



Development of 3D Variational Assimilation System Based on GRAPES-CUACE Adjoint Model (GRAPES-CUACE-3D-Var V1.0) and Its Application in Emission Inversion

Chao Wang^{1,2}, Xingqin An¹, Qing Hou¹, Zhaobin Sun³, Yanjun Li¹, Jiangtao Li¹

5 ¹ Institute of Atmospheric Composition, Chinese Academy of Meteorological Sciences, Beijing 100081

² Department of Atmospheric and Oceanic Sciences, Fudan University, Shanghai 200438

³ Institute of Urban Meteorology, China Meteorological Administration, Beijing 100089

Correspondence to: Xingqin An (anxq@cma.gov.cn)

Abstract. The adjoint method is known for its efficient calculation of sensitive information. After decades of development, assimilation technology based on adjoint method has gradually become an important tool for emission inversion. On the basis of GRAPES-CUACE aerosol adjoint model, and combined with the optimization algorithm and pollutants observations, the GRAPES-CUACE 3D variational (GRAPES-CUACE-3D-Var) assimilation system was further developed, and was used in the inversion of BC emissions in Beijing-Tianjin-Hebei region. The results show that the newly constructed GRAPES-CUACE-3D-Var assimilation system is reasonable and reliable, and can be applied to the emission inversion in Beijing-Tianjin-Hebei region. Compared to the simulations using the a priori BC emissions, the model simulations driven by the a posterior BC emissions in two inversion schemes are in better agreement with measurements. The correlation coefficient between the simulations and the observations is increased from 0.2 before the inversion to 0.7 and 0.64, respectively, and the NMSE is reduced from 0.38 to 0.22 and 0.24, respectively, and the NMB is decreased from 51.53% to 43.37% and 40.90%, respectively, in the two inversion schemes. The spatial distributions of the a posterior BC emissions in the two inversion schemes are consistent with the distributions of the a priori BC emissions. The high-value areas are mainly located in the south of Beijing, Tianjin, central and southern Hebei, and northern Shandong. On the whole, the inversion scheme with a large observation ratio has better optimization effect. The observation information of the target time has a great influence on the a posterior BC emissions in a short period before the target time, and the influence decreases with the reverse time sequence.

25 1 Introduction

Adjoint method is an efficient sensitivity analysis tool developed based on adjoint operator theory and numerical model. Based on the idea of inverse simulation, by establishing the adjoint model of atmospheric chemistry model, the sensitivity of the objective function to any control variable (such as emission source) at any time and any spatial position can be obtained only by running the adjoint model once in inverse time sequence, and then various optimization control problems can be



30 solved quickly by using sensitivity information (Errico, 1997; Sandu et al., 2005; Liu, 2005; Carmichael et al., 2008; An et al., 2016). In the 1970s, Marchuk and Skiba (1976) first applied adjoint method to the field of atmospheric environment to optimize pollution sources location. With the development and popularization of adjoint methods in the later period, atmospheric chemical variational assimilation technology with adjoint model and optimization algorithm as its core has been widely applied to atmospheric pollutant emission inversion (Liu, 2005; Henze et al., 2009; Mao et al., 2015; Cao et al., 2018; Huang et al., 2018).

The essence of source inversion using adjoint method is to solve the nonlinear optimization problem constrained by complex partial differential equations. Firstly, the objective function is established based on prior sources and observation data, and the gradient (sensitivity) of the objective function with respect to the emission source parameters is calculated using adjoint model. Then, the emission source parameters are updated according to the optimization algorithm, and the optimal solution of the objective function is obtained through successive iterations, thus obtaining the optimized emission source parameters (Liu, 2005). Adjoint inversion method can make up for the shortcomings of "bottom-up" and other conventional pollution source investigation methods, realize the reverse optimization and rapid update of the active inventory, and reduce the uncertainty of the source inventory. It is of great significance to improve the simulation accuracy of air pollution and formulate scientific and effective air quality management measures (Zeng and Wu, 2018; Zhu et al., 2018; Huang et al., 2018).

At present, atmospheric chemical models such as STEM(Sandu et al., 2005), CAMx(Liu, 2005), CMAQ(Hakami et al., 2007) and GEOS-Chem(Henze et al., 2007), which are widely used in the world, have successively developed corresponding adjoint models. Through forward model, adjoint model and optimization algorithm, a complete atmospheric chemical variational assimilation system has been constructed. Foreign scholars have carried out extensive research on the inversion of pollution sources. The types of pollution sources retrieved include BC, CO, NO_x, dust and PM_{2.5} precursors, etc. The optimized source list covers the spatio-temporal scales from the region to the world and from the monthly average to the daily average (Hakami et al., 2005; Müller and Stavrou, 2005; Pan et al., 2007; Kurokawa et al., 2009; Henze et al., 2009; Wang et al., 2012; Mao et al., 2015; Zhang et al., 2015; Jeong and Park, 2018). Hakami et al. (2005) used the adjoint model of atmospheric chemical transport model STEM-2k1, based on aerosol observation data in Asia Pacific region, combined with optimization algorithm, to inverse artificial sources of BC aerosol in East Asia. The results show that the simulation results based on posterior sources are closer to the observation data. Müller and Stavrou(2005) optimized the global CO and NO_x emissions from man-made and natural sources in 1997 based on IMAGES adjoint model, using ground and satellite observation data of CO and NO₂.Kurokawa et al. (2009) developed a four-dimensional variational assimilation system RC4-NO_x, which is specially used to optimize NO_x emissions, and used this system to retrieve NO_x emissions from eastern China in 1996, 1999 and 2002. Wang et al. (2012) combined MODIS satellite data and used GEOS-Chem adjoint model to inverse the dust source flux in Taklimakan desert. Mao et al. (2015) discussed the effect of model resolution on BC inversion, and believed that the finer the model resolution, the smaller the inversion error. Zhang et al. (2015) shows that there are obvious differences in the distribution of posterior sources based on inversion of different prior source inventories.



Jin (2001) used the accompanying model of the three-dimensional advection model to inverse the SO₂ emission source in
65 Lanzhou. Zhang et al. (2016) optimized the primary source emission of PM_{2.5} in north China during APEC in 2014 on a
daily average time scale based on GEOS-Chem assimilation system. Huang et al. (2018) updated monthly emission source
inventory based on CAMx adjoint model. Cao et al. (2018) used GEOS-Chem adjoint model to inverse the monthly
NMVOC emissions from human activities and biomass combustion in China in 2007.

In recent years, An et al. (2016) have constructed the aerosol adjoint model of atmospheric chemistry model GRAPES-
70 CUACE, and applied it to source tracking and source sensitivity analysis of PM_{2.5} pollution events (Wang et al., 2017; Zhai
et al., 2018; Wang et al., 2018a, 2018b, 2019). Under the current severe pollution situation in the central and eastern regions
and Beijing-Tianjin-Hebei region (Wang et al., 2014a, 2014b; Zhang et al., 2019), it is urgent to develop and construct an
atmospheric chemical variational assimilation system and use it to retrieve pollution source parameters to obtain more
accurate spatial and temporal distribution of pollution sources, improve the accuracy of air pollution simulation, and reduce
75 the uncertainty of control schemes based on simulation results. Therefore, this study further develops the advantages of
GRAPES-CUACE aerosol adjoint model. Based on the "source-concentration" sensitivity (gradient) information provided by
the adjoint model, combined with the optimization algorithm and concentration observation data, a GRAPES-CUACE three-
dimensional variational assimilation system (GRAPES-CUACE-3D-Var) is developed and constructed, and the BC emission
sources in Beijing-Tianjin-Hebei region are inversed using this system.

80 **2 Data and Methods**

2.1 GRAPES-CUACE model

GRAPES-CUACE is an atmospheric chemical model system independently developed by Chinese scientists (Wang et al.,
2010, 2015). It is the combination of the atmospheric chemistry model cuace (cma unified atmospheric chemistry
environmental forecasting system) (Gong and Zhang, 2008) and the meteorological model GRAPES (global-regional
85 assistance and prediction system) (Chen, 2008). GRAPES-CUACE model realizes the on-line coupling of meteorology and
chemistry, using the GRAPES regional mesoscale model GRAPES-Meso as the driving force of meteorological field.
CUACE mainly includes three modules: aerosol module, gas module and thermodynamic equilibrium module. The aerosol
module involves six aerosol components: sulfate, nitrate, sea salt, black carbon, organic carbon and dust. The particle size is
divided into 12 particle size segments: 0.01-0.02 μm , 0.02-0.04 μm , 0.04-0.08 μm , 0.08-0.16 μm , 0.16-0.32 μm , 0.32-0.64
90 μm , 0.64-1.28 μm , 1.28-2.56 μm , 2.56-5.12 μm , 5.12-10.24 μm , 10.24-20.48 μm and 20.48-40.96 μm according to the
observation results of the particle size distribution of desert soil in China using a multi-component multiphase aerosol
particle size separation algorithm (Wang et al., 2010). The core part of the aerosol module is the aerosol physical and
chemical process, which mainly includes detailed aerosol processes such as aerosol source, transportation, moisture
absorption growth, collision, nucleation, condensation, dry and wet deposition, cloud and sub-cloud removal, and aerosol-



95 cloud interaction, etc (Gong and Zhang, 2008; Wang et al., 2010). In addition, the secondary formation process of natural sources, anthropogenic sources and aerosols has been added to the module (Gong and Zhang, 2008).

2.2 Adjoint theory and methods

Let \mathbf{L} be a continuous linear operator defined on Hilbert space, if there is a continuous linear operator \mathbf{L}^* satisfying:

$$\forall \mathbf{x}, \mathbf{y} \in \mathbf{H}, (\mathbf{L}\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{L}^*\mathbf{y}) \quad (1)$$

100 Then \mathbf{L}^* is called the adjoint operator of \mathbf{L} (Ye and Shen, 2006). Where (\cdot, \cdot) denotes inner product. If \mathbf{x}, \mathbf{y} are continuous functions on domain Ω , the inner product is defined as $(\mathbf{x}, \mathbf{y}) = \int_{\Omega} \mathbf{x} \cdot \mathbf{y} d\Omega$; if \mathbf{x}, \mathbf{y} is a discrete vector: $\mathbf{x} = [x_1, x_2, \dots, x_N]$, $\mathbf{y} = [y_1, y_2, \dots, y_N]$, then there are $(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^N x_i \cdot y_i$. When \mathbf{x}, \mathbf{y} is a vector and \mathbf{L} is a matrix independent of \mathbf{x}, \mathbf{y} , we can get $(\mathbf{L}\mathbf{x}, \mathbf{y}) = \mathbf{y}^T \mathbf{L}\mathbf{x} = \mathbf{x}^T \mathbf{L}^T \mathbf{y} = (\mathbf{x}, \mathbf{L}^T \mathbf{y})$, in other words, for matrix linear operators, the adjoint operator is its transposition, i.e. $\mathbf{L}^* = \mathbf{L}^T$; If the component of the matrix is complex, the adjoint operator is the conjugation of its transpose (Liu, 2005).

105 The atmospheric chemistry model can be divided into several small programs, and each small program is abstracted into a vector function $\mathbf{F}: \mathbf{R}^n \rightarrow \mathbf{R}^m$, which can be expressed as:

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) \quad (2)$$

Where \mathbf{x}, \mathbf{y} are n-dimensional independent variables and m-dimensional dependent variables respectively, corresponding to input variables and output variables in the forward model. Assuming that \mathbf{F} is continuously differentiable, the tangent linear program of the original model can be expressed as:

$$d\mathbf{y} = \nabla_{\mathbf{x}} \mathbf{F} \cdot d\mathbf{x} \quad (3)$$

Where $\nabla_{\mathbf{x}} \mathbf{F}$ is the Jacobian matrix of the forward model operator. According to the property of adjoint operator (formula (1)), the transform tangent linear mode (3) can be obtained:

$$d\mathbf{x}^* = \nabla_{\mathbf{x}}^T \mathbf{F} \cdot d\mathbf{y}^* \quad (4)$$

115 Equation (4) is the adjoint model of the forward model (2). Where $d\mathbf{x}^*$ is n-dimensional and $d\mathbf{y}^*$ is m-dimensional. Comparing equations (3) and (4), it can be seen that tangent model and adjoint model exchange the dimensions of input and output, and the operators are transposed (Liu, 2005). The tangent linear model reflects the concentrated influence of the input disturbance on each output disturbance, while the adjoint model reflects the dependence of the output disturbance on each input disturbance (Cheng et al., 2004). It is not difficult to see that the gradient of the tangent mode calculation function needs to be run n times, and the adjoint mode needs to be run m times. When $n \gg m$ (n-dimensional emission source, pollutant concentration of m stations), the calculation efficiency of the adjoint mode is much higher than that of the tangent mode (Liu, 2005).

2.3 GRAPES-CUACE aerosol adjoint model

Based on the GRAPES-CUACE atmospheric chemistry model, An et al. (2016) constructed the adjoint model of GRAPES-
125 CUACE aerosol module, which mainly includes the adjoint of aerosol physical and chemistry process, the adjoint of flux



calculation process of six kinds of particles (sulfate, nitrate, sea salt, black carbon, organic carbon and sand dust), the adjoint of aerosol transmission process and the adjoint of GRAPES-Meso and CUACE interface program (Jin, 2012; Zhai, 2015; An et al., 2016). The constructed GRAPES-CUACE aerosol adjoint model has passed the correctness test (Jin, 2012; Zhai, 2015; An et al., 2016), and has been well applied in BC (Zhai, 2015; An et al., 2016) and PM_{2.5} source tracking (Wang et al., 2017; Zhai et al., 2018; Wang et al., 2018a, 2018b, 2019) in North China.

Fig. 1 shows the operation flow of the GRAPES-CUACE atmospheric chemistry model and its adjoint model, where J is the objective function, and different objective functions can be defined according to the concerned problems. c is the model state variable (such as BC concentration); s is the model control variable (such as emission sources, mainly including VOCs, NO_x, NH₃, SO₂ and PPM_{2.5}, etc.). Firstly, the atmospheric chemistry model of GRAPES-CUACE is integrated forward, and the basic-state values of each integration step are stored. Subsequently, the gradient $\nabla_c J$ of the objective function with respect to the state variable is calculated and used as an adjoint forcing term, the adjoint model is driven in inverse time sequence, and the corresponding ground state value information is read at each integration step, and finally the sensitivity $\nabla_s J$ of the objective function J with respect to the control variable s at any time and at any spatial position is obtained.

2.4 Gradient descent method

The Gradient descent method is the most classical method in the unconstrained optimization algorithm. It uses the negative gradient direction as the descending direction for iterative search, which has the advantages of low storage cost, simple and easy implementation (Song et al., 2012). For unconstrained optimization problems:

$$\min f(x), x \in R^n \quad (5)$$

Where $f(x)$ has a continuous first-order partial derivative and has a unique minimum value point. The calculation process of the gradient descent method is as follows: ① given the first guess value x_0 of variable x , setting tolerance $\varepsilon > 0$, making $k = 0$; ② calculate $\nabla f(x_k)$, if $\|\nabla f(x_k)\| < \varepsilon$, stop and output x_k ; otherwise, let $d_k = -\nabla f(x_k)$; ③ The step λ_k takes a fixed value or is obtained by one-dimensional search method, so that $x_{k+1} = x_k + \lambda_k d_k$, $k=k+1$, turn ②.

2.5 Source inversion principle based on adjoint model

The adjoint inversion method is essentially based on the Bayesian theory framework, that is, the prior distribution of variables is corrected by using observation information. The objective function of source inversion is generally expressed as:

$$J(\mathbf{x}) = \frac{1}{2}\gamma(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{F}(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1}(\mathbf{F}(\mathbf{x}) - \mathbf{y}) \quad (6)$$

Where \mathbf{x}_b and \mathbf{x} are the priori source and the posterior source vector defined in the n -dimensional control space \mathbf{E} , and \mathbf{y} is the observation vector defined in the m -dimensional observation space \mathbf{O} , and \mathbf{F} is the forward model operator (observation operator), and \mathbf{B} and \mathbf{R} are a priori error covariance matrix of $n \times n$ and an observation error covariance matrix of $m \times m$ respectively, and γ is a weighting factor, which is used to adjust the relative proportion of the observation term and the prior



term in the objective function so as to fully retain the observation data and effective information in the priori estimation in the posteriori estimation (Henze et al., 2009; Cao et al., 2018).

The gradient of the objective function J with respect to the variable \mathbf{x} is:

$$160 \quad \nabla J(\mathbf{x}) = \gamma \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \mathbf{F}^T \mathbf{R}^{-1}(\mathbf{H}(\mathbf{x}) - \mathbf{y}) \quad (7)$$

Where \mathbf{F}^T is the adjoint of the forward model. Based on the gradient calculated by the adjoint model, combined with the optimization algorithm, the gradient is used to iteratively update \mathbf{x} , so that the objective function $J(\mathbf{x})$ takes a minimum value, and \mathbf{x} at this time is the reasonable source distribution.

2.6 Data

165 Adequate observation data is the basis for pollution source inversion, and better inversion results can be obtained by using observation data with large space-time density (Müller and Stavrou, 2005; Jeong and Park, 2018). The observation sites of the Ministry of Environmental Protection in the Beijing-Tianjin-Hebei region are relatively dense, but the BC concentration is not involved in the monitored pollution factors, and the BC observation data is relatively scarce. Therefore, the BC concentration of 18 stations in Beijing-Tianjin-Hebei region is estimated as "BC observation concentration" in this study
170 based on the monthly average ratio of BC/PM_{2.5} and the PM_{2.5} concentration monitored by the Ministry of Environmental Protection (Fig. 2). The monthly average ratio of BC/PM_{2.5} in this study is 4.35% according to the proportion of BC/PM_{2.5} in December 2015 in Beijing's southern suburb site and the results of Zhang et al. (2017).

3 Construction of 3D variation assimilation system based on GRAPES-CUACE adjoint model

Based on the GRAPES-CUACE atmospheric chemistry model and its adjoint model, combined with gradient descent
175 method and observation data, the GRAPES-CUACE-3D-Var assimilation system (Fig. 3) was constructed, mainly including GRAPES-CUACE atmospheric chemical model calculation and ground state value storage, observation data and adjoint forcing term processing, adjoint model calculation, gradient extraction, target function calculation and model parameter optimization iteration, etc. Taking the source inversion problem as an example, it is expected to minimize the difference between the simulated concentration and the observed concentration at 18 stations in Beijing-Tianjin-Hebei region (Fig. 2).
180 Firstly, the emission source parameter is defined as independent variable \mathbf{x} , and the objective function is established; Operating the GRAPES-CUACE atmospheric chemical model to t_p time and storing the ground state value of each integration step; Taking the difference between the simulated and observed values of 18 stations at t_p time as the forcing term, the adjoint model is driven in inverse time sequence, and the sensitivity of the objective function to the emission source parameters at each time in $[t_p, t_0]$ time period is obtained. Updating emission source parameters by using an optimization
185 module; Cycle the GRAPES-CUACE-3D-Var assimilation system until the objective function achieves a minimum value, i.e. a more reasonable posterior source distribution is obtained.



4 Inversion of BC emission source in Beijing-Tianjin-Hebei region based on GRAPES-CUACE-3D-Var assimilation system

4.1 Test design

190 The simulation scope of this study is in the central and eastern region of China (105°E-125°E, 32.25°N-43.25°N) (Fig. 2), and the horizontal grid resolution is 0.5°×0.5°, including 41×23 grids, vertically divided into 31 layers. The integration step is 300s, and the meteorological field uses FNL data at 6h intervals. The emission sources uses the INTEX-B2006 list of 0.5°×0.5° (Zhang et al., 2009), in which the anthropogenic emissions sources in North China have been updated with 2010 emission basic data and emission factors (Lv, 2015). The simulation period is from 20: 00 on November 27 to 19: 00 on
195 December 1, 2015 (BJT) and the first 3 days serve as the start-up period for meteorological and chemical transmission simulation to eliminate the difference between the idealized initial concentration and the actual concentration. The inversion area is the Beijing-Tianjin-Hebei region (113°E-119.5°E, 35.25°N-42.25°N) (black dotted line in Figure 2), and the a priori source is the BC emission source in the Beijing-Tianjin-Hebei region. The target time (t_p) is defined as 07:00 on December 1, 2015 (BJT). By inverting the hourly BC emission sources from 20: 00 on November 30 to 07: 00 on December 1 in Beijing-
200 Tianjin-Hebei region, the simulated concentration of 18 stations in Beijing-Tianjin-Hebei region (Figure 2) at the target time is closer to the observed concentration.

Due to the large difference between the prior sources and the emissions in 2015 (Zheng et al., 2018), this study mainly considers to fully retain the effective information of observation data in the posterior estimation, and defines the objective function as follows:

$$205 \quad J(x) = \frac{1}{2} \sum_{i=1}^{18} (C_{sim,i} - C_{obs,i})^T R^{-1} (C_{sim,i} - C_{obs,i}) \quad (8)$$

Among them, $C_{sim,i}$ and $C_{obs,i}$ are respectively the simulated concentration and observed concentration of BC at each station at the target time (07:00 on December 1, 2015). It is difficult to quantify the covariance matrix R of observation errors, which includes observation instrument errors, physical and chemical process errors involved in numerical model simulation and simulation accuracy errors (Cao et al., 2018). Considering that the calculation of BC concentration by proportional
210 method may cause large errors, so the observation error covariance matrix R takes 500%.

In the source inversion problem, directly taking the emission source intensity as the inverse evolution quantity may lead to negative emission source intensity, which is inconsistent with the actual situation. Therefore, the emission source factor is usually taken as the independent variable for inversion (Mao et al., 2015). Henze et al. (2009) and Jiang et al. (2015) used the logarithmic method to define the emission source factor x :

$$215 \quad x = \ln \left(\frac{s}{s_b} \right) \quad (9)$$

Where s_b is the a priori source emission intensity and s is the source variable that needs to be inverted. In fact, Equation (9) contains the influence of a priori source information on the objective function. In order to compare the influence of the proportion of prior information in (9) on the inversion results, this study designed a set of comparative experiments: $x =$



0.1ln (s/s_b) and $x = 0.01\ln (s/s_b)$, respectively recorded as C_10 and C_100. In the C_10 and C_100 tests, the gradient of
220 the objective function J with respect to the emission factor x is:

$$C_{10}: \nabla_x J = \nabla_s J \cdot s_b \cdot e^{10x} \cdot 10 \quad (10)$$

$$C_{100}: \nabla_x J = \nabla_s J \cdot s_b \cdot e^{100x} \cdot 100 \quad (11)$$

Among them, $\nabla_s J$ is obtained by the adjoint model calculation. Comparing equations (10) and (11), it can be seen that the
proportion of prior information in the C_100 test is relatively large. In contrast, the C_10 test more fully retains the effective
225 information in the observation data.

According to the above test design, the GRAPES-CUACE-3D-Var assimilation system (Fig. 3) is driven by a priori emission,
and the objective function is reduced step by step through cyclic iteration until the difference between the objective function
of the k+1 iteration and the k iteration is less than 1% to reach the convergence condition and the iteration is stopped. The
emission adjustment factor x_{k+1} of the k+1 iteration process is selected as the optimal estimation, and the BC posterior
230 source strength is calculated by using it.

4.2 Adjoint sensitivity analysis

The sensitivity of the objective function to BC priori source at any time in the inversion period can be obtained by running
the GRAPES-CUACE-3D-Var assimilation system. Figure 4 shows the hourly sensitivity distribution from 20: 00 November
30 to 07:00 December 1, 2015. The sensitivity in the southern part of Beijing is negative, indicating that increasing the BC
235 emission source intensity in this area will reduce the difference between the simulated concentration and the observed
concentration. The sensitivity of Tianjin, southern Hebei and northwestern Shandong is positive, indicating that reducing BC
emission source intensity in these areas can improve the consistency of simulation and observation.

At the target time (07:00 on December 1st), the sensitivity value (absolute value) in the inversion area is relatively large, and
the sensitivity value (absolute value) gradually decreases along with the inverse time sequence. At the same time, the
240 sensitivity coverage continues to expand to 11 hours before the target time (20:00 on November 30), and the sensitivity
distribution has been extended to Shanxi, Inner Mongolia and Shandong-Henan borders outside the inversion area. This
phenomenon shows that in a short period before the target time (December 1, 05:00-07:00), BC emission source intensity in
the inversion area has the greatest influence on the target function, and its influence gradually decreases with the inverse
time sequence, while BC emission source intensity outside the inversion area starts to have some influence on the target
245 function.

4.3 BC concentration comparison before and after inversion

The objective function values of the C_10 and C_100 tests decreased with the increase of iteration times, and the objective
function values of the two tests decreased from the initial 2288.5 (the 0th) to 1296.0 (the 18th) and 1295.5 (the 19th) (fig. 5a),
respectively. From the point of view of a single site, the BC posterior source simulation optimized by the C_10 and C_100



250 tests is obviously better than the BC prior source simulation. At the target time (07:00 on December 1, 2015), the simulated concentration of each site based on the BC posterior source is closer to the observed concentration (fig. 5b), and the correlation coefficient between the BC simulated concentration and the observed concentration is increased to 0.7 and 0.64 respectively from 0.2 before optimization (fig. 6a). The normalized mean square error (NMSE) decreased from 0.38 before optimization to 0.22 and 0.24 respectively (fig. 6b), and the normalized mean deviation (NMB) decreased from 51.53% to
255 43.37% and 40.90% (fig. 6c). In general, the optimization results of the C_10 and C_100 tests are relatively reasonable, and the simulation results of each station have been improved to a certain extent at the target time (07:00 on December 1, 2015). Meanwhile, the results of C_10 and C_100 tests are slightly different. The objective function value of the C_100 test reaches the convergence condition in the moderate fluctuation decline. The objective function value of the C_10 test dropped sharply at the first iteration, rebounded and rose during the second iteration, and then the fluctuation dropped to a minimum (fig. 5a).
260 From the correlation coefficient (fig. 6a) and NMSE (fig. 6b), the optimization effect of the C_10 test is better than that of the C_100 test, which is because the C_10 test better captures the BC concentration of site 5 (fig. 5b) at the target time (07:00 on December 1, 2015).

4.4 Comparison of BC emission sources before and after inversion

Fig. 7 shows the average BC prior sources and posterior sources and their difference distribution in Beijing-Tianjin-Hebei
265 region from 20:00 on November 30 to 07:00 on December 1, 2015. It can be seen that the distribution of BC posterior sources (figs. 7b,e) and prior sources (figs. 7a,c) retrieved from the C_10 and C_100 tests are relatively consistent, and the high-value areas are mainly located in the south of Beijing, Tianjin, central and southern Hebei and northern Shandong. Judging from the difference, the posterior source intensity of BC in Tianjin, central and southern Hebei and northern Shandong is lower than the prior source intensity (fig. 7c,f), which is consistent with the actual situation, that is, the BC
270 emissions of each province show a downward trend year by year (Zheng et al., 2018). However, the BC posterior source intensity in southern Beijing has increased compared with the prior source intensity, which is inconsistent with the actual situation (Zheng et al., 2018). Because Beijing area is in the process of high PM_{2.5} concentration at the target time (07:00 on December 1, 2015). Zhang et al. (2017) showed that the BC ratio decreased with the increase of PM_{2.5} concentration. Therefore, the average monthly ratio of BC/PM_{2.5} used in this research will lead to the "observed concentration" of BC at
275 site 5 (Figure 2) being too high (relatively, the simulated value is low), thus causing the emission of posterior sources in Beijing area to increase. Therefore, this phenomenon reflects that the source inversion test based on the GRAPES-CUACE-3D-Var assimilation system can make full use of the observation information to correct the prior source distribution. If the BC concentration at site 5 at the target time (07:00 on December 1, 2015) calculated in this study is the actual observation concentration, then Figures 7c and F show that both C_10 and C_100 tests have made reasonable adjustments to BC source
280 emissions in southern Beijing based on the observation value.

In addition, the posterior sources retrieved by the C_10 and C_100 tests are slightly different. Compared with the C_10 inversion source, the C_100 test has a low inversion value for the BC source intensity in southern Beijing (Fig. 7c, f). This is



because the proportion of observations in the C₁₀ test is larger than that of the C₁₀₀ test (Formula (10) and (11)), and the C₁₀ test makes full use of the BC observation concentration of the site 5 at the target to correct the BC prior source in
285 Beijing. When the difference between the simulated concentration and the observed concentration of BC at Site 5 is reduced, the BC source intensity in southern Beijing will naturally increase. Figure 8 shows the timely varying difference distribution of the C₁₀ inversion source and the C₁₀₀ inversion source. It can be seen that in the shorter period before the target time (03:00-07:00 on December 1st), the difference of BC source intensity in the south of Beijing is large, and the difference is small before that. This phenomenon indicates that the observation data of the target moment has a great influence on the
290 posterior source in the short period before the target time, and its influence gradually decreases with the inverse time sequence.

5 Summary and discussion

Based on the GRAPES-CUACE aerosol adjoint model, combined with gradient descent method and observation data, the GRAPES-CUACE-3D-Var assimilation system is constructed, mainly including GRAPES-CUACE forward model
295 calculation and ground state value storage, observation data and adjoint forcing term processing, adjoint model calculation, gradient extraction, target function calculation and model parameter optimization iteration, etc. Based on this system, taking the difference of BC simulation and observation concentration at 18 stations in Beijing-Tianjin-Hebei region on December 1, 2015 (BJT) as the objective function, and considering the influence of the proportion of observation data on the optimization effect, the design comparison tests C₁₀ and C₁₀₀ inversed BC emission sources in Beijing-Tianjin-Hebei region (113 °E-
300 119.5 °E, 35.25 °N-42.25 °N).

The results show that: The newly constructed GRAPES-CUACE-3D-Var assimilation system is reasonable and reliable, and can be initially applied to BC source inversion in Beijing-Tianjin-Hebei region. Both C₁₀ and C₁₀₀ experimental inversions of BC posterior source simulation results are better than BC prior source simulation results. The correlation coefficient between BC simulated concentration and observed concentration is increased to 0.7 and 0.64 respectively from
305 0.2 before inversion, NMSE is decreased to 0.22 and 0.24 respectively from 0.38 before inversion, NMB is decreased to 43.37% and 40.90% from 51.53%. The BC posterior sources retrieved by C₁₀ and C₁₀₀ tests are consistent with the distribution of prior sources. The high value areas are mainly located in the south of Beijing, Tianjin, central and southern Hebei and northern Shandong. On the whole, the observation proportion in the C₁₀ test is larger, and its optimization effect is better than that in the C₁₀₀ test. The observation information at the target time has a great influence on the posterior
310 source in a short period before the target time, and its influence gradually decreases with the inverse time sequence. The research ideas and experimental results in this paper have laid a foundation for the development of a complete GRAPES-CUACE four-dimensional variational assimilation system. However, although BC/PM_{2.5} ratio method can be used in this study to obtain BC concentration at various stations in Beijing-Tianjin-Hebei region, there is still a certain difference between the actual BC observation concentration and the inversion result, resulting in a certain deviation from the real



315 source emission. At the same time, the gradient descent method slows down its convergence speed when approaching the
minimum value, while the L-BFGS quasi-Newton method (LIMITED-MEMORY BFGS QUASI-NEWTON METHOD) is
an efficient optimization algorithm with high computational accuracy, fast convergence speed and small memory footprint,
which is suitable for solving large-scale optimization problems (Liu and Nocedal, 1989). In the following research, the
constructed GRAPES-CUACE-3D-Var assimilation system will be developed into a complete GRAPES-CUACE-4D-Var
320 assimilation system by using time-continuous observation data combined with L-BFGS quasi-Newton method, and more
accurate pollution source emissions will be retrieved based on this. In addition, taking into account many factors such as air
quality standard, pollution source reduction ratio, economic loss reduction and residents' health benefits, the GRAPES-
CUACE-4D-Var assimilation system is used to solve the PM_{2.5} and O₃ optimization control scheme, so as to provide basis
for optimal control of air pollution.

325 **Code and data availability.** The GRAPES-CUACE atmospheric chemistry model used in this study was distributed by the
Numerical Weather Prediction Center of Chinese Meteorology Administration (<http://nwpc.nmc.cn>) together with the
Institute of Atmospheric Composition of the Chinese Academy of Meteorological Sciences (<http://www.camscma.cn>). The
model was run on an IBM PureFlex System (AIX) with an XL Fortran Compiler. The code of GRAPES-CUACE aerosol
adjoint model is available online at doi:10.5194/gmd-9-2153-2016-supplement. The code of optimization can be downloaded
330 as a Supplement to this article. The observations are available online at <http://www.mee.gov.cn/>.

Author contributions. XA envisioned and oversaw the project. XA and CW designed and developed the GRAPES-
CUACE-3D-Var assimilation system, and prepared the paper. CW designed the experiments and carried out the simulations
with contributions from all other co-authors. QH and ZS provided the observation data used in the study. YL and JL
processed the data and prepared the data visualization. All authors reviewed the paper.

335 **Competing interests.** The authors declare that they have no conflict of interest.

Acknowledgements. This work was supported by the National Key Research and Development Program of China
(2017YFC0210006), and the National Natural Science Foundation of China (41975173 and 91644223). We appreciate Dr.
Lin Zhang from the Numerical Forecast Center of the China Meteorological Administration for providing technical support
in optimization algorithm.



340 References

- An, X. Q., Zhai, S. X., Jin, M., Gong, S., and Wang, Y.: Development of an adjoint model of GRAPES–CUACE and its application in tracking influential haze source areas in north China, *Geosci. Model Dev.*, 9, 2153-2165, doi: 10.5194/gmd-9-2153-2016, 2016.
- Cao, H., Fu, T. M., Zhang, L., Henze, D. K., Miller, C. C., Lerot, C., Abad, G. G., Smedt, I. D., Zhang, Q., Roozendael, M.
345 V. and Hendrick, F.: Adjoint inversion of Chinese non-methane volatile organic compound emissions using space-based observations of formaldehyde and glyoxal, *Atmos. Chem. Phys.*, 18, 15017-15046, doi: 10.5194/acp-2017-1136, 2018.
- Carmichael, G. R., Sandu, A., Chai, T., Daescu, D. N., Constantinescu, E. M. and Tang, Y.: Predicting air quality: Improvements through advanced methods to integrate models and measurements, *J. Comput. Phys.*, 227, 3540-3571, doi: 10.1016/j.jcp.2007.02.024, 2008.
- 350 Chen, D., Xue, J., Yang, X., Zhang, H., Shen, X., Hu, J., Wang, Y., Ji, L. and Chen, J.: New generation of multi-scale NWP system (GRAPES): general scientific design, *Chin. Sci. Bull.*, 53, 3433-3445, doi: 10.1007/s11434-008-0494-z, 2008.
- Cheng, Q., Zhang, L., and Wang, B.: Model adjointisation and its costs, *Sci. China Info. Sci.*, 47:587-611, 2004.
- Errico, R. M.: What is an adjoint model, *Bull. Amer. Meteorol. Soc.*, 78, 2577-2592, doi: 10.1175/1520-0477(1997)078<2577:WIAAM>2.0.CO;2, 1997.
- 355 Gong, S. L. and Zhang, X. Y.: CUACE/Dust—an integrated system of observation and modeling systems for operational dust forecasting in Asia, *Atmos. Chem. Phys.*, 8, 2333-2340, doi: 10.5194/acp-8-2333-2008, 2008.
- Hakami, A., Henze, D. K., Seinfeld, J. H., Chai, T., Tang, Y., Carmichael, G. R. and Sandu, A.: Adjoint inverse modeling of black carbon during the Asian Pacific Regional Aerosol Characterization Experiment, *J. Geophys. Res.-Atmos.*, 110(D14), doi: 10.1029/2004JD005671, 2005.
- 360 Hakami, A., Henze, D. K., Seinfeld, J. H., Singh, K., Sandu, A., Kim, S., Byun, D. and Li, Q.: The adjoint of CMAQ, *Environ. Sci. Technol.*, 41, 7807-7817, doi: 10.1021/es070944p, 2007.
- Henze, D. K., Hakami, A. and Seinfeld, J. H.: Development of the adjoint of GEOS-Chem, *Atmos. Chem. Phys.*, 7, 2413-2433, doi: 10.5194/acp-7-2413-2007, 2007.
- Henze, D. K., Seinfeld, J. H. and Shindell, D. T.: Inverse modeling and mapping US air quality influences of inorganic PM_{2.5}
365 precursor emissions using the adjoint of GEOS-Chem, *Atmos. Chem. Phys.*, 9, 5877-5903, doi: 10.5194/acp-9-5877-2009, 2009.
- Huang, S. X., Liu, F., Sheng, L., Cheng, L. J., Wu, L., Li, J.: On adjoint method based atmospheric emission source tracing. *Chin. Sci. Bull.*, 63, 1594-1605, doi:10.1360/N972018-00196, 2018.
- Jeong, J. I. and Park R. J.: Efficacy of dust aerosol forecasts for East Asia using the adjoint of GEOS-Chem with ground-
370 based observations, *Environ. Pollut.*, 234, 885-893, doi: 10.1016/j.envpol.2017.12.025, 2018.
- Jiang, Z., Jones, D. B. A., Worden, H. M. and Henze, D. K.: Sensitivity of top-down CO source estimates to the modeled vertical structure in atmospheric CO, *Atmos. Chem. Phys.*, 15, 1521-1537, doi: 10.5194/acp-15-1521-2015, 2015.



- Jin, J. J.: Study on the inversion of source strength of air pollutants by variational method, Master dissertation, Lanzhou University, Lanzhou, 1-48, 2001 (in Chinese).
- 375 Jin, M.: Estimation of lightning-produced NO_x and sensitivity analysis of emission inversion by adjoint model, Master dissertation, Nanjing University of Information Science & Technology, Nanjing, 1-65, 2012 (in Chinese).
- Kurokawa, J. I., Yumimoto, K., Uno, I. and Ohara, T.: Adjoint inverse modeling of NO_x emissions over eastern China using satellite observations of NO₂ vertical column densities, *Atmos. Environ.*, 43, 1878-1887, doi: 10.1016/j.atmosenv.2008.12.030, 2009.
- 380 Liu, D. C. and Nocedal, J.: On the limited memory BFGS method for large scale optimization, *Math. Program.*, 45, 503-528, doi: 10.1007/BF01589116, 1989.
- Liu, F.: Adjoint model of Comprehensive Air quality Model CAMx – construction and application, Post-doctoral research report, Peking University, Beijing, 1-101, 2005 (in Chinese).
- Lv, Y.: The Development and Application of Air Pollutant Emission Source Processing System, Master dissertation, Xi'an University of Architecture and Technology, Xi'an, 1-59, 2015 (in Chinese).
- 385 Mao, Y. H., Li, Q. B., Henze, D. K., Jiang, Z., Jones, D. B. A., Kopacz, M., He, C., Qi, L., Gao, M., Hao, W. M. and Liou, K. N., Estimates of black carbon emissions in the western United States using the GEOS-Chem adjoint model, *Atmos. Chem. Phys.*, 15, 7685-7702, doi: 10.5194/acp-15-7685-2015, 2015.
- Marchuk, G. I. and Skiba, I.: Numerical analysis of the conjugate problem for a model of thermal interaction of the atmosphere with oceans and continents, *Izvestiya-Atmospheric and Ocean Physics*, 12, 459-469, 1976.
- 390 Müller, J. F. and Stavrakou, T.: Inversion of CO and NO_x emissions using the adjoint of the IMAGES model, *Atmos. Chem. Phys.*, 5, 1157-1186, doi: 10.5194/acp-5-1157-2005, 2005.
- Pan, L., Chai, T., Carmichael, G. R., Tang, Y., Streets, D., Woo, J. H., Friedli, H. R. and Radke, L. F.: Top-down estimate of mercury emissions in China using four-dimensional variational data assimilation, *Atmos. Environ.*, 41, 2804-2819, doi: 10.1016/j.atmosenv.2006.11.048, 2007.
- 395 Sandu, A., Daescu, D. N., Carmichael, G. R. and Chai, T.: Adjoint sensitivity analysis of regional air quality models, *J. Comput. Phys.*, 204, 222-252, doi: 10.1016/j.jcp.2004.10.011, 2005.
- Song, J., Wang, X. and Feng, X.: Optimization Method, Xi'an University of Electronic Science and Technology Press, Xi'an, 2012 (in Chinese).
- 400 Wang, C., An, X., Zhai, S. and Sun, Z.: The application of an adjoint model in tracking influential haze source areas of pollution episodes, *China Environ. Sci.*, 37:1283–1290, 2017 (in Chinese).
- Wang, C., An, X., Zhai, S., and Sun, Z.: Tracking a severe pollution event in Beijing in December 2016 with the GRAPES-CUACE adjoint model, *J. Meteorol. Res.*, 32, 49-59, doi: 10.1007/s13351-018-7062-5, 2018a.
- Wang, C., An, X., Zhai, S., Hou, Q. and Sun, Z.: Tracking sensitive source areas of different weather pollution types using GRAPES-CUACE adjoint model, *Atmos. Environ.*, 175, 154-166, doi: 10.1016/j.atmosenv.2017.11.041, 2018b.
- 405



- Wang, C., An, X., Zhang, P., Sun, Z., Cui, M. and Ma, L.: Comparing the impact of strong and weak East Asian winter monsoon on PM_{2.5} concentration in Beijing, *Atmos. Res.*, 215, 165-177, doi: 10.1016/j.atmosres.2018.08.022, 2019.
- Wang, H., Gong, S. L., Zhang, H. L., Chen, Y., Shen, X., Chen, D., Xue, J., Shen, Y., Wu, X. and Jin, Z.: A new-generation sand and dust storm forecasting system GRAPES_CUACE/Dust: Model development, verification and numerical simulation, *Chin. Sci. Bull.*, 55, 635-649, doi: 10.1007/s11434-009-0481-z, 2010.
- Wang, H., Xue, M., Zhang, X. Y., Liu, H. L., Zhou, C. H., Tan, S. C., Che, H. Z., Chen, B. and Li, T.: Mesoscale modeling study of the interactions between aerosols and PBL meteorology during a haze episode in Jing-Jin-Ji (China) and its nearby surrounding region-Part 1: Aerosol distributions and meteorological features, *Atmos. Chem. Phys.*, 15, 3257-3275, doi: 10.5194/acp-15-3257-2015, 2015.
- 415 Wang, J., Xu, X., Henze, D. K., Zeng, J., Ji, Q., Tsay, S. C. and Huang, J.: Top-down estimate of dust emissions through integration of MODIS and MISR aerosol retrievals with the GEOS-Chem adjoint model, *Geophys. Res. Lett.*, 39, L08802, doi: 10.1029/2012GL051136, 2012.
- Wang, Y., Yao, L., Wang, L., Liu, Z., Ji, D., Tang, G., Zhang, J., Sun, Y., Hu, B. and Xin, J.: Mechanism for the formation of the January 2013 heavy haze pollution episode over central and eastern China, *Sci. China Earth Sci.*, 57: 14-25, doi: 10.1007/s11430-013-4773-4, 2014a.
- 420 Wang, Z., Li, J., Wang, Z., Yang, W., Tang, X., Ge, B., Yan, P., Zhu, L., Chen, X., Chen, H. and Wand, W.: Modeling study of regional severe hazes over Mid-Eastern China in January 2013 and its implications on pollution prevention and control, *Sci. China Earth Sci.*, 57, 3-13, doi: 10.1007/s11430-013-4793-0, 2014b.
- Ye, Q., Shen, Y.: *Practical Mathematical Manual*, Science Press, Beijing, 2006 (in Chinese).
- 425 Zeng, Q. C. and Wu, L.: Optimal reduction of anthropogenic emissions for air pollution control and the retrieval of emission source from observed pollutants I. Application of incomplete adjoint operator, *Sci. China Earth Sci.*, 61, 951-956, doi: <https://doi.org/10.1007/s11430-017-9199-2>, 2018.
- Zhai, S., An, X., Zhao, T., Sun, Z., Wang, W., Hou, Q., Guo, Z. and Wang, C.: Detection of critical PM_{2.5} emission sources and their contributions to a heavy haze episode in Beijing, China, using an adjoint model, *Atmos. Chem. Phys.*, 18, 6241-430 6258, doi: 10.5194/acp-18-6241-2018, 2018.
- Zhai, S.: Development of the adjoint of GRAPES-CUACE aerosol module and model application to air pollution optimal control problems, Master dissertation, Chinese Academy of Meteorological Sciences, Beijing, 1-81, 2015 (in Chinese).
- Zhang, C., Cheng, X., Zhao, T., Xu, X., Wu, Y., Zhang, R., Cai, W., Su, H., Wang, Y.: Impact of meteorological conditions on high black carbon concentrations in urban area of Beijing in different seasons, *Acta. Sci. Circums.*, 37, 2255-2264, 2017 435 (in Chinese).
- Zhang, L., Henze, D. K., Grell, G. A., Bousserez, N., Zhang, Q., Torres, O., Ahn, C., Lu, Z., Cao, J. and Mao, Y.: Constraining black carbon aerosol over Asia using OMI aerosol absorption optical depth and the adjoint of GEOS-Chem, *Atmos. Chem. Phys.*, 15, 10281-10308, doi: 10.5194/acp-15-10281-2015, 2015.



440 Zhang, L., Shao, J., Lu, X., Zhao, Y., Hu, Y., Henze, D.K., Liao, H., Gong, S. 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, doi: 10.1021/acs.est.6b03010, 2016.

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. and Chen, D.: Asian emissions in 2006 for the NASA INTEX-B mission, Atmos. Chem. Phys., 9, 5131-5153, doi: 10.5194/acp-9-5131-2009, 2009.

445 Zhang, X., Xu, X., Ding, Y., Liu, Y., Zhang, H., Wang, Y. and Zhong, J.: The impact of meteorological changes from 2013 to 2017 on PM_{2.5} mass reduction in key regions in China, Sci. China Earth Sci., 49, 1-18, doi: 10.1007/s11430-019-9343-3, 2019.

Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., Li, H., Li, X., Peng, L., Qi, J. and Yan, L.: Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions, Atmos. Chem. Phys., 18, 14095-14111, doi: 450 10.5194/acp-18-14095-2018, 2018.

Zhu, J., Tang, X., Wang, Z., and Wu, L.: A review of air quality data assimilation methods and their application, Chinese J. Atmos. Sci., 42, 607–620, doi: 10.3878/j.issn.1006-9895.1802.17260, 2018 (in Chinese).

455

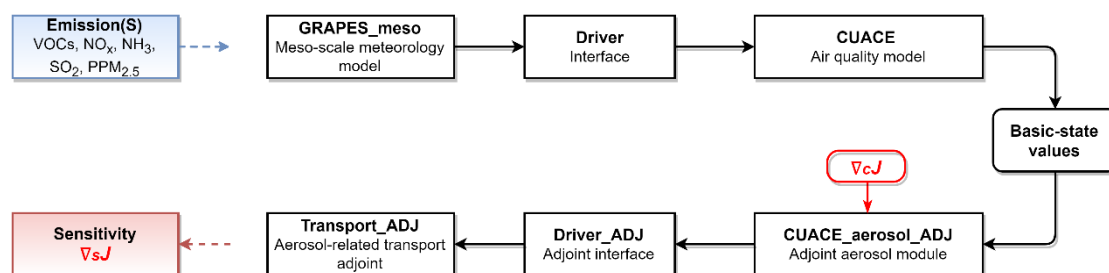
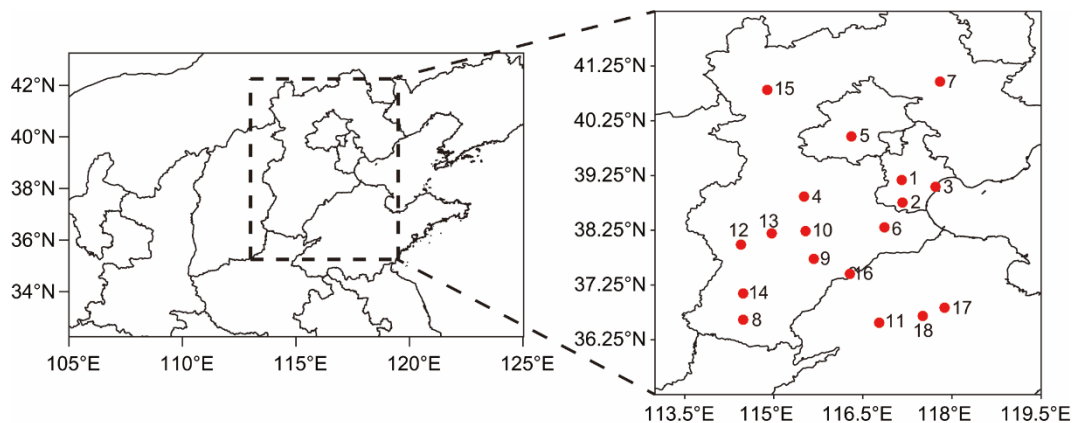


Figure 1: Running process of GRAPES-CUACE chemical model and its adjoint model.



460 **Figure 2: Simulation domain (left) and emission inversion region and observation stations (right).**

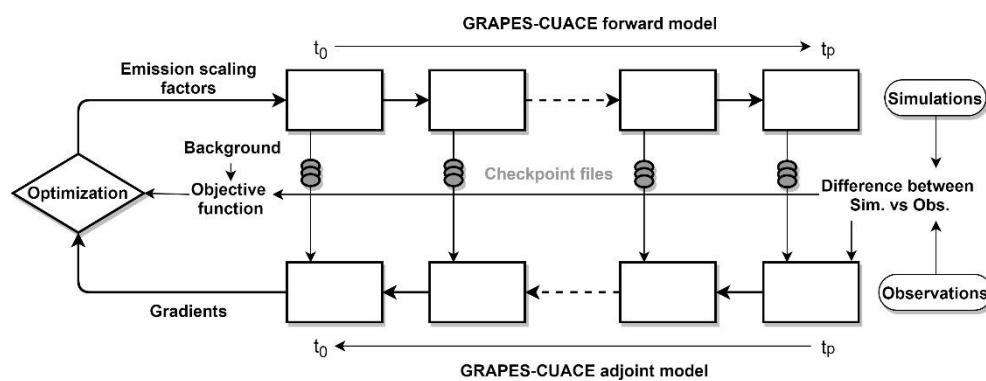
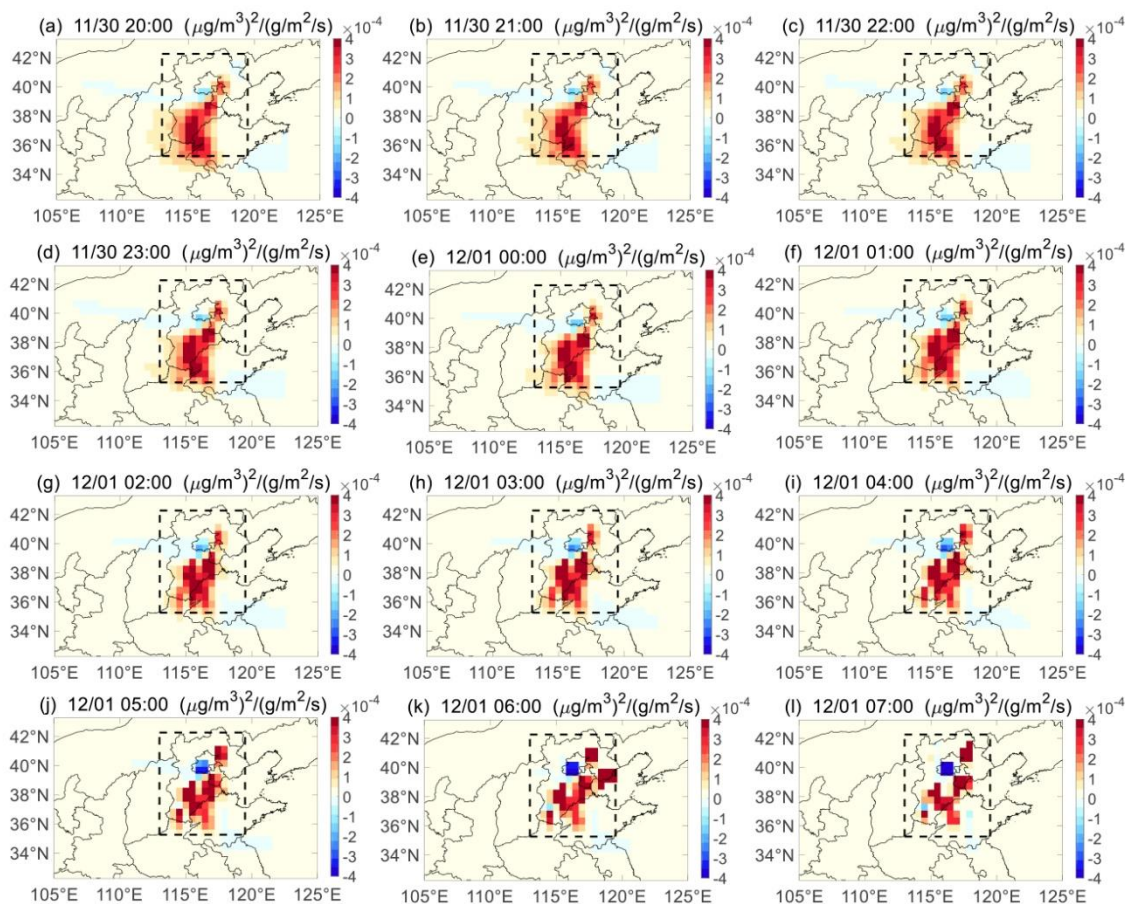
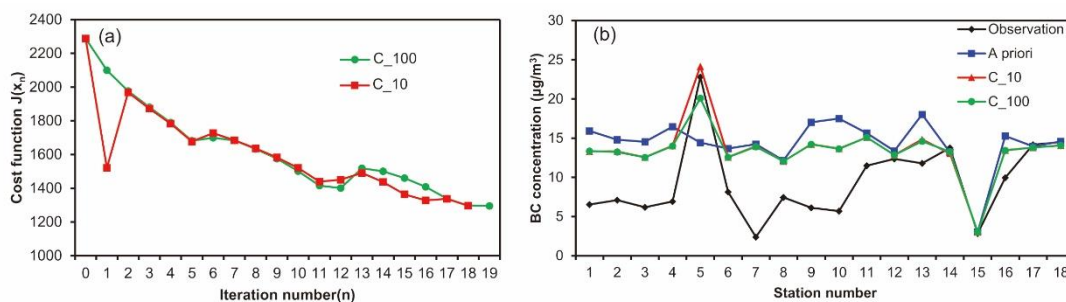


Figure 3: GRAPES-CUACE-3D-Var assimilation system.



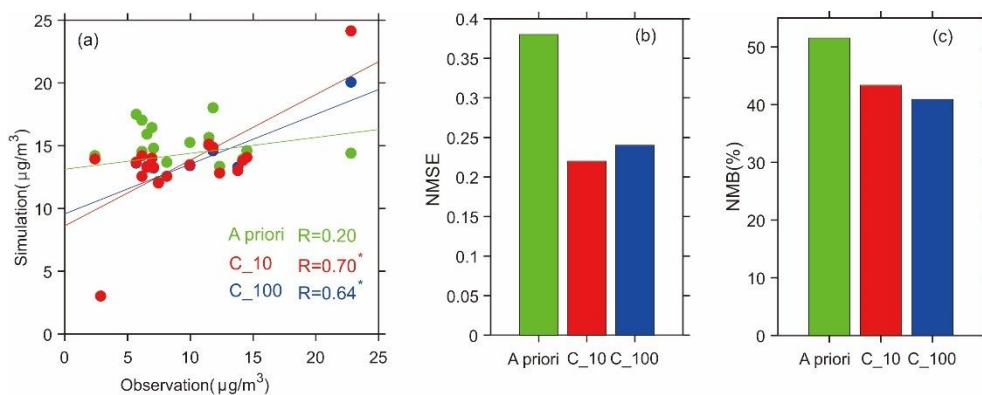
465

Figure 4: The distributions of sensitivity of the cost function to the a priori BC emissions from 20:00 30th Nov. 2015 to 07:00 1st Dec. 2015.



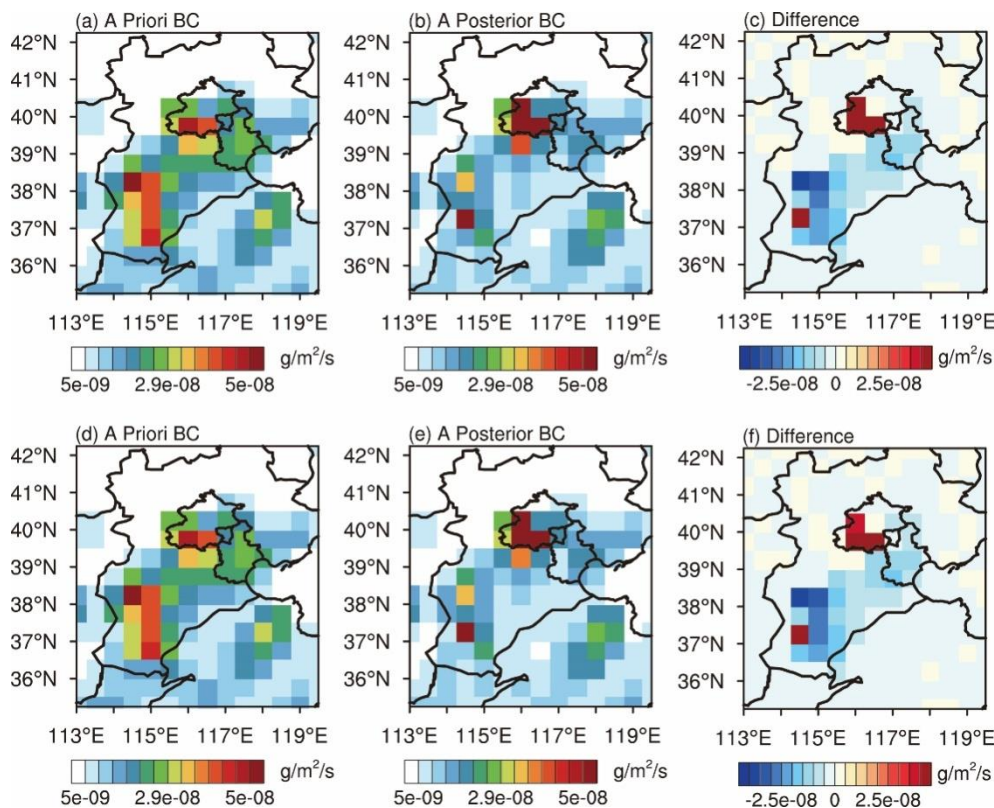
470

Figure 5: (a) Cost function reduction of ExperimentC_10 and ExperimentC_100, (b) BC concentrations observed (black diamond) and simulated with a priori (purple square) and a posterior BC emissions of ExperimentC_10 (red triangle) and ExperimentC_100 (green circle) at each station at the objective time point (07:00 1st Dec. 2015).



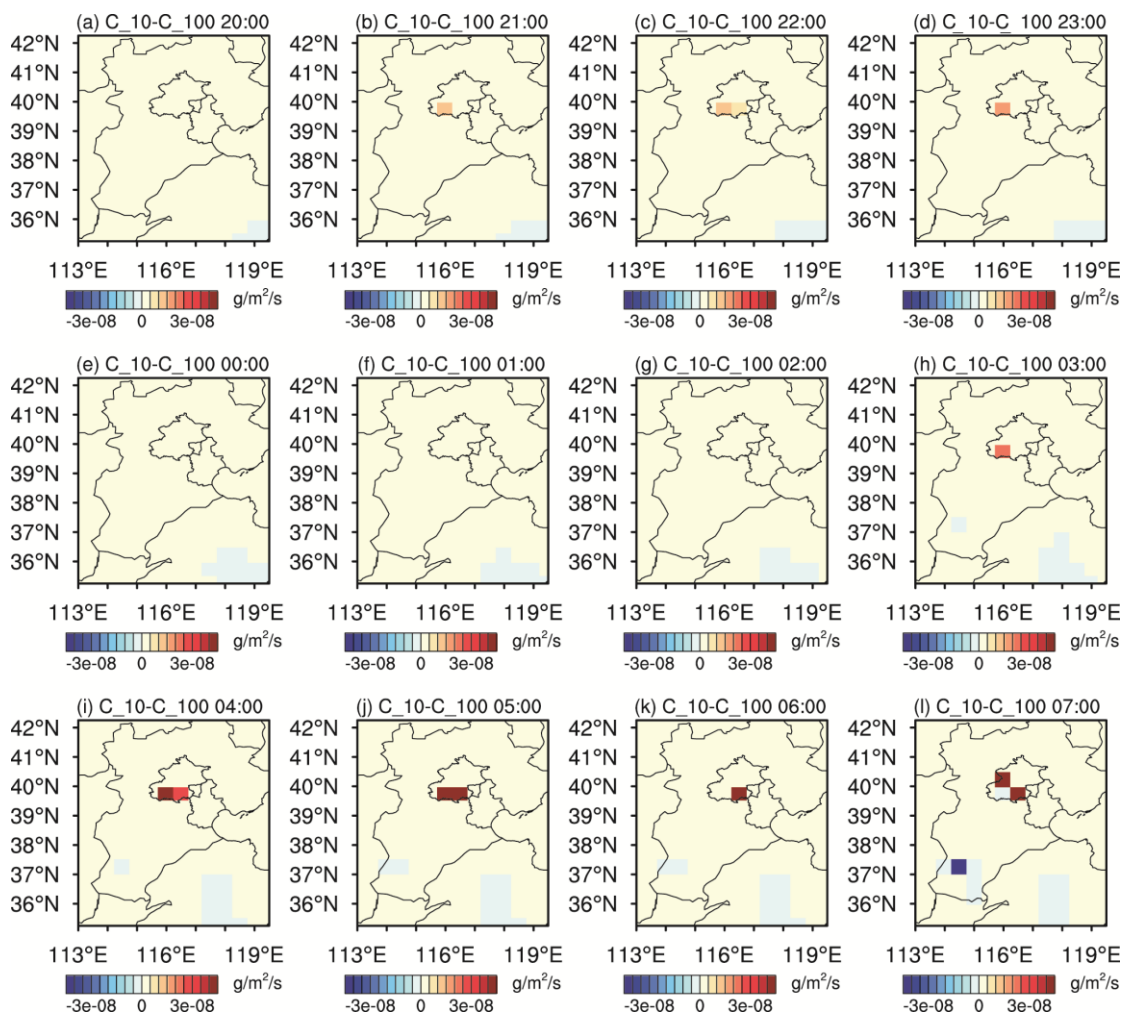
475

Figure 6: (a) Scatter plots, (b) NMSE and (c) NMB statistics of BC concentrations observed and simulated with a priori and a posterior BC emissions of ExperimentC_10 and Experiment C_100 at eighteen stations at the objective time point (07:00 1st Dec. 2015). Note: one star superscript in (a) indicates the values are significantly correlated at the 0.01 level on both sides.



480

Figure 7: Time-averaged distributions of (a, d) a priori and a posterior BC emissions of (b) Experiment C_10 and (e) Experiment C_100 and (c, f) their differences (A Posterior-A Prior) from 20:00 30th Nov. to 07:00 1st Dec. 2015 in Beijing-Tianjin-Hebei area.



485 **Figure 8:** Hourly distributions of the differences between a posterior BC emissions of Experiment C_10 and Experiment C_100 (C_10-C_100) from 20:00 30th Nov. to 07:00 1st Dec. 2015 in Beijing-Tianjin-Hebei area.