. B. A., Kopacz, M., Liu, J., Henze, D. K., and Heald, C.: Quantifying the impact of model errors on top-down estimates of carbon monoxide emissions using satellite observations, *J. Geophys. Res.*, 116, D15306, <https://doi.org/10.1029/2010JD015282>, 2011.
- Kerry, A., Gerhard, R., and Mike, D. F.: Fire-growth modelling using meteorological data with random and systematic perturbations, *Int. J. Wildland Fire*, 16, 174–182, 2007.
- Koohkan, M. R., Bocquet, M., Roustan, Y., Kim, Y., and Seigneur, C.: Estimation of volatile organic compound emissions for Europe using data assimilation, *Atmos. Chem. Phys.*, 13, 5887–5905, <https://doi.org/10.5194/acp-13-5887-2013>, 2013.
- Kopacz, M., Jacob, D. J., Henze, D. K., Heald, C. L., Streets, D. G., and Zhang, Q.: Comparison of adjoint and analytical Bayesian inversion methods for constraining Asian sources of carbon monoxide using satellite (MOPITT) measurements of CO columns, *J. Geophys. Res.*, 114, D04305, <https://doi.org/10.1029/2007JD009264>, 2009.
- Li, M., Liu, H., Geng, G., Hong, C., Liu, F., Song, Y., Tong, D., Zheng, B., Cui, H., Man, H., Zhang, Q., He, K.: Anthropogenic emission inventories in China: a review, *Natl. Sci. Rev.*, 4, 834–866, 2017.
- Li, Z., Zang, Z., Li, Q. B., Chao, Y., Chen, D., Ye, Z., Liu, Y., and Liou, K. N.: A three-dimensional variational data assimilation system for multiple aerosol species with WRF/Chem and an application to $\text{PM}_{2.5}$ prediction, *Atmos. Chem. Phys.*, 13, 4265–4278, <https://doi.org/10.5194/acp-13-4265-2013>, 2013.
- Liu, F., Hu, F., and Zhu, J.: Solving the optimal layout problem of multiple industrial pollution sources using the adjoint method, *Science China: Earth Sciences*, 35, 64–71, 2005.
- Liu, J., Mauzerall, D. L., Chen, Q., Zhang, Q., Song, Y., Peng, W., Klimont, Z., Qiu, X. H., Zhang, S. Q., Hu, M., Lin, W. L., Smith, K. R., and Zhu, T.: Air pollutant emissions from Chinese households: A major and underappreciated ambient pollution source, *P. Natl. Acad. Sci. USA*, 113, 7756–7761, 2016.
- Liu, Z., Liu, Q., Lin, H. C., Schwartz, C. S., Lee, Y. H., and Wang, T.: Three-dimensional variational assimilation of MODIS aerosol optical depth: implementation and application to a dust storm over East Asia, *J. Geophys. Res.*, 116, D23206, <https://doi.org/10.1029/2011JD016159>, 2011.
- Lucas, D. D., Simpson, M., Cameron-Smith, P., and Baskett, R. L.: Bayesian inverse modeling of the atmospheric transport and emissions of a controlled tracer release from a nuclear power plant, *Atmos. Chem. Phys.*, 17, 13521–13543, <https://doi.org/10.5194/acp-17-13521-2017>, 2017.
- Ma, J. and van Aardenne, J. A.: Impact of different emission inventories on simulated tropospheric ozone over China: a regional chemical transport model evaluation, *Atmos. Chem. Phys.*, 4, 877–887, <https://doi.org/10.5194/acp-4-877-2004>, 2004.
- Manning, A. J., Doherty, S. O., Jones, A. R., Simmonds, P. G., Derwent, R. G.: Estimating UK methane and nitrous oxide emissions from 1990 to 2007 using an inversion modeling approach, *J. Geophys. Res.*, 116, D02305, <https://doi.org/10.1029/2010JD014763>, 2011.
- Martin, R. V., Jacob, D. J., Chance, K., Palmer, P. I., and Evans, M. J.: Global inventory of nitrogen oxide emissions constrained by space-based observations of NO_2 columns, *J. Geophys. Res.*, 108, 4537, <https://doi.org/10.1029/2003JD003453>, 2003.
- Michalakes, J., Dudhia, J., Gill, D., Henderson, T., Klemp, J., Skamarock, W., and Wang, W.: The Weather Research and Forecast Model: Software Architecture and Performance, in: *Proceedings of the 11th ECMWF Workshop on the Use of High Performance Computing In Meteorology*, Reading, UK, edited by: Mozdzyński, G., 25–29 October 2004.

- Mijling, B. and van der A, R. J.: Using daily satellite observations to estimate emissions of short-lived air pollutants on a mesoscopic scale, *J. Geophys. Res.*, 117, D17302, <https://doi.org/10.1029/2012JD017817>, 2012.
- Miyazaki, K., Eskes, H. J., and Sudo, K.: Global NO_x emission estimates derived from an assimilation of OMI tropospheric NO₂ columns, *Atmos. Chem. Phys.*, 12, 2263–2288, <https://doi.org/10.5194/acp-12-2263-2012>, 2012.
- Napelenok, S. L., Pinder, R. W., Gilliland, A. B., and Martin, R. V.: A method for evaluating spatially-resolved NO_x emissions using Kalman filter inversion, direct sensitivities, and space-based NO₂ observations, *Atmos. Chem. Phys.*, 8, 5603–5614, <https://doi.org/10.5194/acp-8-5603-2008>, 2008.
- National Center for Atmospheric Research: NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory [data set], <https://doi.org/10.5065/D6M043C6>, 2021.
- Otte, T. L. and Pleim, J. E.: The Meteorology-Chemistry Interface Processor (MCIP) for the CMAQ modeling system: updates through MCIPv3.4.1, *Geosci. Model Dev.*, 3, 243–256, <https://doi.org/10.5194/gmd-3-243-2010>, 2010.
- Pagowski, M., Grell, G. A., McKeen, S. A., Peckham, S. E., and Devenyi, D.: Threedimensional variational data assimilation of ozone and fine particulate matter observations: some results using the weather research and forecasting-chemistry model and grid-point statistical interpolation, *Q. J. Roy. Meteor. Soc.*, 136, 2013–2024, 2010.
- Parrish, D. F. and Derber, J. C.: The National Meteorological Center's spectral statistical interpolation analysis scheme, *Mon. Weather Rev.*, 120, 1747–1763, 1992.
- Peng, L., Zhang, Q., Yao, Z., Mauzerall, D. L., Kang, S., Du, Z., Zheng, Y., Xue, T., and He, K.: Underreported coal in statistics: A survey-based solid fuel consumption and emission inventory for the rural residential sector in China, *Appl. Energ.*, 235, 1169–1182, 2019.
- Peng, Z., Liu, Z., Chen, D., and Ban, J.: 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, 4837–4855, <https://doi.org/10.5194/acp-17-4837-2017>, 2017.
- Sportisse, B.: A review of current issues in air pollution modeling and simulation, *Computat. Geosci.*, 11, 159–181, 2007.
- Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendaal, M., Kanakidou, M., Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J. P.: The continental source of glyoxal estimated by the synergistic use of spaceborne measurements and inverse modelling, *Atmos. Chem. Phys.*, 9, 8431–8446, <https://doi.org/10.5194/acp-9-8431-2009>, 2009.
- Streets, D. G., Bond, T. C., Carmichael, G. R., Fernandes, S. D., Fu, Q., He, D., Klimont, Z., Nelson, S. M., Tsai, N. Y., Wang, M. Q., Woo, J. H., and Yarber, K. F.: An inventory of gaseous and primary aerosol emissions in Asia in the year 2000, *J. Geophys. Res.*, 108, 8809–8820, 2003.
- Tang, X., Zhu, J., Wang, Z. F., and Gbaguidi, A.: Improvement of ozone forecast over Beijing based on ensemble Kalman filter with simultaneous adjustment of initial conditions and emissions, *Atmos. Chem. Phys.*, 11, 12901–12916, <https://doi.org/10.5194/acp-11-12901-2011>, 2011.
- Tang, X., Zhu, J., Wang, Z., Gbaguidi, A., Lin, C., Xin, J., Song, T., and Hu, B.: Limitations of ozone data assimilation with adjustment of NO_x emissions: mixed effects on NO₂ forecasts over Beijing and surrounding areas, *Atmos. Chem. Phys.*, 16, 6395–6405, <https://doi.org/10.5194/acp-16-6395-2016>, 2016.
- Tang, X. Y., Zhang, Y. H., and Shao, M.: *Atmospheric Environmental Chemistry*, 2nd edn., Higher Education Press, Beijing, 447–449, 2006 (in Chinese).
- The University of North Carolina (UNC): SMOKE v3.6 user's manual, The institute for the Environment, Chapel Hill, 520 pp., 2014.
- Wang, C., An, X. Q., Zhai, S. X., Hou, Q., and Sun, Z. B.: Tracking sensitive source areas of different weather pollution types using GRAPES-CUACE adjoint model, *Atmos. Environ.*, 175, 154–166, 2018.
- Wang, L. T., Zhang, P., Yang, J., Zhao, X. J., Wei, W., Su, J., Cheng, D. D., Liu, X., Y., Han, G., G., and Wang, H. J.: Application of CMAQ-DDM-3D in the source analysis of fine particulate matter (PM_{2.5}), *Acta Scientiae Circumstantiae*, 33, 1355–1361, 2013 (in Chinese).
- Wang, P., Wang, H., Wang, Y. Q., Zhang, X. Y., Gong, S. L., Xue, M., Zhou, C. H., Liu, H. L., An, X. Q., Niu, T., and Cheng, Y. L.: Inverse modeling of black carbon emissions over China using ensemble data assimilation, *Atmos. Chem. Phys.*, 16, 989–1002, <https://doi.org/10.5194/acp-16-989-2016>, 2016.
- Wang, Y., McElroy, M. B., Martin, R. V., Streets, D. G., Zhang, Q., and Fu, T.-M.: Seasonal variability of NO_x emissions over east China constrained by satellite observations: Implications for combustion and microbial sources, *J. Geophys. Res.*, 112, D06301, <https://doi.org/10.1029/2006JD007538>, 2007.
- Wang, Q., Huo, H., He, K., Yao, Z., and Zhang, Q.: Characterization of vehicle driving patterns and development of driving cycles in Chinese cities, *Transport. Res. D-Tr. E.*, 13, 289–297, 2008.
- Wang, Y. S., Yao, L., Wang, L., L., Liu, Z. R., Ji, D. S., Tang, G. Q., Zhang, J. K., Sun, Y., Hu, B., and Xin, J. Y.: Mechanism for the formation of the January 2013 heavy haze pollution episode over central and eastern China, *Science China, Earth Sciences*, 57, 14–25, 2014 (in Chinese).
- Willmott, C. J., Robeson, S. M., and Matsuura, K. A.: A refined index of model performance, *Int. J. Climatol.*, 32, 2088–2094, 2012.
- Xu, X. D., Xie, L. A., Cheng, X. H., Xu, J. M., Zhou, X. J., and Ding, G. A.: Application of an Adaptive Nudging Scheme in Air Quality Forecasting in China, *J. Appl. Meteorol. Climatol.*, 47, 2105–2114, 2008.
- Yang, Q., Wang, Y. H., Zhao, C., Liu, Z., William, I., Gustafson, J., and Shao, M.: NO_x emission reduction and its effects on ozone during the 2008 Olympic Games, *Environ. Sci. Technol.*, 45, 6404–6410, 2011.
- Zang, Z. L., Li, Z. J., Pan, X. B., Hao, Z. L., and You, W.: Aerosol data assimilation and forecasting experiments using aircraft and surface observations during CalNex, *Tellus B*, 68, 29812, <https://doi.org/10.3402/tellusb.v68.29812>, 2016.
- Zhang, Q., Streets, D. G., He, K., Wang, Y., Richter, A., Burrows, J. P., Uno, I., Jang, C. J., Chen, D., Yao, Z., and Lei Y.: NO_x emission trends for China, 1995–2004: The view from the ground and the view from space, *J. Geophys. Res.*, 112, D22306, <https://doi.org/10.1029/2007JD008684>, 2007.

- Zhang, L., Shao, J. Y., Lu, X., Zhao, Y. H., Hu, Y. Y., Henze, D., Liao, H., Gong, S., L., and Zhang, Q.: Sources and processes affecting fine particulate matter pollution over North China: An adjoint analysis of the Beijing APEC period, *Environ. Sci. Technol.*, 50, 8731–8740, 2016.
- Zhang, Q., Geng, G. N., Wang, S. W., Andreas, R., and He, K. B.: Satellite remote sensing of changes in NO_x emissions over China during 1996–2010, *Chin. Sci. Bull.*, 57, 2857–2864, 2012.
- Zhang, Q., Streets, D. G., Carmichael, G. R., He, K. B., Huo, H., Kannari, A., Klimont, Z., Park, I. S., Reddy, S., Fu, J. S., Chen, D., Duan, L., Lei, Y., Wang, L. T., and Yao, Z. L.: Asian emissions in 2006 for the NASA INTEX-B mission, *Atmos. Chem. Phys.*, 9, 5131–5153, <https://doi.org/10.5194/acp-9-5131-2009>, 2009.
- Zhao, B., Wang, P., Ma, J. Z., Zhu, S., Pozzer, A., and Li, W.: A high-resolution emission inventory of primary pollutants for the Huabei region, China, *Atmos. Chem. Phys.*, 12, 481–501, <https://doi.org/10.5194/acp-12-481-2012>, 2012.
- Zhao, H. Y., Zhang, Q., Guan, D. B., Davis, S. J., Liu, Z., Huo, H., Lin, J. T., Liu, W. D., and He, K. B.: Assessment of China's virtual air pollution transport embodied in trade by using a consumption-based emission inventory, *Atmos. Chem. Phys.*, 15, 5443–5456, <https://doi.org/10.5194/acp-15-5443-2015>, 2015.
- Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., Li, H., Li, X., Peng, L., Qi, J., Yan, L., Zhang, Y., Zhao, H., Zheng, Y., He, K., and Zhang, Q.: Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions, *Atmos. Chem. Phys.*, 18, 14095–14111, <https://doi.org/10.5194/acp-18-14095-2018>, 2018.
- Zhu, J. and Wang, P.: Ensemble kalman smoother and ensemble kalman filter approaches to the joint air quality state and emission estimation problem, *Chinese Journal of Atmospheric Sciences*, 30, 871–882, 2006.
- Zhu, J., Tang, X., Wang, Z. F., and Wu, L.: A review of air quality data assimilation methods and their application, *Chinese Journal of Atmospheric Sciences*, 42, 607–620, 2018.