

Response to Interactive comments from Anonymous Referee #1

Referee comments are in black. Author responses are in blue.

General comments:

This study developed a new inverse modeling approach based on DDM-3D and 3DVAR, which was applied to assimilate surface SO₂ and NO₂ concentration measurements to optimize SO₂ and NO_x emissions in the Beijing-Tianjin-Heibei region. The a posteriori emission inventory is applied to air quality forecast, and results show the simulated surface SO₂, NO₂, and O₃ concentrations using the a posteriori emission inventory are in better agreement with observations than a priori. I will recommend its publication after the following comments are addressed.

We thank the anonymous referee for his/her insightful and constructive comments. Below are our point-to-point responses in detail.

Specific comments:

1. Inversion modeling tools are summarized in the 3rd paragraph in the introduction, but how to classify them is open to discussion. Is it appropriate to consider adjoint modeling and sensitivity analysis as two inversion modeling tools? In the manuscript Stavrakou et al. (2009) and Zhai et al. (2018) are considered as adjoint modeling. The two references did use adjoint modeling, but actually (1) Stavrakou et al. (2009) used adjoint modeling to calculate sensitivity, which was future used to minimize a 4DVAR cost function; (2) Zhai et al. (2018) only used adjoint modeling to calculate sensitivity without constrain emissions. Thus, Stavrakou et al. (2009) should be considered as 4DVAR, and Zhai et al. (2018) should not be cited. I agree that Henze et al. (2009) and Jiang et al. (2011) should be classified as 4DVAR; the two papers used GEOS-Chem adjoint model to calculate the sensitivity of cost function with respect to emissions scale factors, and the sensitivity were used to minimize cost functions. Considering sensitivity analysis as the name of inversion modelling is inappropriate as many approaches (such as 4DVAR) need to calculate sensitivity. Moreover, Mijling et al. (2012) was considered as sensitivity analysis in the manuscript, but actually it should be Kalman filter. I did not read all the references in the paragraph and could not help classify them; I suggest reading them carefully and classifying them properly.

Many thanks for suggestions. It's right that considering the sensitivity analysis as one type of inversion modelling is inappropriate, we revise improper references and delete the type of sensitivity analysis according to suggestions.

2. Line 86: The reason that GEOS-Chem is not used urban air quality forecast is its spatial resolution is too coarse, and I suggest add this explanation.

Corrected.

3. Line 88-90: Why it is stated EnKF and 4DVAR are absent of sensitivity analysis of the source-receptor relationship? Adjoint model (for example, GEOS-Chem adjoint) is used to calculate the source-receptor relationship in 4DVAR.

Thanks for the suggestion. This sentence is inappropriate and we delete it.

4. Line 115-118: Are there any references to show that multiple receptors will result in high computational costs for adjoint model? My experience is whether assimilating one species or multiple species, the computational time difference for calculating the cost functions with respect to emissions is very small.

Thanks for the comment. Differences of computational costs between multiple receptors and multiple species are larger, because sensitivity of the source-receptor relationship at all grids is calculated one receptor by one receptor, while sensitivity coefficients at all grids will be computed simultaneously for one species or multiply species. We add the related references in the revised manuscript.

5. How the uncertainties in Table 1 are calculated? If they come from other paper, please add references.

We calculate the uncertainties based on activities and emission factors from the references using the Monte-Carlo method. We add the related descriptions and references in the revised version.

6. I suggest add some details of random perturbation method that is used to generate the 30 sets of inventories.

Thanks for the suggestion. Firstly we calculate the uncertainties of SO₂ and NO_x emission sources, and we obtain the probability distribution of uncertainties for four sections of emission sources, respectively. Then we conduct thirty times of random perturbation on uncertainties of four sections of emission sources according to the probability distributions using the same perturbation coefficients for every perturbation. Lastly we calculate the total emission rates using random uncertainties of four sections for 30 sets of inventories, respectively. We add the detailed descriptions and the reference about random perturbation method in the revised manuscript.

7. Line 193: What “24-h strengths of ES for each month” mean?

It means that 24 hours diurnal variation of ES for every month. We revised it.

8. How a priori SO₂ and NO_x emissions are vertically distributed in the model?

We set 32 vertical levels in the CMAQ model and half of them is lower than 2km and other levels are located from 2km height to the top of atmosphere.

9. Please check Eq. (3) careful. Which p should or not be capitalized?

Thanks. P and ϵ are all revised to small letter.

Technical corrections:

1. Line 74: Henze et al., 2008 -> Henze et al., 2009. And the year is also wrong in the reference list.

Thanks. Revised.

2. Please check Eq. (3) and corresponding descriptions carefully. Which p should or not be capitalized?

Thanks. P and ϵ are all revised to small letter. The corresponding descriptions are corrected.

Response to Interactive comments from Anonymous Referee #2

Referee comments are in black. Author responses are in blue.

Comments to “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”

General comments:

Timely precise emissions of air pollutants are crucial for air quality prediction and mitigation. The authors present a newly developed emission inversion method based on the combination three-dimensional decoupled direct (DDM-3D) and 3DVAR data assimilation techniques. The emission inversion method is applied to update the SO₂ and NO_x emissions over the Beijing-Tianjin-Hebei region during a heavy haze period. Their results demonstrate the newly updated emissions are reasonable and helpful to the prediction of the air pollutants including O₃. The manuscript is well-organized and scientifically sound. Therefore, I recommend accepting it after minor revision.

We thank the anonymous referee for his/her insightful and constructive comments. Below are our point-to-point responses in detail.

Specific comments:

L180 Please describe the random perturbation method more detail.

Thanks for the suggestion. Firstly we calculate the uncertainties of SO₂ and NO_x emission sources, and we obtain the probability distribution of uncertainties for four sections of emission sources, respectively. Then we conduct thirty times of random perturbation on uncertainties of four sections of emission sources according to the probability distributions using the same perturbation coefficients for every perturbation. Lastly we calculate the total emission rates using random uncertainties of four sections for 30 sets of inventories, respectively. We add the detailed descriptions and the reference about random perturbation method in the revised manuscript.

L203 Matrix **D** can not represent the impacts of local emissions at one grid on other grids. It should be **C**.

Thanks. It's mistake. Revised.

L270 Please change the Ls here and in formula (5), since you also use Ls in Line 214. It's the same variable to Ls in line 214.

Figure 6 What are the reasons for the large discrepancies of the simulation and

sensitivity coefficient over December 29?

Thanks for the comment. 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 constraint effects of SO₂ and NO_x emission sources with the newly developed method in this study. We add the related descriptions in the discussion part of the revised manuscript.

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

5 Xinghong Cheng¹, Zilong Hao², Zengliang Zang², Zhiquan Liu³, Xiangde Xu¹,
| [Shuisheng Wang⁴](#), Yuelin Liu⁵⁴, Yiwen Hu⁶⁵, Xiaodan Ma⁶⁵

1. State Key Lab of Severe Weather & Key Laboratory for Atmospheric Chemistry, Chinese Academy of Meteorological Sciences, Beijing 100081, China

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

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

[4. GZ Source Clear Tech. Co., Ltd., Guangzhou 510630, China](#)

[54](#). College of Architecture and Environment, Sichuan University, Chengdu 610065, China

[65](#). Nanjing University of Information Science and Technology, Nanjing 210044, China

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

Abstract. We develop a new inversion method which is suitable for linear and nonlinear emission sources (ES) modeling, based on the three-dimensional decoupled direct (DDM-3D) sensitivity analysis module in the Community Multiscale Air Quality (CMAQ) model and the three-dimensional variational (3DVAR) data assimilation technique. We established the explicit observation operator

20 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 December 27-30, 2016. Results show that the spatial

25 distribution of sensitivities of SO₂ and NO_x ES to their concentrations, as well as the BEC matrix of ES, are reasonable. Using ~~the 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 ~~the a~~*priori* ES in the BTH region. The temporal variations in regionally averaged SO₂, NO₂, and O₃ modelled

concentrations using ~~the~~-*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; reducing uncertainties of ES and its dynamic updating; providing 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 SO_2 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 for investigating the causes of air pollution, and atmospheric chemical transport model (ACTM). Uncertainties in ES are a major factor in determining the simulated and forecast accuracy of the ACTM, and these uncertainties can greatly affect the design of ES control strategies (Tang, 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 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 et al., 1996; Streets et al., 2003; Zhang et al., 2009, 2012; Cao et al., 2011; Zhao et al., 2012; Zhao et al., 2015; and 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 et al., 2004). Moreover, real-time updates of emission inventories are difficult to achieve because of its rapid spatiotemporal variations

due to high-speed urbanization, and a delay in the release of statistical data of approximately 1–2 years.

60 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
65 improve the spatiotemporal distribution of ES. With the development of data assimilation technology and ACTM, constraining the strength of ES using ACTM 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
70 inverse modeling (Manning et al., 2011), adjoint modeling (Liu et al., 2005; ~~Stavrakou et al., 2009~~; Koohkan et al., 2013; Zhang et al., 2016; ~~Zhai et al., 2018, and~~ Wang et al., 2018), Bayes estimation theory (Kopacz et al., 2009), ensemble Kalman-filtering (EnKF, Zhu et al., 2006, 2018; Barbu et al., 2009; Tang et al., 2011, 2016; Miyazaki et al., 2012; Mijling et al., 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., ~~2008~~2009; Stavrakou et al., 2009; Corazza et al., 2011; Jiang et al., 2011), an adaptive nudging scheme in the CMAQ model (Xu et al., 2008; Cheng et al., 2010), inversion algorithms combining pollutant dispersion models and a Monte
75 Carlo simulation (Yang et al., 2013), ~~and sensitivity analysis (Fu et al., 2007; Hu et al., 2009; Mijling et al., 2012)~~. Results show that using an inversion modeling approach to retrieve the spatial distribution of
80 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 “online” Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) are used as the forward model to relate the SO₂ emissions to the simulated concentration, and efficiently update the emissions based on the 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 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. ~~In addition, the EnKF and 4DVAR methods exhibit difficulty in accurate inversion of ES in real applications due to the absence of sensitivity analysis of the source receptor (S-R) relationship.~~

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 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 ES, it is necessary to construct an inversion model that can satisfy the aforementioned requirements. Firstly, although 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 hour). Secondly, the assimilation effect of 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 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 modelled results. Thus, the top-down 3DVAR constraint methods for ES based on the first- or high-order DDM-3D sensitivity analysis techniques can maintain the coordination between assimilated field of ES and simulated concentration of air pollutants.

115 The primary methods used to calculate the S-R 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 and
 Heald et al., 2004), NH₃ (Gilliland et al., 2003), and EC (Hu et al., 2009). The adjoint method is a
 backward-sensitivity calculation method, while brute force and DDM-3D are forward-sensitivity
 120 calculation methods. For inverse modeling of pollution sources with 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
 125 (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 Particulate Matter; HDDM3D/PM)
 in the newly released version of 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
 130 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 non-linear relationship.

135 The results of inverse modeling are very sensitive to uncertainties in the ES of NO_x, NH₄ and
 inorganic aerosols (Zhang et al., 2016). Impact of uncertainties in the ES on the assimilation effects
 need 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 BEC matrix of ES based on
 multiple sets of ES. We developed a new inverse modeling approach for ES that combines the
 140 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: Section 2 describes
 the inversion model and presents results of sensitivity analysis and the BEC; Section 3 provides details

of the WRF-CMAQ model, and configurations and experiments of simulation; Section 4 presents the results of the control and experiment simulations with ~~the~~ *a priori* and *a constrained posteriori* ES, respectively; finally, the discussions and conclusions are provided in Section 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 x 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 December 27-30, 2016 and the Multi-resolution Emission Inventory for China (MEIC) data for 2012, with 1 °x1 ° and 0.25 °x0.25 ° 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 et al., 2010). The boundary conditions for chemical trace gases consisted of idealized, northern hemispheric, mid-latitude profiles based upon output from the National Oceanic Atmospheric Administration (NOAA) Agronomy Lab Regional Oxidant model. The model simulation started on December 27, 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 ~~the~~ *a priori* MEIC data for 2012, and an experiment run with *a constrained posteriori* ~~constrained~~ ES.

Hourly measurements of SO₂, NO₂, and O₃ concentrations at 129 stations during December 27–30, 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 Figure 1.

170 3. Inverse modeling method

3.1 Constructing BEC matrix

To construct the BEC matrix for the inversion model, we combined the National Meteorological Center's (NMC) technique (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:

175 (1) Determine the total errors of ES from a priori bottom-up inventory.

Uncertainty analyses of ES require detailed information of activities and emission factors from ~~the a~~ priori MEIC emission inventory. The relevant data collected in the China Environmental Yearbook are limited, and do not satisfy with the requirements of uncertainty analysis. Therefore, we used the available research results relating to SO₂ and NO_x ES, and conducted uncertainty analysis for four types of major
180 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 determining the error ranges in total emission rates of SO₂ and NO_x (Table 1). Uncertainties in SO₂ industry and power plant ES are slightly greater than those for NO_x, while the opposite is true for emissions from the residential and traffic sectors.

185 (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 SO₂ 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 ES based on uncertainty analysis for four sections, respectively. Then we conduct thirty times of random perturbation on
190 uncertainties of four sections of ES according to the probability distributions using the same perturbation coefficients for every perturbation. Lastly we calculate thirty total emission rates using random uncertainties of four sections for every sets of inventories, respectively.

(3) Process the 3-D gridded ES as input to the CMAQ model.

We used the SMOKE model, national population and road network distribution data in 2016, the
195 temporal apportionment coefficients in the BTH region (Zhang et al., 2007, 2009, Simpson et al., 2003, and Wang et al., 2010), and the CB05-ae06-aq chemical species data in the CMAQ model, to process thirty sets of nationwide emission inventories into 3-D 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 3-D gridded ES.

200 Finally, the NMC method was used to calculate the BEC matrix of the 3-D gridded ES for each month, including horizontal and vertical correlation coefficients and standard deviations. The background error is defined as the difference between thirty sets of 3-D gridded ES generated by the random perturbation method, and the 3-D gridded background ES directly processed from the original MEIC emission inventory with the SMOKE model, at every hour ([24 hours diurnal variation of ES for every month](#)~~24-h strengths of ES for each month~~).
205

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:

$$\mathbf{B} \approx \frac{1}{2} \langle (e_t - e_b)(e_t - e_b)^T \rangle, \quad (1)$$

210 where e_t is the perturbation field and e_b is the background field of *a priori* ES. Eq. (1) can be written as follows:

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

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 SD of
215 all state variables in the 3-D grids. ~~\mathbf{D}~~ \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 thirty sets of 3-D gridded ES, and the SD of the BEC matrix for SO₂ and NO_x ES at 08:00 local time in December. SO₂ and NO_x ES
220 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 SO₂ 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
225 SO₂ and NO_x ES 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 ES increases with grid distance and this is related to the regional pollution characteristics of SO_2 . The vertical distributions of the SDs in **B** for SO_2 and NO_x ES vary with height: the SD of SO_2 ES are larger on the fourth and eighth model levels than on other levels; while for 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 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):

$$S_j = P_j \frac{\partial C}{\partial P_j} = P_j \frac{\partial C}{\partial (\varepsilon_j P_j)} = \frac{\partial C}{\partial \varepsilon_j} \quad S_j = p_j \frac{\partial C}{\partial p_j} = p_j \frac{\partial C}{\partial (\varepsilon_j p_j)} = \frac{\partial C}{\partial \varepsilon_j}, \quad (3)$$

where S_j is the sensitivity of the pollutant j to the parameter P_j , P_j is the *a priori* ES of the pollutant j , C is the concentration of the pollutant j , and ε_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 3-D gridded *a priori* ES from MEIC in 2012 to calculate the sensitivity coefficients of SO_2 and NO_2 concentrations with respect to ES during the “heavy haze” episode of December 27–30, 2016. Figure 4 shows the spatial distribution of 96-h averaged sensitivity coefficients for SO_2 and NO_2 concentrations with respect to ES during December 27–30, 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., SO₂ and NO₂ concentrations in those areas are greatly affected by the SO₂ and NO_x ES.

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 ES and receptor concentrations as follows:

$$\mathbf{H} = \frac{\partial \mathbf{C}}{\partial \mathbf{E}} = \frac{\partial \mathbf{C}}{\partial (\varepsilon_j E_0)} = \frac{S_j}{E_0}, \quad (4)$$

where \mathbf{H} is the observation operator matrix, E are ~~the~~a *posteriori* ES, which can be written as the product of the perturbation coefficient and *a priori* ES during the assimilation window time, and E_0 are ~~the~~a *priori* ES. For primary pollutants such as SO₂ and NO₂, S_j 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 PM_{2.5} and O₃, S_j 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 SO₂ and NO_x ES.

3.4 Observational error covariance

We firstly performed quality control on the observed SO₂ and NO₂ concentration data. This process involved three steps:

- (1) Redundant data removal, and matching the density of observation data to the model grid. For some grids with more than one observation station, we used the average of those stations.
- (2) Extrema control, i.e., filtering out data exceeding three times of SD of observation data.
- (3) Anomaly removal, 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. Representation error, ε_r , can be expressed as follows (Pagowski et al., 2010):

$$\varepsilon_r = \gamma \varepsilon_o \sqrt{\frac{\Delta x}{L_s}} \quad (5)$$

where γ is the amplification factor, which is used to adjust the instrument error, ε_o is related to the SO₂ and NO₂ 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 km, 5 km, and 10 km, respectively (Zang et al., 2016). Finally, the total observation error for the SO₂ and NO₂ concentrations, ε , is written as:

$$\varepsilon = \sqrt{\varepsilon_o^2 + \varepsilon_r^2}. \quad (6)$$

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 inversing variable of an *a posteriori* ES. The optimal inversion of ES for SO₂ and NO_x are obtained using Eq. (7). The 3DVAR solves for the minimum value of $J(e)$ to determine the inversing 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 receptor's concentrations; the overall framework is shown in Figure 5.

310 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 ~~the a~~ hourly *posteriori* ES of SO₂ and NO₂ using measurements from 45 and 129 stations, respectively, on December 27, 2016. We validated simulations from the control and experiment run using observational data during December 28–30, 2016.

315 Figure 6 shows time series of hourly, regional averaged SO₂ and NO₂ simulations from the control run, observations, and sensitivity coefficients at 45 stations in the BTH region during December 27–30, 2016. The trends in modelled concentrations and sensitivity coefficients of SO₂ and NO₂ concentrations with respect to ES are consistent, therefore the sensitivity coefficients can reasonably reflect the impacts of the ES on concentrations. However, simulated SO₂ and NO₂ concentrations with ~~the a~~ underestimated *priori* ES are all significantly lower than observations during the heavy haze period. Thus, it is important to improve ~~the a~~ *priori* ES using the inversion model.

Figures 7 and 8 show the spatial distributions of 24-h averaged emission rates from ~~the a~~ *priori* and *a posteriori* ES of SO₂ and NO₂, and their increments on December 27, 2016. Emission rates of ~~the a SO₂ and NO₂ posteriori of SO₂ and NO₂~~ ES in the major cities and surrounding areas clearly increase. 325 Compared with ~~the a~~ *priori* ES, the maximum strengths of SO₂ and NO₂ ES increase by approximately 17% and 500%, respectively. Therefore, the strengths of SO₂ and NO₂ in ~~the a~~ *priori* ES were greatly underestimated, especially for NO₂.

Using the WRF-CMAQ ~~model~~ and *a posteriori* ES, we simulated concentrations of SO₂, NO₂, and O₃ in the BTH region during December 28–30, 2016, and validated these simulations with 330 measurements from 45 stations. Figures 9 and 10 show the spatial distributions of 72-h averaged SO₂ and NO₂ concentrations simulated with ~~the a~~ *priori* and *a posteriori* ES, increments, and their observations. In general, SO₂ and NO₂ concentrations simulated using ~~the a~~ *posteriori* ES are closer to observations than ~~the a~~ *priori* ES, and regional differences in improvements for SO₂ and NO₂ exist. For SO₂, the improvement is noticeable in the BTH region. However, the simulated concentrations in 335 Beijing with ~~the a~~ *posteriori* ES are overestimated. This may be related to greater uncertainties in SO₂ sources and the impacts of regional transport from surrounding areas. For NO₂, simulated differences with ~~the a~~ *priori* and *a posteriori* ES are significant in major cities such as Beijing, Tianjin, Shijiazhuang, Baoding, Xingtai, Handan, and Jinan. The simulated concentrations of NO₂ using ~~the a~~

posteriori ES are more consistent with measurements, while those with ~~the-a~~*a priori* ES are significantly underestimated.

We also investigated temporal variations in regionally-averaged SO₂, NO₂, and O₃ concentrations simulated using ~~the-a~~*a priori* and *a posteriori* ES, and observations from the 45 stations over the BTH region during December 28–30, 2016 (Figure 11). In general, simulated SO₂, NO₂, and O₃ concentrations using ~~the-a~~*a posteriori* ES are closer to measurements, while the SO₂ and NO₂ concentrations simulated by ~~the-a~~*a priori* ES are significantly lower than observations, and the modelled O₃ concentrations are obviously higher than measurements. In addition, the peak of SO₂ simulations with ~~the-a~~*a posteriori* ES are close to measurements, but the peak of NO₂ and valley of O₃ simulations are lower and higher than observations, respectively. This may be related to the absence of inverse modeling of volatile organic compound (VOC) ES and uncertainties of sensitivity coefficients calculation. In this study, we used only first-order sensitivity coefficients, but the relationship between ES of precursors of O₃ such as VOCs and NO_x, and their receptor's concentrations are nonlinear, and O₃ is generated from both NO_x and VOCs ES. Therefore, higher-order sensitivity coefficients are necessary for inverse modeling of ES of NO_x and VOCs.

To further assess the simulated accuracy of SO₂, NO₂, and O₃ concentrations, we calculated the following statistics (Willmott et al., 2011): correlation coefficient (R), root-mean-squared error (RMSE), mean bias (MB), normalized mean bias (NMB), and index of agreement (IOA; see Table 2). Except that R of NO₂ and O₃ decrease and RMSE of O₃ increases using the constrained ES, other statistics show improvements. Especially, MB and NMB of three pollutants decline significantly and IOA are closer to 1.0, which means that modelled results of three pollutants are more consistent with observations. R between SO₂ simulation and observation shows a slight improvement when using ~~the-a~~*a posteriori* ES, whereas R decreases for NO₂ and O₃, and it may be related with the absence of constraint of VOCs ES.

5. Summary and conclusions

We developed a new inverse approach of ES by combining the sensitivity analysis technique between ES and receptor's concentration, and the 3DVAR method. Our approach is suitable for solving the linear or nonlinear inversion problems for ES, and it compute fastly and obtain the relatively accurate real-time dynamic updates of ES. First, we used the sensitivity analysis tool in the CMAQ model to

construct the explicit observation operator matrix between ES and 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 December 27–30, 2016. Results show that the observation operators between SO₂ and NO₂ ES, and their concentrations, as well as spatial distributions of the BEC matrix are both reasonable. Using the 3DVAR inversion model, ~~the-a~~_a priori SO₂ and NO₂ ES improved obviously during the heavy haze process, especially for NO₂ ES. The spatial distributions of SO₂ and NO₂ concentrations simulated using ~~the-a~~_a posteriori ES are more consistent with measurements than ~~the-a~~_a priori ES, especially in major cities over the BTH region. Simulation errors of SO₂, NO₂ and O₃ concentrations with ~~the-a~~_a posteriori ES significantly decrease, whereas simulations of three pollutants using ~~the-a~~_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-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 VOCs ES using the high-order sensitivity analysis technique, and assessing improvement effects of ~~the~~_a posteriori ES with the 3DVAR inversion method and CMAQ model.

Data availability. The NCEP-FNL reanalysis data are publicly available at <http://rda.ucar.edu/datasets/ds083.2/>. The SO₂, NO₂, and O₃ measurements are available at <http://113.108.142.147:20035/emcpublish>.

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 theoretical direction for establishing the inversion

model. [SW performed the uncertainty analysis of ES.](#)

Competing interests. The authors declare that they have no conflicts of interest.

Acknowledgements. 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 work was supported jointly by the Fundamental Research Funds for Central Public-interest Scientific Institution from 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).

References

Bai, N. B., Zhou, X. J.: Estimation of CO₂, SO₂ and NO_x gridded emission sources with a resolution of 1°×1° in Changes of atmospheric ozone and its impact on climate and environment in China, Beijing: China Meteorological Press, 145-150, 1996.

Barbu, A. L., Segers, A. J., Schaap, M., Heemink, A. W., Builtjes, P. J. H.: A multi-component data assimilation experiment directed to sulphur dioxide and sulphate over Europe, Atmos. Environ., 43 (9), 1622-1631, 2009.

Bergamaschi, P., Hein, R., Heimann, M., Crutzen, M., P. J.: Inverse modeling of the global CO cycle: 1. Inversion of CO mixing ratios, J. Geophys. Res., 105(D2), 1909–1927, 2000.

Cao, G. L., Zhang, X. Y., Gong, S. L., Gong, S. L., An, X. Q., Wang, Y. Q.: Emission inventories of primary particles and pollutant gases for China, Chin. Sci. Bull., 56, doi:10.1007/s11434-011-4373-7, 2011.

Cheng, X., Xu, X., 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., Su, H.: Reactive nitrogen chemistry in aerosol water as a source of sulfate during haze events in China, Sci. Adv., 2, e1601530, 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.](#)

- 430 [Phys., 19, 8619–8650, https://doi.org/10.5194/acp-19-8619-2019, 2019.](https://doi.org/10.5194/acp-19-8619-2019)
- Cohan, D., Hu, Y., Hakami, A., Odman, M. T., Russell, A.: Implementation of a direct sensitivity method into CMAQ, Models-3 User's Workshop, RTP, North Carolina, October 22, 2002.
- Corazza, M., Bergamaschi, P., Vermeulen, A. T., Aalto, T.: Inverse modelling of European N₂O emissions: Assimilating observations from different networks, *Atmos. Chem. Phys.*, 11 (5), 2381-2398, 2011.
- 435 [Dai, T., Cheng, Y. M., Goto, D., Li, Y. R., Tang, X., Shi, G. Y., 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.](https://doi.org/10.5194/acp-21-4357-2021)
- Dennis, R., Byun, D., Novak, J.: The next generation of integrated air quality modeling: EPA's Models-3, *Atmos. Environ.*, 30, 1925-1938, 1996.
- 440 Elbern, H., Schmidt, H., Talagrand, O., Ebel, A.: 4D-variational data assimilation with an adjoint air quality model for emission analysis, *Environmental Modelling & Software*, 15 (6), 539-548, 2000.
- Elbern, H., Strunk, A., Schmidt, H., Talagrand, O.: Emission rate and chemical state estimation by 4-dimensional variational inversion, *Atmos. Chem. Phys.*, 7, 3749-3769, 2007.
- 445 Enting, I. G.: *Inverse Problems in Atmospheric Constituent Transport*, Cambridge: Cambridge University Press, 2002.
- ~~Fu, T. M., Jacob, D. J., Palmer, P. I., Chance, K., Wang, Y. X., Barletta, B., Blake, D. R., Stanton, J. C., Pilling, M. J.: Space-based formaldehyde measurements as constraints on volatile organic compound emissions in east and south Asia and implications for ozone, *J. Geophys. Res.*, 112, D06312, doi:10.1029/2006JD007853, 2007.~~
- 450 [Frey, H.C., Zheng, J., Zhao, Y., Li, S., Zhu, Y.: Technical Documentation of the AuvTool Software for Analysis of Variability and Uncertainty, Prepared by North Carolina State University for the Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC. February 2002](#)
- 455 Gilliland, A. B., Dennis, R. L., Roselle, S. J., Pierce, T. E.: Seasonal NH₃ emission estimates for the eastern United States based on ammonium wet concentrations and an inverse modeling method, *J. Geophys. Res.*, 2003, 108(D15), 4477, doi:10.1029/2002JD003063, 2003.

Gilliland, A. B., Wyat, A. K., Pinder, R. W., Dennis, R. L.: Seasonal NH₃ emissions for the continental united states: Inverse model estimation and evaluation, *Atmos. Environ.*, 40 (26), 4986-4998, 2006.

460 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., Nehrkorn, T.: Comparative inverse analysis of satellite (MOPITT) and aircraft (TRACE-P) observations to estimate Asian sources of carbon monoxide, *J. Geophys. Res.*, 2004, 109, D23306, doi:10.1029/2004JD005185, 2004.

Henze, D. K., Seinfeld, J. H., Shindell, D. T.: 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, ~~2008~~2009.

Hong, C., Zhang, Q., He, K., Guan, D., Li, M., Liu, F., Zheng, B.: Variations of China's emission estimates: response to uncertainties in energy statistics, *Atmos. Chem. Phys.*, 17, 1227-1239, 2017.

Hu, Y., Odman, M.T., 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, doi: 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., Bruns, E. A., Crippa, M., Ciarelli, G., Piazzalunga, A., Schwikowski, M., Abbaszade, G., Schnelle-Kreis, J., Zimmermann, R., An, Z., Szidat, S., Baltensperger, U., Haddad, I. E., 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., Heald, C.: Quantifying the impact of model errors on top-down estimates of carbon monoxide emissions using satellite observations, *J. Geophys. Res.*, 116, D15306, 2011.

Kerry, A., Gerhard, R., Mike, D. F.: Fire-growth modelling using meteorological data with random and systematic perturbations, *International Journal of Wildland Fire*, 16(2), 174-182, 2007.

Koohkan, M. R., Bocquet, M., Roustan, Y., Kim, Y., Seigneur, C.: Estimation of volatile organic compound emissions for Europe using data assimilation, *Atmos. Chem. Phys.*, 13 (12), 5887-5905, 2013.

- Kopacz, M., Jacob, D.J., Henze, D.K., Heald, C. L., Streets, D. G., 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, 2009.
- Li, M., Liu, H., Geng, G., Hong, C., Liu, F., Song, Y., Tong, D., Zheng, B., Cui, H., Man, H., Zhang, Q.,
490 He, K.: Anthropogenic emission inventories in China: a review, *National Science Review*, 4, 834-866, 2017.
- Li, Z., Zang, Z., Li, Q.B., Chao, Y., Chen, D., Ye, Z., Liu, Y., Liou, K.N.: A three dimensional variational data assimilation system for multiple aerosol species with WRF/Chem and an application to PM2.5 prediction, *Atmos. Chem. Phys.*, 13, 4265-4278, 2013.
- 495 Liu, F., Hu, F., Zhu, J.: Solving the optimal layout problem of multiple industrial pollution sources using the adjoint method, *Science China: Earth Sciences*, 35 (1), 64-71, 2005.
- Liu, J., Mauzerallb, 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, *Proc. Natl. Acad. Sci. USA*,
500 113 (28), 7756-7761, 2016.
- Liu, Z., Liu, Q., Lin, H. C., Schwartz, C.S., Lee, Y. H., 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, 2011.
- Ma, J., Aardenne, J. A. V.: Impact of different emission inventories on simulated tropospheric ozone over
505 China: a regional chemical transport model evaluation, *Atmos. Chem. Phys.*, 4, 877-887, 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, doi:10.1029/2010JD014763, 2011.
- Martin, R.V., Jacob, D.J., Chance, K., Palmer, P.I., Evans, M.J.: Global inventory of nitrogen oxide
510 emissions constrained by space-based observations of NO2 columns, *J. Geophys. Res.*, 108, 4537, 2003.
- ,Michalakes, J., J. Dudhia, D. Gill, T. Henderson, J. Klemp, W. Skamarock, W. Wang.: The Weather Reseach and Forecast Model: Software Architecture and Performance, *Proceedings of the 11th*

- ECMWF Workshop on the Use of High Performance Computing In Meteorology, 25-29 October
515 2004, Reading U.K. Ed. George Mozdzynski, 2004.
- Mijling, B., 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 (D17), D17302,
doi:10.1029/2012JD017817, 2012.
- Miyazaki, K., Eskes, H. J., Sudo, K.: Global NO_x emission estimates derived from an assimilation of
520 OMI tropospheric NO₂ columns, *Atmos. Chem. Phys.*, 12, 2263-2288, 2012.
- Napelenok, S. L., Pinder, R. W., Gilliland, A. B., 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, 2008.
- Otte, T. L., Pleim, J. E.: The Meteorology-Chemistry Interface Processor (MCIP) for the CMAQ
525 modeling system: updates through MCIPv3.4.1, *Geosci. Model Dev.*, 3, 243-256, 2010.
- Pagowski, M., Grell, G.A., McKeen, S.A., Peckham, S.E., 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. R. Meteorol.
Soc.*, 136, 2013-2024, 2010.
- 530 Parrish, D.F., Derber, J.C.: The National Meteorological Center's spectral statistical interpolation
analysis scheme, *Mon. Weather Rev.*, 120, 1747-1763, 1992.
- Peng, Z., Liu, Z., Chen, D., Ban, J. M.: 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, 2017.
- 535 [Peng, L., Zhang, Q., Yao, Z., Mauzerall, D. L., Kang, S., Du, Z., Zheng, Y., Xue, T., 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.](#)
- Sportisse, B.: A review of current issues in air pollution modeling and simulation, *Computational
Geosciences*, 11 (2), 159-181, 2007.
- 540 Stavrou, T., Müller, J. F., De, S. I., Roozendaal, V.: The continental source of glyoxal estimated by the
synergistic use of spaceborne measurements and inverse modeling, *Atmos. Chem. Phys.*, 9 (21),

8431-8446, 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., Yarber, K.F.: An inventory of gaseous and primary aerosol emissions in Asia in the year 2000, *J. Geophys. Res.*, 108(D21), 8809-8820, 2003.

Tang X, Zhu J, Wang Z F, 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 (24), 12901-12916, 2011.

Tang, X., Zhu, J., Wang, Z. F., Gbaguidi, A., Lin, C. Y., Xin, J. Y., Song, T., 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 (10), 6395-6405, 2016.

Tang, X.Y., Zhang, Y. H., Shao, M.: *Atmospheric Environmental Chemistry* (second edition) (Chin. Ver.). Beijing: Higher Education Press, 447-449, 2006.

The University of North Carolina.: SMOKE v3.6 user's manual, Chapel Hill: The institute for the Environment, 1-520, 2014.

Wang, C., An, X. Q., Zhai, S. X., Hou, Q., 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., Wang, H. J.: Application of CMAQ-DDM-3D in the source analysis of fine particulate matter (PM_{2.5}), *Acta. Scientiae Circumstantiae* (Chin. Ver.), 33(5), 1355-1361, 2013.

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., Cheng, Y. L.: Inverse modeling of black carbon emissions over China using ensemble data assimilation, *Atmos. Chem. Phys.*, 16 (2), 989-1002, 2016.

Wang, Y. S., Yao, L., Wang, L., L., Liu, Z. R., Ji, D. S., Tang, G. Q., Zhang, J. K., Sun, Y., Hu, B., Xin, J. Y.: Mechanism for the formation of the January 2013 heavy haze pollution episode over central and eastern China, *Science China. Earth Sciences* (Chin. Ver.), 57, 14-25, 2014.

Wang, Y., McElroy, M. B., Martin, R. V., Streets, D. G., Zhang, Q., 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, doi,10.1029/2006JD007538, 2007.

- 570 Willmott, C. J., Robeson, S. M. and Matsuura, K. A.: A refined index of model performance, International Journal of Climatology, 32, 2088-2094, 2012.
- Xu, X. D., Xie, L. A., Cheng, X. H., Xu., J. M., Zhou, X. J., Ding, G. A.: Application of an Adaptive Nudging Scheme in Air Quality Forecasting in China, J. Appl. Meteorol. Climatol., 47, 2105-2114, 2008.
- 575 Yang, Q., Wang, Y. H., Zhao, C., Liu, Z., William, I., Gustafson, J., Shao, M.: NOX emission reduction and its effects on ozone during the 2008 Olympic Games, Environ. Sci. Technol., 45 (15), 6404-6410, 2011.
- Zang, Z. L., Li, Z.J., Pan, X.B., Hao, Z.L., You, W.: Aerosol data assimilation and forecasting experiments using aircraft and surface observations during CalNex, Tellus B, 68, 29812,2016.
- 580 ~~Zhai, S. X., An, X. Q., Zhao, T. L., Sun, Z. B., Wang, W., Hou, Q., Guo, Z. Y., Wang, C.: Detection of critical PM2.5 emission sources and their contributions to a heavy haze episode in Beijing, China, using an adjoint model, Atmos. Chem. Phys., 18, 6241–6258,2018.~~
- Zhang, L., Shao, J. Y., Lu, X., Zhao, Y. H., Hu, Y. Y., Henze,D., Liao, H., Gong, S., L., Zhang, Q.: Sources and processes affecting fine particulate matter pollution over North China: An adjoint
- 585 analysis of the Beijing APEC period, Environ. Sci. Technol., 50 (16), 8731-8740, 2016.
- Zhang, Q., G. N. Geng, S. W. Wang, S.W., Andreas, R., HE, K. B.: Satellite remote sensing of changes in NOx 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., Yao Z. L.: Asian emissions in 2006
- 590 for the NASA INTEX-B mission, Atmos. Chem. Phys., 9, 5131-5153, 2009.
- Zhao, B., Wang, P., Ma, J. Z., Zhu, S., Pozzer, A., Li, W.: A high-resolution emission inventory of primary pollutants for the Huabei region, China, Atmos. Chem. Phys., 12, 481-501, 2012.
- Zhao, H. Y., Zhang, Q., Guan, D. B., Davis, S. J., Liu, Z., Huo, H., Lin, J. T., Liu, W. D., He, K. B.: Assessment of China's virtual air pollution transport embodied in trade by using a consumption-based
- 595 emission inventory, Atmos. Chem. Phys., 15(10), 5443-5456, 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., Zhang, Q.: Trends in China's anthropogenic emissions since 2010

[as the consequence of clean air actions, Atmos. Chem. Phys., 18, 14095-14111, 2018.](#)

Zhu, J., Tang, X., Wang, Z. F., Wu, L.:A review of air quality data assimilation methods and their
application, Chinese Journal of Atmospheric Sciences, 42 (3), 607-620, 2018.

Zhu, J., 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(5),
871-882, 2006.

|

640

Table 1. Uncertainty of NO_x and SO₂ 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%)

645

Table 2. Statistics for simulated SO₂, NO₂, and O₃ from control and experiment runs using ~~the~~a*priori* and a*posteriori* inversed ES at 45 stations in the BTH region during December 28–30, 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

650

655

660

665

670

675

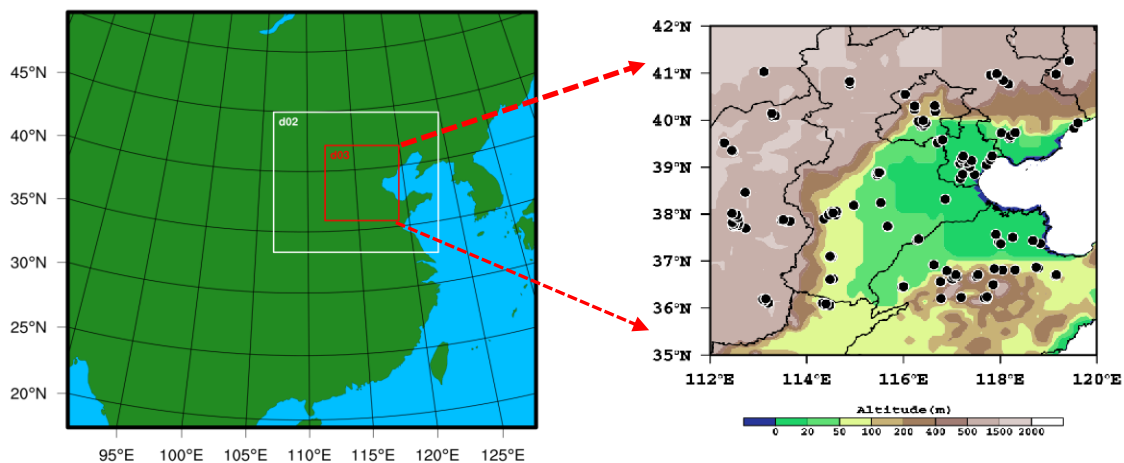


Fig. 1 (a) Domain of the WRF-CMAQ model and (b) location of environmental monitoring stations in the innermost domain over the BTH region.

680

685

690

695

700

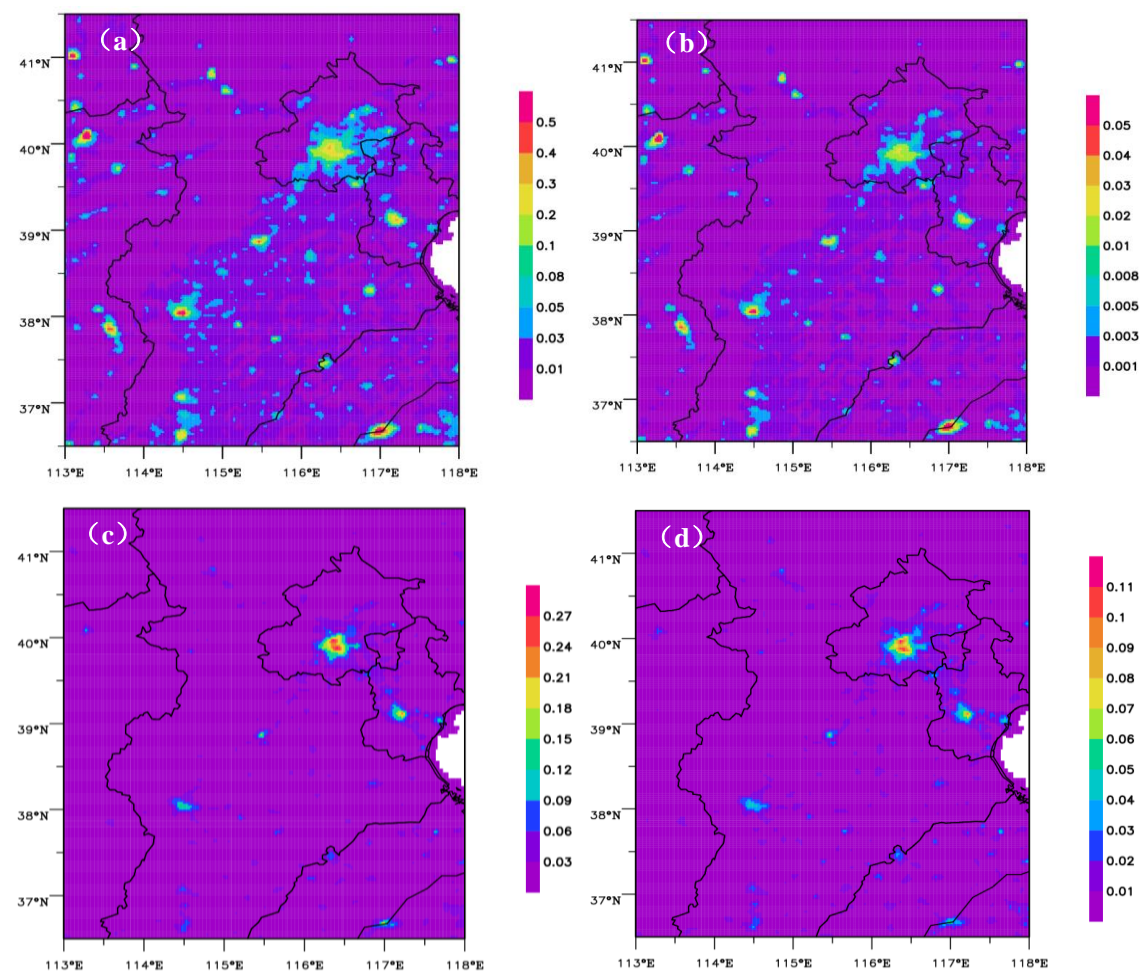


Fig. 2 Spatial distributions of (a) averaged emission rates of SO₂ ES, (b) standard deviation in the BEC of SO₂ ES, (c) averaged emission rates of NO_x ES, (d) standard deviation in the BEC of NO_x ES at 08:00 local time in December 2012. Unit: mole/s.

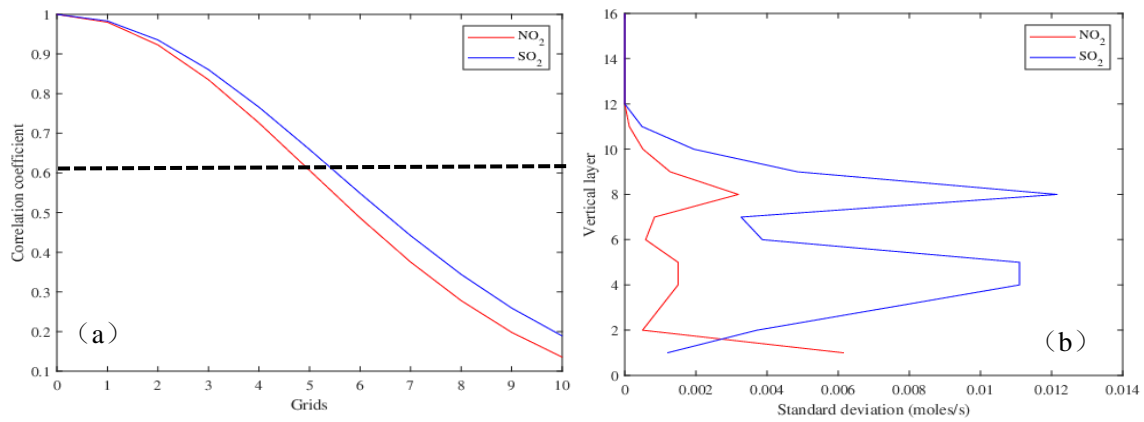


Fig. 3 (a) Horizontal correlation coefficients with increasing grid distance, and (b) vertical profiles of standard deviations in the BEC of SO₂ and NO₂ ES in December 2012. Dashed line is the baseline of horizontal correlation scale.

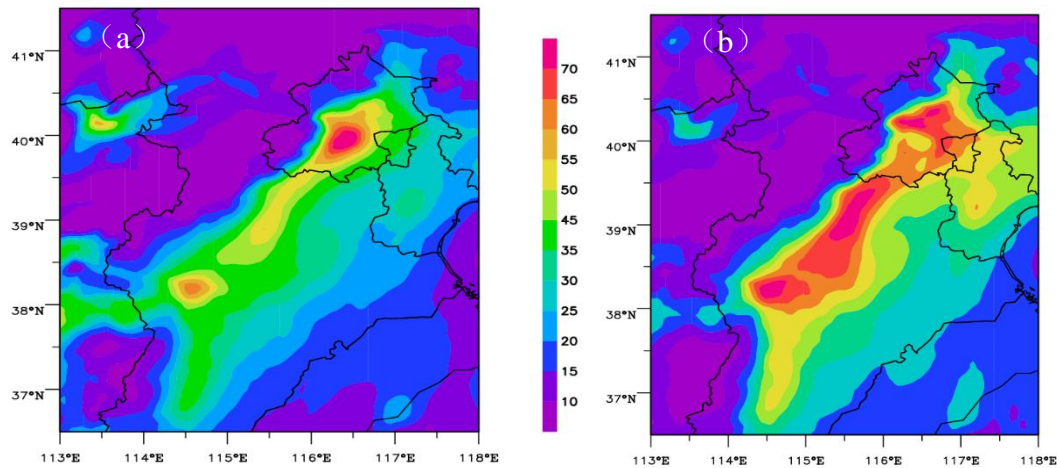


Fig. 4 Spatial distributions of 96-h averaged sensitivity coefficients ($\mu\text{g m}^{-3}$) of (a) SO₂ and (b) NO_x concentrations with respect to SO₂ and NO_x ES during December 27–30, 2016.

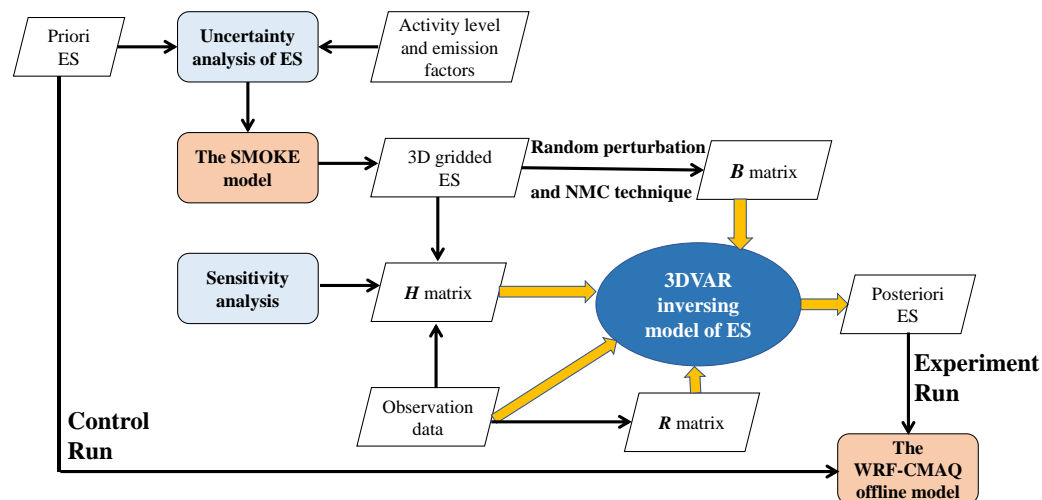


Fig. 5 Flowchart of the 3DVAR inversion model of ES and simulation experiments.

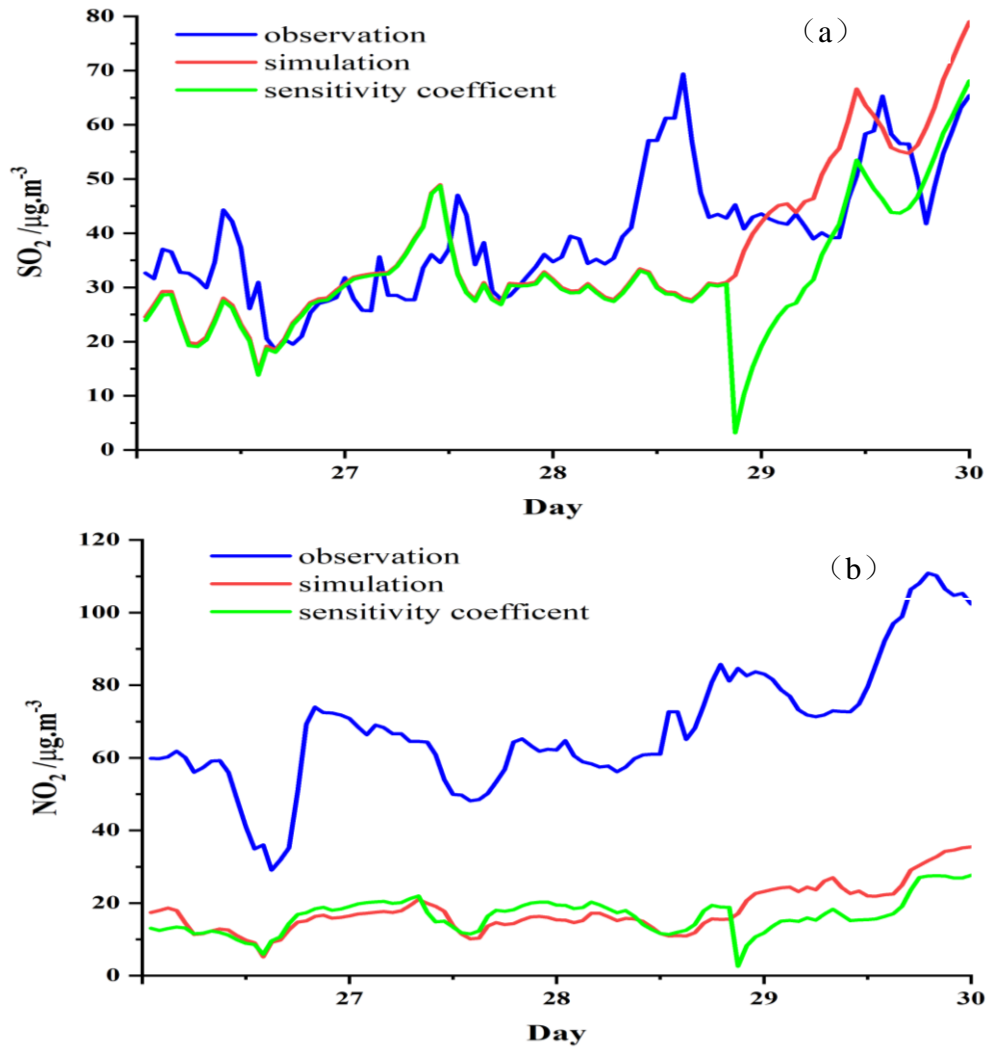


Fig. 6 Time series of hourly, regionally-averaged (a) SO_2 and (b) NO_2 simulations with the *a priori* ES, observations, and the first-order sensitivity coefficients between the ES and receptor's concentration at 45 stations over the BTH region during December 27-30, 2016.

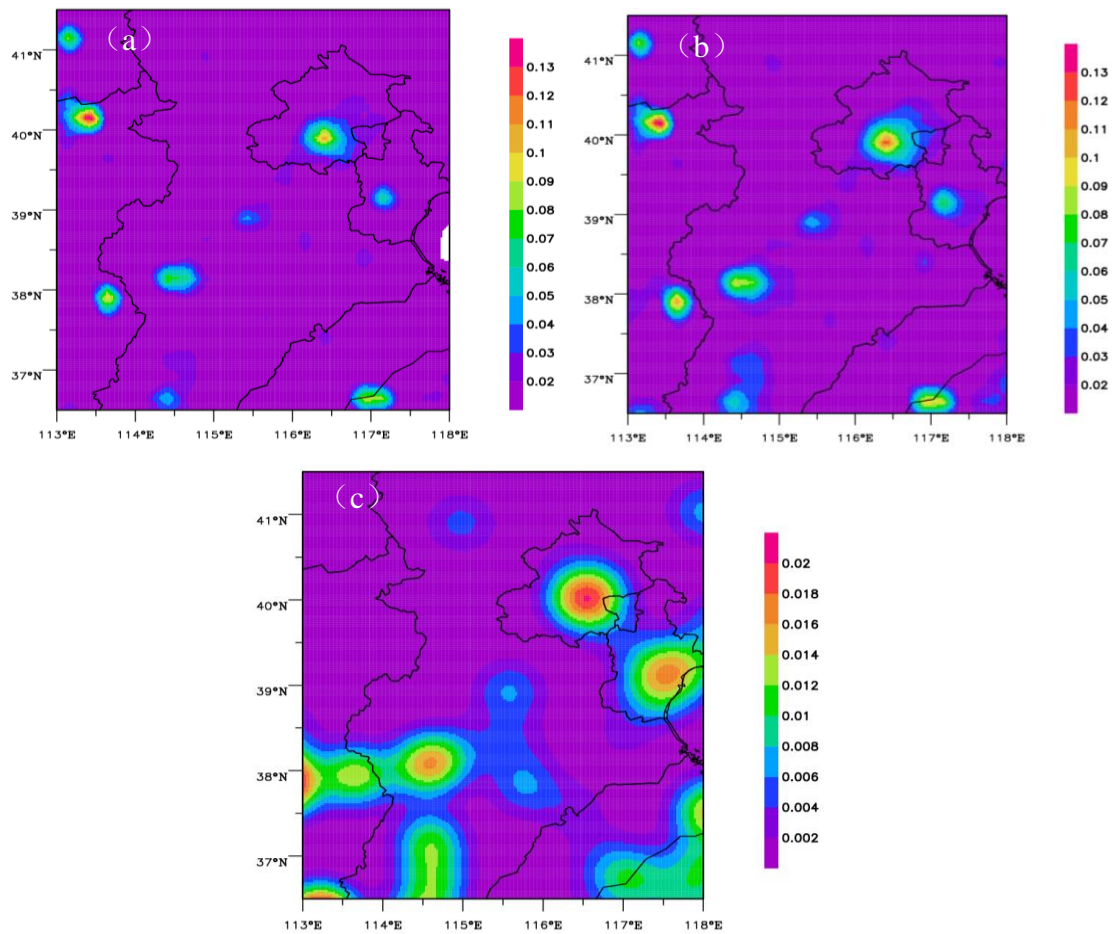
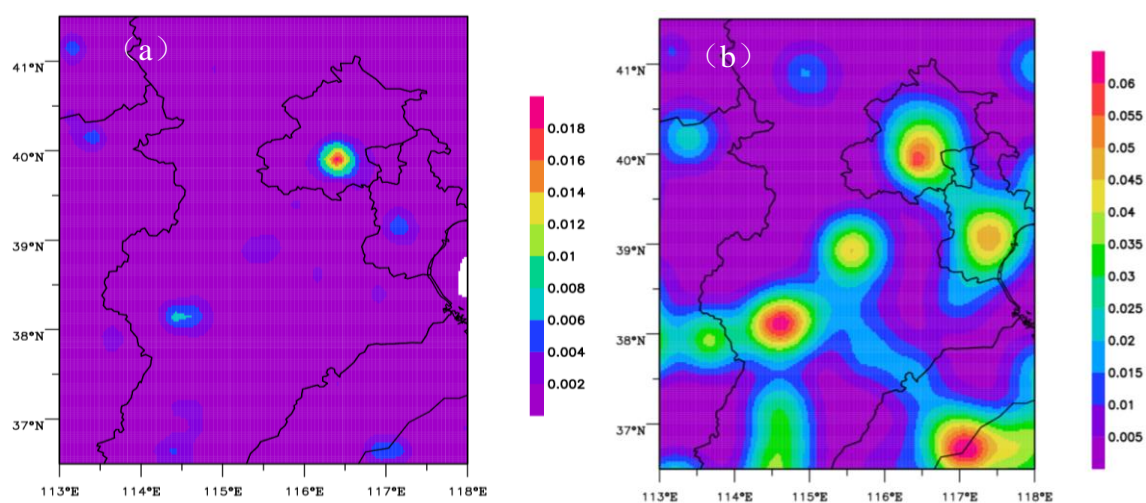


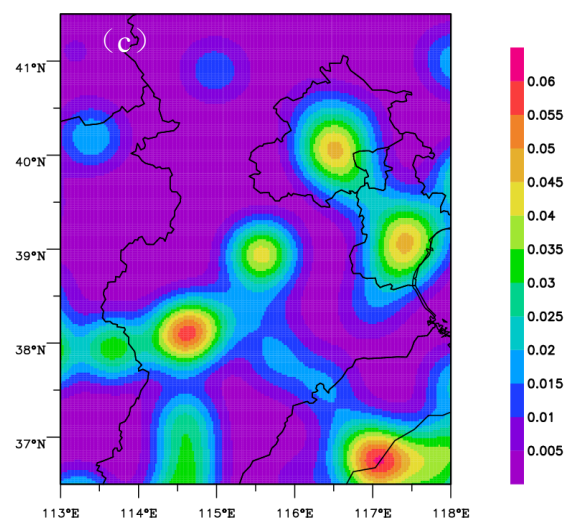
Fig. 7 Spatial distributions of 24-h averaged emission rates for SO_2 (mole/s) from the (a) *a priori* and (b) *a posteriori* ES, and (c) the increment on December 27, 2016.

815

820



825



830 **Fig. 8** Same to Fig.7 except for NO_2 .

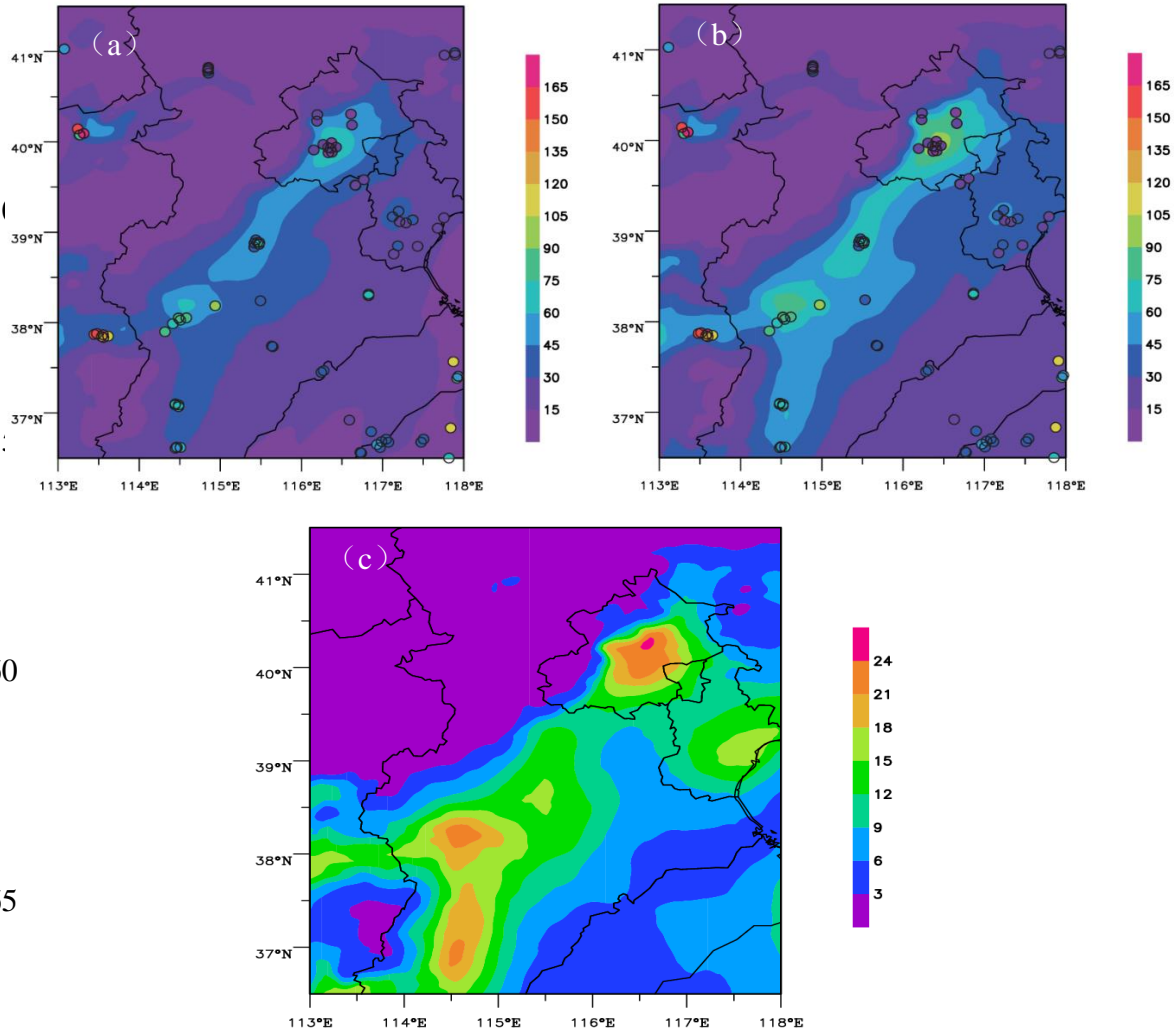
835

840

845

850

855



860

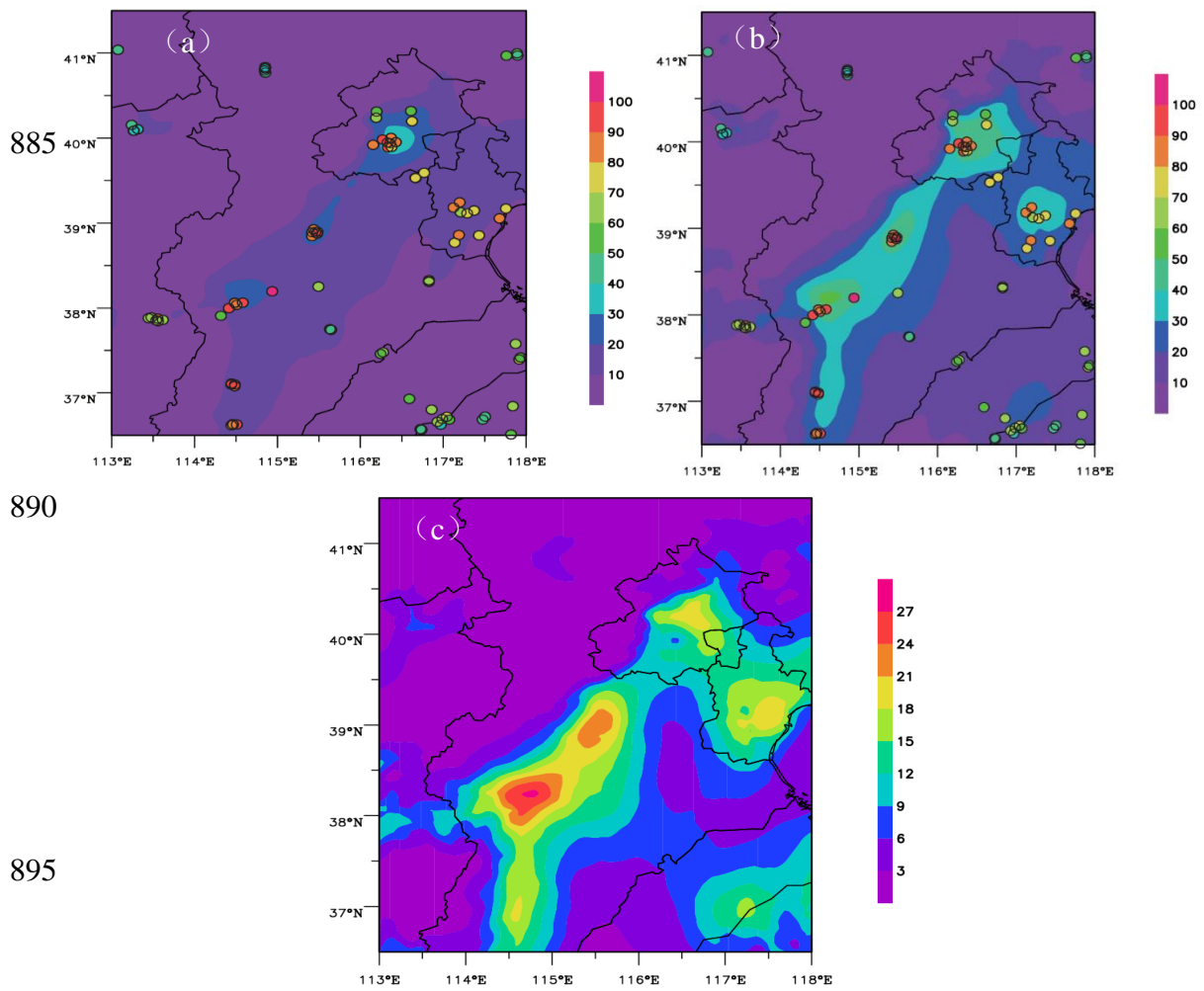
865

870

Fig. 9 Spatial distribution of 72-h averaged SO₂ concentration simulated with the (a) *a priori* and (b) *a posteriori* ES, and (c) the increment during December 28-30, 2016. Color solid dot denotes the measurements.

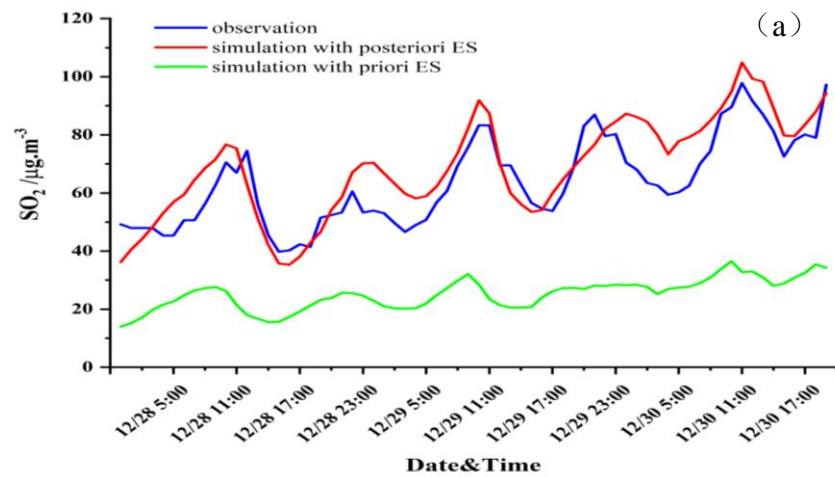
875

880

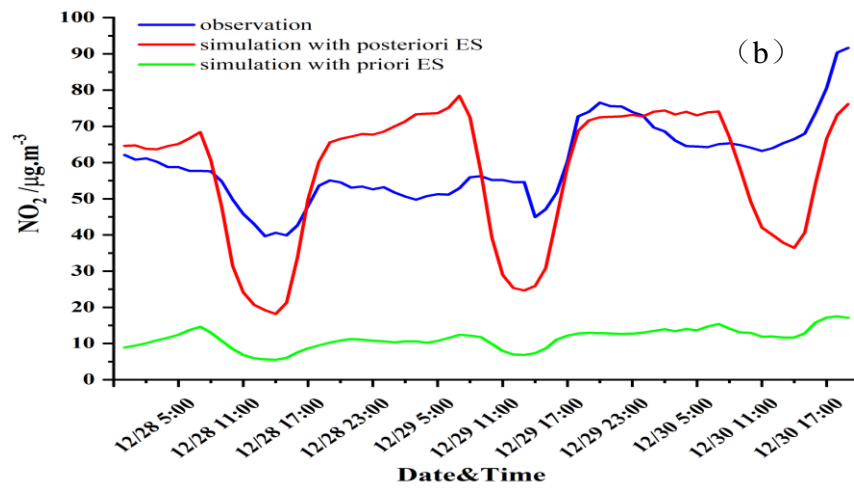


910

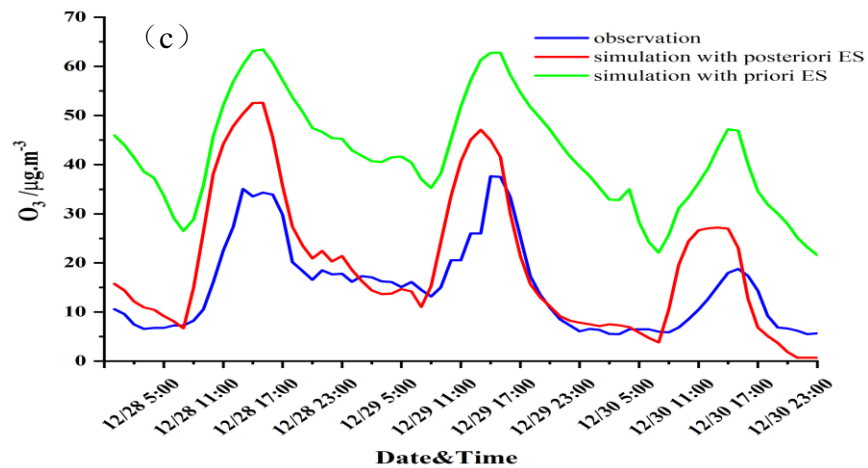
915



920



925



930

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

