Development of Four Dimensional Variational Assimilation System Based on GRAPES-CUACE Adjoint Model (GRAPES-CUACE-4D-Var V1.0) and Its Application in Emission Inversion

Chao Wang[1,2], Xingqin An[1], Qing Hou[1], Zhaobin Sun[3], Yanjun Li[1], Jiangtao Li[1]

[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. We developed a four-dimensional variational (4D-Var) data assimilation system for the GRAPES-CUACE atmospheric chemistry model. In this study, we adapted it for assimilating surface black carbon (BC) concentrations and optimizing its daily emissions in North China on July 4th 2016, when high BC concentrations observed in Beijing.

1 Introduction

Three-dimensional (3-D) atmospheric chemical transport models (CTMs) are important tools for air quality research, which are used not only for predicting spatial and temporal distributions of air pollutants, but also for providing sensitivities of air pollutant concentrations with respect to various parameters (Hakami, et al., 2007). Among several methods of sensitivity analysis, the adjoint method is known to be an efficient means of calculating the sensitivities of a cost function with respect to a large number of input parameters (Sandu et al., 2005; Hakami et al., 2007; Henze et al., 2007; Zhai et al., 2018). The sensitivity information provided by the adjoint approach can be applied to a variety of optimization problems, such as formulating optimized pollution-control strategies, inverse modelling and variational data assimilation (Liu, 2005; Hakami, et al., 2007).

Four-dimensional variational (4D-Var) data assimilation, which is an important application of adjoint models, provides insight into various model inputs, such as initial conditions and emissions (Liu et al., 2005; Yumimoto and Uno, 2006). In the past decades, many scholars have successively developed adjoint models of various 3-D CTMs and the 4D-Var data assimilation systems to optimize model parameters. Elbern et al. (1999, 2000, 2001, 2007) constructed the adjoint of the EURAD CTM and performed inverse modelling of emissions and chemical data assimilation. Sandu et al. (2005) built the adjoint of the comprehensive chemical transport model STEM-III and conducted the data assimilation in a twin experiments framework as well as the assimilation of a real data set, with the control variables being O_3 or NO_2. Hakami et al. (2005) adapted the adjoint model of STEM-2k1 model for assimilating black carbon (BC) concentrations and recovery of its emissions. Liu (2005) and Huang et al. (2018) developed the adjoint of CAMx model and further expanded it into an air

quality forecasting and pollution-control decision support system. Müller and Stavrakou (2005) constructed an inverse modelling framework based on the adjoint of the global model IMAGES and used it to optimize the global annual CO and NOx emissions for the year 1997. More recently, the CMAQ community (Hakami et al., 2007) built the adjoint of CMAQ model and its 4D-Var assimilation scheme, which were used to optimize NOx emissions (Kurokawa et al., 2009; Resler et al., 2010) and ozone initial state (Park et al., 2016). The adjoint of GEOS-Chem model and its 4D-Var assimilation system firstly developed by Henze et al. (2007, 2009) have been applied in a number of studies to improve aerosol (Wang et al., 2012; Mao et al., 2015; Jeong and Park, 2018), CO (Jiang et al., 2015) and NMVOC (Cao et al., 2018) emissions estimates. Zhang et al. (2016) applied the 4D-Var assimilation system using the adjoint model of GEOS-Chem with the fine 1/4°× 5/16° horizontal resolution to optimize daily aerosol primary and precursor emissions over North China. These researches have laid good foundations for developing adjoint models of CTMs and optimizing model parameters. However, only a few of these adjoint models and their 4D-Var assimilation systems have been widely applied to regional air pollution in China. The development and applications of adjoint models of 3-D CTMs and their 4D-Var data assimilation systems are still limited in China. Further research and more attention are required.

Nowadays, several mega urban agglomerations in China, such as the Beijing-Tianjin-Hebei region, the Yangtze River Delta region, and the Fenwei Plain, are still suffering from severe air pollution (Zhang et al., 2019; Xiang et al., 2019; Haque et al., 2020; Zhao et al., 2020). Previous studies have shown that emission-reduction strategies, which are mainly based on the results of atmospheric chemistry simulations, play an important role in reducing pollutant concentrations and improving air quality (Zhang et al., 2016; Zhai et al., 2016). The emission inventory is an important basic data for atmospheric chemistry simulation, and its uncertainty will affect the accuracy of air pollution simulation, which in turn will affect the accuracy of pollution-control measures based on the model results (Huang et al., 2018). In order to improve the accuracy of atmospheric chemistry simulation and the feasibility of the pollution-control strategies, the emission data obtained by the "bottom-up" method needs to be inverted, which can be done through the atmospheric chemical variational assimilation system, to reduce the impact of emission uncertainty.

GRAPES-CUACE is an atmospheric chemistry model system developed by the Chinese Academy of Meteorological Sciences (CAMS)(Gong and Zhang, 2008; Zhou et al., 2008, 2012; Wang et al., 2010, 2015). GRAPES (Global/Regional Assimilation and PrEdiction System) is a numerical weather prediction system built by China Meteorological Administration (CMA), and it can be used as a global model (GRAPES-GFS) or as a regional mesoscale model (GRAPES-Meso) (Chen et al., 2008; Zhang and Shen, 2008). CUACE (CMA Unified Atmospheric Chemistry Environmental Forecasting System) is a unified atmospheric chemistry model constructed by CAMS to study both air quality forecasting and climate change(Gong and Zhang, 2008; Zhou et al., 2008, 2012). Using the meteorological fields provided by GRAPES-Meso, the GRAPES-CUACE model has been realized the online coupling of meteorology and chemistry (Gong and Zhang, 2008; Zhou et al., 2008, 2012; Wang et al., 2010, 2015). The GRAPES-CUACE model not only plays an important role in the scientific researches on air pollution in China (Gong and Zhang, 2008; Zhou et al., 2008, 2012; Wang et al., 2010, 2015, 2018), but

also has been officially put into operation since 2014 at National Meteorological Center of CMA for providing guidance for air quality forecasting over China (Ke, 2019).

Recently, An et al. (2016) constructed the aerosol adjoint module of GRAPES-CUACE model, which has been subsequently applied in tracking influential BC and $PM_{2.5}$ source areas in north China (Zhai et al., 2018; Wang et al., 2018a, 2018b, 2019). However, these applications of GRAPES-CUACE aerosol adjoint model are still limited to sensitivity analysis, and the sensitivity information is not fully used to solve various optimization problems mentioned above. At the same time, considering the current severe pollution situation in mega urban agglomerations in China, more accurate emission data are urgently required to formulate reasonable and effective pollution-control strategies. In this study, we developed a new 4D-Var data assimilation system on the basis of the GRAPES-CUACE adjoint model, which was adapted for assimilating surface BC concentrations and optimizing its daily emissions in North China on July 4th 2016, when high BC concentrations observed in Beijing.

2 Methodology

2.1 Forward model description

2.1.1 GRAPES-Meso

GRAPES-Meso is a real-time operational weather forecasting model used by China Meteorological Administration (Chen et al., 2008; Zhang and Shen, 2008). The GRAPES-Meso model uses fully compressible non-hydrostatic equations as its model core. The vertical coordinates adopt the height-based, terrain-following coordinates, and the horizontal coordinates use the spherical coordinates of equal longitude-latitude grid points. The horizontal discretization adopts an Arakawa-C staggered grid arrangement and a central finite-difference scheme with second-order accuracy, while the vertical discretization adopts the vertically staggered variable arrangement proposed by Charney-Phillips. The time integration discretization uses a semi-implicit and semi-Lagrangian temporal advection scheme. The large-scale transport processes (both horizontal and vertical) for all gases and aerosols in GRAPES-CUACE are calculated by the dynamic framework of GRAPES-Meso, which implements the quasi-monotone semi-Lagrangian (QMSL) semi-implicit scheme on each grid (Wang et al., 2010). The physical processes principally involve micro-physical precipitation, cumulus convection, radiative transfer, land surface and boundary layer processes. Each physical process incorporates several schemes and can also be tailored by the user (Xu et al., 2008). The major physical options that we selected include the WSM6 cloud microphysics scheme (Hong et al., 2006), the Betts-Miller-Janjic cumulus convection scheme (Betts et al., 1986; Janjic, 1994); the RRTM(Rapid Radiative Transfer Model; Mlawer et al., 1997) long-wave radiation scheme, the short-wave scheme based on Dudhia (1989); the Monin-Obukhov surface layer scheme (Monin et al., 2009); the MRF planetary boundary layer scheme (Mellor et al., 1974; Andre et al., 1978); and the Noah land surface scheme (Chen et al., 1996).

2.1.2 CUACE

The atmospheric chemistry model CUACE mainly includes three modules: the aerosol module(module_ae_cam), the gaseous chemistry module(module_gas_radm) and the thermodynamic equilibrium module(module_isopia) (Gong and Zhang, 2008; Zhou et al., 2008, 2012; Wang et al., 2010, 2015). The interface program that connects CUACE and GRAPES-Meso is called aerosol_driver.F. This program transmits the meteorological fields calculated in GRAPES-Meso and the emission data processed as needed to each module of CUACE. The physical and chemical processes of 66 gas species and 7 aerosol species (sulfate, nitrate, sea salt, black carbon, organic carbon, siol dust and ammonium) in the atmosphere are comprehensively considered in the CUACE model (Wang et al., 2015).

CUACE adopts CAM (Canadian Aerosol Module; Gong et al., 2003) and RADM II (the second-generation Regional Acid Deposition Model; Stockwell et al., 1990) as its aerosol module and gaseous chemistry module, respectively. CAM involves six types of aerosols: sulfate(SF), nitrate(NI), sea salt(SS), black carbon(BC), organic carbon(OC) and siol dust(SD), which are segregated into 12 size bins with diameter ranging from 0.01 to 40.96 μm according to the multiphase multicomponent aerosol particle size separation algorithm (Gong et al., 2003; Zhou et al.,2008, 2012; Wang et al., 2010, 2015). CAM also calculates the vertical diffusion trend of aerosol particles by solving the vertical diffusion equation. The core of CAM is the aerosol physical and chemical processes, including hygroscopic growth, coagulation, nucleation, condensation, dry deposition/sedimentation, below-cloud scavenging, and aerosol activation, which is coherently integrated with the gaseous chemistry in CUACE(Gong et al., 2003; Zhou et al.,2008, 2012; Wang et al., 2010, 2015). The gas chemistry provides the production rates of sulphate aerosols and secondary organic aerosols (SOA), and meanwhile generates an oxidation background for aqueous phase aerosol chemistry, in which sulphate transformation changes the distributions of SO_2 in clouds (Zhou et al., 2012). Both nucleation and condensation are considered for sulphate aerosol formation depending on the atmospheric state after gaseous H_2SO_4 formed from the oxidation of sulphurous gases such as SO_2, H_2S and DMS(Zhou et al., 2012). Second organic aerosols as generated from gaseous precursors are partitioning among different bins through condensation using the same approach as the gaseous H_2SO_4 condensation to sulphate (Zhou et al., 2012). Given that the nitrates(NI) and ammonium(AM) formed through the gaseous oxidation are unstable and prone to further decomposition back to their precursors, CUACE adopts ISORROPIA to calculate the thermodynamic equilibrium between them and their gas precursors (West et al., 1998; Nenes et al., 1998a, 1998b; Zhou et al., 2012). ISORROPIA contains 15 equilibrium reactions, and the main species include gas phase: NH_3, HNO_3, HCL, H_2O, liquid phase: NH_4^+, Na^+, H^+, Cl^-,NO_3^-, SO_4^{2-}, HSO_4^-, OH^-, H_2O, and solid phase: $(NH_4)_2SO_4$, NH_4HSO_4, $(NH_4)_3H(SO_4)_2$, NH_4NO_3, NH_4Cl, NaCl, $NaNO_3$, $NaHSO_4$, Na_2SO_4 (Nenes et al., 1998a).

The emissions used in this study are based on statistical data of anthropogenic emissions reported from government agencies for 2007 (Cao et al., 2011). Emission source types included residences, industry, power plants, transportation, biomass combustion, livestock and poultry breeding, fertilizer use, waste disposal, solvent use, and light industrial product manufacturing (Cao et al., 2011; Zhai et al., 2018). These emission data were transformed through the Sparse Matrix

Operator Kernel Emissions (SMOKE) module into hourly gridded off-line data for 32 species, including BC, OC, SF, NI, fugitive dust particles and 19 non-methane volatile organic compounds (VOC), CH_4, NH_3, CO, CO_2, SO_x and NO_x, at three vertical levels (non-point source on the ground, middle elevation point source at 50m and high elevation point source at 120m), as required by the GRAPES-CUACE model. Furthermore, natural sea salt and natural sand/dust emissions were also calculated on-line in the model (Zhou et al., 2012; Zhai et al., 2018).

2.2 Adjoint model

2.2.1 Adjoint theory

Assuming that **L** is a linear operator defined in the Hilbert space **H**, if there is another linear operator $\mathbf{L}^*$ satisfying:

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

Then $\mathbf{L}^*$ is called the adjoint operator of **L** (Ye and Shen, 2006). Where $(.,.)$ denotes the inner product in **H**. If x, y are continuous functions on a domain Ω, the inner product is defined as $(x, y) = \int_\Omega x \cdot y d\Omega$; if x, y are discrete vectors, $x = [x_1, x_2, \dots, x_N]$, $y = [y_1, y_2, \dots, y_N]$, then the inner product is $(x, y) = \sum_{i=1}^{N} x_i \cdot y_i$. When x and y denote vectors and **L** is a matrix (independent of x and y), we can obtain:

$$(Lx, y) = y^T Lx = x^T L^T y = (x, L^T y). \quad (2)$$

In other words, for a matrix-type linear operator, the adjoint operator is its transpose: $L^* = L^T$ (Liu, 2005).

An atmospheric chemistry model can be viewed as a numerical operator $\mathbf{F}: \boldsymbol{R}^n \rightarrow \boldsymbol{R}^m$, which can be expressed as:

$$Y = F(X), \quad (3)$$

where $X \in R^n$ and $Y \in R^m$ are vectors representing the input and output variables in the atmospheric chemistry model, respectively. If F is differentiable, then the differential of Y (δY) can be denoted by the differential of X (δX), and the tangent linear model (TLM) of the atmospheric chemistry model can be expressed as:

$$\delta Y = \nabla_X F \cdot \delta X, \quad (4)$$

where $\delta X \in R^n$ and $\delta Y^* \in R^m$ are input and output variables in the TLM, respectively. And $\nabla_X F$ is the Jacobian matrix.

According to EQ. (1) and (2), the adjoint model of the TLM can be expressed as:

$$\delta X^* = \nabla_X^T F \cdot \delta Y^*, \quad (5)$$

where $\delta Y^* \in R^m$ and $\delta X^* \in R^n$ are input and output variables in the adjoint model, respectively. Comparing EQ. (4) and (5), it can be seen that the dimensions of input and output are exchanged between the TLM and the adjoin model, and the operator in EQ. (5) is the transpose of the operator in EQ. (4) (Liu, 2005). It is easy to see that the gradient (sensitivity) of the objective function with respect to input variables can be obtained through n times TLM simulations or m times adjoint simulations. When n>>m (such as n-dimensional emission sources and m-dimensional pollutant concentrations), the calculation efficiency of the adjoint model is much higher than that of the TLM (Liu, 2005).

2.2.2 GRAPES-CUACE aerosol adjoint

The GRAPES-CUACE aerosol adjoint model was constructed by An et al. (2016) based on the adjoint theory (Ye and Shen, 2006; Liu, 2005) and the CUACE aerosol module, which mainly includes the adjoint of physical and chemical processes and flux calculation processes of six types of aerosols (SF, NI, SS, BC, OC and SD) in CAM module, the adjoint of interface programs that connect GRAPES-Meso and CUACE, and the adjoint of aerosol transport processes.

As described in An et al. (2016), after the construction of the adjoint model is completed, its accuracy must be verified to confirm its reliability. Since the adjoint model is built on the basis of the TLM, the validity of the TLM must be ensured before the accuracy of the adjoint model is tested. The verification formula of tangent linear codes can be expressed as:

$$\text{Index} = \lim_{\delta X \to 0} \frac{F(X+\delta X)-F(X)}{\delta X F'(X)} = 1.0, \quad (6)$$

where the denominator is the TLM output, and the numerator is the difference between the output value of the original model with input $X + \delta X$ and input X. It is necessary to decrease the value of δX by an equal ratio and repeat the calculation of the above formula. If the result approaches 1.0, the tangent linear codes are correct. It was verified that all input variables in the model, such as the concentration value of pollutants (xrow) and the particle's wet radius (rhop), have passed the TLM test.

The adjoint codes can be validated on the basis of the correct tangent linear codes. The adjoint codes and the tangent linear codes need to satisfy Eq.(2) for all possible combinations of X and Y. In Eq.(2), L and L^* represent the tangent linear process and the adjoint process, respectively. To simplify the testing process, the adjoint input is the tangent linear output: Y = L(X). Thus, Eq.(5) can be expressed as:

$$(\nabla F \cdot dX, \nabla F \cdot dX) = (dX, \nabla^T F(\nabla F \cdot dX)). \quad (7)$$

By substituting dX into the tangent linear codes, the output value $\nabla F \cdot dX$ can be obtained and the left part of the equation can be computed. Then, taking $\nabla F \cdot dX$ as the input of the adjoint codes, the adjoint output $\nabla^T F(\nabla F \cdot dX)$ can be obtained and the right part of the equation can be calculated. On condition that the left and right sides of EQ.(7) are equal within the range of machine errors, the constructed adjoint model is validated. It was verified that all input variables in the model have passed the adjoint test. Taking the pollutant concentration variable (xrow) as an example, both sides of EQ.(7) produce values with 14 identical significant digits or more. This result is within the range of machine errors, so the values of the left and the right sides are considered equal. Thus, the pollutant concentration variable (xrow) has passed the adjoint test.

After the TLM and the adjoint model were verified, the GRAPES–CUACE aerosol adjoint model was constructed. The operation flowchart of the adjoint model is shown in Fig. 1. *J* is the objective function, which can be defined according to the concerned problems. *c* and *s* represent state variables (such as BC concentration) and control variables (such as emission sources, mainly including VOCs, NO_x, NH_3, SO_2 and $PPM_{2.5}$) in the model, respectively. First of all, the GRAPES-CUACE atmospheric chemistry model should be integrated to store the basic-state values of the unequilibrated variables in checkpoint files. The intermediate values are recalculated or saved in stack using PUSH&POP method during the adjoint operating process. Subsequently, the gradient of *J* with respect to *c* ($\nabla_c J$) as well as the saved basic-state values are taken as

input data for the adjoint backward integration. Finally, the sensitivity of J with respect to s ($\nabla_s J$) can be obtained. A full description of the construction, framework, and operational flowchart of the GRAPES-CUACE aerosol adjoint model can be found in An et al. (2016).

2.3 L-BFGS method

Limited-memory BFGS (L-BFGS) is an optimization algorithm in the family of quasi-Newton methods that approximates the Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS) using a limited amount of computer memory (Liu and Nocedal, 1989). The L-BFGS-B algorithm extends L-BFGS to solve large nonlinear optimization problems subject to simple bounds on the variables (Byrd et al., 1995; Zhu et al., 1997), which can be expressed as:

$$\min f(x), x \in R^n \tag{8.1}$$

$$\text{subject to } l \le x \le u \tag{8.2}$$

where f is a nonlinear function, the vectors l and u represent lower and upper bounds on the variables, and the number of variables n is assumed to be large. The algorithm is also appropriate and efficient for solving unconstrained problems in which the variables have no bounds. With the supply of the objective function f and its gradient g, but with no requirement of knowledge about the Hessian matrix of the objective function f, the algorithm can be useful for solving large problems where the Hessian matrix is difficult to compute or is dense.

The brief procedure of the L-BFGS-B algorithm is as follows. At each iteration, a limited memory BFGS approximation to the Hessian is updated. The limited memory BFGS matrix is used to define a quadratic model of the objective function f. A search direction d_k is computed by a two-stage approach. First, use the gradient projection method to identify a set of active variables, such as variables that will be held at their bounds. Then, the quadratic model is approximately minimized with respect to the free variables. The search direction is defined to be the vector leading from the current iterate to this approximate minimizer. Finally, a line search is performed along the search direction d_k to compute a step length λ_k, and the variables are updated through $x_{k+1} = x_k + \lambda_k d_k$. The L-BFGS-B algorithm has three termination criteria: the number of iterations reaches the set maximum value; the change of the objective function in consecutive iterations is relatively small; and the modulus of the projected gradient is small enough.

3 Description of GRAPES-CUACE-4D-Var

The new 4D-Var data assimilation system, GRAPES-CUACE-4D-Var, was constructed on the basis of GRAPES-CUACE atmospheric chemistry model, GRAPES-CUACE aerosol adjoint model and L-BFGS-B method. A schematic diagram of GRAPES-CUACE-4D-Var is shown in Fig. 2. The main parts of GRAPES-CUACE-4D-Var include GRAPES-CUACE atmospheric chemistry simulation, during which the basic-state values of the unequilibrated variables in checkpoint files are saved, observations and adjoint forcing term processing, GRAPES-CUACE aerosol adjoint model simulation, gradient

extraction, cost function calculation and optimization. The details of cost function, observations and optimization of emission inversion are as follows.

3.1 Cost function

Based on Bayesian theory and the assumption of Gaussian error distributions (Rodgers, 2000), the cost function of the emission inversion is generally defined as follow:

$$J(x) = \frac{1}{2}\gamma(x - x_{\mathrm{b}})^{\mathrm{T}}\mathrm{B}^{-1}(x - x_{\mathrm{b}}) + \frac{1}{2}\sum_{i=0}^{p}(F_i(x) - y_i)^{\mathrm{T}}\mathrm{R}^{-1}(F_i(x) - y_i) \quad (9)$$

where x, which we sought to optimize, generally represents the state vector of emissions or their scaling factors, x_{b} is the priori estimate of x, B is the error covariance estimate of x_{b}, F is the forward model, y is the vector of measurements that are distributed during the time interval [t_0, t_p], R is the observation error covariance matrix, and γ is the regularization parameter.

In this study, we followed the method in Henze et al. (2009), and defined x as the state vector of scaling factors of BC emissions:

$$x = \ln\left(\frac{s}{s_b}\right) \quad (10)$$

where s is the state vector of the daily gridded emissions of BC at three vertical levels (non-point source on the ground, middle elevation point source at 50m and high elevation point source at 120m), s_b is the priori estimate of s. Thus, the priori estimate of x (x_{b}) is equal to zero. According to Cao et al. (2011), the uncertainty of priori BC emissions used in this study is 76.2%. Therefore, we assigned the priori error covariance matrix (B) to be diagonal and the uncertainty to be 76.2% for BC emissions. Due to the lack of information to completely construct a physically representative B, the regularization parameter γ is introduced to balance the background and observation terms in the cost function. As described in Henze et al. (2009), an optimal value of γ can be found with the L-curve method (Hansen, 1998). Here, we followed this method, and obtained $\gamma = 0.01$ through several emission inversions with a range of γ (10, 1, 0.1, 0.01, 0.001, 0.0001, 0.00001).

3.2 Observations

The surface measurements of BC were collected from China Atmosphere Watch Network (CAWNET). The CAWNET was established by China Meteorological Administration to monitor the BC surface mass concentration over China in 2004, and had 54 monitoring stations in the summer of 2016. The monitoring of BC was conducted by an aethalometer (Model AE 31, Magee Scientific Co., USA), which uses a continuous optical gray scale measurement method to produce real-time BC data (Gong et al., 2019). In this study, we used the recommended mass absorption coefficient for the instrument at an 880-nm wavelength with 24-hour mean values of BC during July 1-31, 2016 at 5 representative stations of CAWNET in North China (Fig.S1).

The surface $PM_{2.5}$ concentrations were obtained from the public website of China Ministry of Ecology and Environment (MEE) (http://www.mee.gov.cn/). The network started to release real-time hourly concentrations of SO_2, NO_2, CO, ozone

(O_3), $PM_{2.5}$ and PM_{10} in 74 major Chinese cities since January 2013, which further increased to 338 cities in 2016. The $PM_{2.5}$ data were collected by the TEOM1405-F monitor, which draws ambient air through a sample filter at constant flow rate, continuously weighing the filter and calculates the near real-time mass concentration of the collected particulate matter. We used hourly surface $PM_{2.5}$ concentrations for July 1-31, 2016 at 48 cities in North China, including 12 cities in the Beijng-Tianjin-Hebei region (Fig.S1). Here, we have averaged $PM_{2.5}$ concentrations at several monitoring sites in each city to represent a regional condition.

To improve the performance of emission inversion, adequate observations are needed for constraining the model. Due to the limited BC monitoring sites in North China, we used the surface $PM_{2.5}$ concentrations at 48 cities described above and the $BC/PM_{2.5}$ ratio to obtain the hourly BC concentrations for July 1-31, 2016 at 48 cities in North China. The detailed calculation process can be found in the Supplement.

The observation error covariance matrix (R), which is difficult to quantify, generally includes contributions from the measurement error, the representation error, and the forward model error (Henze et al., 2009; Zhang et al., 2016; Cao et al., 2018). And there is also a certain error in calculating the BC concentration based on the $BC/PM_{2.5}$ ratio. To reflect the possibly large uncertainties of the observation, we assumed R to be diagonal and with error of 100%.

3.3 Optimization

Minimization of the cost function EQ. (9) is performed through optimization. Starting from an initial guess (x equal to zero), the forward model simulates BC concentrations at each integration step during the time interval [t_0, t_p], and the adjoint model which is driven by the discrepancy between simulated and observed BC concentrations calculates the gradients of the cost function with respect to the scaling factors of BC emission (Fig.2). Subsequently, the gradients are supplied to the L-BFGS-B optimization routine (Byrd et al., 1995; Zhu et al., 1997) to minimize the cost function iteratively (Fig.2). At each iteration, the improved estimates of the scaling factors are implemented and the forward and adjoint models are integrated.

3.4 Setup of emission inversion experiment

The simulation domain in this study is North China (105°E-125°E, 32.25°N-43.25°N, Fig. S1), covering 41 × 23 horizontal grids with resolution of 0.5° × 0.5°, and vertically divided into 31 layers with integration time step of 300 s. The FNL data at a 6-h interval is used as meteorological input. The priori emission used here is the daily gridded BC emission at three vertical levels (non-point source on the ground, middle elevation point source at 50m and high elevation point source at 120m) mentioned above. The assimilation window is from 20:00 Beijing Time (BJT) 3 July to 19:00 BJT 4 July 2016. The hourly BC concentrations at 36 cities during this time interval are adapted for the emission inversion, and the BC concentrations of the remaining 12 cities are used for validation of the inversion effect (Fig. S1). The simulation is initialized at 20:00 BJT 30 June, the first three days are set as the spin-up time. The convergence criterion used in the optimization is that the objective function decreases by less than 1% in consecutive iterations. And the scaling factors of BC emission have no upper and lower bounds here.

4 Results and discussion

5 Conclusions

285

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
290 adjoint model is available online at doi:10.5194/gmd-9-2153-2016-supplement. The code of optimization can be downloaded 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-4D-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
295 processed the data and prepared the data visualization. All authors reviewed the paper.

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
300 in optimization algorithm.

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.

Andre, J. C., Demoor, G., Lacarrere, P., Therry, G. and Duvachat, R.: Modeling the 24-hour evolution of the mean and turbulent structures of the planetary boundary layer. J. Atmos. Sci., 35, 1861-1883. doi: 10.1175/1520-0469(1978)035<1861: Mtheot>2.0.Co;2 ,1978.

Arakawa A., Lamb V. R.: Computational Design of the Basic Dynamical Processes of the UCLA General Circulation Model. Methods in Computational Physics, General Circulation Models of the Atmosphere. New York: Academic Press, 1997.

Betts, A.K., Miller, M.J.: A new convective adjustment scheme Part II: Single column tests using GATE wave, BOMEX, and arctic air-mass data sets. Q. J. R. Meteorol. Soc. 112, 693–709, doi: 10.1002/qj.49711247308, 1986.

Byrd, R.H., Lu, P., Nocedal, J. and Zhu, C.: A limited memory algorithm for bound constrained optimization. SIAM Journal on scientific computing, 16(5), pp.1190-1208. doi:10.1137/0916069, 1995.

Cao, G. L., Zhang, X. Y., Gong, S. L., An, X. Q., and Wang, Y. Q.: Emission inventories of primary particles and pollutant gases for China, Chinese Sci. Bull., 56, 781–788, https://doi.org/10.1007/s11434-011-4373-7, 2011.

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. 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.

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.

Chen, F., Mitchell, K., Schaake, J., Xue, Y. K., Pan, H. L., Koren, V., Duan, Q. Y., Ek, M. and Betts, A.: Modeling of land surface evaporation by four schemes and comparison with FIFE observations. Journal of Geophysical Research-Atmospheres, 101, 7251-7268. doi: 10.1029/95jd02165, 1996.

Elbern, H.; Schmidt,H. A four-dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling. J. Geophys. Res., Atmos., 104, 18583- 18598. doi:10.1029 / 1999JD900280, 1999.

Elbern, H.; Schmidt, H.; Talagrand, O.; Ebel, A. 4D-variational data assimilation with an adjoint air qualitymodel for emission analysis. Environ. Modell. Software, 15, 539-548, https://doi.org/10.1016/S1364-8152(00)00049-9, 2000.

Elbern, H.; Schmidt, H. Ozone episode analysis by four- dimensional variational chemistrydata assimilation. J. Geophys. Res., Atmos., 106, 3569-3590, 2001.

Elbern, H., Strunk, A., Schmidt, H. and Talagrand, O.: Emission rate and chemical state estimation by 4-dimensional variational inversion, Atmos. Chem. Phys., 7, 3749-3769, 2007.

Gong, S. L., Barrie, L. A., Blanchet, J.-P., Salzen, K. v., Lohmann,U., Lesins, G., Spacek, L., Zhang, L. M., Girard, E., and Lin, H.: Canadian Aerosol Module: A size-segregated simulation of atmospheric aerosol processes for climate and air quality models, 1, Module development, J. Geophys. Res., 108, 4007, https://doi.org/10.1029/2001JD002002, 2003.

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.

Gong, T., Sun, Z., Zhang X., Zhang, Y.,Wang, S., Han, L., Zhao, D., Ding, D., and Zheng, C.: Associations of black carbon and PM2.5 with daily cardiovascular mortality in Beijing, China[J]. Atmospheric Environment, 2019. doi:10.1016/j.atmosenv.2019.116876, 2019.

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.

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.

Haque, M. M., Fang, C., Schnelle-Kreis, J., Abbaszade, G., Liu ,X. Y., Bao, M. Y., Zhang, W. Q. and Zhang, Y. L.: Regional haze formation enhanced the atmospheric pollution levels in the Yangtze River Delta region, China: Implications for anthropogenic sources and secondary aerosol formation. Science of the Total Environment 728, doi:10.1016/j.scitotenv.2020.138013, 2020.

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 PM2.5 precursor emissions using the adjoint of GEOS-Chem, Atmos. Chem. Phys., 9, 5877-5903, doi: 10.5194/acp-9-5877-2009, 2009.

Hong, S. and Lim, J. J.: The WRF Single-Moment 6-Class Microphysics Scheme (WSM6). Asia-pac. J. Atmos. Sci., 42, 129-151, 2006.

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.

Janjić, Z.I.: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. Mon. weather rev., 122, 927-945, doi:10.1175 / 1520-0493（1994）122 <0927：TSMECM> 2.0.CO; 2, 1994.

Jeong, J. I. and Park R J.: Efficacy of dust aerosol forecasts for East Asia using the adjoint of GEOS-Chem with ground-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.

Ke H. 2019. Construction and application of a real-time emission model of open biomass burning. Master dissertation, Chinese Academy of Meteorological Sciences, Beijing, 1-57, 2019 (in Chinese).

Kurokawa, J. I., Yumimoto, K., Uno, I. and Ohara, T.: Adjoint inverse modeling of NOx 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.

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).

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.

Mellor, G. L., and Yamada, T.:A Hierarchy of Turbulence Closure Models for Planetary Boundary Layers. J. Atmos. Sci., 31, 1791-1806.doi: 10.1175/1520-0469(1974)031<1791:Ahotcm>2.0.Co;2 , 1974.

Monin, A.S. and Obukhov,A.M.: Basic laws of turbulent mixing in the surface layer of the atmosphere. Contrib. Geophys. Inst. Acad. Sci., USSR, (151), 163–187 (in Russian).

Müller, J. F. and Stavrakou, T.: Inversion of CO and NOx emissions using the adjoint of the IMAGES model, Atmos. Chem. Phys., 5,1157-1186, doi: 10.5194/acp-5-1157-2005, 2005.

Nenes, A., Pandis, S. N., and Pilinis, C.: ISORROPIA: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols. Aquatic geochemistry, 4(1), 123-152.doi:10.1023/a:1009604003981, 1998a.

Nenes, A., Pilinis, C. and Pandis, S.: Continued development and testing of a new thermodynamics aerosol module for urban and regional air quality models. Atmos. Environ. 33, 1553-1560.https://doi.org/10.1016/S1352-2310(98)00352-5, 1998b.

Resler, J., Eben, K., Jurus, P. and Liczki, J.: Inverse modeling of emissions and their time profiles. Atmos. Pollut. Res., 1, 288-295, doi:10.5094/apr.2010.036, 2010.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Sci., Singapore, 2000.

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.

Stockwell, W. R., Middleton, P., Change, J. S., and Tang, X.:The second generation regional acid deposition model chemical mechanism for regional air quality modeling. J. Geophys. Res. 95, 16343–16376, 1990.doi:10.1029/JD095iD10p16343, 1990.

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.

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 PM2.5 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.

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.

West, J. J., Pilinis, C., Nenes, A. and Pandis, S. N.: Marginal direct climate forcing by atmospheric aerosols. Atmos. Environ. 32, 2531-2542. doi:10.1016/s1352-2310(98)00003-x, 1998.

Xiang, S. L., Liu, J. F., Tao, W., Yi, K., Xu, J. Y. , Hu, X. R., Liu, H. Z., Wang, Y. Q., Zhang, Y. Z., Yang, H. Z., Hu, J. Y., Wan, Y., Wang, X. J., Ma, J. M., Wang, X. L. and S.: Control of both PM2.5 and O-3 in Beijing-Tianjin-Hebei and the surrounding areas. Atmos. Environ., 224, doi:10.1016/j.atmosenv.2019.117250, 2020.

Xu, G., Chen, D., Xue, J., Sun, J., Shen, X., Shen, Y., Huang, L., Wu, X., Zhang, H., and Wang, S.: The program structure de- signing and optimizing tests of GRAPES physics, Chinese Sci. Bull., 53, 3470–3476, https://doi.org/10.1007/s11434-008-0418-y, 2008.

Ye, Q., Shen, Y.: Practical Mathematical Manual, Science Press, Beijing, 2006 (in Chinese).

Yumimoto, K. and Uno, I.: Adjoint inverse modeling of CO emissions over Eastern Asia using four-dimensional variational data assimilation. Atmos. Environ., 40, 6836-6845, doi:10.1016/j.atmosenv.2006.05.042, 2006.

Zhai, S., An, X., Zhao, T., Sun, Z., Wang, W., Hou, Q., Guo, Z. and 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, doi: 10.5194/acp-18-6241-2018, 2018.

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.

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, R. and Shen, X.: On the development of the GRAPES – a new generation of the national operational NWP system in China, Chinese Sci. Bull., 53, 3429–3432, doi: 10.1007/s11434-008-0462-7, 2008.

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 PM2.5 mass reduction in key regions in China, Sci. China Earth Sci., 49, 1-18, doi: 10.1007/s11430-019-9343-3, 2019.

Zhao, Z. J., Liu, R., and Zhang, Z. Y.: Characteristics of Winter Haze Pollution in the Fenwei Plain and the Possible Influence of EU During 1984-2017. Earth Space Sci., 7, doi:10.1029/2020ea001134, 2020.

Zhou, C.H., Gong, S.L., Zhang, X.Y., Wang, Y.Q., Niu, T., Liu, H.L., Zhao, T.L., Yang, Y.Q. and Hou, Q.: Development and evaluation of an operational SDS forecasting system for East Asia: CUACE/Dust, Atmos. Chem. Phys., 8, 787–798, 2008.

Zhou, C. H., Gong, S. L., Zhang, X. Y., Liu, H. L., Xue, M., Cao, G. L., An, X. Q., Che, H. Z., Zhang, Y. M., and Niu, T.: Towards the improvements of simulating the chemical and optical properties of Chinese aerosols using an online coupled model – CUACE/Aero, Tellus B, 64, 18965, https://doi.org/10.3402/tellusb.v64i0.18965, 2012.

Zhu, C., Byrd, R.H., Lu, P. and Nocedal, J.: Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization, ACM T. Math. Software (TOMS), 23, 550-560, doi:10.1145/279232.279236, 1997.

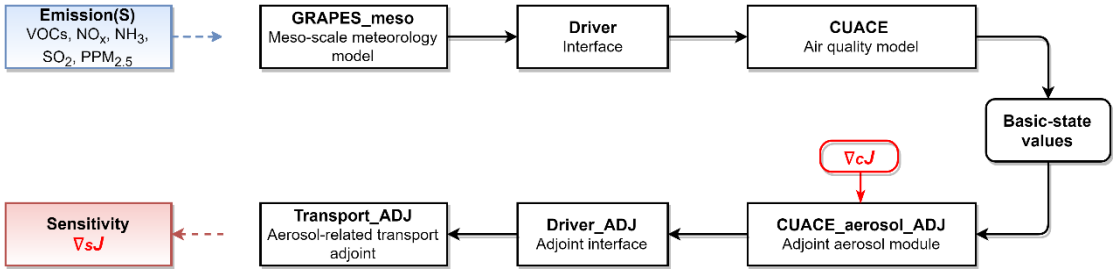

Figure 1: Running process of GRAPES-CUACE atmospheric chemistry model and its adjoint model.

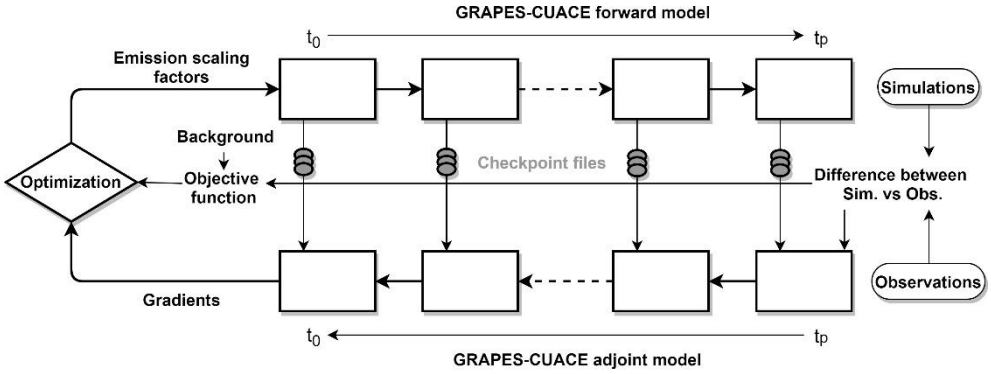

Figure 2: GRAPES-CUACE-4D-Var assimilation system.