Geosci. Model Dev. Discuss., 8, 7313–7345, 2015
www.geosci-model-dev-discuss.net/8/7313/2015/
doi:10.5194/gmdd-8-7313-2015

Geoscientific Model Development Discussions
Open Access

Tracking influential haze source areas in North China using an adjoint model, GRAPES–CUACE

X. Q. An[1], S. X. Zhai[1], M. Jin[2], S. L. Gong[1], and Y. Wang[1]

[1]Chinese Academy of Meteorological Sciences, China Meteorological Administration, Beijing, China
[2]Wuhan Meteorological Service, Wuhan, China

Received: 11 April 2015 – Accepted: 31 July 2015 – Published: 28 August 2015

Correspondence to: X. Q. An (anxq@cams.cma.gov.cn)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Based upon the adjoint theory, the adjoint of the aerosol module in the atmospheric chemical modeling system GRAPES–CUACE (Global/Regional Assimilation and PrEdiction System coupled with the CMA Unified Atmospheric Chemistry Environment) was developed and tested for its correctness. Through statistic comparison, BC (black carbon aerosol) concentrations simulated by GRAPES–CUACE were generally consistent with observations from Nanjiao (one urban observation station) and Shangdianzi (one rural observation station) stations. To track the most influential emission-sources regions and the most influential time intervals for the high BC concentration during the simulation period, the adjoint model was adopted to simulate the sensitivity of average BC concentration over Beijing at the highest concentration time point (referred to as the Objective Function) with respect to BC emission amount over Beijing–Tianjin–Hebei region. Four types of regions were selected based on administrative division and sensitivity coefficient distribution. The adjoint model was used to quantify the effects of emission-sources reduction in different time intervals over different regions by one independent simulation. Effects of different emission reduction strategies based on adjoint sensitivity information show that the more influential regions (regions with relatively larger sensitivity coefficients) do not necessarily correspond to the administrative regions, and the influence effectiveness of sensitivity-oriented regions was greater than the administrative divisions. The influence of emissions on the objective function decreases sharply approximately for the pollutants emitted 17–18 h ago in this episode. Therefore, controlling critical emission regions during critical time intervals on the basis of adjoint sensitivity analysis is much more efficient than controlling administrative specified regions during an experiential time period.

GMDD

8, 7313–7345, 2015

1 Introduction

In the large-scale scientific and engineering calculation fields, derivative calculation exists everywhere. Solving a nonlinear optimal problem requires calculating the gradient, Hessian Matrix or higher-order reciprocal form (Cheng and Zhang, 2009). The traditional Finite Difference Method is aimed at some basic state, changing the concerned input variable values in proper order, obtaining the difference in output variables, and determining sensitivities of output variables to that input variable. This method usually creates truncation errors and is costly, being used only in the case of few input variables. The DDM (Decoupled Direct Method), which makes use of the TLM (Tangent Linear Model), is an improve of the Finite Difference Method, but is still limited in cases of few input variables. Comparatively, the adjoint method is an efficient sensitivity analysis approach, suitable for calculating the parametric sensitivities of complex numerical model systems and solving various optimal problems on the basis of sensitivity information. An adjoint model can work out the sensitivity of every variable in each time period and each simulation grid for the objective function at one simulation, which is much more efficient than the Finite Difference Method and the DDM. The adjoint method is used to calculate the derivatives of meromorphic functions on the basis of machine precision; thus, it has higher calculation precision and costs less, being propitious to large-scale nonlinear complex calculation and playing a significant role in meteorological and environmental fields. Based on the adjoint operator theory and the development of numerical models, the adjoint method is applied more and more for the inversion of pollution sources and other calculations that involve substantial input parameters. Through this method, the tangent linear model and adjoint model of the original model can be obtained on the basis of the traditional Finite Difference Method along with adjoint equation theory. The principle is to build the objective function by using the difference between the modelled and the observed, then to calculate the gradient (sensitivity) of the objective function of to the model input parameter by using the adjoint model; this gradient can be then used as a decreasing step length, cor-

GMDD
8, 7313–7345, 2015

recting the input values until the objective function reaches the minimum value through continuous iteration processes; therefore obtaining satisfactory input parameter values (Wang, 2000).

The adjoint method has a unique advantage for the complex multi-parameter system. Only through one simulation can it work out the sensitivity or gradient of the objective function to all of the input parameters (Liu, 2005) and quickly solve various types of optimal control and inversion problems by using the gradient information (Chen et al., 1998; Liu and Hu, 2003). Marchuk et al. (1976, 1986) first applied the adjoint method to the atmospheric environment field by using the method in the optimal control and reasonable site selection of pollution sources. They cleverly utilized the conjugation property of the adjoint operator, thus avoiding the pollutant transmission problems in the repeated problem-solving and greatly lessening the calculation amount. Skiba et al. (2000, 2002, 2003) developed Marchuk's method and applied it to solving the problems of atmospheric environment control. At present, some scientists have developed adjoint models for air quality models and conducted sensitivity analyses and assimilation based on these adjoint models. These adjoint models include the European air pollution dispersion model of the University of Cologne, Germany (EURAD model) (Elbern et al., 2000), which is mainly used in the simulation of large areas, the air quality model STEM-III (Sandu et al., 2005) and the atmospheric chemical transmission model CAMx (Liu et al., 2007) of the United States, etc. The adjoint of gaseous processes in CMAQ was already developed, which included the chemical conversion and the transmission processes of 72 active species (Hakami et al., 2007). On this basis, the adjoint of the aerosol processes in CMAQ will also be developed; and the CMAQ adjoint will hopefully be the first coupled gas–aerosol, regional scale adjoint model to explicitly describe aerosol mass composition and size distribution (Turner, 2010, 2014). Resler et al. (2010) present a version of the 4-D-var (four-dimensional variational) method and successfully used the adjoint of CMAQ optimized diurnal profiles of NO_2 emissions. Sfetsos et al. (2013) applied the CMAQ adjoint in Athens surface O_3 concentration-concentration and concentration-sources sensitivity analysis. The GEOS–Chem ad-

GMDD
8, 7313–7345, 2015

joint was generated both manually and automatically and contains the secondary formation processes of inorganic aerosols (Henze et al., 2007). Using the 4-D-var method in the GEOS–Chem adjoint model, Henze et al. (2009) constrained emissions estimates through assimilation of sulphate and nitrate aerosol measurements from the IMPROVE network. Zhang et al. (2009) quantified source contributions to O_3 pollution at two adjacent sites on the US west coast in spring 2006 by using the GEOS–Chem chemical transport model and its adjoint. García-Chan et al. (2014) utilized the adjoint method in optimizing the location and management of a new industrial plant and displayed the application of the adjoint method in optimal control problems. F. Paulot et al. (2014) inverse modeled the NH_3 emissions in the United States, European Union, and China by using the GEOS–Chem adjoint for assimilating observational data.

Some scientists consider the distribution of population density as well as pollutants exposure – healthy reaction relationships in the objective function. For example, Pappin et al. (2013) calculated health benefit influences on Canada and the United States separately from emissions of individual source locations in Canada and the United States, by estimating a certain reduction in anthropogenic emissions of NO_x and VOCs. Zhao et al. (2013) calculated and discussed effective emissions controlling strategies under a warming climate with regards to the reduction of O_3 concentration and short-term mortality due to O_3 exposure. Koo et al. (2013) quantified the health risk from intercontinental pollution by using the GEOS–Chem adjoint model.

GRAPES–CUACE is an on-line coupling of the atmospheric model GRAPES (Global-Regional Assimilation and Prediction system) (Xue and Chen, 2008) and the aerosol model CUACE (CMA Unified Atmospheric Chemistry Environmental Forecasting System) (Gong, 2003). GRAPES is a numerical weather prediction system developed by Chinese scientists under the organization of the China Meteorological Administration (CMA). CUACE is an air quality forecasting and climate research system developed by the Chinese Academy of Meteorological Science (CAMS). In this research, the adjoint model of the GRAPES–CUACE aerosol module has been developed and

GMDD
8, 7313–7345, 2015

used in receptor-source sensitivity analysis to lay the foundation of further emission inverse modeling.

2 Methodology

2.1 GRAPES–CUACE introduction

The GRAPES–CUACE aerosol module involves six types of particles, including sulphate, organic carbon, black carbon, nitrate, sea-salt and soil dust, which are divided into 12 sections by using the multiphase multicomponent aerosol particle size separation algorithm (Gong et al., 2003). The mass conservation equation of the size-distributed multiphase multicomponent aerosols can be expressed as:

$$\frac{\partial X_{ip}}{\partial t} = \left.\frac{\partial X_{ip}}{\partial t}\right|_{\text{TRANSPORT}} + \left.\frac{\partial X_{ip}}{\partial t}\right|_{\text{SOURCES}} + \left.\frac{\partial X_{ip}}{\partial t}\right|_{\text{CLEAR AIR}} + \left.\frac{\partial X_{ip}}{\partial t}\right|_{\text{DRY}} + \left.\frac{\partial X_{ip}}{\partial t}\right|_{\text{IN-CLOUD}} + \left.\frac{\partial X_{ip}}{\partial t}\right|_{\text{BELOW-CLOUD}}$$

where the rate of change of mixing ratio of dry particle mass constituent p in size range i has been divided into components (or tendencies) for transport, sources, clear air, dry deposition/sedimentation, in-cloud and below-cloud processes.

This module involves the vertical diffusion process of aerosols in the atmosphere. By solving the vertical diffusion equation, the vertical diffusion trend of aerosol particles is calculated. In addition, the data of emission sources are put into the module, which include both anthropogenic and natural emission sources. The aerosol physical and chemical processes section is the core of this module, including some primary aerosol processes in the atmosphere: aerosol emission, moisture absorption increase, collision, coring, condensation, dry deposition, gravity setting, sub-cloud cleanup, aerosol

GMDD
8, 7313–7345, 2015

activation, interaction between aerosols and clouds, transmission of sulphate in clouds and clear sky.

2.2 Aerosol adjoint construction and validation

2.2.1 General theory of adjoint method

Building an adjoint model for a forward model is a very complex task. To speed things up and reduce mistakes, the whole model is divided into many small programs. Abstract one small program as a vector function $F : \boldsymbol{R}^n \rightarrow \boldsymbol{R}^m$, which can be expressed as:

$$\boldsymbol{Y} = \mathbf{F}(\boldsymbol{X}) \tag{1}$$

where $\boldsymbol{X}$ is a series of n-dimensional independent variables and $\boldsymbol{Y}$ is m-dimensional dependent variables, representing the input variables and output variables of the original programs, respectively. Assuming that the function $\mathbf{F}$ is continuously differentiable at given spots of X_0, the tangent linear model (TLM) can be expressed as:

$$\mathrm{d}\boldsymbol{Y} = \nabla_{X_0}\mathbf{F} \cdot \mathrm{d}\boldsymbol{X} \tag{2}$$

where

$$\nabla_{X_0}\mathbf{F} = \begin{pmatrix} \nabla^{\mathrm{T}}_{X_0}\mathbf{F}_1 \\ \nabla^{\mathrm{T}}_{X_0}\mathbf{F}_2 \\ \vdots \\ \nabla^{\mathrm{T}}_{X_0}\mathbf{F}_m \end{pmatrix} = \begin{pmatrix} \frac{\partial \mathbf{F}_1}{\partial \boldsymbol{X}_1} & \frac{\partial \mathbf{F}_1}{\partial \boldsymbol{X}_2} & \cdots & \frac{\partial \mathbf{F}_1}{\partial \boldsymbol{X}_n} \\ \frac{\partial \mathbf{F}_2}{\partial \boldsymbol{X}_1} & \frac{\partial \mathbf{F}_2}{\partial \boldsymbol{X}_2} & \cdots & \frac{\partial \mathbf{F}_2}{\partial \boldsymbol{X}_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial \mathbf{F}_m}{\partial \boldsymbol{X}_1} & \frac{\partial \mathbf{F}_m}{\partial \boldsymbol{X}_2} & \cdots & \frac{\partial \mathbf{F}_m}{\partial \boldsymbol{X}_n} \end{pmatrix}$$

Based on the mathematical formula:

$$(\mathbf{L}\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}, \mathbf{L}^*\boldsymbol{Y}) \tag{3}$$

GMDD
8, 7313–7345, 2015

the adjoint expression of Eq. (2) will be:

$$d\boldsymbol{X}^* = \nabla_{X_0}^{\mathrm{T}} \mathbf{F} \cdot d\boldsymbol{Y}^* \tag{4}$$

where $d\boldsymbol{X}^*$ is of n dimensions and $d\boldsymbol{Y}^*$ is of m dimensions. Comparing Eq. (2) with Eq. (4), it is seen that these two formulas exchange dimensions between input and output with the transposition of their gradient factors. Obviously, the computing cost of the tangent linear model is proportional to the numbers of concerned input variables because the tangent linear model requires as many calculations as the number of input variables concerned with gradient information. Conversely, the adjoint model can obtain all this information through one calculation. When the number of output variables is much smaller than the concerned input variables, the superiority of the adjoint method is demonstrated.

2.2.2 Validation of tangent linear model

After the adjoint model is built, its correctness must be verified to confirm its reliability. The adjoint model is a concomitant of the tangent linear model (TLM). Thus, the validity of the tangent linear model must be ensured before the correctness of the adjoint model is tested. If all of the codes are tested as a whole, then error locations will be difficult to sense. To reduce this difficulty, both the tangent linear model and the adjoint model are divided into smaller sections, which are then tested separately. After these sections are confirmed, the assembled TLM and adjoint model will be tested.

Supposing that the code of every small section is regarded as $\boldsymbol{Y} = \mathbf{F}(\boldsymbol{X})$, then the Taylor expansions of $\mathbf{F}(\boldsymbol{X} + \delta\boldsymbol{X})$ at point X are:

$$\mathbf{F}(\boldsymbol{X} + \delta\boldsymbol{X}) = \mathbf{F}(\boldsymbol{X}) + \delta\boldsymbol{X}\mathbf{F}'(\boldsymbol{X}) + \frac{1}{2}(\delta\boldsymbol{X})^2\mathbf{F}''(\boldsymbol{X}) + \ldots + o(\delta\boldsymbol{X})^n\mathbf{F}^{(n)}(\boldsymbol{X}) \tag{5}$$

After transformation:

$$\frac{\mathbf{F}(\boldsymbol{X} + \delta\boldsymbol{X}) - \mathbf{F}(\boldsymbol{X})}{\delta\boldsymbol{X}\mathbf{F}'(\boldsymbol{X})} = 1 + \frac{1}{2}\delta\boldsymbol{X}\frac{\mathbf{F}''(\boldsymbol{X})}{\mathbf{F}'(\boldsymbol{X})} + \ldots + o(\delta\boldsymbol{X})^{n-1}\frac{\mathbf{F}^{(n)}(\boldsymbol{X})}{\mathbf{F}'(\boldsymbol{X})} \tag{6}$$

GMDD
8, 7313–7345, 2015

When $\delta \boldsymbol{X}$ approaches zero, the limit for the above equation is calculated as:

$$\text{Index} = \lim_{\delta \boldsymbol{X} \to 0} \frac{\mathbf{F}(\boldsymbol{X} + \delta \boldsymbol{X}) - \mathbf{F}(\boldsymbol{X})}{\delta \boldsymbol{X} \mathbf{F}'(\boldsymbol{X})} = 1.0 \tag{7}$$

In which the denominator is the TLM output, and the numerator is the difference between the output value of the original model with input $\boldsymbol{X} + \delta \boldsymbol{X}$ and input $\boldsymbol{X}$. To calculate the limit of the above equation repeatedly, we only need to decrease $\delta \boldsymbol{X}$ by an equal-ratio value. If the result approaches 1.0, it reflects that the tangent linear codes are correct. Generally speaking, the decrease of $\delta \boldsymbol{X}$ causes the limit value to approach 1.0, but due to the machine rounding error, the limit values might decrease first and then increase, appearing as a parabola. The validation results are displayed in Table 1.

2.2.3 Validation of the adjoint model

After all of the tangent linear codes have passed the testing, the adjoint codes can then be tested on this basis. The adjoint codes and tangent linear codes need to satisfy Eq. (3):

$$(\mathbf{L}(\boldsymbol{X}), \boldsymbol{Y}) = (\boldsymbol{X}, \mathbf{L}^*(\boldsymbol{Y}))$$

In which $\mathbf{L}$ represents the tangent linear process and $\mathbf{L}^*$ the adjoint process. To simplify the testing process, the adjoint input is set as the tangent linear output: $\boldsymbol{Y} = \mathbf{L}(\boldsymbol{X})$. Then, the above equation can be expressed as:

$$(\nabla \mathbf{F} \cdot \mathrm{d}\boldsymbol{X}, \nabla \mathbf{F} \cdot \mathrm{d}\boldsymbol{X}) = \left(\mathrm{d}\boldsymbol{X}, \nabla^{\mathrm{T}} \mathbf{F}(\nabla \mathbf{F} \cdot \mathrm{d}\boldsymbol{X})\right) \tag{8}$$

By putting $\mathrm{d}\boldsymbol{X}$ into the tangent linear codes, the output value $\nabla \mathbf{F} \cdot \mathrm{d}\boldsymbol{X}$ can be obtained and the left part of the equation can be computed. Then, taking $\nabla \mathbf{F} \cdot \mathrm{d}\boldsymbol{X}$ as the input of the adjoint codes, we obtain its output value $\nabla^{\mathrm{T}} \mathbf{F}(\nabla \mathbf{F} \cdot \mathrm{d}\boldsymbol{X})$ and then calculate the right part of the equation. As long as the result of the equation is approximately equal

GMDD
8, 7313–7345, 2015

(within the error range), the constructed adjoint model is considered to have passed the validation. The validation results of pollutant concentration variable xrow are shown in Table 2.

Seen from Table 2, both sides of the equation have produced values with 14 identical significant digits or more. This result is within the range of computer errors, so the values of the left and the right sides are considered to be equal, so the pollutant concentration variable xrow has passed the adjoint testing. Due to limited space, only the adjoint testing result of xrow is presented here. In fact, when performing the actual validations, all of the parameters are tested, respectively; although some parameters only have 11–12 identical significant digits, indicating lower precision, they are still considered to be within the permitted scope. Till now, all of the model variables have passed the adjoint testing.

2.2.4 Construction of the adjoint of GRAPES–CUACE aerosol model

After each part and the assembled TLM and adjoint model have been verified, the GRAPES–CUACE adjoint model is constructed. The structures and parameters passing flowchart are shown in Fig. 1. GRAPES–CUACE is an on-line meteorological chemistry modeling system, and aerosol transport processes are from GRAPES.

2.3 Sensitivity analysis

To conduct sensitivity analysis and solve environmental optimization problems, we tend to take into account various factors, including air quality standard, economic loss, health benefit, emissions reduction enforceable ratios range, suitable locations for factories, etc. Hence, a reasonable evaluation function J is needed, which includes one or several of the above factors as independent variables or/and as controlling conditions. In the adjoint method, such a function is called the objective function. We can define various types of objective functions based on different purposes. An objective

GMDD
8, 7313–7345, 2015

function is always a simple function of output parameters (e.g., $J = J(\boldsymbol{Y})$) compared with a complex atmospheric chemistry modeling system $\boldsymbol{Y} = \mathbf{F}(\boldsymbol{X})$.

The adjoint input, also called the forcing term (Fig. 1), is the gradient of J with respect to model output $\boldsymbol{Y}$: $\nabla_Y J$, which is relatively easy to obtain. The adjoint output, also called aiming sensitivity information, is the gradient of J with respect to any model parameter $\boldsymbol{X}$: $\nabla_X J$. To endow a definite physical meaning to sensitivity information, we define the sensitivity coefficient as the product of one model parameter $\boldsymbol{X}$ and $\nabla_X J$: $\boldsymbol{X} \cdot \nabla_X J$. This sensitivity coefficient has the same unit as the objective function J.

When controlling a severe pollution event, J is often defined as the concentration of a concerned pollutant at the time with the most serious pollution. Then, the inverse adjoint method can be used to locate where and when the emissions should have the greatest influence.

In emission inventory optimization problems, J is often defined as the discrepancy between the simulated and observed values. Running the adjoint model once, the gradients (sensitivity) of the objective function to emission amount can be obtained, and then, by using the gradient information iteratively, the optimal solution of the objective function is determined. In this adjoint sensitivity analysis research, we use the adjoint method to locate the most influential emission sources regions and the most influential emission time periods.

2.4 Model setup

In this study, the GFS reanalysis data, which are collected 6 times a day with $1^\circ \times 1^\circ$ resolution, are used for initial and boundary conditions in the GRAPES–CUACE modeling system, and INTEX–B2006 ($0.5^\circ \times 0.5^\circ$) is used as the emission sources. With a horizontal resolution of $0.5^\circ \times 0.5^\circ$, the simulation domain covers Northeast China (105–125° E, 32.25–42.25° N), as shown in Fig. 2. Our analysis mainly focuses on the Beijing–Tianjin–Hebei (BTH) region. The entire simulation period is from 20:00 BT (Beijing Time) 28 June 2008 to 20:00 BT 4 July 2008, and the first 72 h are regarded as the spin-up time.

GMDD
8, 7313–7345, 2015

2.5 Observations

The data used in this paper are hourly black carbon aerosol (BC) average concentrations from the Beijing Meteorological Observatory Nanjiao Station and Shangdianzi Station. The Nanjiao Station (39.8° N, 116.47° E) is located in the Atmospheric Observation Test Base in the southern suburb of Beijing. It is next to the Beijing urban area in the north and close to Fifth Ring Road in the south, where traffic flows are relatively large. The Shangdianzi Station (40.65° N, 117.12° E) is at the Shangdianzi village of Miyun County in northeastern Beijing. This station is a regional atmospheric background station, around which there is no obvious industrial pollution and few human activities, i.e., it has a better ecological environment. The locations of the two stations are shown in Fig. 2.

3 Results and discussion

BC is an important component of atmospheric aerosols. It is produced mainly by the imperfect combustion of carbonaceous materials, such as fossil fuels and biomass raw materials. Its sources include anthropogenic and natural emission sources. Natural sources (e.g., volcanic eruption and forest fires) are occasional and regional, contributing little to the long-term background BC concentration in the atmosphere (Nagamoto et al., 1993). Comparatively, many human activities increase the concentration of the BC aerosols; so anthropogenic sources are the primary sources for BC. Cao et al. (2006) and Streets et al. (2001) noted that the vast majority of BC emission in China is created by the untreated raw coal, honeycomb briquettes, and biomass fuels that people use in their daily lives.

3.1 High BC concentration episode and model validation

The simulated ground BC concentration distributions from 20:00 BT 3 July to 11:00 BT 4 July are shown in Fig. 3. These six graphs demonstrate the formation and transporta-

GMDD
8, 7313–7345, 2015

tion processes of this high BC concentration episode over Beijing. At 20:00 BT 3 July, two small spots of high BC concentration appear around Shijiazhuang and southern Beijing. Then, at 23:00 BT 3 July, these two high BC concentration spots are obviously enlarged, and are almost connected, extending to northern Xingtai, eastern Baoding, Langfang and Tianjin. At 2:00 BT 4 July, high BC concentration area develops around Beijing, Tianjin, southern Hebei and the Henan province. Then, it gets enlarged and intensified continuously during the subsequent hours until 11:00 BT 4 July, when the influenced scope begins to narrow due to enhanced dispersion and vertical movement in the boundary layer. However, the BC concentration over Beijing still remains at a relatively higher level.

Figure 4 shows hourly variation of ground level BC concentration in Beijing. It is easy to notice that during the first 2 simulated days, the BC concentration value reaches its peak at approximately 2:00 BT 2 and 3 July and its lowest value at approximately 15:00 BT. This result is closely affected by the diurnal height variation of the boundary layer, atmospheric stability and diffusion conditions. Different from the previous 2 days, the highest BC concentration value on 4 July, i.e., 15.7 $\mu g\,m^{-3}$, occurs at 11:00 BT. This might be because, on 4 July, the atmospheric condition is more stable and the pollutant diffusion condition is unsatisfactory, thus leading to BC accumulation.

The model results are compared with observation data in Fig. 5. The correlation coefficients of the simulated and the observed BC concentrations at Shangdianzi and Nanjiao station are 0.65 and 0.54, respectively. So the general variation trends of the simulated and observed BC concentrations are consistent, except that the simulated BC concentration values are bigger than the corresponding observed ones. Overall, the model results are acceptable.

3.2 Objective function and sensitivity coefficient definitions

As mentioned above, the adjoint method can provide information about influences of location-specific sources on the function called “objective function”. To determine the area and time period of the most important emission sources that induce the high BC

GMDD
8, 7313–7345, 2015

concentration over Beijing at 11:00 BT 4 July 2008 (Fig. 4), we define the objective function J as average BC concentration over Beijing at 11:00 BT 4 July 2008.

The adjoint input, also regarded as the forcing term, is $\partial J/\partial C$. C represents the pollutant concentration, such as BC concentration, at the objective time point. The direct output from the adjoint model is the gradient of J with respect to any model parameter var: $\partial J/\partial \mathrm{var}$. If Q is emission intensity, we define the emission sensitivity coefficient Φ as:

$$\Phi = Q\frac{\partial J}{\partial Q}$$

In this way, the emission sensitivity coefficient Φ has the same unit with J and has a specific physical meaning. The bigger the sensitivity coefficient value is, the larger the influence of BC emission in that area has on J. If BC emission is cut by N %, the value of J will decrease by $N\% \cdot \Phi$, which means that the average BC concentration over Beijing at the objective time point will decrease by $N\% \cdot \Phi$.

3.3 Distribution of adjoint sensitivity

When controlling air quality by cutting down emissions, we tend to cut emissions over a certain period of time, e.g., starting to cut emissions 1–3 days ahead of the predicted severe pollution day. Based on this practical concept, sensitivity coefficients at every model backward integral time step are added from the objective time point (highest BC concentration: 11:00 BT 4 July 2008) to a certain preceding time point, as illustrated in Fig. 6. Figure 6 shows a spatial–temporal cumulative effect from BC emissions to the objective function J.

As shown in Fig. 6, sensitivity coefficients accumulate along an inverse time series. When sensitivity coefficients from the previous 1 h until the objective time point are added, only the Tongzhou and Daxing districts in Beijing have sensitivity coefficients of 0.05–0.1 $\mu g\,m^{-3}$. When sensitivity coefficients are added during the last 6 h, the influential area is remarkably enlarged, with a maximum value of 0.3–0.4 $\mu g\,m^{-3}$. As

GMDD
8, 7313–7345, 2015

the hours ahead of the objective time points considered extend, this influenced area is continually enlarged and intensified. When it reaches the 16 h period of time, as shown in Fig. 6d, the more critical area expands to Langfang and Baoding of Hebei province, and the maximum value is approximately 0.7 $\mu g\,m^{-3}$, which indicates that reducing BC emission at the ratio of N % from 19:00 BT 3 July to the objective time point over this grid cell could result in about an average $N\% \cdot 0.7\,\mu g\,m^{-3}$ decrease of BC concentration over Beijing, the objective region, at 11:00 BT 4 July 2008, the objective time point. However, along with this accumulation procedure, the expansion of the influential region scope and the increase of its sensitivity coefficients begin to slow down. Only a tiny difference between 24 h of accumulation (Fig. 6f) and 48 h of accumulation (Fig. 6g) is observed. This phenomenon reflects that emissions at a very early time, such as more than 24 h ahead of the objective time point, have little influence on J. When a heavy pollution event needs to be controlled by reducing emissions, the time period with most significant influence should be scientifically determined to cut emissions both effectively and economically.

3.4 Time series of sensitivity coefficients over different regions

Adjoint sensitivity analysis is a powerful compliment to forward methods. While forward techniques are source-based, backward methods provide receptor-based sensitivity information. Under this conception, we use the adjoint method to locate the most influential emission sources area and the most influential emission time period.

Four types of regions are defined according to administrative division and sensitivity coefficients distribution (Table 3 and Fig. 7). BTH refers to the administrative Beijing–Tianjin–Hebei region, which covers 105 grid cells and is approximately 318 000 km^2; BJ represents administrative Beijing, which contains 10 grid cells and covers an area of around 30 000 km^2. InR-1 (Influential Region 1) has 7 grid cells, occupying about 21 000 km^2, which is smaller than that of BJ, whose sensitivity coefficient values are obviously bigger than others; InR-2 (Influential Region 2) covers InR-1 and 10 more grid

GMDD
8, 7313–7345, 2015

cells with secondary large coefficient values, having 17 grid cells in total and covering approximately 51 000 km^2.

To compare the effects of emission-sources reduction at different time points in the 4 regions, we add BC emission sensitivity coefficients vertically and extract their inverse time series values (Fig. 8). Figure 8a is the inverse time series of sensitivity coefficients at every 5 min integration time step. It reflects the influence of BC emission on the objective function J at each model integration time step ahead of the objective time point. Figure 8b shows the time cumulative sensitivity coefficients, which reveal the decrease of J due to BC emission reduction over a certain period of time ahead of the most polluted time point.

In Fig. 8a, the sensitivity coefficients of BTH, InR-1 and BJ reach their peak values at 18:00 BT 3 July, while that InR-2 at 17:00 BT 3 July, and then all decrease sharply along backward time sequence. This phenomenon indicates that emissions emitted 17–18 h before the most serious pollution time point have rapid decreasing effects on J along the inverse time sequence axis. Correspondingly, in Fig. 8b, the time cumulative sensitivity coefficients obviously slow down their increasing trend at 18:00 BT 3 July. This phenomenon shows that cutting emissions before the predicted pollution episode can have better effects on air quality control than doing so after severe pollution events occur. In addition, it also shows that the emission reduction start-up time point should be scientifically determined based on adjoint sensitivity or other information to increase the efficiency of air quality control.

Then we compared the preceding 18 h, 17:00 BT 3 July to 11:00 BT 4 July, cumulative sensitivity coefficients of the above 4 regions (Table 4), given it that the sensitivity coefficient on 17:00 BT 3 July is still relatively high (for BTH, InR-1, and BJ). From Table 4, the simulated SC (sensitivity coefficient) of BTH is 7.3 μg m^{-3}, meaning that a reduction of N % BC emissions over BTH will cause an $N\% \cdot 7.3\,\mu g\,m^{-3}$ decrease of average BC concentration in Beijing on 11:00 BT 4 July. In general, it is obvious that reducing emissions over the whole BTH region will contribute most positively to air quality control in Beijing, followed by InR-2, InR-1 and BJ respectively. However, in

GMDD
8, 7313–7345, 2015

the 4 regions, the SC/Grid (sensitivity coefficient per grid) value of InR-1 is the largest, meaning cutting emissions of InR-1 has the most obvious effectiveness on decreasing BC concentration in Beijing. The SC/Grid of BTH is the smallest, and InR-2 equals BJ in between. BTH covers an area which is 6.2 times that of InR-2, but the SC and SC/Grid of Inr-2 are 80 % and 5.0 times of BTH (Table 4). A similar phenomenon is found between BJ and InR-1. InR-1 accounts for only 70 % of the area of BJ, but the SC and SC/Grid of InR-1 are 1.2 and 1.6 times of BJ.

4 Conclusions

On the basis of adjoint theory and methods, this paper has constructed an adjoint model for an aerosol module of the atmospheric chemical model GRAPES–CUACE and tested the correctness of the model. At the same time, the GRAPES–CUACE model and its aerosol adjoint were adopted to perform a numerical simulation and a receptor-source sensitivity test. Compared with the BC aerosol observations from the Nanjiao and Shangdianzi Station, the hourly trends of BC concentration were similar, with correlation coefficients of 0.65 and 0.54, respectively.

The GRAPES–CUACE adjoint model simulated the sensitivity of concentration on emission and was adopted to track the most influential emission-sources regions and most influential time intervals for the high BC concentration. Four types of regions were selected and compared based on administrative division and adjoint sensitivity coefficient distribution. The result of the aerosol adjoint model suggested that the influence effectiveness of sensitivity-oriented regions was greater than the administrative divisions. For the case studied in this paper, emissions from 17–18 h ahead of the objective time point had a much larger influence than emissions emitted earlier. It is found that in order to increase emission reduction efficiency, influential regions should be located scientifically (e.g., according to adjoint sensitivity coefficients distribution) rather than by administrative divisions.

GMDD
8, 7313–7345, 2015

Code availability

We use the GRAPES–CUACE as distributed by Numerical Weather Prediction Center of Chinese Meteorology Administration (http://nwpc.cma.gov.cn) together with Institute of Atmospheric Composition of Chinese Academy of Meteorological Sciences (http://cadata.cams.cma.gov.cn). The model runs on IBM PureFlex System (AIX) with XL Fortran Compiler. The CUACE-ADJ code can be requested from the corresponding author or downloaded as a Supplement to this article.

The Supplement related to this article is available online at doi:10.5194/gmdd-8-7313-2015-supplement.

Acknowledgements. This study was supported by the National Basic Research Program of China "973" (2011CB403404), and partially supported by the CMA Innovation Team for Haze-fog Observation and Forecasts. We appreciate Lin Zhang, Feng Liu, Qiang Cheng, Hongliang Zhang and Min Xue for providing technical support in adjoint model construction. Thanks also to the developers of the GRAPES–CUACE aerosol model.

References

Cao, G., Zhang, X., Wang, Y., Che, H., and Chen D.: Inventory of black carbon emission from China, Adv. Clim. Ch. Res., 2, 259–264, 2006 (in Chinese).

Cao, G., Zhang, X., Gong, S., An, X., and Wang, Y.: Emission inventories of primary particles and pollutant gases for China, Chin. Sci. Bull., 56, 261–268, 2011.

Chen, H., Hu, F., Zeng, Q., and Chen, J.: Some practical problems of optimizing emissions from pollution sources in air, climatic and environ, Res., 3, 163–172, 1998 (in Chinese).

Cheng, Q., Zhang, H., and Wang, B.: Algorithms of automatic differentiation, Math. Num. Sin., 33, 15–36, 2009 (in Chinese).

Elbern, H., Schmidt, H., Talagrand, O., and Ebel, A.: 4-D-variational data assimilation with an adjoint air quality model for emission analysis, Environ. Modell. Softw., 15, 539–548, 2000.

GMDD

8, 7313–7345, 2015

García-Chan, N., Alvarez-Vázquez, L., Martínez, A., and Vázquez-Méndez, M.: On optimal location and management of a new industrial plant: numerical simulation and control, J. Frankl. Inst., 351, 1356–1371, 2014.
Gong, S. L., Barrie, L. A., Blanchet, J.-P., von Salzen, K., 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, doi:10.1029/2001JD002002, 2003.
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, 2007.
Henze, D. K., Hakami, A., and Seinfeld, J. H.: Development of the adjoint of GEOS–Chem, Atmos. Chem. Phys., 7, 2413–2433, doi:10.5194/acp-7-2413-2007, 2007.
Henze, D. K., Seinfeld, J. H., and Shindell, D. T.: Inverse modeling and mapping US air quality influences of inorganic PM 2.5 precursor emissions using the adjoint of GEOS–Chem, Atmos. Chem. Phys., 9, 5877–5903, doi:10.5194/acp-9-5877-2009, 2009.
Koo, J., Wang, Q., Henze, D. K., Waitz, I. A., and Barrett, S. R.: Spatial sensitivities of human health risk to intercontinental and high-altitude pollution, Atmos. Environ., 71, 140–147, doi:10.1016/j.atmosenv.2013.01.025, 2013.
Liu, F.: Adjoint model of Comprehensive Air quality Model CAMx – construction and application, Post-doctoral Research Report, College of Environmental Sciences and Engineering, Peking University, China, 102 pp., 2005 (in Chinese).
Liu, F. and Hu, F.: Inversion of diffusion on coefficients and effect of related difference schemes, J. Appl. Meteorol. Sci., 14, 331–338, 2003 (in Chinese).
Liu, F., Zhang, Y., Su, H., and Hu, J.: Adjoint model of atmospheric chemistry transport model CAMx: construction and Application, Acta Sci. Natru. Uni. Peki., 43, 764–770, 2007 (in Chinese).
Maruk, G.: Mathematical Models in Environmental Problems, Elsevier Science Publishers, New York, 1986.
Maruk, G. and Skiba, Y.: Numerical calculation of the conjugate problem for a model of the thermal interaction of the atmosphere with the oceans and continents, Izv. Atmos. Oceanic Phys., 12, 279–284, 1976.
Nagamoto, F. and Zhou, M.: Aeolian transport of aerosol black carbon from China to the ocean, Atmos. Environ., 28, 3251–3260, 1994.

GMDD
8, 7313–7345, 2015

Pappin, A. J. and Hakami, A.: Source attribution of health benefits from air pollution abatement in Canada and the United States: an adjoint sensitivity analysis, Environ. Health Persp., 121, 572–579, doi:10.1289/ehp.1205561, 2013.

Paulot, F., Jacob, D. J., Pinder, R. W., Bash, J. O., Travis, K., and Henze, D. K.: Ammonia emissions in the United States, European Union, and China derived by high-resolution inversion of ammonium wet deposition data: interpretation with a new agricultral emissions inventory (MASAGE_NH3), J. Geophys. Res. Atmos., 119, 4343–4364, doi:10.1002/2013JD021130, 2014.

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

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, 2005.

Sfetsos, A., Vlachogiannis, D., and Gounaris, N.: An investigation of the factors affecting the ozone concentrations in an urban environment, Atmos. Clim. Sci., 3, 11–17, 2013.

Skiba, Y. N. and Parra-Guevara, D.: Industrial pollution transport, Part 1: formulation of the problem and air pollution estimates, Part 2: control of industrial emissions, Env. Model. Ass., 5, 169–175, 177–184, 2000.

Skiba, Y. N. and Davydova-Belitskaya, V.: Air pollution estimates in Guadalajara City, Env. Model. Ass., 7, 153–162, 2002.

Skiba, Y. N. and Davydova-Belitskaya, V.: On the estimation of impact of vehicular emissions, Ecol. Model., 166, 169–184, 2003.

Streets, D., Gupta, S., and Waldho, S.: Black carbon emissions in China, Atmos. Environ., 35, 4281–4296, 2001.

Turner, M.: Inverse modeling of NO_x and NH_3 precursor emissions using the adjoint of CMAQ, Research Prelim Report, Department of Mechanical Engineering, University of Colorado, 18 p., available at: http://mattdturner.com/research/styled-3/ (last access: 27 August 2015), 2010.

Turner, M.: Constraining Aerosol Health Impacts with Sensitivity Analysis using the Adjoint of CMAQ, Comprehensive Exam Report, Department of Mechanical Engineering, University of Colorado at Boulder, available at: http://mattdturner.com/research/styled-3/, January 2014.

Wang, X.: Toward the objective analysis, four-dimensional assimilation and adjoint method, J. PLA Univ. Sci. Tech., 1, 67–74, 2000 (in Chinese).

GMDD
8, 7313–7345, 2015

Xu, L., Wang, Y., Chen, Z., Luo, Y., and Ren, W.: Progress of black carbon aerosol research I: emission, removal and concentration, Adv. Earth Sci., 21, 352–359, 2006 (in Chinese).
Xue, J. and Chen, D.: Scientific design and application of numerical predicting system GRAPES, Science Press, Beijing, 2008.
Yin, Y., Cui, Z., and Zhang, H.: Numerical simulations of mass distribution of aerosol over China in 2006, Trans. Atmos. Sci., 32, 595–603, 2009 (in Chinese).
Zhang, L., Jacob, D. J., Kopacz, M., Henze, D. K., Singh, K., and Jaffe, D. A.: Intercontinental source attribution of ozone pollution at western US sites using an adjoint method, Geophys. Res. Lett., 36, L11810, doi:10.1029/2009gl037950, 2009.
Zhao, S., Pappin, A. J., Morteza Mesbah, S., Joyce Zhang, J., MacDonald, N. L., and Hakami, A.: Adjoint estimation of ozone climate penalties, Geophys. Res. Lett., 40, 5559–5563, 2013.
Zhu, J., Zeng, Q., Guo, D., and Liu, Z.: Optimal control of sedimentation in navigation channels, J. Hydraul. Eng., 125, 750–759, 1999.
Zhu, J. and Zeng, Q.: A mathematical theory frame for atmospheric pollution control, Sci. China Ser. D, 32, 864–870, 2002.

GMDD
8, 7313–7345, 2015

Table 1. Validation results of tangent linear model.

a	Index (xrow)	Index (rhop)
1.00000000000	0.961383789	1.064836676
0.10000000000	0.996231252	1.005283209
0.01000000000	0.999622785	1.000526942
0.00100000000	0.999962182	1.000052673
0.00010000000	0.9999995320	1.000005301
0.00001000000	0.999995319	1.000000848
0.00000100000	0.999974073	1.000001471
0.00000010000	0.998912182	1.000034692
0.00000001000	0.996789129	1.000189939
0.00000000100	0.913747381	1.002300501

GMDD
8, 7313–7345, 2015

Table 2. Validation results of adjoint model.

Integral step	VALTGL	VALADJ
1	$0.253071834334799587 \times 10^{-11}$	$0.253071834334799587 \times 10^{-11}$
2	$0.138781684963437701 \times 10^{-7}$	$0.138781684963437635 \times 10^{-7}$
3	$0.197243288646595624 \times 10^{-6}$	$0.197243288646595703 \times 10^{-6}$
4	$0.285995663142418833 \times 10^{-6}$	$0.285995663142418833 \times 10^{-6}$
5	$0.138094513716334626 \times 10^{-6}$	$0.138094513716334599 \times 10^{-6}$
6	$0.158774915826234477 \times 10^{-6}$	$0.158774915826234609 \times 10^{-6}$
7	$0.205383106884893541 \times 10^{-6}$	$0.205383106884893673 \times 10^{-6}$
8	$0.113356629291541069 \times 10^{-6}$	$0.113356629291540963 \times 10^{-6}$
9	$0.151566991405230902 \times 10^{-6}$	$0.151566991405230823 \times 10^{-6}$
10	$0.174929034468917025 \times 10^{-6}$	$0.174929034468917104 \times 10^{-6}$
11	$0.333573941572600298 \times 10^{-6}$	$0.333573941572600616 \times 10^{-6}$
12	$0.185912861066765391 \times 10^{-6}$	$0.185912861066765523 \times 10^{-6}$

GMDD
8, 7313–7345, 2015

Table 3. Information of 4 emission reduction regions.

Region	Number of Grid cells	Area (km^2)
BTH	105	318 000
BJ	10	30 000
InR-1	7	21 000
InR-2	17	51 000

GMDD
8, 7313–7345, 2015

Table 4. 18 h (17:00 BT 3 July–11:00 BT 4 July) cumulative SC and SC/Grid over 4 emission reduction regions.

Regions SC	BTH (μgm^{-3})	BJ (μgm^{-3})	InR-1 (μgm^{-3})	InR-2 (μgm^{-3})	InR-2/BTH	InR-1/BJ
SC	7.3	3.5	4.0	5.9	0.8	1.2
SC/Grid	0.07	0.35	0.58	0.35	5.0	1.6

SC: Sensitivity coefficient
SC/Grid: Sensitivity coefficient per simulation grid

GMDD
8, 7313–7345, 2015

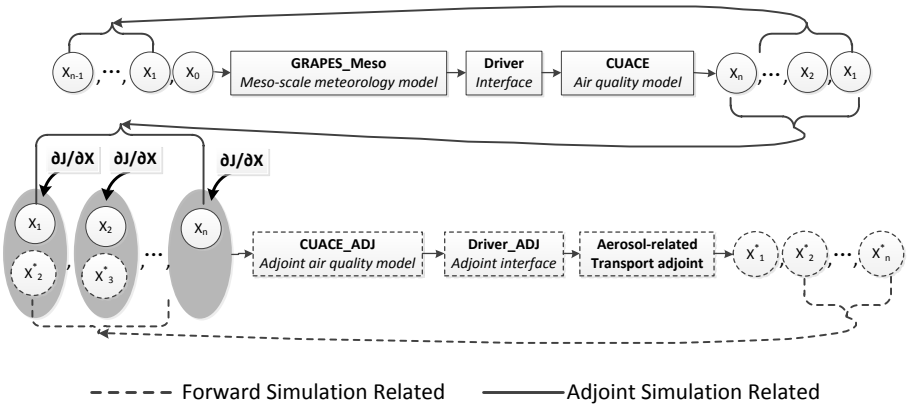

Figure 1. Frame work of GRAPES–CUACE, CUACE–ADJ and the flowchart of parameters transmission. (ADJ is short for adjoint; X_n, X_{n+1} represent model parameters after n, $n+1$ GRAPES–CUACE integral time steps, respectively; X_n^*, X_2^* represent, correspondingly, X_n's adjoint $\partial J/\partial X_n$ and X_2's adjoint $\partial J/\partial X_2$, where J is the objective function; $\partial J/\partial X$ are forcing terms, they are added at every backward integral step if have nonzero values; structures and variables in solid line frames are related to forward simulation; structures and variables in dashed frames are related to adjoint simulation.)

GMDD
8, 7313–7345, 2015

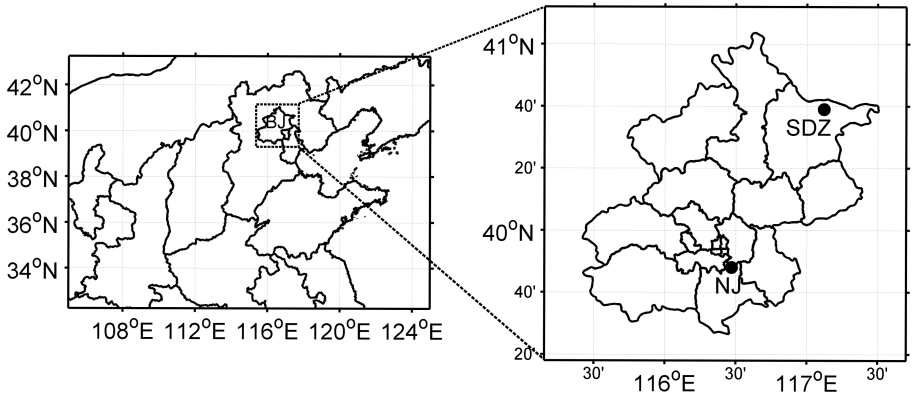

Figure 2. Left: model domain settings (left); right: the locations of Nanjiao (NJ) and Shangdianzi (SDZ) observation sites.

GMDD
8, 7313–7345, 2015

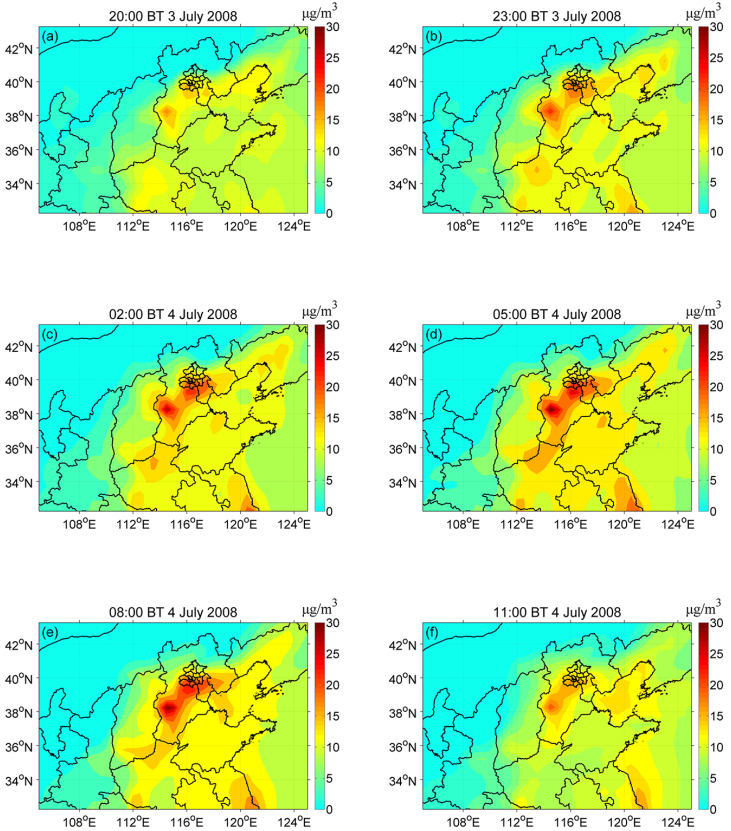

Figure 3. BC concentration distribution at ground level (Units: $\mu g\,m^{-3}$).

GMDD
8, 7313–7345, 2015

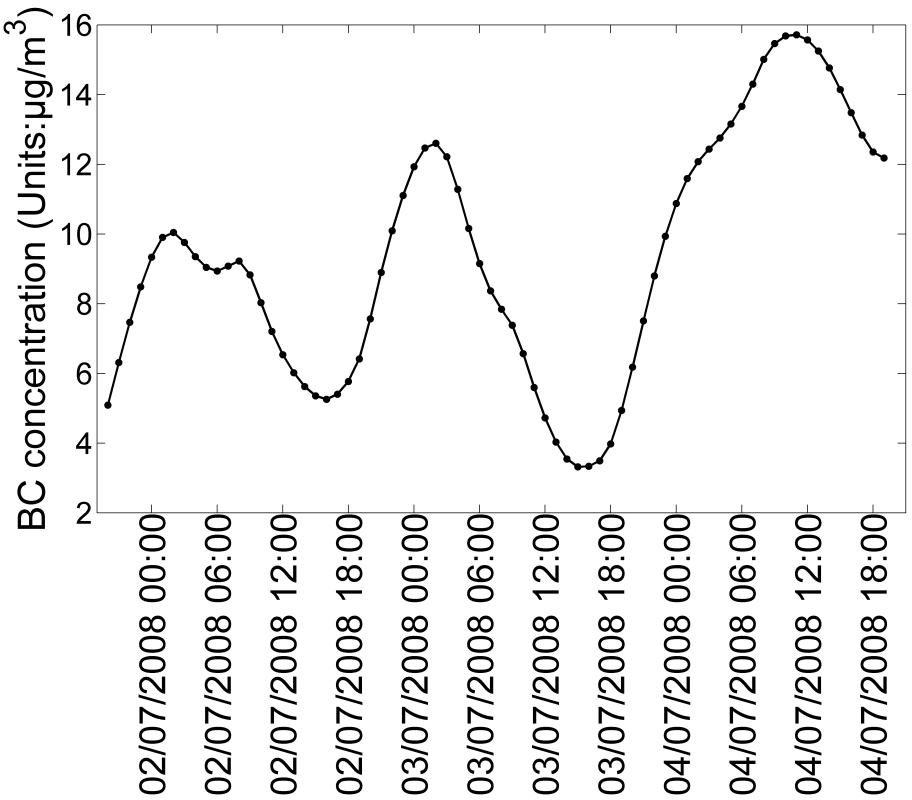

Figure 4. Hourly variation of ground BC concentration in Beijing.

GMDD
8, 7313–7345, 2015

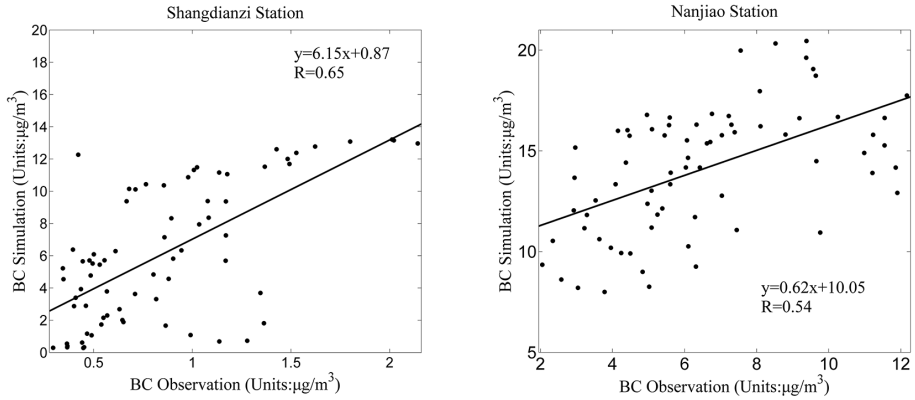

Figure 5. Correlation coefficients of simulated and observed BC concentration at Shangdianzi and Nanjiao Station.

GMDD
8, 7313–7345, 2015

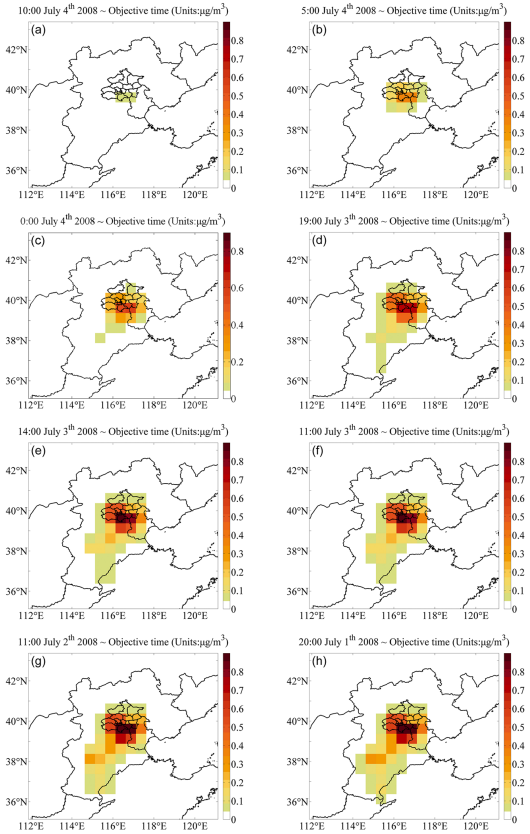

Figure 6. Cumulative sensitivity coefficient distribution. **(a)**–**(e)** are 1, 6, 11, 16, 21 h cumulative sensitivity coefficients, **(f)**–**(g)** are 24, 48 h cumulative coefficients, and **(h)** is the last backward simulation time step.

GMDD
8, 7313–7345, 2015

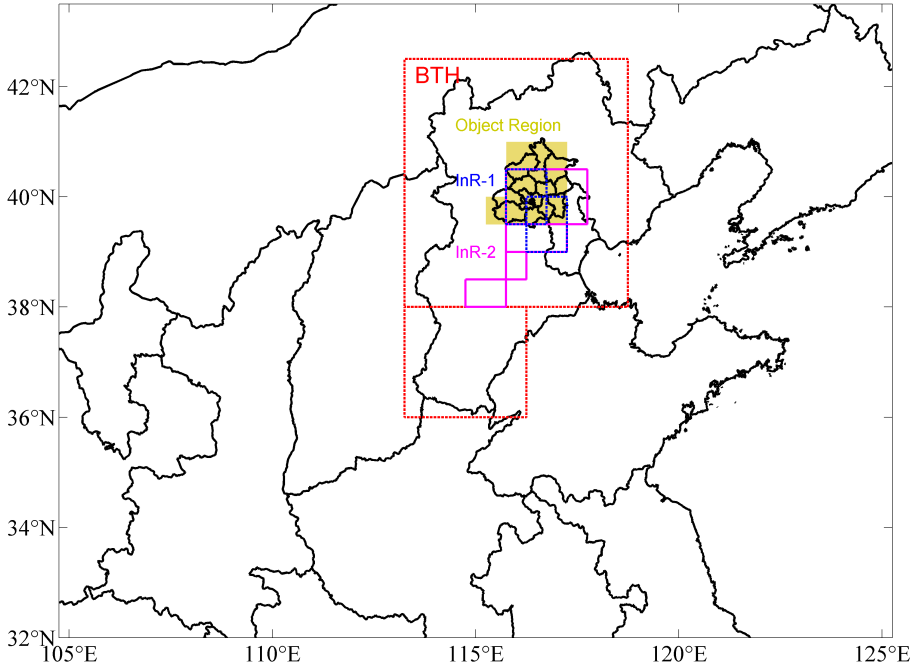

Figure 7. Different influential regions. BTH: Red dashed frame; InR-1: Blue dashed frame; InR-2: Pinkish red solid frame; Object Region: yellow shadow.

GMDD
8, 7313–7345, 2015

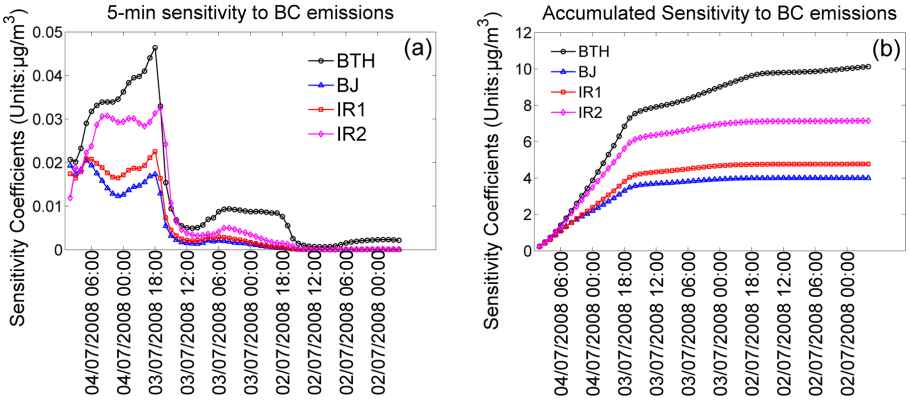

Figure 8. (a) Sensitivity coefficients at each 5 min integration time step along inverse time sequence; **(b)** Cumulative sensitivity coefficients along inverse time series.

GMDD
8, 7313–7345, 2015