



Development of four-dimensional variational assimilation system based on the GRAPES–CUACE adjoint model (GRAPES–CUACE-4D-Var V1.0) and its application in emission inversion

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

¹Institute of Atmospheric Composition and Environmental Meteorology, Chinese Academy of Meteorological Sciences, Beijing 100081, China

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

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

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

Received: 7 November 2019 – Discussion started: 3 February 2020

Revised: 21 November 2020 – Accepted: 2 December 2020 – Published: 22 January 2021

Abstract. In this study, a four-dimensional variational (4D-Var) data assimilation system was developed based on the GRAPES–CUACE (Global/Regional Assimilation and Prediction System – CMA Unified Atmospheric Chemistry Environmental Forecasting System) atmospheric chemistry model, GRAPES–CUACE adjoint model and L-BFGS-B (extended limited-memory Broyden–Fletcher–Goldfarb–Shanno) algorithm (GRAPES–CUACE-4D-Var) and was applied to optimize black carbon (BC) daily emissions in northern China on 4 July 2016, when a pollution event occurred in Beijing. The results show that the newly constructed GRAPES–CUACE-4D-Var assimilation system is feasible and can be applied to perform BC emission inversion in northern China. The BC concentrations simulated with optimized emissions show improved agreement with the observations over northern China with lower root-mean-square errors and higher correlation coefficients. The model biases are reduced by 20 %–46 %. The validation with observations that were not utilized in the assimilation shows that assimilation makes notable improvements, with values of the model biases reduced by 1 %–36 %. Compared with the prior BC emissions, which are based on statistical data of anthropogenic emissions for 2007, the optimized emissions are considerably reduced. Especially for Beijing, Tianjin, Hebei, Shandong, Shanxi and Henan, the ratios of the optimized emissions to prior emissions are 0.4–0.8, indicating that the BC emissions in these highly industrialized regions

have greatly reduced from 2007 to 2016. In the future, further studies on improving the performance of the GRAPES–CUACE-4D-Var assimilation system are still needed and are important for air pollution research in China.

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 condi-

tions and emissions (Liu, 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 and Schmidt (1999, 2001), Elbern et al. (2000, 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-experiment 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) applied the adjoint model of the STEM-2k1 model for assimilating black carbon (BC) concentrations and the recovery of its emissions. Liu (2005) and Huang et al. (2018) developed the adjoint of the 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 NO_x emissions for the year 1997. More recently, the CMAQ (Community Multiscale Air Quality Modeling System) team (Hakami et al., 2007) built the adjoint of CMAQ model and its 4D-Var assimilation scheme, which were used to optimize NO_x emissions (Kurokawa et al., 2009; Resler et al., 2010) and ozone initial state (Park et al., 2016). The adjoints of the GEOS-Chem model and its 4D-Var assimilation system first 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 (non-methane volatile organic compound) (Cao et al., 2018) emission estimates. Zhang et al. (2016) applied the 4D-Var assimilation system using the adjoint model of GEOS-Chem with the fine $1/4^\circ \times 5/16^\circ$ horizontal resolution to optimize daily aerosol primary and precursor emissions over northern China. This research has 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., 2020; 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 represents important basic data for atmospheric chemistry simulation, and its uncer-

tainty 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 optimized, 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 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 research on air pollution in China (Gong and Zhang, 2008; Zhou et al., 2008, 2012; Wang et al., 2010, 2015) but has also been officially in operation since 2014 at the 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 the GRAPES–CUACE model, which was subsequently applied in tracking influential BC and $\text{PM}_{2.5}$ source areas in northern 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 applied for assimilating surface BC concentrations and optimizing its daily emissions in northern China on 4 July 2016, when a pollution event occurred in Beijing. The following part is divided into four sections. Section 2 introduces the data and methods, Sect. 3 describes the GRAPES–CUACE-4D-Var assimilation system, Sect. 4 presents the results and discussions, and the conclusions are found in Sect. 5.

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 (Charney and Phillips, 1953). 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 microphysical 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 and Lim, 2006), the Betts–Miller–Janjic cumulus convection scheme (Betts and Miller, 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 and Obukhov, 1954), the MRF (medium-range forecast) planetary boundary layer scheme (Hong and Pan, 1996) 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, soil 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), BC, organic carbon (OC) and soil 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 or 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 sulfate aerosols and secondary organic aerosols (SOAs) and meanwhile generates an oxidation background for aqueous-phase aerosol chemistry, in which sulfate transformation changes the distributions of SO_2 in clouds (Zhou et al., 2012). Both nucleation and condensation are considered for sulfate aerosol formation depending on the atmospheric state after gaseous H_2SO_4 formed from the oxidation of sulfurous gases such as SO_2 , H_2S and DMS (dimethyl sulfide) (Zhou et al., 2012). Secondary organic aerosols as generated from gaseous precursors are partitioned into different bins through condensation using the same approach as the gaseous H_2SO_4 condensation to sulfate (Zhou et al., 2012). Given that the NIs 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, b; Zhou et al., 2012). ISORROPIA contains 15 equilibrium reactions, and the main species include the 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 ($(\text{NH}_4)_2\text{SO}_4$, NH_4HSO_4 , $(\text{NH}_4)_3\text{H}(\text{SO}_4)_2$, 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) mod-

ule into hourly gridded off-line data for 32 species, including BC, OC, SF, NI, fugitive dust particles and 19 non-methane volatile organic compounds (VOCs), CH₄, NH₃, CO, CO₂, SO_x and NO_x, at three vertical levels (non-point source on the ground, middle-elevation point source at 50 m and high-elevation point source at 120 m), as required by the GRAPES–CUACE model. Furthermore, natural sea salt and natural sand or dust emissions were also calculated online in the model (Zhou et al., 2012; Zhai et al., 2018).

2.2 Adjoint model

2.2.1 Adjoint theory

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

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

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

$$(\mathbf{L}\mathbf{x}, \mathbf{y}) = \mathbf{y}^T \mathbf{L}\mathbf{x} = \mathbf{x}^T \mathbf{L}^T \mathbf{y} = (\mathbf{x}, \mathbf{L}^T \mathbf{y}). \quad (2)$$

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

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

$$\mathbf{Y} = \mathbf{F}(\mathbf{X}), \quad (3)$$

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

$$\delta\mathbf{Y} = \nabla_{\mathbf{X}} \mathbf{F} \cdot \delta\mathbf{X}, \quad (4)$$

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

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

$$\delta\mathbf{X}^* = \nabla_{\mathbf{X}}^T \mathbf{F} \cdot \delta\mathbf{Y}^*, \quad (5)$$

where $\delta\mathbf{Y}^* \in R^m$ and $\delta\mathbf{X}^* \in R^n$ are input and output variables in the adjoint model, respectively. Comparing Eqs. (4) and (5), it can be seen that the dimensions of input and output are exchanged between the TLM and the adjoint 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 $n \gg m$ (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 the 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

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

where the denominator is the TLM output, and the numerator is the difference between the output value of the original model with input $\mathbf{X} + \delta\mathbf{X}$ and input \mathbf{X} . It is necessary to decrease the value of $\delta\mathbf{X}$ 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 \mathbf{X} and \mathbf{Y} . In Eq. (2), \mathbf{L} and \mathbf{L}^* represent the tangent linear process and the adjoint process, respectively. To simplify the testing process, the adjoint input is the tangent linear output: $\mathbf{Y} = \mathbf{L}(\mathbf{X})$. Thus, Eq. (5) can be expressed as

$$(\nabla\mathbf{F} \cdot d\mathbf{X}, \nabla\mathbf{F} \cdot d\mathbf{X}) = (d\mathbf{X}, \nabla^T \mathbf{F}(\nabla\mathbf{F} \cdot d\mathbf{X})). \quad (7)$$

By substituting $d\mathbf{X}$ into the tangent linear codes, the output value $\nabla\mathbf{F} \cdot d\mathbf{X}$ can be obtained and the left part of the equation can be computed. Then, taking $\nabla\mathbf{F} \cdot d\mathbf{X}$ as the input of the adjoint codes, the adjoint output $\nabla^T \mathbf{F}(\nabla\mathbf{F} \cdot d\mathbf{X})$ 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 problems concerned. 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 $\text{PPM}_{2.5}$) 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 the PUSH&POP method, which pushes the intermediate values into a continuous memory space and pops them out where needed, 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-B method

The limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm (L-BFGS) is an optimization algorithm in the family of quasi-Newton methods that approximates the 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

$$\min f(\mathbf{x}), \mathbf{x} \in R^n, \quad (8a)$$

$$\text{subject to } \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \quad (8b)$$

where f is a nonlinear function, the vectors \mathbf{l} and \mathbf{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 \mathbf{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 approxima-

tion 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 \mathbf{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 \mathbf{d}_k to compute a step length λ_k , and the variables are updated through $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \mathbf{d}_k$. 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 the GRAPES–CUACE atmospheric chemistry model, the GRAPES–CUACE aerosol adjoint model and the 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 follows:

$$J(\mathbf{x}) = \frac{1}{2} \gamma (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} \sum_{i=0}^p (\mathbf{F}_i(\mathbf{x}) - \mathbf{y}_i)^T \mathbf{R}^{-1} (\mathbf{F}_i(\mathbf{x}) - \mathbf{y}_i), \quad (9)$$

where \mathbf{x} , which we sought to optimize, generally represents the state vector of emissions or their scaling factors, \mathbf{x}_b is the prior estimate of \mathbf{x} , \mathbf{B} is the error covariance estimate of \mathbf{x}_b , \mathbf{F} is the forward model, \mathbf{y} is the vector of measurements that are distributed during the time interval $[t_0, t_p]$, \mathbf{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 \mathbf{x} as the state vector of scaling fac-

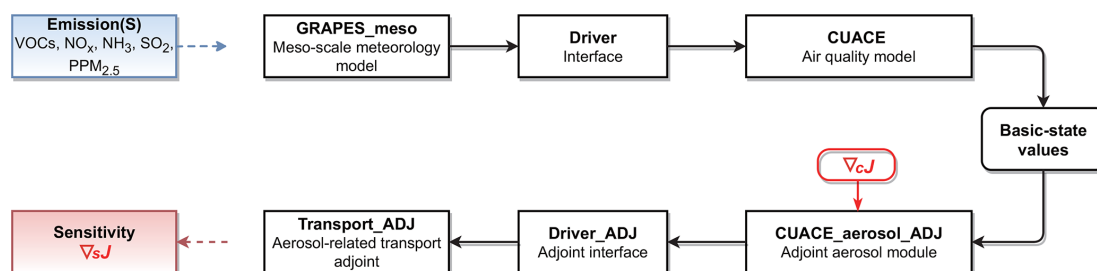


Figure 1. Running process of GRAPES–CUACE atmospheric chemistry model and its adjoint model.

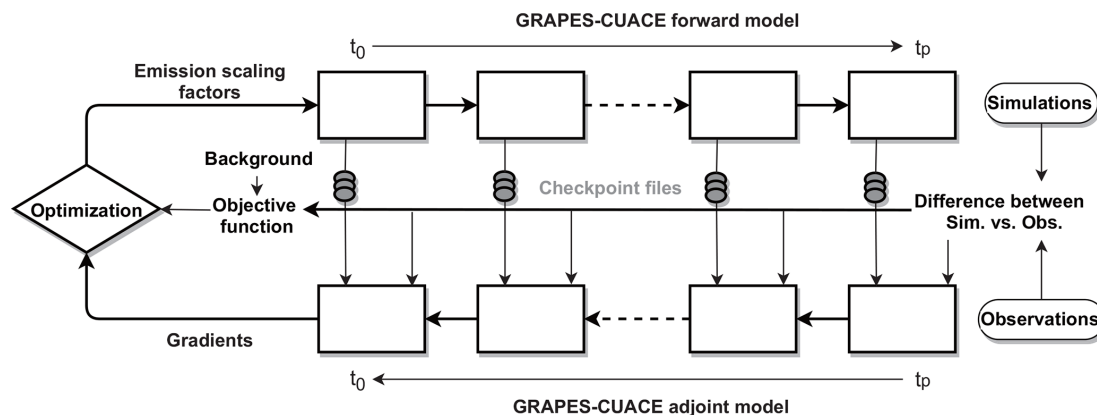


Figure 2. GRAPES–CUACE-4D-Var assimilation system.

tors of BC emissions:

$$\mathbf{x} = \ln \left(\frac{\mathbf{s}}{\mathbf{s}_b} \right), \quad (10)$$

where \mathbf{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 50 m and high-elevation point source at 120 m) and \mathbf{s}_b is the prior estimate of \mathbf{s} . Thus, the prior estimate of $\mathbf{x}(\mathbf{x}_b)$ is equal to 0. According to Cao et al. (2011), the uncertainty of prior BC emissions used in this study is 76.2 %. Therefore, we assigned the prior error covariance matrix (\mathbf{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 \mathbf{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 $\gamma = 0.0001$ through several emission inversions with a range of γ (10, 1, 0.1, 0.01, 0.001, 0.0001, 0.00001, 0.000001, 0.0000001).

3.2 Observations

The surface measurements of BC were collected from the China Atmosphere Watch Network (CAWNET). The CAWNET was established by the 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 greyscale 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 h mean values of BC during 1–31 July 2016 at five representative stations of CAWNET in northern China (Fig. S1 in the Supplement).

The surface $\text{PM}_{2.5}$ concentrations were obtained from the public website of the China Ministry of Ecology and Environment (MEE) (<http://www.mee.gov.cn/>, last access: 14 January 2021). The network started to release real-time hourly concentrations of SO_2 , NO_2 , CO, ozone (O_3), $\text{PM}_{2.5}$ and PM_{10} in 74 major Chinese cities in January 2013, which further increased to 338 cities in 2016. The $\text{PM}_{2.5}$ 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 $\text{PM}_{2.5}$ concentrations for 1–31 July 2016 at 48 cities in northern China, including 12 cities in the Beijing–Tianjin–Hebei region (Fig. S1). Here, we have averaged $\text{PM}_{2.5}$ 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 northern China, we used the surface $\text{PM}_{2.5}$ concentrations at 48 cities described above and the $\text{BC}/\text{PM}_{2.5}$ ratio to obtain the hourly BC concentrations for 1–31 July 2016 at 48 cities in northern China. The detailed calculation process can be found in the Supplement.

The observation error covariance matrix (\mathbf{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 $\text{BC}/\text{PM}_{2.5}$ ratio. To reflect the possibly large uncertainties of the observation, we assumed \mathbf{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 (\mathbf{x} equal to 0), 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 northern China (105° – 125° E, 32.25° – 43.25° N; Fig. S1), covering 41×23 horizontal grids with a resolution of $0.5^{\circ} \times 0.5^{\circ}$ and vertically divided into 31 layers with an integration time step of 300 s. The National Centers for Environmental Prediction Final (FNL) Analysis dataset at a 6 h interval is used as meteorological input. The prior emission used here is the daily gridded BC emission at three vertical levels (non-point source on the ground, middle-elevation point source at 50 m and high-elevation point source at 120 m) mentioned above. The results calculated by the $\text{BC}/\text{PM}_{2.5}$ ratio show that the BC concentration in Beijing was high on 4 July 2016. So the assimilation window is from 20:00 CST, 3 July, to 19:00 CST, 4 July 2016. The hourly BC concentrations at 36 cities during this time interval are used 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 CST, 30 June; the first 3 d are set as the spin-up time. The convergence criterion used in the optimization is that the objective function decreases by less

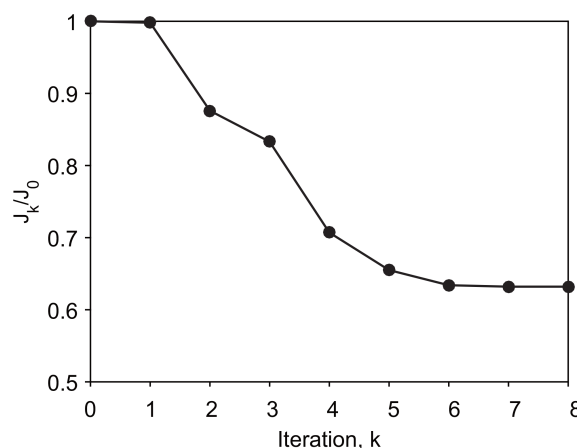


Figure 3. Cost function reduction.

than 1 % in consecutive iterations. According to the maximum estimation range of the prior emissions, here the upper and lower bounds of the scaling factors of BC emissions are $\ln(1.6)$ and $\ln(0.4)$, respectively.

4 Results and discussion

4.1 Comparisons between the simulated and observed concentrations

The progression of the cost function at iteration k (J_k/J_0) during the optimization procedure is shown in Fig. 3. The cost function quickly reduces and reaches the convergence criterion after eight iterations, with values of the converged cost function reduced by 37 %.

Figure 4 shows the spatial distribution of observed and simulated daily BC concentrations on 4 July 2016. In general, the results simulated with the prior emission reflect the distribution characteristics of BC concentration in northern China to a certain extent, with high values mainly located in Beijing and central and southern Hebei and low values mainly located in Inner Mongolia and eastern Shandong. However, the differences between the simulated and the observed BC concentrations are considerable, and almost all are over-predictions. The optimized (posterior) emissions compensate for the over-predictions and largely reduce the model biases. For instance, the model biases for BC in Beijing, Tianjin, Shijiazhuang, Jinan, Taiyuan and Zhengzhou are reduced by 46 % (from 5.4 to 2.9 $\mu\text{g}/\text{m}^3$), 26 % (from 6.6 to 4.9 $\mu\text{g}/\text{m}^3$), 29 % (from 18.4 to 13.1 $\mu\text{g}/\text{m}^3$), 20 % (from 6.6 to 5.3 $\mu\text{g}/\text{m}^3$), 34 % (from 4.1 to 2.7 $\mu\text{g}/\text{m}^3$) and 20 % (from 6.9 to 5.5 $\mu\text{g}/\text{m}^3$), respectively (Fig. 4a, b). The results simulated with the optimized emission also show improved agreement with the observations over northern China with lower root-mean-square errors and higher correlation coefficients (Fig. 4a, b).

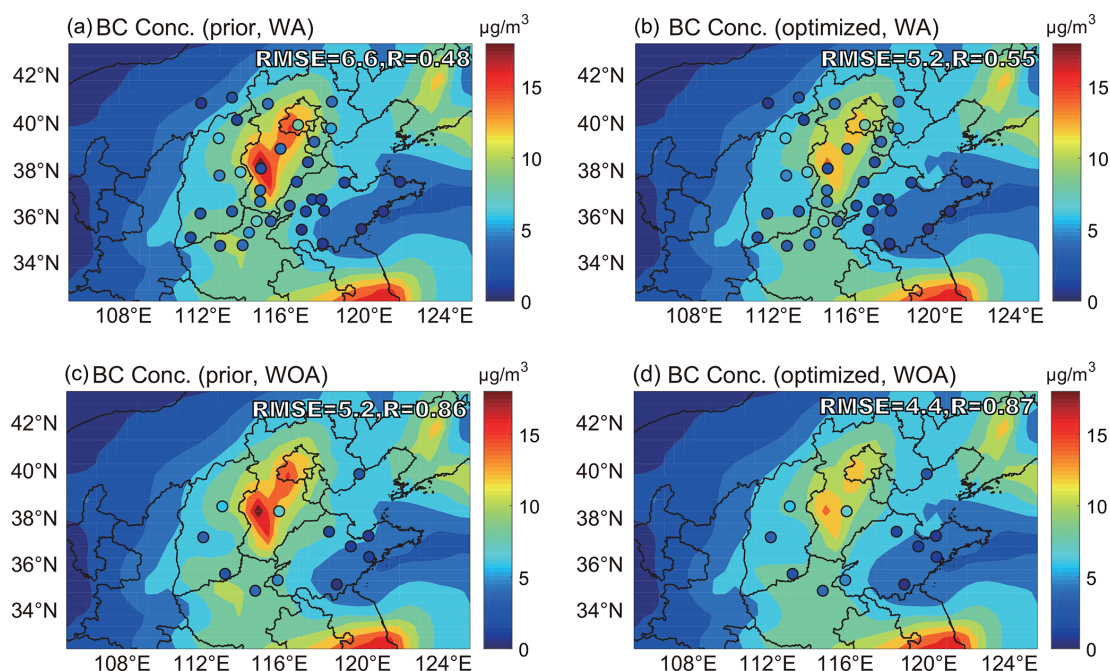


Figure 4. The spatial distribution of observed and simulated daily BC concentrations on 4 July 2016. The observations (circles) are over-plotted over model simulations with the (a, c) prior and (b, d) optimized emissions. The observations at 36 cities in (a) and (b) were used in the assimilation, and the observations at 12 cities in (c) and (d) were not used in the assimilation. The root-mean-square error (RMSE) and correlation coefficient (R) between observation and simulation are shown as insets. The observed BC concentrations were calculated by the BC/PM_{2.5} ratio method.

It is crucial to validate the assimilation results by observations that were not utilized in the assimilation. The BC concentrations at 12 cities were used for validation (Fig. 4c, d). Assimilation compensates for over-predictions, reduces the root-mean-square errors (from 5.2 to 4.4) and improves the correlation coefficients (from 0.86 to 0.87). The model biases for BC in Hengshui and Yizhou are reduced by 28 % (from 7.2 to 5.2 $\mu\text{g}/\text{m}^3$) and 36 % (from 3.9 to 2.5 $\mu\text{g}/\text{m}^3$), respectively. The improvements of the remaining 10 cities are also notable, with values of the model biases reduced by 1 %–20 %.

4.2 Comparisons between the prior and optimized BC emissions

Figure 5 shows the spatial distributions of the prior and optimized daily BC emissions, which are at three vertical levels (non-point source on the ground, middle-elevation point source at 50 m and high-elevation point source at 120 m) as required by the GRAPES–CUACE model. The BC emissions on the ground are mostly non-point sources and at 50 m are concentrated middle-elevation sources; therefore they are distributed in a large area (Fig. 5a, c), while the BC emissions at 120 m are mainly from a few high-elevation point sources, so they are scattered in the area (Fig. 5e). It can be seen that the distributions of optimized BC emissions at three vertical levels are relatively consistent with those of prior emissions

(Fig. 5b, d, f). However, the optimized emissions are considerably reduced (Fig. 5b, d, f). Especially for the regions where observation sites are located, such as southern Beijing, Tianjin, central and southern Hebei, northwest Shandong, central Shanxi, and northern Henan, BC emissions decrease significantly. As for the regions where observation sites are not located, such as Liaoning and Jiangsu, BC emissions are almost unchanged. The reason for this phenomenon is discussed in Sect. 4.3.

In order to analyse the differences between the prior and optimized BC emissions in Beijing–Tianjin–Hebei–Shanxi–Shandong–Henan region in more depth and detail, we calculated the ratio of the optimized emissions to prior emissions (optimized emissions divide by prior emissions), as shown in Fig. 6. In general, assimilation has the largest reduction in the non-point source on the ground (Fig. 6a), followed by the middle-elevation point source at 50 m (Fig. 6b), and the smallest reduction in the high-elevation point source at 120 m (Fig. 6c). This is related to the intensity of BC emissions at three vertical levels. The intensity of BC emissions on the ground (Fig. 5a) is about 1–2 orders of magnitude higher than that of the middle-elevation point source at 50 m (Fig. 5c) and the high-elevation point source at 120 m (Fig. 5e). In other words, the BC emissions on the ground have the most significant effect on the overall cost function and therefore reduced most during the progression of assimilation.

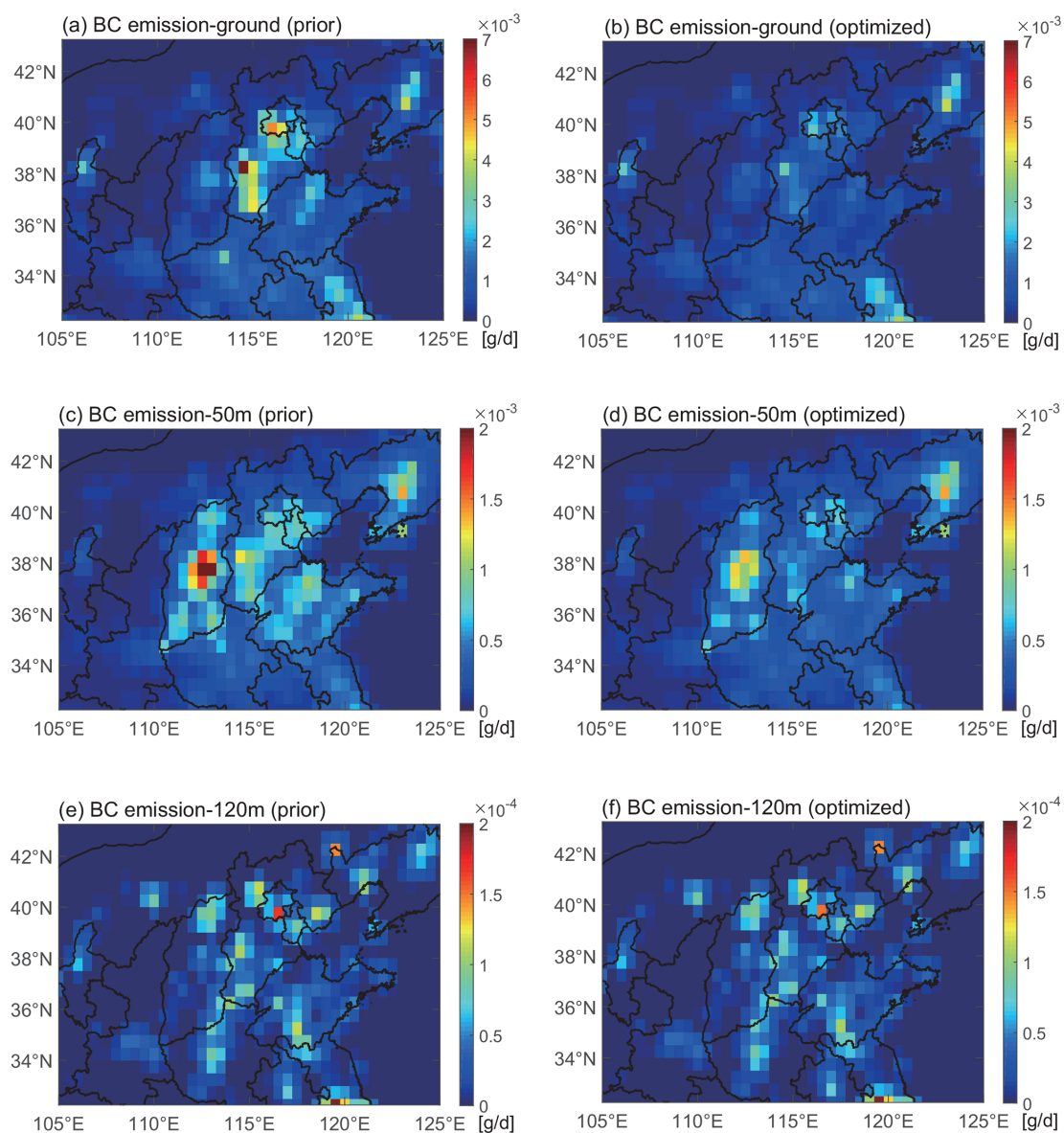


Figure 5. Spatial distributions of the (a, c, e) prior and (b, d, f) optimized daily BC emissions. The emissions are at three vertical levels: (a, b) non-point source on the ground, (c, d) middle-elevation point source at 50 m and (e, f) high-elevation point source at 120 m, as required by the GRAPES–CUACE model.

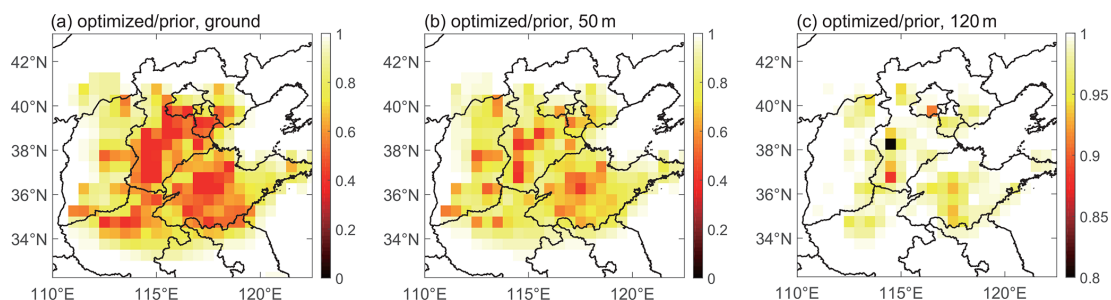


Figure 6. The ratio of optimized emissions to prior emissions (optimized emissions divide by prior emissions). (a) Non-point source on the ground, (b) middle-elevation point source at 50 m and (c) high-elevation point source at 120 m.

Table 1. Anthropogenic BC emissions in China by province in 2006^a and 2016^b (units: Gg/year).

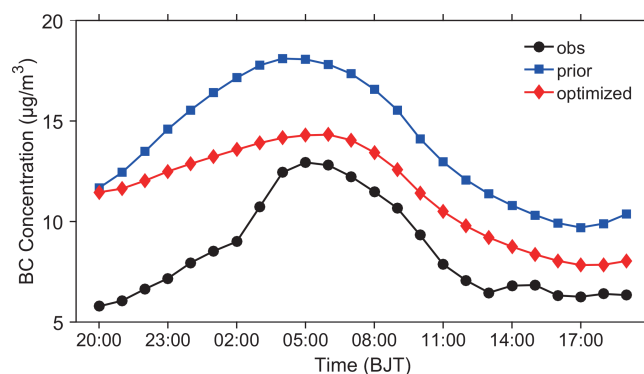
Province	2006	2016
Beijing	19.00	7.70 (0.41)
Tianjin	15.00	10.81 (0.72)
Hebei	137.00	97.99 (0.72)
Shanxi	139.00	69.26 (0.50)
Shandong	132.00	108.82 (0.82)
Henan	133.00	82.42 (0.62)
Total China	1811.00	1315.28 (0.73)

Note: numbers in parentheses represent emission ratios relative to 2006. ^a Source: Zhang et al. (2009). ^b Source: MEIC (Multi-resolution Emission Inventory for China), <http://meicmodel.org/> (last access: 14 January 2021).

The prior BC emissions used in this study are based on statistical data of anthropogenic emissions for 2007 (Cao et al., 2011), and the BC observations used for assimilation are from 2016, so the ratios of the optimized emissions to prior emissions can reflect the changes in BC emissions from 2007 to 2016 to a certain extent. From the perspective of each province, we can see that the ratios of the optimized emissions to prior emissions in Beijing, Tianjin, Hebei, Shanxi, Shandong and Henan are 0.4–0.8, 0.4–0.7, 0.4–0.8, 0.6–0.8, 0.4–0.8 and 0.5–0.8, respectively. This indicates that the BC emissions in these highly industrialized regions have greatly reduced from 2007 to 2016, which is consistent with previous studies (Zheng et al., 2018). Table 1 lists the anthropogenic BC emissions in China by province in 2006 and 2016 and the emission ratios of 2016 relative to 2006. According to previous research, the emission ratios of 2016 relative to 2006 are 0.41–0.82 in Beijing–Tianjin–Hebei–Shanxi–Shandong–Henan region and 0.73 over China. It can be seen that the ratios of the optimized emissions to prior emissions calculated in this study are within a reasonable range, which also shows that the newly constructed GRAPES–CUACE-4D-Var assimilation system can obtain reasonable BC emissions based on the observations.

4.3 Discussion

Although Sect. 4.1 shows that assimilation reduces the model biases and improves all statistical values at each site, there are still over-predictions to a certain degree. In addition to the emission, the initial concentration is also an important factor that affects the BC simulation. Here, we take Beijing as an example to analyse the influence of the initial concentration on the BC simulation. Figure 7 shows the comparison of observed and simulated (with the prior and optimized BC emissions, respectively) BC concentrations in Beijing from 20:00, 3 July, to 19:00, 4 July 2016. At the initial moment (20:00, 3 July 2016), the initial concentration for simulation ($11.5 \mu\text{g}/\text{m}^3$) was about 2 times higher than the

**Figure 7.** Comparison of observed and simulated (with the prior and optimized BC emissions, respectively) BC concentrations in Beijing from 20:00, 3 July, to 19:00, 4 July 2016. The observed BC concentrations were calculated by the BC/PM_{2.5} ratio method.

observations ($5.7 \mu\text{g}/\text{m}^3$). With such a high initial concentration, the simulated BC concentrations with prior emissions were significantly higher than the observed concentrations at all times. The simulations with optimized emissions compensated for the over-predictions, but the BC concentrations were still higher than the observations in the first few hours (from 20:00, 3 July, to 02:00, 4 July 2016). And the model bias during this time period was $5.2 \mu\text{g}/\text{m}^3$. As the influence of the initial concentration on the simulation gradually weakened, the simulated BC concentrations with optimized emissions were largely reduced and much closer to the observations from 03:00 to 19:00, 4 July 2016, and the model biases during this time period also decreased, with the value of $1.9 \mu\text{g}/\text{m}^3$. This indicates that for short-term simulation, the influence of initial concentration on the simulation is not negligible. In further work, it is also of great significance to use the GRAPES–CUACE-4D-Var assimilation system to optimize the initial concentration to improve the simulation effect.

As described in Sect. 4.2, after assimilation, the BC emissions in the regions where observation sites are located decrease significantly, while in the regions where observation sites are not located, such as Liaoning and Jiangsu, BC emissions are almost unchanged. This may be due to the short period of the observation data used for assimilation. In this study, since the GRAPES–CUACE-4D-Var assimilation system has not yet implemented parallel computing, the assimilation window was set for 24 h. In such a short period of time, the pollutants emitted from Liaoning and Jiangsu may not have been transported to the regions where observation sites are located, thus having little impact on the BC concentrations in these regions. In other words, the emissions from Liaoning and Jiangsu have little effect on the overall cost function, so there is little change during the progression of assimilation. Therefore, implementing parallel computing of the GRAPES–CUACE-4D-Var assimilation system and per-

forming emission inversion for a longer period (i.e. 1 month) is another important task in the future.

5 Conclusions

In this study, we developed a new 4D-Var data assimilation system for the GRAPES–CUACE atmospheric chemistry model (GRAPES–CUACE-4D-Var) and applied it for assimilating surface BC concentrations and optimizing its daily emissions in northern China on 4 July 2016, when a pollution event occurred in Beijing. The main conclusions are as follows.

- The newly constructed GRAPES–CUACE-4D-Var assimilation system is feasible and can be applied to perform BC emission inversion in northern China.
- The BC concentrations simulated with optimized emissions show improved agreement with the observations over northern China with lower root-mean-square errors and higher correlation coefficients. The model biases are reduced by 20 %–46 %. The validation of assimilation results with observations that were not utilized in the assimilation shows that assimilation compensates for over-predictions, reduces the root-mean-square errors and improves the correlation coefficients. The improvements are also notable, with values of the model biases reduced by 1 %–36 %.
- Compared with the prior BC emissions, the optimized emissions are considerably reduced. Especially for Beijing, Tianjin, Hebei, Shandong, Shanxi and Henan, the ratios of the optimized emissions to prior emissions are 0.4–0.8, indicating that the BC emissions in these highly industrialized regions have greatly reduced from 2007 to 2016, which is consistent with previous studies.

In the following work, implementing parallel computing of the GRAPES–CUACE-4D-Var assimilation system and performing emission inversion for a longer period is an important task. Apart from the emissions, the initial concentration is also an important factor for short-term simulation. It is of great significance to use the GRAPES–CUACE-4D-Var assimilation system to optimize the initial concentration to improve the simulation effect.

Meanwhile, several mega urban agglomerations in China are facing atmospheric compound pollution with high PM_{2.5} and O₃ concentrations (Li et al., 2019; Zhang et al., 2019; Xi et al., 2020; Haque et al., 2020; Zhao et al., 2020). To improve air quality, it is urgent to formulate reasonable and effective emission-reduction measures. Therefore, further studies on expanding the function of the GRAPES–CUACE-4D-Var assimilation system and taking into account factors such as air quality standards, the proportion of emissions that can be reduced, the economic cost and residents' health benefits

of emission reduction are crucial for formulating optimized pollution-control strategies for PM_{2.5} and O₃ in China.

Code and data availability. The GRAPES–CUACE atmospheric chemistry model used in this study was distributed by the National Meteorological Center of the Chinese Meteorology Administration (2021, <http://www.nmc.cn>) together with the Institute of Atmospheric Composition and Environmental Meteorology of the Chinese Academy of Meteorological Sciences (2021, <http://www.camscma.cn>). The model was run on an IBM PureFlex System (AIX) with an XL Fortran Compiler. The code of the GRAPES–CUACE aerosol adjoint model is available online at <https://doi.org/10.5194/gmd-9-2153-2016-supplement> (An et al., 2016). The code of GRAPES_CUACE_4D_Var_driver.F can be downloaded as a Supplement to this article. The observations are available online at <http://www.mee.gov.cn/> (Ministry of Ecology and Environment of the People's Republic of China, 2021).

Supplement. The supplement related to this article is available online at: <https://doi.org/10.5194/gmd-14-337-2021-supplement>.

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 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. We thank Lin Zhang from the Numerical Forecast Center of the China Meteorological Administration for providing technical support with the optimization algorithm.

Financial support. This research has been supported by the National Key Research and Development Program of China (grant no. 2017YFC0210006) and the National Natural Science Foundation of China (grant nos. 41975173 and 91644223).

Review statement. This paper was edited by Augustin Colette and reviewed by two anonymous referees.

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, <https://doi.org/10.5194/gmd-9-2153-2016>, 2016.
- Andre, J. C., Demoor, G., Lacarrere, P., Therry, G., and Duvauchet, R.: Modeling the 24-hour evolution of the mean and turbulent structures of the planetary boundary layer, *J. Atmos. Sci.*, 35, 1861–1883, [https://doi.org/10.1175/1520-0469\(1978\)035<1861:Mtheot>2.0.CO;2](https://doi.org/10.1175/1520-0469(1978)035<1861:Mtheot>2.0.CO;2), 1978.
- Betts, A. K. and 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. Roy. Meteor. Soc.*, 112, 693–709, <https://doi.org/10.1002/qj.49711247308>, 1986.
- Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C.: A limited memory algorithm for bound constrained optimization, *SIAM J. Sci. Comput.*, 16, 1190–1208, <https://doi.org/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., De Smedt, I., Zhang, Q., van Roozendaal, M., Hendrick, F., Chance, K., Li, J., Zheng, J., and Zhao, Y.: Adjoint inversion of Chinese non-methane volatile organic compound emissions using space-based observations of formaldehyde and glyoxal, *Atmos. Chem. Phys.*, 18, 15017–15046, <https://doi.org/10.5194/acp-18-15017-2018>, 2018.
- Charney, J. G. and Phillips, N. A.: Numerical integration of the quasi-geostrophic equations for barotropic and simple baroclinic flows, *J. Meteorol.*, 10, 71–99, [https://doi.org/10.1175/1520-0469\(1953\)010<0071:NIOTQG>2.0.CO;2](https://doi.org/10.1175/1520-0469(1953)010<0071:NIOTQG>2.0.CO;2), 1953.
- 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, *Chinese Sci. Bull.*, 53, 3433–3445, <https://doi.org/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, *J. Geophys. Res.-Atmos.*, 101, 7251–7268, <https://doi.org/10.1029/95jd02165>, 1996.
- Dudhia, J.: Numerical study of convection observed during the winter monsoon experiment using a meso-scale two-dimensional model, *J. Atmos. Sci.*, 46, 3077–3107, [https://doi.org/10.1175/1520-0469\(1989\)046<3077:NSOCOD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2), 1989.
- Elbern, H. and Schmidt, H.: A four-dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling, *J. Geophys. Res.-Atmos.*, 104, 18583–18598, <https://doi.org/10.1029/1999JD900280>, 1999.
- Elbern, H. and Schmidt, H.: Ozone episode analysis by four-dimensional variational chemistry data assimilation, *J. Geophys. Res.-Atmos.*, 106, 3569–3590, 2001.
- Elbern, H., Schmidt, H., Talagrand, O., and Ebel, A.: 4D-variational data assimilation with an adjoint air quality model for emission analysis, *Environ. Modell. Softw.*, 15, 539–548, [https://doi.org/10.1016/S1364-8152\(00\)00049-9](https://doi.org/10.1016/S1364-8152(00)00049-9), 2000.
- Elbern, H., Strunk, A., Schmidt, H., and Talagrand, O.: Emission rate and chemical state estimation by 4-dimensional variational inversion, *Atmos. Chem. Phys.*, 7, 3749–3769, <https://doi.org/10.5194/acp-7-3749-2007>, 2007.
- 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, <https://doi.org/10.5194/acp-8-2333-2008>, 2008.
- 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, T., Sun, Z., Zhang, X., Zhang, Y., Wang, S., Han, L., Zhao, D., Ding, D., and Zheng, C.: Associations of black carbon and PM_{2.5} with daily cardiovascular mortality in Beijing, China, *Atmos. Environ.*, 214, 116876, <https://doi.org/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, D14301, <https://doi.org/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, <https://doi.org/10.1021/es070944p>, 2007.
- Hansen, P. C.: Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion, Society for Industrial and Applied Mathematics, Philadelphia, USA, 1998.
- 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, *Sci. Total Environ.*, 728, 138013, <https://doi.org/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, <https://doi.org/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} precursor emissions using the adjoint of GEOS-Chem, *Atmos. Chem. Phys.*, 9, 5877–5903, <https://doi.org/10.5194/acp-9-5877-2009>, 2009.
- Hong, S. and Lim, J. J.: The WRF Single-Moment 6-Class Microphysics Scheme (WSM6), *Asia-pac. J. Atmos. Sci.*, 42, 129–151, 2006.
- Hong, S. Y. and Pan, H. L.: Nonlocal boundary layer vertical diffusion in a medium-range forecast model, *Mon. Weather Rev.*, 124, 2322–2339, [https://doi.org/10.1175/1520-0493\(1996\)124<2322:NBLVDI>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<2322:NBLVDI>2.0.CO;2), 1996.
- Huang, S. X., Liu, F., Sheng, L., Cheng, L. J., Wu, L., and Li, J.: On adjoint method based atmospheric emission source tracing, *Chinese Sci. Bull.*, 63, 1594–1605, <https://doi.org/10.1360/N972018-00196>, 2018.
- Janjić, Z. I.: The step-mountain eta coordinate model: Further developments of the convection, viscous sub-

- layer, and turbulence closure schemes, *Mon. Weather Rev.*, 122, 927–945, [https://doi.org/10.1175/1520-0493\(1994\)122<0927:TSMECM>2.0.CO;2](https://doi.org/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, <https://doi.org/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, <https://doi.org/10.5194/acp-15-1521-2015>, 2015.
- Ke, H.: 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 NO_x emissions over eastern China using satellite observations of NO₂ vertical column densities, *Atmos. Environ.*, 43, 1878–1887, <https://doi.org/10.1016/j.atmosenv.2008.12.030>, 2009.
- Li, K., Jacob, D. J., Liao, H., Shen, L., Zhang, Q., and Bates, K. H.: Anthropogenic drivers of 2013–2017 trends in summer surface ozone in China, *P. Natl. Acad. Sci. USA*, 116, 422–427, <https://doi.org/10.1073/pnas.1812168116>, 2019.
- Liu, D. C. and Nocedal, J.: On the limited memory BFGS method for large scale optimization, *Math. Program.*, 45, 503–528, <https://doi.org/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, <https://doi.org/10.5194/acp-15-7685-2015>, 2015.
- Ministry of Ecology and Environment of the People's Republic of China: Air Quality, available at: <http://www.mee.gov.cn/>, last access: 15 January 2021.
- Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, *J. Geophys. Res.-Atmos.*, 102, 16663–16682, <https://doi.org/10.1029/97JD00237>, 1997.
- 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, 1954 (in Russian).
- Müller, J.-F. and Stavrou, T.: Inversion of CO and NO_x emissions using the adjoint of the IMAGES model, *Atmos. Chem. Phys.*, 5, 1157–1186, <https://doi.org/10.5194/acp-5-1157-2005>, 2005.
- Nenes, A., Pandis, S. N., and Pilinis, C.: ISORROPIA: A new thermodynamic equilibrium model for multiphase multi-component inorganic aerosols, *Aquat. Geochem.*, 4, 123–152, <https://doi.org/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](https://doi.org/10.1016/S1352-2310(98)00352-5), 1998b.
- Park, S. Y., Kim, D. H., Lee, S. H., and Lee, H. W.: Variational data assimilation for the optimized ozone initial state and the short-time forecasting, *Atmos. Chem. Phys.*, 16, 3631–3649, <https://doi.org/10.5194/acp-16-3631-2016>, 2016.
- Resler, J., Eben, K., Jurus, P., and Liczki, J.: Inverse modeling of emissions and their time profiles, *Atmos. Pollut. Res.*, 1, 288–295, <https://doi.org/10.5094/apr.2010.036>, 2010.
- Rodgers, C. D.: Inverse methods for atmospheric sounding—Theory and practice, Ser. on Atmos. Oceanic and Planet. Phys., Vol. 2, Singapore, <https://doi.org/10.1142/9789812813718>, 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, <https://doi.org/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, <https://doi.org/10.1029/JD095iD10p16343>, 1990.
- The Chinese Academy of Meteorological Sciences: Scientific research, available at: <http://www.camscma.cn/>, last access: 15 January 2021.
- The National Meteorological Center: Numerical forecast, available at: <http://www.nmc.cn/>, last access: 15 January 2021.
- 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, <https://doi.org/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, <https://doi.org/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 PM_{2.5} concentration in Beijing, *Atmos. Res.*, 215, 165–177, <https://doi.org/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, *Chinese Sci. Bull.*, 55, 635–649, <https://doi.org/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, <https://doi.org/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, <https://doi.org/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, [https://doi.org/10.1016/S1352-2310\(98\)00003-x](https://doi.org/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

- Tao, S.: Control of both PM_{2.5} and O₃ in Beijing-Tianjin-Hebei and the surrounding areas, *Atmos. Environ.*, 224, 117259, <https://doi.org/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 designing and optimizing tests of GRAPES physics, *Chinese Sci. Bull.*, 53, 3470–3476, <https://doi.org/10.1007/s11434-008-0418-y>, 2008.
- Ye, Q. and 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, <https://doi.org/10.1016/j.atmosenv.2006.05.042>, 2006.
- Zhai, S., An, X., Liu, Z., Sun, Z., and Hou, Q.: Model assessment of atmospheric pollution control schemes for critical emission regions, *Atmos. Environ.*, 124, 367–377, <https://doi.org/10.1016/j.atmosenv.2015.08.093>, 2016.
- 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, 6241–6258, <https://doi.org/10.5194/acp-18-6241-2018>, 2018.
- 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, <https://doi.org/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., Chen, D., Duan, L., Lei, Y., Wang, L. T., and Yao, Z. L.: Asian emissions in 2006 for the NASA INTEX-B mission, *Atmos. Chem. Phys.*, 9, 5131–5153, <https://doi.org/10.5194/acp-9-5131-2009>, 2009.
- 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, <https://doi.org/10.1007/s11434-008-0462-7>, 2008.
- Zhang, X., Xu, X., Ding, Y., Liu, Y., Zhang, H., Wang, Y. m 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, <https://doi.org/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, e2020EA001134, <https://doi.org/10.1029/2020ea001134>, 2020.
- Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., Li, H., Li, X., Peng, L., Qi, J., Yan, L., Zhang, Y., Zhao, H., Zheng, Y., He, K., and Zhang, Q.: Trends in China’s anthropogenic emissions since 2010 as the consequence of clean air actions, *Atmos. Chem. Phys.*, 18, 14095–14111, <https://doi.org/10.5194/acp-18-14095-2018>, 2018.
- 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, <https://doi.org/10.5194/acp-8-787-2008>, 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 on-line 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. Softw.*, 23, 550–560, <https://doi.org/10.1145/279232.279236>, 1997.