Atmos. Chem. Phys. Discuss., 15, 20851–20879, 2015
www.atmos-chem-phys-discuss.net/15/20851/2015/
doi:10.5194/acpd-15-20851-2015

Atmospheric Chemistry and Physics Discussions

Inverse modeling of black carbon emissions over China using ensemble data assimilation

P. Wang, H. Wang, Y. Q. Wang, X. Y. Zhang, S. L. Gong, M. Xue, C. H. Zhou, H. L. Liu, X. Q. An, T. Niu, and Y. L. Chen

Institute of Atmospheric Composition, Key Laboratory of Atmospheric Chemistry (LAC) of China Meteorological Administration (CMA), Chinese Academy of Meteorological Sciences (CAMS), Beijing, 100081, China

Received: 24 March – Accepted: 26 June 2015 – Published: 4 August 2015

Correspondence to: P. Wang (pwang@cams.cma.gov.cn)

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

Abstract

Emissions inventories of black carbon (BC), which are traditionally constructed using a "bottom-up" approach based on activity data and emissions factors, are considered to contain a large level of uncertainty. In this paper, an ensemble optimal interpolation (EnOI) data assimilation technique is used to investigate the possibility of optimally recovering the spatially resolved emissions bias of BC. An inverse modeling system for emissions is established for an atmospheric chemistry aerosol model and two key problems related to ensemble data assimilation in the top-down emissions estimation are discussed: (1) how to obtain reasonable ensembles of prior emissions; and (2) establishing a scheme to localize the background-error matrix. An experiment involving a one month simulation cycle with EnOI inversion of BC emissions is performed for January 2008. The bias of the BC emissions intensity in China at each grid point is corrected by this inverse system. The inversed emission over China in January is 240.1 Gg, and annual emission is about 2750 Gg, which is over 1.8 times of bottom-up emission inventory. The results show that, even though only monthly mean BC measurements are employed to inverse the emissions, the accuracy of the daily model simulation improves. Using top-down emissions, the average root-mean-square error of simulated daily BC is decreased by nearly 30 %. These results are valuable and promising for a better understanding of aerosol emissions and distributions, as well as aerosol forecasting.

1 Introduction

Black carbon (BC) refers to light-absorbing carbon aerosols produced by all kinds of incomplete combustion processes (Bond and Bergstrom, 2006). It is an important component of atmospheric particulate matter, affecting weather, climate and air quality, and therefore attracts much attention among the scientific community. Its absorptive nature, which directly causes reductions in incoming shortwave solar radiation, is a key con-

ACPD
15, 20851–20879, 2015

tributor to climate forcing by aerosols (Liousse et al., 1993; Menon et al., 2002; Hansen et al., 2005; Ramanathan and Carmichael, 2008). BC aerosols have been shown to act as cloud condensation nuclei when they become hydrophilic, affecting cloud microphysical properties and rainfall processes. The lifetime of BC is about 3–10 days, and it can be transported far from its source to affect remote and pristine areas (Hansen et al., 1988; Hara et al., 2008). Its light-absorbing properties reduce atmospheric visibility (Wolff, 1981). Qiu and Yang (2000) showed that BC contributes to the considerable degradation in optical depths and visibility noted in northern China. Furthermore, from the human health perspective, these particles, which are generally sub-micron in size, contribute greatly to the threat of pulmonary diseases, as they can penetrate into the lungs while also carrying a variety of toxic elements with them. Therefore, an accurate picture of the distribution and variation of BC is crucial to our understanding of climate change and pollutant dynamics, and ultimately helps us to develop better policies to tackle associated environmental problems.

However, there is considerable uncertainty involved in the estimation of the distribution of BC and its contribution BC emissions inventory, which is traditionally constructed from the "bottom up" approach based on activity data and emissions factors, is considered to have large uncertainty (Cao et al, 2006). The overall uncertainty of BC over all of Asia is estimated at about 360 % (Streets et al., 2003). Emissions of BC are difficult to determine under the best of circumstances, largely because of the uncertainty in quantifying the fraction of total particulate matter that is elemental carbon of less than approximately 1 µm in diameter. This fraction is highly sensitive to the fuel type and combustion conditions (Wehner et al., 1999), necessitating a detailed treatment of emission factors by fuel, sector, and the degree of emissions control. Such problems are particularly compounded in China, where no statistics are available on the types of combustor and particulate controls, neither in terms of the prevalence, nor the performance, of each type of device.

Inverse modeling is a powerful approach to observation-based inferences about atmospheric model inputs (e.g. emissions). Hakami et al. (2005) developed a 4DVar in-

ACPD
15, 20851–20879, 2015

verse modeling method for the recovery of BC emissions, and sizable improvements were found at sub-regional levels. However, the domain-wide emissions inventory did not change significantly because measurements at four observation sites only were used to inverse and assimilate. Employing a multiple linear regression model, Fu et al. (2012) derived a "top-down" emissions estimate of annual BC and organic carbon (OC). They discovered that, when emissions of BC increased from 1.92 to 3.05 Tg yr^{-1}, the average model simulation of annual mean BC increased from 1.1 to 1.9 µg m^{-3}, which showed better agreement with the observed amount (2.5 µg m^{-3}).

The ensemble Kalman filter (EnKF), introduced by Evensen (1994), a technology based on ensemble forecasting and Kalman filter theory, has been successfully employed in atmospheric chemistry analyses, such as dust storm and aerosol data assimilation (Lin et al., 2008; Sekiyama et al., 2010; Tang et al., 2013). EnKF has some advantages over 4DVar insofar as it does not require the reconstruction of an adjoint model, which is technically difficult and cumbersome for the complex chemical transport model. However, the algorithm is highly sensitive to ensemble size (Mitchell et al., 2002), and therefore tends to be computationally demanding and has limited use in large-scale and on-line atmospheric chemical transport models. Moreover, the EnKF method assumes that the probability density functions (PDFs) of the initial conditions, emissions and observations are Gaussian in their distributions. When there is large bias, problems such as filer-divergence will lead to analysis failure.

In this paper, an ensemble optimal interpolation (EnOI) data assimilation method is used to investigate the possibility of optimally correcting the spatially resolved emissions bias of BC. The background-error covariances are estimated using the ensemble, but the model only needs a single forecast, allowing the use of a larger ensemble than EnKF. The preliminary results for the inversed emissions for BC are presented in this paper. The details of the methodology are described in Sect. 2, followed by a description of the model and the observations used in Sect. 3. The inverse modeling results are presented in Sect. 4, and a summary and discussion is provided in Sect. 5.

ACPD
15, 20851–20879, 2015

2 Methodology

2.1 Inverse theory and formulation

Air quality models can be generally written as

$$\begin{aligned} \boldsymbol{C}(x,t_{k+1}) &= \mathbf{M}_k \boldsymbol{C}(x,t_k) + \boldsymbol{Q}(x,t_{k+1}) + \boldsymbol{q}(x,t_{k+1}) \\ \boldsymbol{Q}(x,t_{k+1}) &= \boldsymbol{Q}^b(x,t_{k+1}) + \boldsymbol{Q}'(x,t_{k+1}) \end{aligned}, \quad (1)$$

where $\boldsymbol{C}$ is the vector of pollutant concentrations, x is the spatial location, $\boldsymbol{Q}$ represents emissions, $\mathbf{M}_k$ is the model time forward operator (could be nonlinear), and $\boldsymbol{q}$ is the model error, a random variable.

The observations are assumed to be also available at time $t_0, t_1, \cdots, t_K$:

$$\boldsymbol{y}(t_k) = \mathbf{H}_k[\boldsymbol{C}(x,t_k)] + \boldsymbol{r}_k, \quad (2)$$

where $\boldsymbol{r}$ represents observation errors. The operator $\mathbf{H}$ in Eq. (2) is the projection from the whole model domain to observation locations.

Without observations, we can carry out a simulation with a given initial condition $\boldsymbol{C}^b(x,t_0)$, emission inventory $\boldsymbol{Q}^b(x,t_k),(k=1,\cdots,\boldsymbol{K})$, and ignore model error $q^b(x,t_k) = 0,(k=1,\cdots,K)$ to solve Eq. (1) and to obtain a numerical solution $\boldsymbol{C}^b(x,t_k),(k=0,1,\cdots,\boldsymbol{K})$. The estimation problem discussed here is that, given observations (2), can we determine a better estimation of $\boldsymbol{C}(x,t_k)$ and $\boldsymbol{Q}(x,t_k)$ than $\boldsymbol{C}^b(x,t_k)$ and $\boldsymbol{Q}^b(x,t_k)$, which is also called the background or the first guess. Generally, this problem can be approached from the perspective of how to find the PDFs of the emissions $\boldsymbol{Q}(x,t_k)$ and associated model (1) solution $\boldsymbol{C}(x,t_k)$, given observations $\boldsymbol{y}(t_k)$.

Zhu and Wang (2006) introduced the formulation of this estimation theory given the PDFs of the initial condition, emissions and observations are Gaussian distributed. The inverse method seeks an optimal estimate of the emission that is consistent with both the observation and priori constraints of source by minimizing the following cost

ACPD
15, 20851–20879, 2015

function:

$$J(\boldsymbol{C}_k,\boldsymbol{Q}_k)=\iint\begin{bmatrix}\boldsymbol{C}_k(x)-\boldsymbol{C}_k^f(x)\\ \boldsymbol{Q}_k(x)-\boldsymbol{Q}_k^b(x)\end{bmatrix}^T\begin{pmatrix}\mathbf{W}_{Ck}(x,x') & \mathbf{W}_{CQk}^T(x,x')\\ \mathbf{W}_{CQk}(x,x') & \mathbf{W}_{Qk}(x,x')\end{pmatrix}\begin{bmatrix}\boldsymbol{C}_k(x')-\boldsymbol{C}_k^f(x')\\ \boldsymbol{Q}_k(x')-\boldsymbol{Q}_k^b(x')\end{bmatrix}\mathrm{d}x\mathrm{d}x' + [\boldsymbol{y}_k-\mathbf{H}_k(\boldsymbol{C}_k)]^T\mathbf{W}_o[\boldsymbol{y}_k-\mathbf{H}_k(\boldsymbol{C}_k)] \quad (3)$$

Based on the Kalman filter, we have the following analysis equations to minimizing the cost function:

$$\begin{pmatrix}\boldsymbol{C}_k^a\\ \boldsymbol{Q}_k^a\end{pmatrix}=\begin{pmatrix}\boldsymbol{C}_k^f\\ \boldsymbol{Q}_k^f\end{pmatrix}+\mathbf{K}\left(\boldsymbol{y}_k-H_k(\boldsymbol{C}_k^f)\right), \quad (4)$$

$$\mathbf{K}=\begin{pmatrix}\mathbf{V}_{Ck}(x,x')\mathbf{H}_k^T\\ \mathbf{V}_{CQk}(x,x')\mathbf{H}_k^T\end{pmatrix}\left(\mathbf{H}_k\mathbf{V}_{Ck}(x,x')\mathbf{H}_k^T+\mathbf{V}_o\right)^{-1}$$

2.2 Ensemble approach

To solve Eq. (4), we can use the ensemble method. First, we generate the N ensemble of emission, and then integrate model (1) N times from t_0 to t_1.The resulting N model outputs $C_1^{(j)-}$ and N emission ensembles $Q_1^{(j)-}$ form a joint state–emissions ensemble,

$$\mathbf{A}_1=\begin{pmatrix}C_1^{(1)-} & C_1^{(2)-} & \cdots & C_1^{(N)-}\\ Q_1^{(1)-} & Q_1^{(2)-} & \cdots & Q_1^{(N)-}\end{pmatrix}. \quad (5)$$

This joint ensemble can be used to construct the joint error covariance

$$\begin{pmatrix}\mathbf{V}_{Ck}(x,x') & \mathbf{V}_{CQk}^T(x,x')\\ \mathbf{V}_{CQk}(x,x') & \mathbf{V}_{Qk}(x,x')\end{pmatrix}=\frac{1}{N-1}\mathbf{A'}_1\mathbf{A'}_1^T, \quad (6)$$

where $\mathbf{A}'_1$ denotes the ensemble anomaly referenced to the ensemble mean.

ACPD
15, 20851–20879, 2015

The analysis includes updating each $C_1^{(j)-}$ and $Q_1^{(j)-}$ in ensemble form as follows:

$$\begin{pmatrix} C_1^{(j)+} \\ Q_1^{(j)+} \end{pmatrix} = \begin{pmatrix} C_1^{(j)-} \\ Q_1^{(j)-} \end{pmatrix} + \mathbf{K}\left(y_1^{(j)} - \mathbf{H}_1(C_1^{(j)-})\right), \tag{7}$$

where $y_1^{(j)}, j = 1, \cdots, N$ are the ensembles of perturbed observation samples.

Then, the estimated emission is

$$Q_1^a = \frac{1}{N}\sum_{j=1}^{N} Q_1^{(j)+}. \tag{8}$$

2.3 Ensemble Optimal Interpolation (EnOI)

The EnOI analysis is computed by solving an equation similar to Eq. (4), written as

$$\begin{pmatrix} C^a \\ Q^a \end{pmatrix} = \begin{pmatrix} C^f \\ Q^b \end{pmatrix} + \mathbf{K}\left(y - \mathbf{H}C^f\right), \tag{9}$$

where the analysis is computed in the stationary ensemble of model states and emissions sampled during a long-term integration. Once the ensemble is created, only one single model state and emission are updated. The method greatly reduces the computation cost and provides a sub-optimal solution compared to the EnKF approach.

3 Model description and BC measurements

The model used in this study is an online-coupled chemical weather forecasting system, GRAPES/CUACE, which consists of two components: (1) GRAPES, which refers to a mesoscale meteorological model, the Global/Regional Assimilation and Prediction

ACPD
15, 20851–20879, 2015

System, developed by the China Meteorological Administration (CMA). It produces meteorological fields (winds, turbulence, precipitation etc.) to drive the CUACE chemistry model. (2) CUACE, the Chinese Unified Atmospheric Chemistry Environment model, which includes emissions, transport, dry and wet depositions, and removal both in and below clouds of both gases and aerosols (Gong et al., 2008; Wang et al., 2009; Zhou et al., 2012). These two parts are coupled online. Using the national official basic information of emission sources published in 2005 based on the bottom-up inventory developed by Cao et al. (2005), the Emission Subsystem (EMIS) provides hourly gridded offline emission intensities of 32 species, including seven categories of aerosol species (sea salt, sand/dust, BC, OC, sulfates, nitrates, and ammonium salts) in the aerosol modules over the model domain. The basic gridded BC emissions inventory is based on energy consumption and activity information for various emission sectors: industry, residential, transport, power generation, agriculture, biomass burning, and others.

The model domain for this study is approximately (70–140°, 17–57° E), as shown in Fig. 1. The computational domain consists of 140 × 80 × 32 grid cells, with a horizontal resolution of 0.5°, and 31 vertical layers from the ground to 200 hPa. The simulation period is the month of January 2008.

The near-real time (NRT) data used in this work are the surface daily and hourly BC concentrations collected from over 30 CMA Atmosphere Watch Network (CAWNET) stations. The locations of these sites are shown in Fig. 1. Observations considered for assimilation are shown as red circles, and observations considered for verification are shown as blue stars. Information of the stations is presented in Table 1. The rural stations are typically located some 100 km away from local pollutant sources or nearby major cities, and at moderate height above the area's local elevation. At the urban stations, the sampling heights are 50–100 m higher than the city's average elevation. This enables the production of samples that are representative of the region, rather than the immediate locality.

ACPD
15, 20851–20879, 2015

The BC observations are obtained using Aethalometer (Model AE-31, Magee Scientific, Berkeley, California, USA) instruments, which measure optically-absorbing filterable aerosol material on a 5 min time base (Zhang et al., 2008) at seven wavelengths of 370, 470, 520, 590, 660, 880 and 950 nm. The BC concentrations used in this study are derived from optical absorption at 880 nm.

4 Inverse modeling of BC emissions over China

4.1 Experiment design

The monthly mean of simulated BC is put into the equation as C, and the y also uses the monthly mean of the BC observations. To obtain the ensemble for calculating the K, we make perturbations of the emissions (30 members), put them in the model, and run the model for 24 h. Subsequently, we obtain 24 × 30 ensembles. The reason we use monthly mean data is to reduce the effects of other factors of model error, such as meteorology.

4.2 Ensemble strategy

A good ensemble system should satisfy at least two conditions: (1) the ensemble mean should be close to the truth; (2) the ensemble spread should be a reasonable representation of the root-mean-square error (RMSE) between the ensemble and the truth. Because the atmospheric components are usually log-normal distributed, we form the perturbation of the emissions according the following PDF:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}} \frac{1}{x}. \quad (10)$$

Here, we use the procedure proposed by Evensen (2004) to generate the pseudo-random fields, which is able to compute them with the mean equating to zero, variance equating to one, and a specified covariance that determines the smoothness

ACPD
15, 20851–20879, 2015

of the fields. Using the two-dimensional fast Fourier transform (FFT), we obtain two-dimensional perturbations of the emissions, and then use singular vector decomposition (SVD) to obtain the optimal ensembles. In this paper, we compare two schemes: Scheme A, which only considers the spatial correlation,

$$\boldsymbol{Q}'_i(x,t) = \boldsymbol{Q}(x,t) + q_i(x) \cdot \sigma_Q(x,t), \tag{11}$$

and Scheme B, which considers both the spatial and temporal correlation,

$$\begin{aligned} \boldsymbol{Q}'_i(x,t) &= \boldsymbol{Q}(x,t) + q_i(x,t) \cdot \sigma_Q(x,t) \\ q_i(x,t) &= q_i(x,t-1) + \sqrt{1-\alpha^2} \cdot w_i(x) \end{aligned}. \tag{12}$$

To test whether the sample strategies are good, we compare the ensemble of the emissions and BC concentrations. Figure 2 shows the ensemble of the emissions for the two schemes. Although the two schemes produce the perturbations of the emissions with the same variance, the ensemble of the emissions by Scheme B has a much larger spread then Scheme A. Figure 3 shows the ensemble of BC concentration simulations computed by the corresponding emissions, with the red dots representing the observations. It is clear that the ensemble simulation of Scheme B can include the observation more reasonably.

4.3 Localization in the inverse modeling

The inverse modeling of the BC emissions in this study uses the ensemble member to calculate the background-error covariance matrix between the concentrations and emissions. The accuracy of the matrix is a very important factor for the inverse result. Since we cannot obtain as many ensembles as model dimensions, we should employ appropriate techniques that eliminate the effects of sample error and that localize the impact of an observation to a subset of the model state variables. Figure 4a shows the correlation of the background-error between the sites concentrations and emissions. The figure shows that the site concentration is highly correlated with the emissions

ACPD
15, 20851–20879, 2015

near the site, which is expected. However, there are some emissions from remote regions that are also correlated with the site concentration, i.e., spurious correlations. Localization is an essential tool for an ensemble-based assimilation to adequately span the model sub-space.

Usually, a typical implementation of localization involves the multiplication of the ensemble-based covariance by a correlation function, so the gain matrix is re-expressed as Eq. (13). We use the distance-dependent covariance localization scheme,

$$\mathbf{K} = (\mathbf{C} \circ \mathbf{P}^b)\mathbf{H}^T[\mathbf{H}(\mathbf{C} \circ \mathbf{P}^b)\mathbf{H}^T + \mathbf{R}]^{-1}, \quad (13)$$

where $\mathbf{C}$ is a correlation matrix and the operation denoted by the open circle is an element-by-element matrix multiplication (also called a Schur product).

Here, we use an elliptic function with e folding scales of L_a (major axis) and L_b (minor axis). When the distance-dependent covariance localization scheme is used, these covariances are artificially reduced to near zero with e folding length-scales of L_a and L_b. The localization scales are set by calculating the shape of the covariance matrix.

Figure 4c shows the correlation of the background-error between the site concentrations and emissions. Using this localization scheme, the spurious correlations far away from the site concentration is reduced, but the pattern of correlation around the site remains. Therefore, the inversion provides a more accurate estimation of emissions.

4.4 Results of BC emissions bias correction

A month-long simulation is conducted, driven by the current Chinese bottom-up emissions inventories for BC, and the results are compared with surface-observed BC concentrations (Fig. 5). The comparison of model results to observations evaluates the bottom-up inventories. The simulated average monthly mean BC concentration for all assimilated and verified sites is 1.152 $\mu g\,m^{-3}$, which is nearly 79 % lower than the observed value of 5.238 $\mu g\,m^{-3}$. Moreover, the model fails to capture the spatiotemporal

ACPD
15, 20851–20879, 2015

variability in the BC observations and underestimates the BC concentrations at almost all assimilated sites except site SD. In the central eastern China, where the region is considered as main contributor of BC emission, simulations are lower than observation. Moreover, sites in northwestern China , such as Gansu and Xinjiang province, where the region are considered to have low emission because their relatively underdeveloped economy, also present large bias with observations. This indicates that the bottom-up BC emissions are underestimated and spatiotemporally misrepresented.

We use the 27 BC monthly observations to inverse the emissions. Figure 6 shows the bottom-up or prior emissions (E1) and inversed emissions (E2). There are significant increases in the E2 emissions over many regions of China, including eastern China, central china, Sichuan Basin, and northwestern China. Only region around Beijing and northeast part of Heibei province presents a bit of decrease in E2, because model simulation at the assimilated site SD is higher than the observation. The results show that the basic emissions produced by the bottom-up method are underestimated in most regions of China in January, not only in eastern China and central china where the rural population density and economic level are high, but also in northwestern China which have lower rural population densities and lower economic level.

We compare our inversed provincial and national emission of BC with bottom-up inventories (Table 2). The five largest contributions by province in E1 are from Heibei, Shangdong, Henan, shanxi and Sichuan. In E2 by the inversion, the five largest BC emission province are same with E1, but emissions are greatly enhanced in many provinces. We find emission from half of provinces in China as to be increased by a factor of over 2, such as Xinjiang, Gansu, Sichuan province. The resulting estimate emission over China in January is 240.1 Gg, and lead to estimation of annual emission being about 2750 Gg (= 2.75 Tg), which is over 1.8 times of bottom-up emission inventory (1.5 Tg).

Previous bottom-up estimation of Chinese emission range from 1.05 to 1.92 Tg (Streets et al., 2003; Zhang et al., 2009; Qin and Xie, 2011). Our inversed emission are nearly 40 % higher than these inventory. Fu et al. (2012) discussed why the previous

ACPD
15, 20851–20879, 2015

bottom up emission were so similar low. One possible reason is that these bottom-up inventories are based on same statistics data source. By top-down regression method they estimate the annual BC emission is 3.05 ± 0.78 Tg, which is consistent with annual emission of our estimation.

Table‘3 lists the monthly mean BC concentrations at the observation sites. It shows that there are large errors between the monthly BC concentration simulations driven by the bottom-up emissions (E1) and observations. Most of them feature a positive bias. After EnOI inversion, the model simulation for most observation sites in China is much closer to the observation, even for the verification sites. The error percentage decreases from 78.97 to 39.54 %, which is an almost 50 % decrease.

We investigate the variations in daily BC concentrations in January 2008 (Fig. 7). The red line shows the observations; the blue line is the model simulation driven by bottom-up emissions (E1); and the green line is the model simulation driven by inversed emissions (E2). Even though we only employ the monthly mean BC measurements to inverse the emissions, the accuracy of the daily model simulation is also improved. Take Zhenzhou site (ZZ) for example, driven by inversed emission inventory, the daily simulation concentration and its variation are more consistent with observation. Observations in GL site and LA site exhibited a peak during 2–8 January, but simulation with E1 does not present, whereas simulation with E2 does. We calculate the RMSE between the daily model simulation and observation (Fig. 8). The blue bar is the RMSE between the daily model simulation driven by bottom-up emissions (E1) and the observation, and the red bar is the RMSE between the daily model simulation driven by the inversed emissions (E2) and the observation. As we can see, almost all the RMSEs decrease, with the average RMSE dropping from 5.08 to 3.47, which is a decrease of about 31.56 %.

From the Figures and tables above, it is apparent that the biases of the BC emission intensities in China at each grid point are corrected by the EnOI inversion system. Where the bias is large, the RMSE decreases significantly. However, if there is small bias, such as in the ZhuZhang (XG) site, the correction is tiny, and the RMSE de-

ACPD
15, 20851–20879, 2015

creases only slightly. We also find that, near Beijing, emissions are considered to be overestimated, and this is because the observation in the SD site has negative bias. However, after the inversion, the BC concentration at the verification site (BJ) also increases, that reason is the emissions around Beijing and southwest of Heibei province were underestimated, and been corrected. The accuracy of the SD site changes negligibly. We believe that the reason is that the horizontal grid resolution of 50 km not high enough to distinguish between two very close observation sites with different variation. This problem could be solved if we use a model and emissions inventory with higher resolution.

5 Summary and discussion

In this study an inverse modeling system is developed for BC emissions in an online coupled chemical weather forecasting system, GRAPES/CUACE, using the inexpensive EnOI methodology. The emissions sampling strategy is discussed and improved. With its time correlation strategy, the ensemble forecast can have a larger spread to include the observations. The effect of localization in the analysis is also studied. With reasonable localization, the effects of sample error and spurious correlations are reduced.

BC aerosols in China are simulated and compared to surface measurements, with the goal of deriving top-down BC emissions bias estimates. We conduct a month-long simulation for January 2008, driven by the current Chinese bottom-up BC emissions inventories. Comparison of the model results to observations at background and rural sites present large biases in model simulation. The simulated average monthly mean BC concentration from all assimilated sites is 1.152 $\mu g\,m^{-3}$, which is 78 % lower than the observed amount (5.238 $\mu g\,m^{-3}$), indicating the BC source is underestimated on a national scale, likely due to being short of accurate and detailed activity statistics, such as small-scale industrial activity, domestic combustion, and transportation. There are also missing sources in northwestern provinces including Gansu, Xinjiang province,

ACPD
15, 20851–20879, 2015

which have lower rural population densities and lower economic level. This indicates that the bottom-up BC emissions are not only too low, but spatiotemporally misrepresented. The monthly mean of the simulated BC is put into the Eq. (9) as the background forecast, and the monthly mean of the BC observations is considered as the observation. With the inversion of EnOI, the bias of the BC emissions intensity in China at each grid point is corrected by the inverse system. We find emission from half provinces in China as to be enhanced by a factor of 1.8. The emission over China in January by our estimation is 240.1 Gg, and annual BC emission is about 2750 Gg. The monthly mean eliminates the meteorological factors and other uncertainties, so the underestimation comes mainly from the emissions, which explains the low biases in the simulated concentrations using inversed emission inventory. Applying top-down emissions inventory, the simulated average monthly mean concentration at all sites is improved to 3.282 $\mu g\,m^{-3}$. Despite only employing the monthly mean BC measurements to inverse the emissions, the accuracy of the daily model simulation also improves. The average RMSE drops from 5.08 to 3.47 (a decrease of approximately 31.56 %), showing that where the bias is large, the RMSE decreases significantly. It is also found that the performance of EnOI may fail under the circumstances when the horizontal grid resolution is not high enough to distinguish between two very close observation sites with different variations. Our study finds that EnOI is a useful and computation-free method to correct the bias of aerosol emissions. These results are valuable and promising for a better understanding of aerosol emissions and distributions, as well as aerosol forecasting. Our top-down EC emission estimates are 50–200 % higher than the basic bottom-up estimates. We plan to explore the use of other data, such as satellite-derived aerosol optical depth observations, to evaluate our results. Many gaps remain in our understanding of Chinese carbonaceous aerosol sources. We suggest that EnOI may provide a practical and cost-effective alternative to the EnKF for correction of the aerosol emissions bias where computational cost is a limiting factor. In future work, we intend to use EnOI to correct the bias of every month's emissions in a year, and then

ACPD
15, 20851–20879, 2015

employ EnKF to optimally recover temporally resolved (e.g. daily timescale) emissions inventories of BC and other aerosols.

Acknowledgements. This work is supported by the National Basic Research Program (973) (2011CB403404), the National Natural Scientific Foundation of China (41305117) and Basic research fund of CAMS (2013Y005).

References

Andreae, M. O. and Merlet, P.: Emission of trace gases and aerosols from biomass burning, Global Biogeochem. Cy., 15, 955–966, doi:10.1029/2000GB001382, 2001.

Bond, T. C. and Bergstrom, R. W.: Light absorption by carbonaceous particles: An investigative review, Aerosol Sci. Technol., 40, 1–41, 2006.

Cao, G., Zhang, X., and Zheng, F.: Inventory of black carbon and organic carbon emissions from China, Atmos. Environ., 40, 6516–6527, doi:10.1016/j.atmosenv.2006.05.070, 2006.

Cao, J. J., Lee, S. C., Chow, J. C. Watson, J. G., Ho, K. F., Zhang, R. J., Jin, Z. D., Shen, Z. X., Chen, G. C., Kang, Y. M., Zou, S. C., Zhang, L. Z., Qi, S. H., Dai, M. H., Cheng, Y., and Hu, K.: Spatial and seasonal distributions of carbonaceous aerosols over China, J. Geophys. Res., 112, D22S11, doi:10.1029/2006JD008205, 2007.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics, J. Geophys. Res., 99, 10143–10162, 1994.

Evensen, G.: The Ensemble Kalman Filter for combined state and parameter estimation. IEEE Control Syst. Mag., 29, 83–104, 2009.

Fu, T.-M., Cao, J. J., Zhang, X. Y., Lee, S. C., Zhang, Q., Han, Y. M., Qu, W. J., Han, Z., Zhang, R., Wang, Y. X., Chen, D., and Henze, D. K.: Carbonaceous aerosols in China: top-down constraints on primary sources and estimation of secondary contribution, Atmos. Chem. Phys., 12, 2725–2746, doi:10.5194/acp-12-2725-2012, 2012.

Gong, S. L. and Zhang, X. Y.: CUACE/Dust – an integrated system of observation and modeling systems for operational dust forecasting in Asia, Atmos. Chem. Phys., 8, 2333–2340, doi:10.5194/acp-8-2333-2008, 2008.

Hakami, A., Henze, D. K., Seinfeld, J. H., Chai, T., Tang, Y., Carmichael, G. R., and Sandu, A.: Adjoint inverse modeling of black carbon during the Asian Pacific Regional Aerosol Characterization Experiment, J. Geophys. Res., 110, D14301, doi:10.1029/2004JD005671, 2005.

ACPD
15, 20851–20879, 2015

Hansen, A. D. A., Bodhaine, B. A., Dutton, E. G., and Schnell, R. C.: Aerosol black carbon measurements at the South Pole: initial results, 1986–1987, Geophys. Res. Lett., 15, 1193–1196, 1988.

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., Russell, G., Aleinov, I., Bauer, M., Bauer, S., Bell, N., Cairns, B., Canuto, V., Chandler, M., Cheng, Y., Del Genio, A., Faluvegi, G., Fleming, E., Friend, A., Hall, T., Jackman, C., Kelley, M., Kiang, N., Koch, D., Lean, J., Lerner, J., Lo, K., Menon, S., Miller, R., Minnis, P., Novakov, T., Oinas, V., Perlwitz, J., Perlwitz, J., Rind, D., Romanou, A., Shindell, D., Stone, P., Sun, S., Tausnev, N., Thresher, D., Wielicki, B., Wong, T., Yao, M., and Zhang, S.: Efficacy of climate forcings, J. Geophys. Res., 110, 1–45, 2005.

Hara, K., Osada, K., Yabuki, M., Hayashi, M., Yamanouchi, T., Shiobara, M., and Wada, M.: Measurement of black carbon at Syowa station, Antarctica: seasonal variation, transport processes and pathways, Atmos. Chem. Phys. Discuss., 8, 9883–9929, doi:10.5194/acpd-8-9883-2008, 2008.

Lin, C., Wang, Z., and Zhu, J.: An Ensemble Kalman Filter for severe dust storm data assimilation over China, Atmos. Chem. Phys., 8, 2975–2983, doi:10.5194/acp-8-2975-2008, 2008.

Liousse, C., Cachier, H., and Jennings, S. G.: Optical and thermal measurements of black carbon aerosol content in different environments: Variation of the specific attenuation cross-section, sigma, Atmos. Environ., 27A, 1203–1211, 1993.

Menon, S., Hansen, J., Nazarenko, L., and Luo, Y.: Climate effects of black carbon aerosols in China and India, Science, 297, 2250–2253, 2002.

Qin, Y. and Xie, S. D.: Spatial and temporal variation of anthropogenic black carbon emissions in China for the period 1980–2009, Atmos. Chem. Phys., 12, 4825–4841, doi:10.5194/acp-12-4825-2012, 2012.

Ramanathan, V. and Carmichael, G.: Global and regional climate changes due to black carbon, Nat. Geosci., 1, 221–227, 2008.

Sekiyama, T. T., Tanaka, T. Y., Shimizu, A., and Miyoshi, T.: Data assimilation of CALIPSO aerosol observations, Atmos. Chem. Phys., 10, 39–49, doi:10.5194/acp-10-39-2010, 2010.

Streets, D. G., Gupta, S., Waldhoff, S. T., Wang, M. Q., Bond, T. C., and Bo, Y.: Black carbon emissions in China, Atmos. Environ., 35, 4281–4296, 2001.

Streets, D. G., Bond, T. C., Carmichael, G. R., Fernandes, S. D., Fu, Q., He, D., Klimont, Z., Nelson, S. M., Tsai, N. Y., Wang, M. Q., Woo, J.-H., and Yarber, K. F.: An inventory of

ACPD
15, 20851–20879, 2015

gaseous and primary aerosol emissions in Asia in the year 2000, J. Geophys. Res., 108, 8809, doi:10.1029/2002JD003093, 2003a.
Streets, D. G., Yarber, K. F., Woo, J.-H., and Carmichael, G. R.: Biomass burning in Asia: Annual and seasonal estimates and atmospheric emissions, Global Biogeochem. Cy., 17, 1099, doi:10.1029/2003GB002040, 2003b.
Tang, X., Zhu, J., Wang, Z. F., Wang, M., Gbaguidi, A., Li, J., Shao, M., Tang, G. Q., and Ji, D. S.: Inversion of CO emissions over Beijing and its surrounding areas with ensemble Kalman filter, Atmos. Environ., 81, 676–686, 2013.
Wang, H., Gong, S. L., Zhang, H. L., Chen, Y., Shen, X. S., Chen, D. H., Xue, J. S., Shen, Y. F., Wu, X. J., and Jin, Z. Y.: A new-generation sand and dust storm forecasting system GRAPES_CUACE/Dust: Model development, verification and numerical simulation, Chin. Sci. Bull., 55, 635–649, doi:10.1007/s11434-009-0481-z, 2009.
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, doi:10.5194/acp-9-5131-2009, 2009
Zhang, X. Y., Wang, Y. Q., Wang, D., Gong, S. L., Arimoto, R., Mao, L. J., and Li, J.: Characterization and sources of regionalscale transported carbonaceous and dust aerosols from different pathways in costal and sandy land areas of China, J. Geophys. Res., 110, D15301, doi:10.1029/2004JD005457, 2005.
Zhang, X. Y., Wang, Y. Q., Zhang, X. C., Guo, W., Gong, S. L., Zhao, P., and Jin, J. L.: Carbonaceous aerosol composition over various regions of China during 2006, J. Geophys. Res., 113, D14111, doi:10.1029/2007JD009525, 2008.
Zhou, C. H., Gong, S. L., Zhang, X. Y., Liu, H. L., Xue, M., Cao, G. L., An, X. Q., Chen, H. Z., Zhang, Y. M., and Niu, T.: Towards the improvements of simulating the chemical and optical properties of Chinese aerosols using an online coupled model – CUACE/Aero, Tellus B, 64, 18965, doi:10.3402/tellusb.v64i0.18965, 2012
Zhu, J. and Wang, P.: Ensemble Kalman smoother and Ensemble Kalman filter aproaches to the joint air quality state and emission estimation problem, Chin. J. Atmos. Science, 30, 871–882, 2006.

ACPD
15, 20851–20879, 2015

Table 1. Observation site information.

num	observation sites	LON	LAT	ALT
1	AKeDaLa (AK)	87.97	47.12	562
2	TaZhong (TZ)	83.67	39	1099.3
3	HaMi (HM)	93.52	42.82	737.2
4	EJiNaQi (EJ)	101.07	41.95	940.5
5	DunHuang (DH)	94.68	40.15	1140
6	WaLiGan (WL)	100.92	36.28	3816
7	ZhuRiHe (ZR)	112.9	42.4	1151.9
8	YuLin (YL)	109.2	38.43	1105
9	YuShe (YS)	112.98	37.07	1041.4
10	LongFengShan (LF)	127.6	44.73	330.5
11	XiLinHaoTe (XL)	116.12	43.95	1003
12	TongLiao (TL)	122.27	43.6	178.7
13	FuShun (FS)	123.95	41.88	163
14	GuCheng (GC)	115.8	39.13	11
15	DaLian (DL)	121.63	38.9	91.5
16	ChengDu (CDu)	104.04	30.65	553
17	ZhuZhang (XG)	99.73	28.02	3580
18	ZhengZhou (ZZ)	113.68	34.78	110
19	XiAn (XA)	108.97	34.43	410
20	GuiLin (GL)	110.3	25.32	164.4
21	LinAN (LA)	119.73	30.3	138.6
22	LuShan (LS)	115.99	29.57	1165
23	NanNing (NN)	108.35	22.82	172
24	PanYu (PY)	113.35	23	131
25	GaoLanShan (GLs)	105.85	36	2161.5
26	ChangDe (CD)	111.71	29.17	565
27	ShangDianZi (SD)	117.12	40.65	293.3
28	ShenYang(SY)	123.41	41.76	110
29	Beijing (BJ)	116.47	39.8	31.3
30	HuiMin (HM)	117.53	37.48	11.7
31	JinSha (JS)	114.2	29.63	330.5

ACPD
15, 20851–20879, 2015

Table 2. Estimates of provincial BC emission (Gg) by bottom-up and inversed method.

Province	base (E1)	inverse (E2)	change (E2/E1)	Province	base (E1)	inverse (E2)	change (E2/E1)
Hebei	12.1	22.3	1.8	Jilin	3.9	7.8	2.0
Shanxi	11.0	17.8	1.6	Zhejiang	1.8	4.7	2.6
ShanDong	10.3	13.6	1.3	Fujian	1.4	2.3	1.6
Henan	9.9	19.3	1.9	Jiangxi	1.9	3.7	1.9
Sichuan	6.4	16.4	2.6	Guangdong	2.7	5.0	1.9
Liaoning	6.3	13.1	2.1	Hainan	0.2	0.2	0.9
Nei Mongol	6.2	11.9	1.9	Xizhang	0.3	0.3	1.0
Jiangsu	5.5	13.2	2.4	Gansu	3.1	8.1	2.6
Hubei	5.3	11.3	2.1	Qinhai	0.5	0.9	1.9
Anhui	4.8	8.2	1.7	Ningxia	1.1	2.8	2.6
Heilongjiang	4.5	6.5	1.4	Xinjiang	3.2	9.4	2.9
Hunan	4.1	10.8	2.6	Chongqin	1.8	2.9	1.6
Guizhou	4.1	5.0	1.2	Beijng	2.8	2.6	0.9
Yunan	3.5	4.0	1.1	Tianjin	1.1	1.3	1.2
Guangxi	3.2	6.4	2.0	Shanghai	0.3	0.7	2.0
Shannxi	3.1	7.7	2.5				
Total in January	126.5	240.1	1.9				
Year	1449.2	2750.6					

ACPD
15, 20851–20879, 2015

Table 3. Model simulations and surface observations of monthly mean BC concentrations at assimilation sites and verification sites (units: $\mu g\,m^{-3}$) and the relative error percentage (= (|model – obs|/obs) × 100 %).

site	Model (E1)	Model E2)	observation	Relative error percentage (E1)	Relative error percentage (E2)
AK	0.07	0.51	0.44	86.9	13.0
TZ	0.04	2.20	1.08	98.4	51.1
HM	0.06	4.90	2.21	98.9	54.9
EJ	0.05	7.84	2.67	99.4	66.0
DH	0.06	3.55	1.02	98.4	71.4
WL	0.13	0.94	1.03	85.7	9.3
ZR	0.14	3.37	1.00	95.7	70.2
YL	0.31	1.88	0.89	83.6	52.9
YS	2.70	6.94	5.56	61.1	19.9
LF	0.58	5.16	2.23	88.8	56.8
XL	0.14	0.93	0.37	84.7	59.8
TL	0.47	7.42	2.97	93.6	59.9
FS	2.00	7.06	4.82	71.7	31.6
GC	3.79	14.24	7.60	73.4	46.7
DL	1.74	4.85	4.13	64.1	14.8
CDu	1.45	9.71	7.14	85.0	26.5
XG	0.20	0.30	0.21	34.8	29.2
ZZ	3.28	10.89	10.68	69.9	2.0
XA	1.02	3.66	3.57	72.1	2.5
GL	0.26	4.99	1.92	94.8	61.5
LA	1.00	6.19	4.19	83.8	32.3
LS	0.73	2.08	2.69	65.0	29.4
NN	0.55	6.36	2.26	91.3	64.5
PY	0.55	8.69	2.15	93.6	75.2
GL	0.46	4.11	1.82	88.8	55.8
CD	0.71	4.38	3.02	83.7	31.0
SD	1.39	0.81	1.38	71.8	70.5
BJ	3.45	11.96	8.68	71.2	27.4
HM	4.27	8.06	6.41	47.0	20.4
JS	3.20	5.35	4.47	40.1	16.4
SY	0.89	3.05	3.14	70.7	2.8
Assi_sites mean	0.88	4.96	2.93	82.19	42.91
Veri_sites mean	2.95	7.10	5.67	57.24	16.76
All_sites mean	1.15	5.24	3.28	78.97	39.54

ACPD
15, 20851–20879, 2015

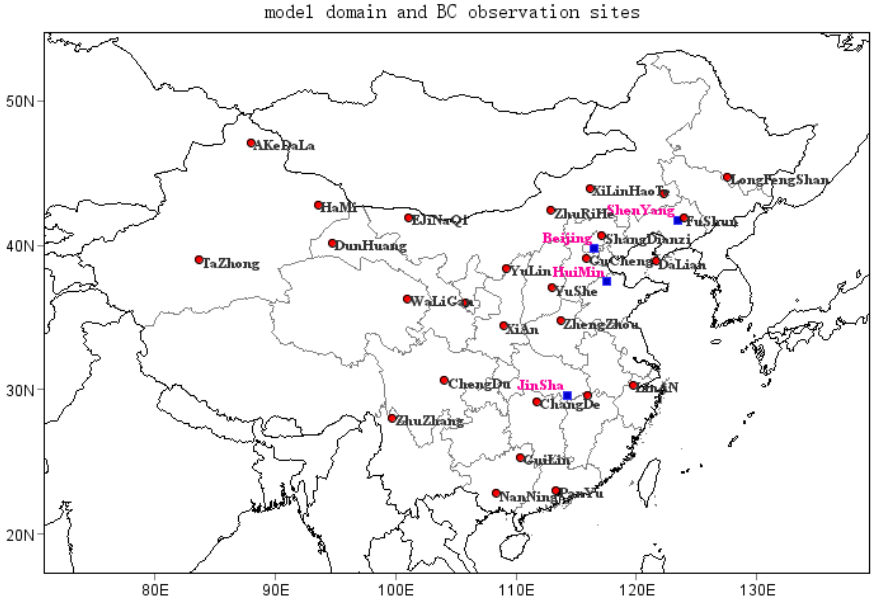

Figure 1. The GRAPES_CUACE model domain and BC observation sites used in this study. Observations considered for assimilation are shown as red points, while observations considered for verification are shown as blue squares. Information about stations is shown in Table 1.

ACPD
15, 20851–20879, 2015

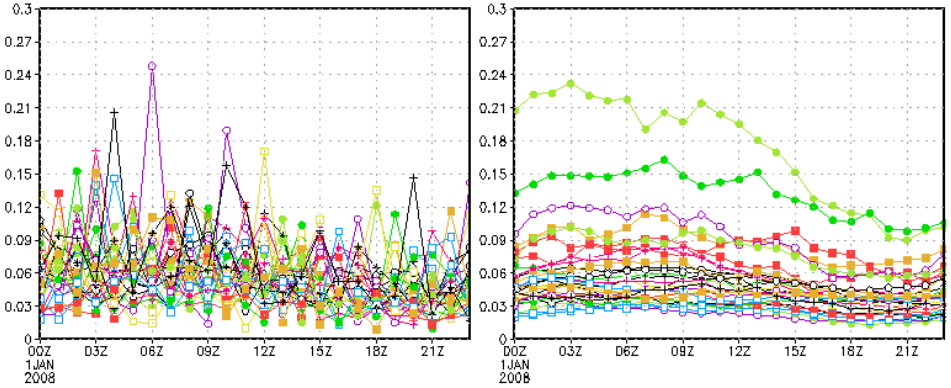

Figure 2. Emission ensembles at grid (116.5° N, 40° E): sample strategy with no time correlation (left); sample strategy with time correlation afa = 0.9 (right) (units: mg s m^{-2}).

ACPD
15, 20851–20879, 2015

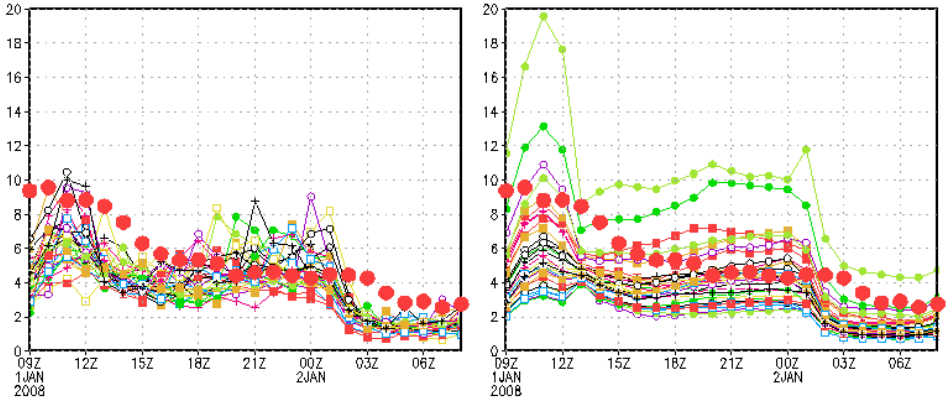

Figure 3. The ensembles of BC model simulation at grid 116.5° N, 40° E. The left one uses sample strategy with no time correlation. The right one uses sample strategy with time correlation afa = 0.9. The red points is the daily variation of BC observations (units: $\mu g\,m^{-3}$).

ACPD
15, 20851–20879, 2015

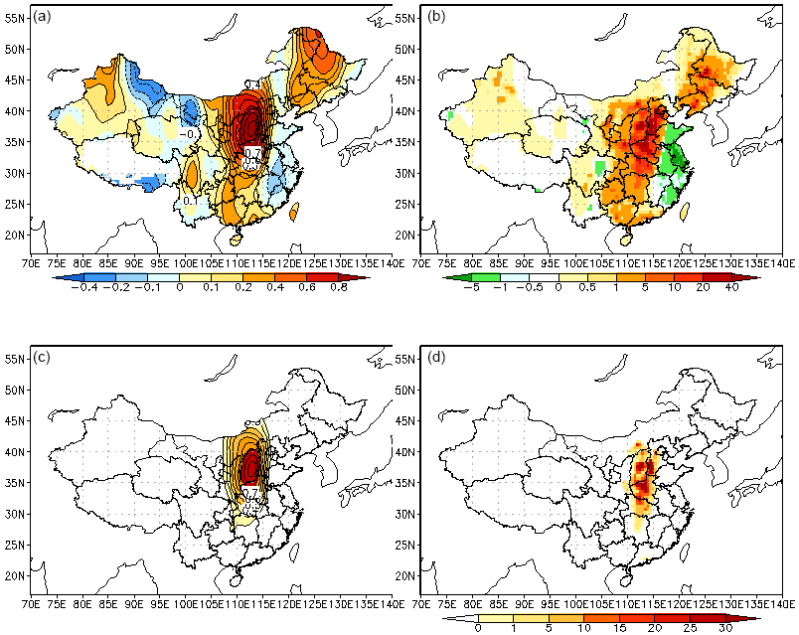

Figure 4. BC emissions inversed by observations from site Zhengzhou (ZZ): **(a)** back-ground error correlation between site ZZ and other grids without localization; **(b)** BC emissions inversed by observations from site ZZ without localization; **(c)** background-error correlation between site ZZ and other grids with localization; **(d)** BC emissions inversed by observations from site ZZ with localization.

ACPD
15, 20851–20879, 2015

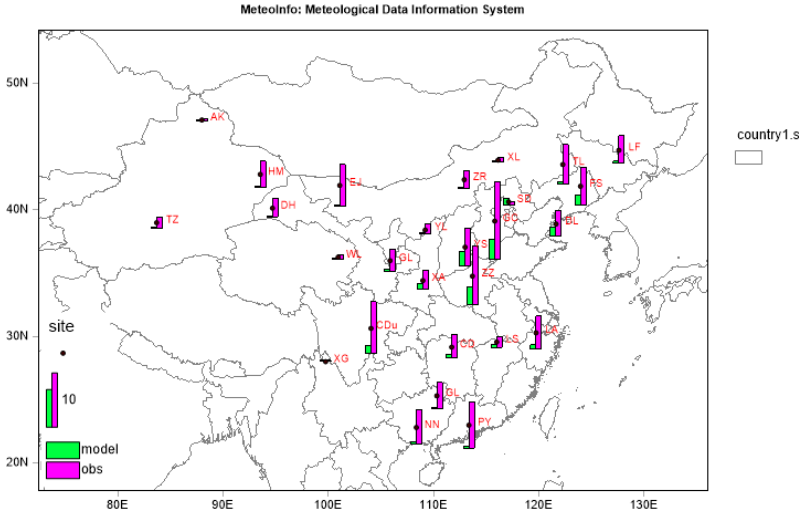

Figure 5. BC monthly mean concentrations for January 2008 at assimilation observation sites. Green bars show the model-simulated BC using the bottom-up emissions inventory; pink bars show the observed surface BC concentrations (units: $\mu g\,m^{-3}$).

ACPD
15, 20851–20879, 2015

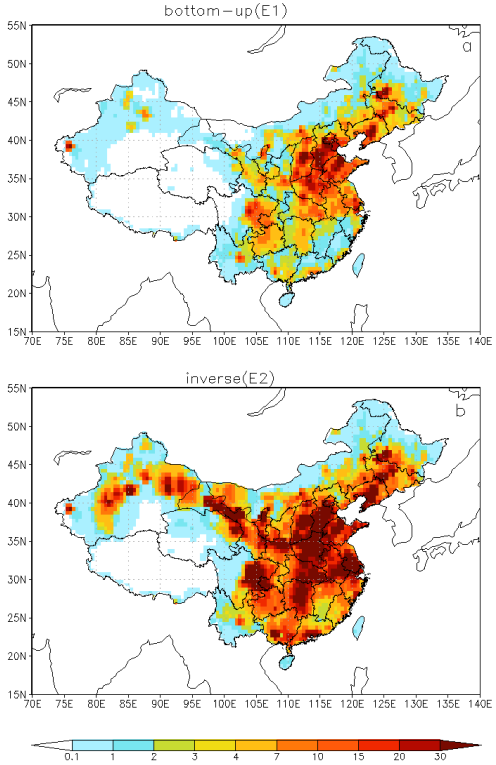

Figure 6. BC emissions from China inversed by 27 observation sites (units: $\mu g\,s\,m^{-2}$): **(a)** bottom-up emissions, E1; **(b)** inversed emissions, E2.

ACPD
15, 20851–20879, 2015

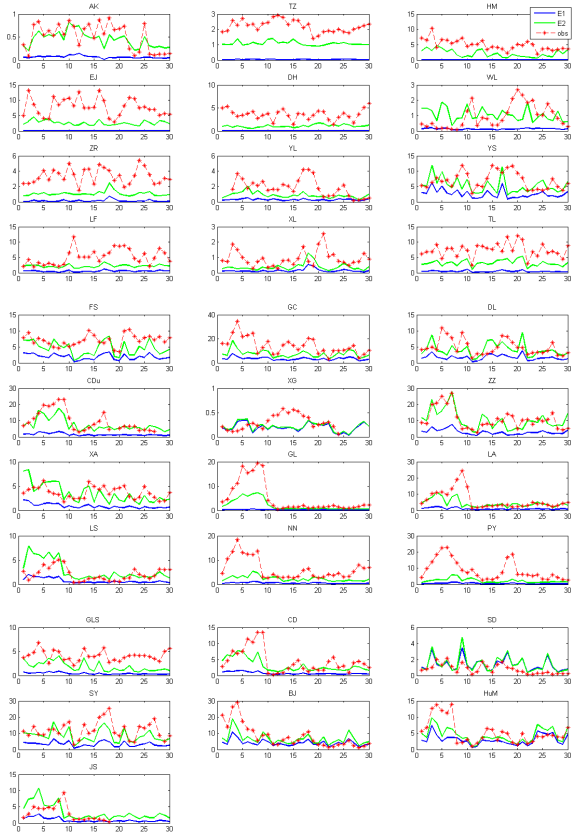

Figure 7. BC daily concentrations in January 2008 (units: $\mu g\,m^{-3}$). The red dotted line shows the observation; the blue line is the model simulation driving by bottom-up emissions (E1); and the green line is the model simulation driven by inversed emissions (E2).

ACPD
15, 20851–20879, 2015

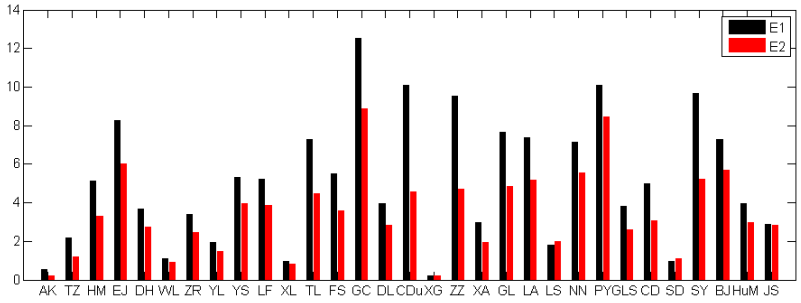

Figure 8. The root-mean-square error (RMSE) between the daily model simulation and observation. The black bars show the RMSE between the daily model simulation driven by bottom-up emissions (E1) and observations, and the red bars show the RMSE between the daily model simulation driven by inversed emissions (E2) and observations. (units: $\mu g\,m^{-3}$).

ACPD
15, 20851–20879, 2015