Oct. 31, 2018.

Atmos. Chem. Phys.

RE: Manuscript Number: acp-2018-768

Dear Editors:

Thank you very much for your kind decision letter on our paper entitled “The impact of multi-species surface chemical observations assimilation on the air quality forecasts in China” (acp-2018-768). We are grateful for the helpful comments from you and the reviewers. We have changed the manuscript according to the reviewer’s suggestions. The main changes include: 1) A simulation using the optimized emissions from 5 to 16 October 2014 were performed to investigate the impact of optimized emissions on chemical simulations; 2) We have rewritten the experimental design in Section 4. All the scientific questions have been resolved in the revised version (Please see details in it). So we hope this manuscript will be published in ACP. We are looking forward to hearing from you soon.

Sincerely Yours,

Zhen Peng

Response to Reviewer #1's comments:

We thank Referee # 1 for his thoughtful comments and suggestions that have helped to improve our manuscript. Our responses to comments (in bold style) and the corresponding changes to the manuscript are detailed below.

Summary and general comments:

This manuscript investigated the application of ensemble Kalman filter (EnKF) for constraining the atmospheric chemical species including PM_{10}, $PM_{2.5}$, SO_2, NO_2, O_3 and CO. The simultaneous assimilation of various surface air quality measurements improved the representation of the initial conditions and emission factors of aforementioned species, as well as their 72-hours forecasts. This investigation on the assimilation of various air quality observations for a severe haze pollution event provides a promising case study for the regional air-quality modeling. I would recommend the minor revision with the considerations of several issues as listed below.

List of minor comments:

1) Section 2.1: Which dataset (reanalysis) did you use for the meteorological initial and boundary conditions? Were the perturbations also added to the meteorology? If not, please add one or two sentences to mention that the uncertainty of the meteorology forecasts is not considered in this study.

The meteorological initial and boundary conditions were provided by the National Centers for Environmental Prediction Global Forecast System (GFS). The temperature, water vapor, velocity, geopotential height and dry surface pressure fields of the meteorological initial and boundary conditions were perturbed by adding Gaussian random noise with a zero mean and static background error covariances (Torn et al., 2006) to generated the 50 ensemble members by WRFDA. We have added these sentences in Line 274-278, Page 10.

2) L107-108: Are emission scaling factors $\boldsymbol{\lambda}$ spatially varying?

Yes, the emission scaling factors $\boldsymbol{\lambda}$ here are spatially varying. In our system, we use the ensemble forecast chemical fields $\mathbf{C}_{i,t}^{\mathrm{f}}$ and the previous DA cycles' analysis scaling factors $\boldsymbol{\lambda}_{i,t-3}^{\mathrm{a}}$, $\boldsymbol{\lambda}_{i,t-2}^{\mathrm{a}}$, $\boldsymbol{\lambda}_{i,t-1}^{\mathrm{a}}$ to evaluate the emission scaling factors $\boldsymbol{\lambda}_{i,t}^{\mathrm{f}}$. Since $\mathbf{C}_{i,t}^{\mathrm{f}}$ were spatially varying, the ensemble concentration ratios $\boldsymbol{\kappa}_{i,t} = \mathbf{C}_{i,t}^{\mathrm{f}} / \overline{\mathbf{C}_t^{\mathrm{f}}}$ were spatially varying too. Thus, $\boldsymbol{\lambda}_{i,t}^{\mathrm{f}} = \frac{1}{4}\left(\boldsymbol{\lambda}_{i,t-3}^{\mathrm{a}} + \boldsymbol{\lambda}_{i,t-2}^{\mathrm{a}} + \boldsymbol{\lambda}_{i,t-1}^{\mathrm{a}} + \boldsymbol{\lambda}_{i,t}^{\mathrm{p}}\right) = \frac{1}{4}\left(\boldsymbol{\lambda}_{i,t-3}^{\mathrm{a}} + \boldsymbol{\lambda}_{i,t-2}^{\mathrm{a}} + \boldsymbol{\lambda}_{i,t-1}^{\mathrm{a}} + (\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}}\right) = \frac{1}{4}\left(\boldsymbol{\lambda}_{i,t-3}^{\mathrm{a}} + \boldsymbol{\lambda}_{i,t-2}^{\mathrm{a}} + \boldsymbol{\lambda}_{i,t-1}^{\mathrm{a}} + \beta\left(\boldsymbol{\kappa}_{i,t} - \overline{\boldsymbol{\kappa}_t}\right) + \overline{\boldsymbol{\kappa}_t}\right)$ were spatially varying.

We have added these sentences in Line 156-158, Page6.

3) L154-156: Why the inflation factors for the chemical species β are different among the variables? Could you please provide the strategy you took to find these values?

Peng et al. (2015) first used the forecast model of scaling operator $\mathbf{M}_{\mathrm{SF}}$ to prepare the ensemble emission scaling factors $\boldsymbol{\lambda}^{\mathrm{f}}$ in order to optimize all CO_2 fluxes as a whole at grid scale. In Peng et al. (2015), the ensemble spread of $\boldsymbol{\kappa}_{i,t} = \mathbf{C}_{i,t}^{\mathrm{f}} / \overline{\mathbf{C}_t^{\mathrm{f}}}$ was very small (ranging from 0 to 0.08 in most area at model-level 1), though the values of the ensemble spread of $\mathbf{C}_{i,t}^{\mathrm{f}}$ after inflation could reach 1 to 14 ppmv in most area at model-level 1. Therefore, covariance inflation was used to keep it at a certain level. After covariance inflation, the ensemble spread of $\lambda_{\mathrm{i,t}}^{\mathrm{a}}$ ranged from 0.1 to 0.8 in most model area for $\beta = 70$. Besides, several sensitive experiments were performed to investigate β (10, 50, 60, 70, 75, 80, 100). The ensemble spread of $\lambda_{\mathrm{i,t}}^{\mathrm{a}}$ ranged from 0.05 to 1.25 for β =60, 70, 75, 80. And the CO_2 DA system worked comparatively well for β =60, 70, 75, 80. The assimilated CO2 fluxes deviated markedly from the "true" CO2 fluxes when the ensemble spread of $\lambda_{\mathrm{i,t}}^{\mathrm{a}}$ were too small for β =10, 50 or when the ensemble spread of $\lambda_{\mathrm{i,t}}^{\mathrm{a}}$ were too large for β =100. Though CO_2 fluxes inversion was another topic, we mentioned it here because this experience was very helpful for us to develop

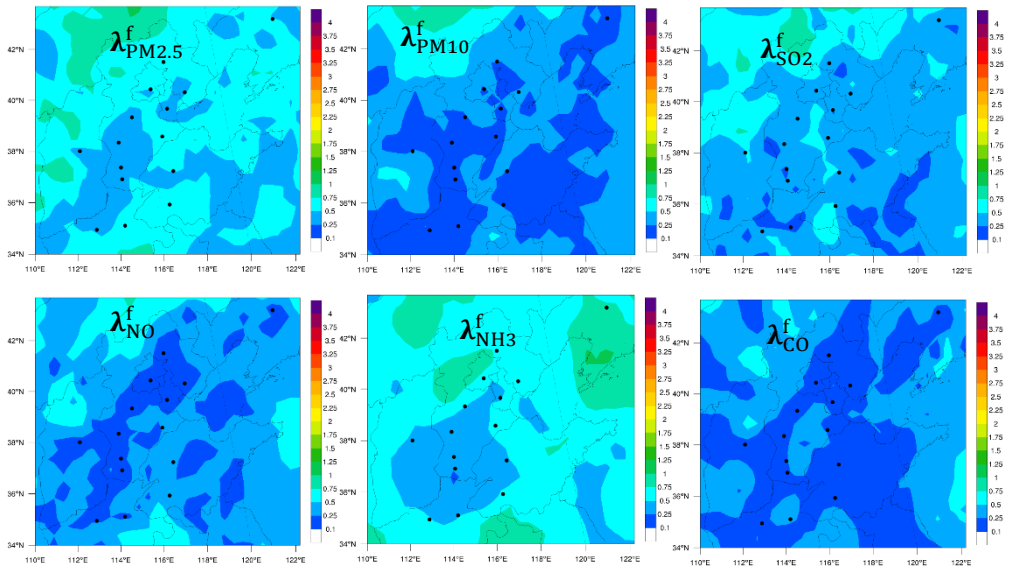
the joint DA system for aerosol.

In Peng et al. (2017), four emission scaling factors, $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{PM2.5}}$, $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{SO2}}$, $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{NO}}$ and $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{NH3}}$, are optimized in Peng et al. (2017) when the pure surface $PM_{2.5}$ observations are assimilated. We use the same inflation factor β to keep the ensemble spreads of $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{PM2.5}}$, $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{SO2}}$, $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{NO}}$ and $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{NH3}}$ at a certain level. Several sensitive experiments were performed to investigate β (1.2, 1.5, 1.8, 2, 2.5). It is seemed that reasonable results can be obtained when the ensemble spread of the emission scaling factors $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{PM2.5}}$ ranged from 0.1 to 1. Finally, $\beta = 1.5$ was chosen in Peng et al. (2017). The area-averaged ensemble spreads of $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{PM2.5}}$, $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{SO2}}$, $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{NO}}$ and $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{NH3}}$ were stably distributed around 0.5, 1.0, 1.5 and 0.8 respectively over the three sub-regions: Beijing–Tianjin–Hebei region, Yangtze River delta and Pearl River delta. It is apparent that the ensemble spread of $\lambda^{\mathrm{f}}_{\mathrm{SO2}}$ and $\lambda^{\mathrm{f}}_{\mathrm{NO}}$ is a little large due to the same β.

Therefore, it is better to choose different inflation factors for different emission scaling factors. We have performed several sensitive experiments to determine the value of β over a 2-day period before the experiments written in the manuscript. The criterion we choose β is to keep the ensemble spread of the scaling factors ranging from 0.1 to 1 in most model area. Finally, β is chosen as 1.3, 1.4, 1.3, 1.2, 1.2, and 1.4 for $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{PM2.5}}$, $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{PM10}}$, $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{SO2}}$, $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{NO}}$, $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{NH3}}$ and $\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{CO}}$ (See ReFig. 1)

Perhaps there are very few negative values for $(\boldsymbol{\kappa}_{\mathrm{i,t}})_{\mathrm{inf}}$ after inflation. A quality control procedure is performed for $(\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}}$ before further appliance. All these negative data were set as 0 in this work. Then $(\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}}$ were re-centered to ensure the ensemble mean values of $(\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}}$ were all 1. Then, another quality control procedure is performed for $\boldsymbol{\lambda}^{\mathrm{a}}_{i,t}$ to keep them positive. Thus, all $\boldsymbol{\lambda}^{\mathrm{f}}_{i,t}$ and $\boldsymbol{\lambda}^{\mathrm{a}}_{i,t}$ could be positive.

We have added these sentences in Line 158-163, 166-169, Page 6.



ReFig. 1. Spatial distribution of the ensemble spread for $\boldsymbol{\lambda}^{f}_{PM2.5}$, $\boldsymbol{\lambda}^{f}_{PM10}$, $\boldsymbol{\lambda}^{f}_{SO2}$, $\boldsymbol{\lambda}^{f}_{NO}$, $\boldsymbol{\lambda}^{f}_{NH3}$ and $\boldsymbol{\lambda}^{f}_{CO}$ at the lowest model level at 0000 UTC 6 October 2014 in the NCP region.

4) **L257-259: How did you perturb the initial conditions, lateral boundary conditions and emissions? In other words, please provide how you estimated the background uncertainty and spatial correlations (i.e. background covariance structures) for the chemical state variables in adding perturbations?**

Before the first DA cycle, a 50-member ensemble of four-day spin-up forecasts was performed, with perturbed meteorological initial conditions (ICs), lateral boundary conditions (LBCs) and emissions, from 0000 UTC 1 October to 2300 UTC 4 October 2014. The perturbed meteorological ICs and LBCs are created by adding Gaussian random noise (Torn et al., 2006) to the temperature, water vapor, velocity, geopotential height and dry surface pressure fields of the products of the National Centers for Environmental Prediction Global Forecast System (GFS) by WRFDA. The perturbed emissions were generated also by adding Gaussian random noise with a standard deviation of 10 percent of the corresponding anthropogenic emissions. The aerosol ICs were zero and the aerosol LBCs were idealized profiles embedded within the WRF/Chem model. They are the same as in Peng et al. (2017). It is noted that the perturbed emissions were only used in the initial part.

In the DA part, the ICs were the analysis of the previous DA cycle, the meteorological LBCs were the perturbed LBCs. The anthropogenic emissions, $\mathbf{E}^{f}_{PM2.5}$, $\mathbf{E}^{f}_{PM10}$, $\mathbf{E}^{f}_{SO2}$, $\mathbf{E}^{f}_{NO}$, $\mathbf{E}^{f}_{NH3}$, $\mathbf{E}^{f}_{CO}$, sulfate $\mathbf{E}^{f}_{SO4}$ and nitrate $\mathbf{E}^{f}_{NO3}$ are calculated by using the forecast emission scaling factors. Other species, such as the organic compounds $\mathbf{E}_{org}$ and elemental compounds $\mathbf{E}_{BC}$, are perturbed by adding Gaussian random noise. Since the emissions are calculated by EQ. (1), their background uncertainties and the spatial correlations are completely dependent on those of the corresponding emission factors. The forecast scaling factors are calculated by EQ. (2) ~ (5). And no other perturbations are added to the scaling factors; no other correlations are assumed for the scaling factors.

The experimental design is the same as in Peng et al (2017). We have rewritten briefly in Section 4 to avoid the repetition (Line 272-293, Page 10-11).

5) L275-279 and Figure 2: This is very promising. I would imagine that the impacts of other sources of uncertainties in air-quality forecast that were not directly considered in this study (such as chemical schemes and parameterizations in forecast model, and meteorology) were indirectly considered through the well-calibrated inflations of state variables. Could you please make a comment about the impacts of these other uncertainty sources in discussion section? I believe it would be helpful for the future readers of this manuscript.

It is true that the impacts of other sources of uncertainties in air-quality forecast (such as chemical schemes and parameterizations in forecast model, and meteorology) were not directly considered through the well-calibrated inflations of state variables. EnKF assimilation is influenced greatly by model errors and observation errors. But it is very difficult to accurately evaluate the uncertainties of models, though the covariance inflation technique was simply applied for all state variables to roughly compensate for model errors. Therefore, we can only obtain suboptimal results through EnKF assimilation.

We have added the above paragraph in Lines 476-482, Page 17.

6) Figure 4: It is not clear to me what "The shaded backgrounds indicate the distribution of the observations, where the top edge represented the 90th percentile and the bottom edge the 10th percentile" means. Does this distribution represent the observation values of individual sites in the Beijing–Tianjin–Hebei (BTH) region? Are other (red, black, pink, blue and light green) thick lines average of all sites in BTH region? The purpose to show these two values together is unclear to me, since the grey shaded line and other thick lines do not seem to be comparable each other. I would recommend to add more explanations about this figure, or to remove the grey shaded lines.

Yes. the grey shaded line represent the distribution of the observation values of individual sites in the Beijing–Tianjin–Hebei (BTH) region. Other (red, black, pink, blue and light green) thick lines represent the average values of all sites in BTH region. No more information could be obtained from the grey shaded line since the average values of observations (red line) were shown. Thus we remove the grey shaded lined in Figure 4.

List of specific comments:

1) L174: Please change "chose" to "chosen".

We have revised the word in Line 188, Page 6.

2) L296: I think "was able to" better fits with this context than "could".

We have changed the word in Line 328, Page 12.

3) Figure 4: The acronyms of "an" and "ct" is not described (although they can be guessed from the figure caption). Could you please add the explanation of those acronyms in the figure captions, such as "the analysis (referred to as "an", pink line)"?

We have changed theese in Line 848-855, Page 32.

4) Figure 11: Please add the explanation of grey shaded lines in the top panels.

We remove the grey shaded lined in Figure 11, similar to Figure 4.

Response to Reviewer #2's comments:

We thank Referee # 2 for his thoughtful comments and suggestions that have helped to improve this manuscript. Our responses to comments (in bold style) and the corresponding changes to the manuscript are detailed below. In particular, we have added a simulation using the optimized emissions from 5 to 16 October 2014 according to his suggestions.

There is not much to criticize about the manuscripts as it relies on the assimilation methodology previously described by Peng et al. (2017). (1) Since the assimilation experiment was conducted over a ten-day period it is uncertain if the conclusions about different performance of forecasts for various species would hold in a general. The most interesting are results on emission factors. (2) Did you encounter negative lambdas and if so what did you do about them? (3) An ultimate test of the optimized emissions would compare a simulation using the optimized emissions with a control. (4) Would an ENFK run with concentrations as state vectors using optimized emissions be identical to the EnKF run with concentrations and emission factors as the state vectors? (5) Link http://113.108.142.147:20035/emcpublish (p. 3) would be a valuable data source on pollution over China for many users but the access requires installation of Microsoft Silverlight a software for watching videos. That seems odd and is not be allowed on government computers. Could that be ameliorated?

(1) Since the assimilation experiment was conducted over a ten-day period it is uncertain if the conclusions about different performance of forecasts for various species would hold in a general.

It is true that only a case was investigated in this work and it is uncertain if the conclusions about different performance of forecasts for various species would hold in a general. More case studies are needed to obtain general results in future works.

We have added the above paragraph in Lines 548-551, Page 19.

(2) Did you encounter negative lambdas and if so what did you do about them?

There are very few negative values for $(\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}}$ after inflation (in Equation 3). A quality control procedure is performed for $(\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}}$ before further appliance. All these negative data were set as 0 in this work. Then $(\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}}$ were re-centered to ensure the

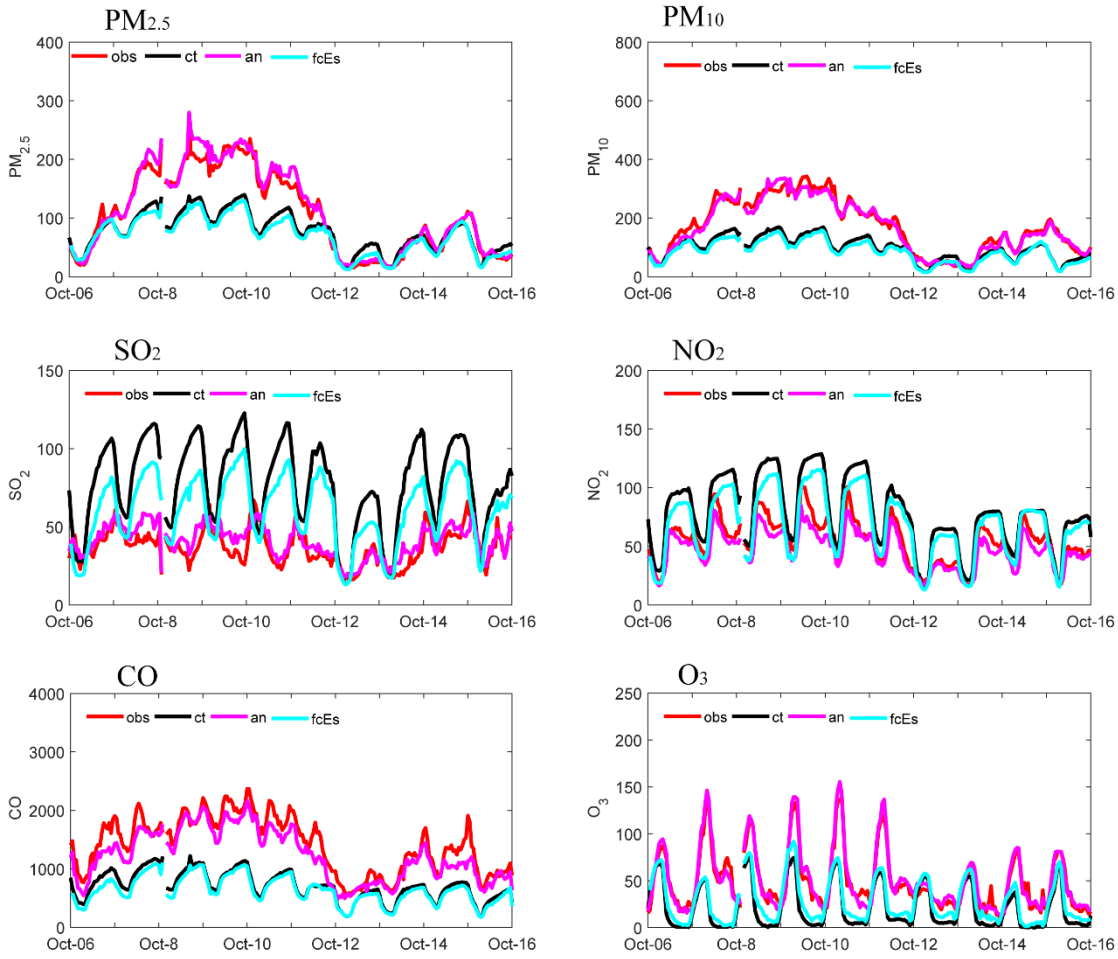
ensemble mean values of $(\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}}$ were all 1. Besides, another quality control procedure is performed for $\boldsymbol{\lambda}_{i,t}^{\mathrm{a}}$ to keep them positive. Thus, all $\boldsymbol{\lambda}_{i,t}^{\mathrm{f}}$ and $\boldsymbol{\lambda}_{i,t}^{\mathrm{a}}$ could be positive.

We have added the above paragraph in Lines 158-163, Page 6.

(3) An ultimate test of the optimized emissions would compare a simulation using the optimized emissions with a control.

We have performed a simulation (fcEs) using the optimized emissions from 5 to 16 October 2014 to investigate the impact of optimized emissions on chemical simulations. Same as the control run, the ICs were the ensemble mean of the spin-up forecasts at 00:00 UTC on 5 October 2014. Thus the difference between the fcEs and the control run is the anthropogenic emissions. The results showed that the fcEs performed very similar to the control run in the whole in the BTH region (ReFig. 1). For $PM_{2.5}$, PM_{10} and CO, the values of the fcEs were a little smaller than those of the control run, which were consistent with the difference of the anthropogenic emissions. For SO_2 and NO_2, fcEs performed much better than the control run in most time though significant systematic overestimation still existed during the nighttime. For O_3, miner improvements were also gained due to the better simulation in fcEs for NO_2.

We have added the above paragraph in Line 443-453, Page 15. For ReFig.1, the cyan line (refer to as "fcEs") was added in Figure 4 to save space.



ReFig. 1. Time series of the hourly pollutant concentrations in the Beijing–Tianjin–Hebei (BTH) region obtained from observations (referred to as "obs", red line), the control run (referred to as "ct", black line), the analysis (referred to as "an", pink line), the simulation only using the optimized emissions (referred to as "fcEs", cyan line). The observations were obtained from the 47 independent sites in the BTH region. The modelled time series were interpolated to the 47 independent sites using the spatial bilinear interpolator method. Units: μg m^{-3}.

(4) Would an EnFK run with concentrations as state vectors using optimized emissions be identical to the EnKF run with concentrations and emission factors as the state vectors?

The optimized emissions are only the results of a mathematical optimum by utilizing observations. They are influenced greatly by model errors and observation errors. If the optimized emissions used in the EnFK experiment run with pure

concentrations as state vectors are identical to the emissions assimilated in the joint EnFK experiment run with concentrations and emission factors (representing emissions) as state vectors, identical results may be obtained.

We have added the above paragraph in Line 116-121, Page 4-5.

(5) Link http://113.108.142.147:20035/emcpublish (p. 3) would be a valuable data source on pollution over China for many users but the access requires installation of Microsoft Silverlight a software for watching videos. That seems odd and is not be allowed on government computers. Could that be ameliorated?

Yes, we agree with the reviewer that the requirement of installation of Microsoft Silverlight software to view the data is odd. There is another website for the data: http://www.resdc.cn/data.aspx?dataid=186. The data can be downloaded by request. If you are interested in the data, please contact the data manager of the website.

The impact of multi-species surface chemical observations assimilation on the air quality forecasts in China

Zhen Peng[1*], Lili Lei[1], Zhiquan Liu[2*], Jianning Sun[1,3], Aijun Ding[1,3], Junmei Ban[2], Dan Chen[4], Xingxia Kou[4], Kekuan Chu[1]

1 School of Atmospheric Sciences, Nanjing University, Nanjing, China

2 National Center for Atmospheric Research, Boulder, Colorado, USA

3 Institute for Climate and Global Change Research, Nanjing University, Nanjing, China

4 Institute of Urban Meteorology, CMA, Beijing, China

Abstract. An Ensemble Kalman Filter data assimilation (DA) system has been developed to improve air quality forecasts using surface measurements of PM_{10}, $PM_{2.5}$, SO_2, NO_2, O_3 and CO together with an online regional chemical transport model, WRF-Chem (Weather Research and Forecasting with Chemistry). This DA system was applied to simultaneously adjust the chemical initial conditions (ICs) and emission inputs of the species affecting PM_{10}, $PM_{2.5}$, SO_2, NO_2, O_3 and CO concentrations during an extreme haze episode that occurred in early October 2014 over the East Asia. Numerical experimental results indicate that ICs play key roles in $PM_{2.5}$, PM_{10} and CO forecasts during the severe haze episode over the North China Plain. The 72-h verification forecasts with the optimized ICs and emissions performed very similarly to the verification forecasts with only optimized ICs and the prescribed emissions. For the first-day forecast, near perfect verification forecasts results were achieved. However, with longer range forecasts, the DA impacts decayed quickly. For the SO_2 verification forecasts, it was efficient to improve the SO_2 forecast via the joint adjustment of SO_2 ICs and emissions. Large improvements were achieved for SO_2 forecasts with both the optimized ICs and emissions for the whole 72-h forecast range. Similar improvements were achieved for SO_2 forecasts with optimized ICs only for just the first 3 h, and then the impact of the ICs decayed quickly. For the NO_2 verification forecasts, both forecasts

performed much worse than the control run without DA. Plus, the 72-h O_3 verification forecasts performed worse than the control run during the daytime, due to the worse performance of the NO_2 forecasts, even though they performed better at night. However, relatively favorable NO_2 and O_3 forecast results were achieved for the Yangtze River delta and Pearl River delta regions.

1 Introduction

Predicting and simulating air quality remains a challenge in heavily polluted regions (Wang et al., 2014; Ding et al. 2016). Chemical data assimilation (DA), which combines observations and model simulations, is recognized as one effective method to improve air quality forecasts. It has been widely used to assimilate aerosol measurements from both ground-based and space-borne platforms, including surface PM_{10} observations (Jiang et al., 2013; Pagowski et al., 2014), surface $PM_{2.5}$ observations (Li et al., 2013; Zhang, 2016), Lidar observations (Yumimoto et al., 2007, 2008), aerosol optical depth products from AERONET (the AErosol RObotic NETwork) (Schutgens et al., 2010a-b, 2012), and from various satellites (Sekiyama et al., 2010; Liu et al., 2011; Dai et al., 2014). These studies indicate that assimilating observations can substantially improve the spatiotemporal variations of aerosol in the simulation and forecasts.

Aerosols are not only primarily emitted, but also with a larger portion secondary formed through reactions with several gaseous-phases precursors and oxidants in the atmosphere (Huang et al., 2014; Nie et al., 2014; Xie et al., 2015). So, observations of trace gases are also useful in assimilating data for aerosol simulations and forecasts. Efforts to assimilate atmospheric-composition observations, like O_3, SO_2, NO, NO_2, CO, and NH_3, have also been made. For example, Elbern et al. (1997, 1999, 2000, 2001, 2007) developed a 4D-VAR (four-dimensional variational) system to assimilate surface measurements of O_3, SO_2, NO and NO_2 to improve air quality forecasts with the joint adjustment of initial conditions (ICs) and emission rates. Later, van Loon et al. (2000) assimilated O_3 in the transport chemistry model LOTOS, based on an Ensemble Kalman Filter (EnKF). Heemink and Segers (2002) attempted to reconstruct NO_x and volatile

organic compound (VOC) emissions for O_3 forecasting by assimilating O_3. Carmichael et al. (2003, 2008a, 2008b) developed 4D-VAR and EnKF systems to assimilate O_3 and NO_2 to improve ICs and emission sources for O_3 forecasting. Hakami et al. (2005) constrained black carbon (BC) emissions during the Asian Pacific Regional Aerosol Characterization Experiment. Henze et al. (2007, 2009) estimated SO_x, NO_x and NH_3 emissions based on a 4D-VAR method by assimilating surface sulfate and nitrate aerosol observations. Other studies have estimated the NO_x (van der et al., 2006, 2017; Mijling et al., 2009, 2012, 2013; Ding. et al., 2015) and SO_2 emissions (van der et al., 2017) based on an extended Kalman filter by assimilating SO_2 and NO_2 retrievals from SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY) and OMI (Ozone Monitoring Instrument). Barbu et al. (2009) applied an EnKF to optimize the emissions and conversion rates using surface measurements of SO_2 and sulfate. McLinden (2016) constrained SO_2 emissions using space-based observations.

In recent years, severe haze pollution episodes have begun to occur more frequently in China, especially in the megacity clusters of eastern China (e.g., Parrish and Zhu, 2009; Sun et al., 2015; Zhang et al., 2015a). Thus, regional haze, especially when accompanied by extremely high $PM_{2.5}$ concentrations, has drawn significant research interest. However, there are large uncertainties involved in the numerical prediction of atmospheric aerosols. During severe haze pollution episodes, air quality models often underestimate the extreme peak mass concentration of particulate matter (Wang et al., 2014). Previous studies have revealed that the assimilation of atmospheric-composition observations can improve air quality forecasts by constraining the uncertainties of both the chemical ICs and emissions (Tang et al., 2010, 2011, 2013, 2016; Miyazaki et al., 2012, 2013, 2014). Peng et al. (2017) demonstrated that significant improvements in forecasting $PM_{2.5}$ can be achieved via the joint adjustment of ICs and source emissions using an EnKF to assimilate surface $PM_{2.5}$ observations.

In 2013, China launched an atmospheric environmental monitoring system that provides real-time and online atmospheric chemical observations, including PM_{10}, $PM_{2.5}$, SO_2, NO_2, O_3, and CO (http://113.108.142.147:20035/emcpublish/). This

dataset provides an opportunity to improve air quality forecasts using DA. However, such fruitful observations are less used in air quality forecast despite of large discrepancy existed between the forecast and observations. But it is now possible to estimate the impact on forecast improvement of simultaneously assimilating various surface observations. Thus, we developed an EnKF system that can simultaneously assimilate surface measurements of PM_{10}, $PM_{2.5}$, SO_2, NO_2, O_3 and CO to correct WRF-Chem (Weather Research and Forecasting model with Chemistry) forecasts using the Goddard Chemistry Aerosol Radiation and Transport (GOCART) aerosol scheme. As an extension to Peng et al. (2017), the impact of simultaneously assimilating various surface aerosol and chemical observations was investigated.

Sections 2 and 3 briefly describe the DA system and observations used in this study, respectively. The experimental design is introduced in Section 4. Finally, the assimilation results are presented in Section 5, before a brief summary in Section 6.

2 DA system

The DA system in this study was the same as the one used in Peng et al. (2017). It can simultaneously analyze the chemical ICs and emissions with the assimilation of surface $PM_{2.5}$ observations. A brief summary of the DA system is introduced here.

In every DA cycle, the ensemble emission scaling factors $\boldsymbol{\lambda}^{\mathrm{f}}$ are first calculated by the forecast model of scaling factors $\mathbf{M}_{\mathrm{SF}}$ (see details of $\mathbf{M}_{\mathrm{SF}}$ in section 2.2). Then, the ensemble forecast emissions $\mathbf{E}^{\mathrm{f}}$ are calculated using the following equation:

$$\mathbf{E}_{i,t} = \boldsymbol{\lambda}_{i,t}\mathbf{E}_t^{\mathrm{p}}, (i = 1, \dots, N), \quad (1)$$

where $\mathbf{E}_t^{\mathrm{p}}$ is the prescribed anthropogenic emission. The ensemble members of chemical fields $\mathbf{C}^{\mathrm{f}}$ are forecasted using WRF-Chem, forced by the forecast emissions $\mathbf{E}^{\mathrm{f}}$ whose ICs are previously analyzed concentration fields. Now, the background of the joint vector, $\mathbf{x}^{\mathrm{f}} = \left[\mathbf{C}^{\mathrm{f}}, \boldsymbol{\lambda}^{\mathrm{f}}\right]^{\mathrm{T}}$, has been produced. Then, the analyzed state vector, $\mathbf{x}^{\mathrm{a}} = [\mathbf{C}^{\mathrm{a}}, \boldsymbol{\lambda}^{\mathrm{a}}]^{\mathrm{T}}$, is optimized using an ensemble square root filter (EnSRF). Finally, the assimilated emissions $\mathbf{E}^{\mathrm{a}}$ can be obtained using equation (1). It is noted that the optimized emissions are only the results of a mathematical optimum by utilizing

observations. If the optimized emissions used in the EnSRF experiment run with pure concentrations as state vectors are identical to the emissions obtained from the joint EnSRF experiment run with concentrations and emission factors (representing emissions) as state vectors, identical results may be obtained.

2.1 WRF-Chem model

The model used to simulate the transport of aerosols and chemical species was the WRF-Chem (Grell et al., 2005). As in Peng et al. (2017), we used version 3.6.1 and the physical and chemical parameterization options are listed in Table 1. The model computational domain covered almost the whole China and the horizontal resolution was 40.5 km. Figure 1b shows our area of interest, the North China Plain (NCP). The model included 57 vertical levels and the model top was 10 hPa.

The hourly prior anthropogenic emissions were based on the Multi-resolution Emission Inventory for China (MEIC) (Li et al., 2014) for October 2010, instead of the regional emission inventory in Asia (Zhang et al., 2009) for the year 2006 in Peng et al. (2017). The reason we chose the MEIC-2010 was that the total emissions are reasonable for cities over the NCP (Zheng et al., 2016). The original resolution of the MEIC-2010 is 0.25 ° × 0.25 °, but has been processed to match the model resolution (40.5 km) (Chen et al., 2016). No time variation was added to maintain objectivity in the prior anthropogenic emissions.

2.2 Forecast model of scaling factors

In this work, the primary sources to be optimized were the emissions of PM_{10}, $PM_{2.5}$, SO_2, NO, NH_3 and CO. The sources of NH_3 were analyzed because they also impact greatly on the aerosols distribution. Thus, the emission scaling factors $\boldsymbol{\lambda}_{i,t}^{\mathrm{f}} = (\boldsymbol{\lambda}_{\mathrm{PM2.5}}^{\mathrm{f}}, \boldsymbol{\lambda}_{\mathrm{PM10}}^{\mathrm{f}}, \boldsymbol{\lambda}_{\mathrm{SO2}}^{\mathrm{f}}, \boldsymbol{\lambda}_{\mathrm{NO}}^{\mathrm{f}}, \boldsymbol{\lambda}_{\mathrm{NH3}}^{\mathrm{f}}, \boldsymbol{\lambda}_{\mathrm{CO}}^{\mathrm{f}})$ were prepared by the forecast model of scaling operator $\mathbf{M}_{\mathrm{SF}}$ before WRF-Chem integration.

We used the same persistence forecast operator $\mathbf{M}_{\mathrm{SF}}$ to forecast $\boldsymbol{\lambda}_{i,t}^{\mathrm{f}}$ as in Peng et al. (2017). The forecast operator was developed by using the ensemble forecast

chemical fields. Thus,

$$\boldsymbol{\kappa}_{i,t} = \frac{\mathbf{C}^{\mathrm{f}}_{i,t}}{\overline{\mathbf{C}^{\mathrm{f}}_{t}}}, (i = 1, \dots, N), \quad (2)$$

$$(\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}} = \beta\left(\boldsymbol{\kappa}_{i,t} - \overline{\boldsymbol{\kappa}_{t}}\right) + \overline{\boldsymbol{\kappa}_{t}}, (i = 1, \dots, N), \quad (3)$$

$$\boldsymbol{\lambda}^{\mathrm{p}}_{i,t} = (\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}}, \quad (4)$$

$$\boldsymbol{\lambda}^{\mathrm{f}}_{i,t} = \frac{1}{4}\left(\boldsymbol{\lambda}^{\mathrm{a}}_{i,t-3} + \boldsymbol{\lambda}^{\mathrm{a}}_{i,t-2} + \boldsymbol{\lambda}^{\mathrm{a}}_{i,t-1} + \boldsymbol{\lambda}^{\mathrm{p}}_{i,t}\right), (i = 1, \dots, N), \quad (5)$$

where $\mathbf{C}^{\mathrm{f}}_{i,t}$ is the ith ensemble member of the chemical fields at time t, and $\overline{\mathbf{C}^{\mathrm{f}}_{t}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{C}^{\mathrm{f}}_{i,t}$ is the ensemble mean; $\boldsymbol{\kappa}_{i,t}$ is the ensemble concentration ratios and $\overline{\boldsymbol{\kappa}_{t}}$ is the ensemble mean of $\boldsymbol{\kappa}_{i,t}$ with values of 1; β is the inflation factor to keep the ensemble spreads of $\boldsymbol{\kappa}_{i,t}$ at a certain level; $\boldsymbol{\lambda}^{\mathrm{a}}_{i,t-1}$, $\boldsymbol{\lambda}^{\mathrm{a}}_{i,t-2}$ and $\boldsymbol{\lambda}^{\mathrm{a}}_{i,t-3}$ are the previous assimilated emission scaling factors. It is noted that $\boldsymbol{\lambda}^{\mathrm{f}}_{i,t}$ are spatially varying because they are calculated by using the spatially varying variables, the forecast chemical fields $\mathbf{C}^{\mathrm{f}}_{i,t}$. Besides, There are very few negative values for $(\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}}$ after inflation. A quality control procedure is performed for $(\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}}$ before further appliance. All these negative data were set as 0 in this work. Then $(\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}}$ were re-centered to ensure the ensemble mean values of $(\boldsymbol{\kappa}_{i,t})_{\mathrm{inf}}$ were all 1. Besides, another quality control procedure is performed for $\boldsymbol{\lambda}^{\mathrm{a}}_{i,t}$ to keep them positive. Thus, all $\boldsymbol{\lambda}^{\mathrm{f}}_{i,t}$ and $\boldsymbol{\lambda}^{\mathrm{a}}_{i,t}$ could be positive.

In this study, the ensemble forecast chemical fields of PM_{25}, PM_{10}, SO_2, NO, NH_3 and CO of the previous assimilation cycle are respectively used to calculate the emission scaling factors $(\boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{PM2.5}}, \boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{PM10}}, \boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{SO2}}, \boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{NO}}, \boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{NH3}}, \boldsymbol{\lambda}^{\mathrm{f}}_{\mathrm{CO}})$. Previous works (Peng et al., 2015, 2017) showed that reasonable results can be obtained when the ensemble spread of the emission scaling factors ranged from 0.1 to 1. In order to keep the ensemble spread of the scaling factors at this level in most model area, β is chosen as 1.3, 1.4, 1.3, 1.2, 1.2, and 1.4 for the ensemble concentration ratios of P_{25}, P_{10}, SO_2, NO, NH_3 and CO, respectively in Equation (3).

Then, the sources $\mathbf{E}^{\mathrm{f}}_{i,t} = (\mathbf{E}^{\mathrm{f}}_{\mathrm{PM2.5}}, \mathbf{E}^{\mathrm{f}}_{\mathrm{PM10}}, \mathbf{E}^{\mathrm{f}}_{\mathrm{SO2}}, \mathbf{E}^{\mathrm{f}}_{\mathrm{NO}}, \mathbf{E}^{\mathrm{f}}_{\mathrm{NH3}}, \mathbf{E}^{\mathrm{f}}_{\mathrm{CO}})$ are calculated

using equation (1).

From the perspective of $PM_{2.5}$ emissions, these include the unspeciated primary sources of $PM_{2.5}$ $\mathbf{E}_{PM2.5}$, sulfate $\mathbf{E}_{SO4}$, and nitrate $\mathbf{E}_{NO3}$. We updated $\mathbf{E}_{PM2.5}$, $\mathbf{E}_{SO4}$ and $\mathbf{E}_{NO3}$ (including the nuclei and accumulation modes) following Peng et al. (2017).

2.3 DA algorithm

The assimilation algorithm employed was the EnSRF proposed by Whitaker and Hamill (2002). The EnKF proposed by Evensen (1994) needs perturbations of observations in practice. Compared to the original EnKF, the EnSRF obviates the need to perturb the observations and avoids additional sampling errors introduced by perturbing observations.

We used the same EnSRF as in Schwartz et al. (2012, 2014). The ensemble member was chosen as 50. The localization radius was chosen as 607.5 km, so EnSRF analysis increments were forced to zero at 607.5 km away from an observation (Gaspari and Cohn, 1999). The posterior (after assimilation) multiplicative inflation factor was chosen as 1.2 for all the concentration analysis.

2.4 State variables

The DA system provides joint analysis of ICs and emissions following Peng et al. (2017). Among them, 16 WRF-Chem/GOCART aerosol variables are included as the state variables. Besides, chemical species, such as SO_2, NO_2 and O_3 are also included because they are the most important gas-phase precursors or oxidants of the secondary inorganic aerosols. CO is also assimilated because CO is an important tracer of combustion sources, as well as a precursor of O_3 beyond NO_2 (Parrish et al., 1991). The state variables of the emission scaling factors are $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_{PM2.5}, \boldsymbol{\lambda}_{PM10}, \boldsymbol{\lambda}_{SO2}, \boldsymbol{\lambda}_{NO}, \boldsymbol{\lambda}_{NH3}, \boldsymbol{\lambda}_{CO})$.

Similar to weak-coupling DA, the DA system simultaneously updates both the ICs and the emissions, but with no cross-variable update, in order to avoid the effects of spurious multivariate correlations in the background error covariance that may develop due to the limited ensemble size and errors in both the model and observations

(Miyazaki et al. 2012).

For the $PM_{2.5}$ observations, the observation operator is expressed as (Schwartz et al., 2012)

$$y^{\mathrm{f}}_{\mathrm{pm25}} = \boldsymbol{\rho}_{\mathrm{d}}[\mathbf{P_{25}} + 1.375\mathbf{S} + 1.8(\mathbf{OC_1} + \mathbf{OC_2}) + \mathbf{BC_1} + \mathbf{BC_2} + \mathbf{D_1} + 0.286\mathbf{D_2} + \mathbf{S_1} + 0.942\mathbf{S_2}], \quad (6)$$

where $\boldsymbol{\rho}_{\mathrm{d}}$ is the dry air density; P_{25} is the fine unspectiated aerosol contributions; S represents sulfate; OC_1 and OC_2 are hydrophobic and hydrophilic organic carbon respectively; BC_1 and BC_2 are hydrophobic and hydrophilic black carbon respectively; D_1 and D_2 are dusts with effective radii of 0.5 and 1.4 μm espectively; S_1 and S_2 are sea salts with effective radii of 0.3 and 1.0 μm espectively. In fact, $PM_{2.5}$ observations were only used to analyze P_{25}, S, OC_1, OC_2 BC_1, BC_2, D_1, D_2, S_1, S_2 and $\boldsymbol{\lambda}_{\mathrm{PM2.5}}$. Since we had no NH_3 observations, $PM_{2.5}$ observations were also used to analyze $\boldsymbol{\lambda}_{\mathrm{NH3}}$ (see Table 2). For other control variables, $PM_{2.5}$ observations were not allowed to alter them.

For the PM_{10} observations, the PM_{10} observation operator is expressed as (Jiang et al., 2013)

$$y^{\mathrm{f}}_{\mathrm{pm10}} = \boldsymbol{\rho}_{\mathrm{d}}[\mathbf{P_{10}} + \mathbf{P_{25}} + 1.375\mathbf{S} + 1.8(\mathbf{OC_1} + \mathbf{OC_2}) + \mathbf{BC_1} + \mathbf{BC_2} + \mathbf{D_1} + 0.286\mathbf{D_2} + \mathbf{D_3} + 0.87\mathbf{D_4} + \mathbf{S_1} + 0.942\mathbf{S_2} + \mathbf{S_3}]. \quad (7)$$

Thus,

$$y^{\mathrm{f}}_{\mathrm{pm10-2.5}} = \boldsymbol{\rho}_{\mathrm{d}}[\mathbf{P_{10}} + \mathbf{D_3} + 0.87\mathbf{D_4} + \mathbf{S_3}], \quad (8)$$

meaning that, in the assimilation experiments, we did not use the PM_{10} observations directly. In equation (13) and (14), P_{10} denotes the coarse-mode unspectiated aerosol contributions; D_3 and D_4 are dusts with effective radii of 2.4 and 4.5 μm respectively; S_3 is sea salt with effective radii of 3.25 μm. We used the $PM_{10\text{-}2.5}$ observations (the differences between the PM_{10} observations and the $PM_{2.5}$ observations, $y^{\mathrm{o}}_{\mathrm{pm10-2.5}} = y^{\mathrm{o}}_{\mathrm{pm10}} - y^{\mathrm{o}}_{\mathrm{pm10}}$) to analyze P_{10}, D_3, D_4, S_3 and $\boldsymbol{\lambda}_{\mathrm{PM10}}$. In addition, $PM_{10\text{-}2.5}$ observations were used to analyze D_5 and S_4, since they are coarse-mode mineral dust and sea salt aerosols. $PM_{10\text{-}2.5}$ observations were not allowed to impact other control variables.

Moreover, as shown in Table 2, SO_2 observations were used to analyze the SO_2 concentration and $\boldsymbol{\lambda}_{SO2}$. NO_2 observations were used to estimate the NO, NO_2 concentration and $\boldsymbol{\lambda}_{NO}$. CO observations were used to analyze the CO concentration and $\boldsymbol{\lambda}_{CO}$. And finally, O_3 observations were only used to analyze the O_3 concentration.

3. Observations and errors

The surface chemical observations used in this study were obtained from the Ministry of Ecology and Environment of China. Altogether, there were 876 observational sites over the model domain (Figure 1). At most sites, one measurement was selected randomly for the assimilation experiment on a 0.1 °× 0.1 °grid. Altogether, 355 stations were kept for the model domain, where 133 assimilation stations were located on the NCP and 40 stations were located in the Beijing–Tianjin–Hebei (BTH) region. Other stations were used for verification purposes: 167 independent stations were located on the NCP and 47 in the BTH region.

The observation error covariance matrix $\mathbf{R}$ included measurement errors and representation errors. We assumed that $\mathbf{R}$ is a diagonal matrix (without observation correlation).

Following Elbern et al. (2007), the measurement error ε_0 is defined as

$$\varepsilon_0 = a + b * \Pi_0, \quad (9)$$

where Π_0 represents the measurements for $PM_{2.5}$, $PM_{10\text{-}2.5}$, SO_2, NO_2, CO or O_3 (units: μg m^{-3}). A value of $a = 1.5$ and $b = 0.0075$ was chosen for $PM_{2.5}$, $PM_{10\text{-}2.5}$, SO_2, and NO_2. For CO, $a = 10$ and $b = 0.0075$.

The representativeness error is defined as

$$\varepsilon_r = r\varepsilon_0\sqrt{\Delta x/\mathrm{L}}, \quad (10)$$

where $r = 0.5$, $\Delta x = 40.5$ km (the model resolution), and $L = 3$ km due to the lack of the information of the station type (Elbern et al., 2007).

Finally, the total error (ε_t) is defined as

$$\varepsilon_t = \sqrt{{\varepsilon_0}^2 + {\varepsilon_r}^2}, \quad (11)$$

In order to ensure data reliability, the observations were subjected to quality

control before DA. Data values larger than a certain threshold were classified as unrealistic and were not assimilated. The threshold values were chosen as 700, 800, 300, 300, 400 and 4000 μg m^{-3} for $PM_{2.5}$, $PM_{10-2.5}$, SO_2, NO_2, O_3 and CO, respectively. In addition, observations leading to innovations exceeding a certain value were also omitted. These threshold values were chosen as 70 μg m^{-3} for $PM_{2.5}$, $PM_{10-2.5}$, SO_2, NO_2 and O_3. Also, 1500 μg m^{-3} was chosen for CO.

4. Experimental design

The DA experiment followed that of Peng et al. (2017), in which the assimilation of pure surface $PM_{2.5}$ measurements with the EnKF was performed to correct finer aerosol variables and associated emissions. The experiment focused on an extreme haze event that occurred in October 2014 over North China.

The 50-member ensemble spin-up forecasts were first performed from 1 to 4 October 2014 using the perturbed meteorological ICs, lateral boundary conditions (LBCs) and emissions. The perturbed meteorological ICs and LBCs are created by adding Gaussian random noise (Torn et al., 2006) to the temperature, water vapor, velocity, geopotential height and dry surface pressure fields of the products of the National Centers for Environmental Prediction Global Forecast System (GFS) by WRFDA. The perturbed emissions were generated also by adding Gaussian random noise with a standard deviation of 10 percent of the corresponding anthropogenic emissions. The aerosol ICs were zero and the aerosol LBCs were idealized profiles embedded within the WRF-Chem model. And both them are not perturbed (Peng et al., 2017).

Then, the observed PM_{10}, $PM_{2.5}$, SO_2, NO_2, O_3 and CO data starting from 5 to 16 October were assimilated hourly to adjust the ICs and the corresponding emissions. the ICs were the analysis of the previous DA cycle. The meteorological LBCs were perturbed. The anthropogenic emissions, $\mathbf{E}_{PM2.5}$, $\mathbf{E}_{PM10}$, $\mathbf{E}_{SO2}$, $\mathbf{E}_{NO}$, $\mathbf{E}_{NH3}$, $\mathbf{E}_{CO}$, sulfate $\mathbf{E}_{SO4}$ and nitrate $\mathbf{E}_{NO3}$ are calculated by using the forecast emission scaling factors. Other species, such as the organic compounds $\mathbf{E}_{org}$ and elemental compounds

290 $\mathbf{E}_{BC}$, are perturbed by adding Gaussian random noise. Since the emissions are calculated
291 by EQ. (1), their background uncertainties and the spatial correlations are completely
292 dependent on those of the corresponding emission factors. The forecast scaling factors
293 are calculated by EQ. (2) ~ (5). And no other perturbations are added to the scaling
294 factors; no other correlations are assumed for the scaling factors.

295 After that, two sets of 72-h forecasts were performed, each at 00:00 UTC from 6
296 to 15 October 2014, with hourly forecasting outputs for the assimilation experiment.
297 These two sets of forecasting experiments were conducted using the ensemble mean of
298 the concentration analysis as the ICs. One set of the experiments was forced by the
299 optimized emissions (denoted as fcICsEs), and the other was forced by the prescribed
300 anthropogenic emissions (denoted as fcICs). The aim was to use the difference between
301 the fcICsEs and fcICs to indicate the impact of the optimized emissions.

302 Moreover, we also run a control experiment. The ICs were based on the ensemble
303 mean of the spin-up forecasts at 00:00 UTC on 5 October 2014. The emissions were
304 the prescribed emissions.

305

306 **5. Results**

307 5.1 Ensemble performance

308 We begin by assessing the ensemble performance for the DA system. Figure 2 shows
309 the time series of the prior total spreads and the prior root-mean-square errors (RMSEs)
310 for $PM_{2.5}$, PM_{10}, and the four trace gases calculated against all observations in the BTH
311 region. It shows that the magnitudes of the total spreads were close to the RMSEs,
312 indicating that the DA system was well calibrated (Houtekamer et al., 2005).

313 Figure 3 shows the area-averaged time series extracted from the ensemble spread
314 of the six emission scaling factors ($\boldsymbol{\lambda}^{f}_{PM2.5}$, $\boldsymbol{\lambda}^{f}_{PM10}$, $\boldsymbol{\lambda}^{f}_{SO2}$, $\boldsymbol{\lambda}^{f}_{NO}$, $\boldsymbol{\lambda}^{f}_{NH3}$ and $\boldsymbol{\lambda}^{f}_{CO}$) in the
315 BTH region. It shows that the ensemble spread of all the scaling factors were very stable
316 throughout the ~10-day experiment period, which indicates that $\mathbf{M}_{SF}$ can generate
317 stable artificial data to generate the ensemble emissions. The value of the emission
318 scaling factors ranged from 0.2 to 0.6, indicating that the uncertainty of the assimilated
319 emissions was about 20%–60%.

320

321 5.2 Forecast improvements

322 In order to evaluate the overall performance of the DA system, time series of the hourly
323 pollutant concentrations from the control run, the analysis, and the first-day forecast of
324 the two forecasting experiments were compared with the independent observations in
325 the BTH region (Figure 4). Besides, model evaluation statistics (Table 3) were
326 calculated against independent observations from 6 to 16 October 2014. In addition,
327 biases and RMSEs were presented as a function of forecast range for the control,
328 analysis, and forecast experiments (Figures 5–7).

329 The control run did not perform very well, although it was able to capture the
330 synoptic variability and reproduce the overall pollutant levels when there was a severe
331 haze event. The statistics show that there were larger systematic biases and RMSEs and
332 a smaller correlation coefficient (CORR) for the control (see Table 3). The biases were
333 −34.1, −77.7, −565.7 and −31 $\mu g \cdot m^{-3}$ for $PM_{2.5}$, PM_{10}, CO, and O_3, respectively, from
334 6 to 16 October—about 29.7%, 44.5%, 42.9% and 53.9% lower than the corresponding
335 observed concentrations. During the severe haze episode from 8 to 10 October in
336 particular, when observed $PM_{2.5}$ were larger than 200 $\mu g \cdot m^{-3}$, the biases reached −90.5,
337 −143.1, −911.8 and −39.1$\mu g \cdot m^{-3}$, respectively—about 44.4%, 51.9%, 49.2% and 55.7%
338 lower than the corresponding observed concentrations, suggesting a significant
339 systematic underestimation of the WRF-Chem simulation. Additionally, a significant
340 overestimation of 48.1 $\mu g \cdot m^{-3}$ was obtained for SO_2—about 145.8% higher than the
341 observed concentrations. As for the NO_2 simulation, WRF-Chem was able to
342 realistically describe the diurnal and synoptic evolution of NO_2 concentrations. The
343 model bias was 22.4 $\mu g \cdot m^{-3}$, which was about 39.7% higher than the observed NO_2.
344 These results were similar to the simulations of Chen et al. (2016). Most of the WRF-
345 Chem settings used here were the same as those used in Chen et al. (2016), except that
346 they used CBMZ (Carbon Bond Mechanism, version Z) and MOSAIC (Model for
347 Simulating Aerosol Interactions and Chemistry) as the gas-phase and aerosol chemical
348 mechanisms.

349 After the assimilation of surface observations, the time series of the hourly

pollutant concentrations from the analysis showed much better agreement with observations than those from the control. The magnitudes of the bias and the RMSEs decreased and the CORRs increased for all six species. The biases were 5.1, −5.6, 8.1, −8.3, −160.4 and 2.1 μg m^{-3} for $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO and O_3, respectively—about 4.4%, −3.2%, 24.5%, −14.7%, −12.17% and 3.7% of the corresponding observed concentrations, indicating that the analysis fields were very close to the observations. The RMSEs were 51.5, 63.4, 27.9, 31.7, 618.9 and 31.1 μg m^{-3}, respectively—about 44.1%, 52.9%, 58.1%, 20.2%, 35.7% and 38.78% lower than the RMSEs of the control run. The CORRs reached 0.891, 0.890, 0.540, 0.557, 0.705 and 0.753, respectively. These statistics indicate that the DA system was able to adjust the chemical ICs efficiently.

The $PM_{2.5}$, PM_{10} and CO concentrations from both sets of forecasting experiments benefitted substantially from the DA procedure, as expected. Smaller biases and RMSEs were obtained for almost the entire 72-h forecast range (see Figures 5–7), as compared with the control run. For the first-day forecast in particular, the model performed almost perfectly. It faultlessly captured the diurnal and synoptic variability of the pollutant (see figure 4), in a manner that was very close to that of the analysis. The overall biases were 6.5, −11.9 and 100.4 μg m^{-3} for $PM_{2.5}$, PM_{10} and CO, respectively; and the RMSEs were 77.8, 98.7 and 805.1 μg m^{-3}, respectively, in fcICsEs24 (see Table 3). In fcICs24, the biases were 8.3, −10.3 and 130.2 μg m^{-3}, respectively; and the RMSEs were 75.1, 95.9 and 838.2 μg m^{-3}, respectively (see Table 3). However, with longer-range forecasts, the impact of DA quickly decayed. The relative reductions in RMSE mostly ranged from 30% to 5% for the second- and third-day forecast. From the perspective of the impact of the assimilated emissions, fcICs performed similarly to fcICsEs for $PM_{2.5}$, PM_{10} and CO, indicating that ICs play key roles in aerosol and CO forecasts during severe haze episodes, while the impact of assimilated emissions seems negligible.

For the SO_2 verification forecast, however, fcICsEs performed much better than both fcICs and the control run. Smaller biases and RMSEs were obtained for almost the entire 72-h forecast range. At nighttime in particular (from 18 to 23 h, 42 to 47 h, and

66 to 73 h), when there was significant systematic overestimation in the control run, both the biases and the RMSEs in fcICsEs were about 30% lower than those of the control run. During the daytime (from 0 to 9 h, 24 to 33 h, and 48 to 57 h), fcICsEs still performed slightly better, although the control run did a near perfect job. As for fcICs, smaller biases and RMSEs were obtained for only the first 3 h. Then, the performance was the same as the control run, indicating that the impact of the ICs had disappeared. These results demonstrate the superiority of the assimilated emissions, and that the joint adjustment of SO_2 ICs and emissions is an efficient way to improve the SO_2 forecast.

The NO_2 DA results for the independent sites showed really poor performance (see Figures 5–7). Smaller biases were gained in the daytime of the experiment trials. At nighttime, however, when the simulated NO_2 deviated considerably from the observations in the control run, the biases of both sets of the validation forecasts became even larger. Besides, almost all the RMSEs of both sets of the validation forecasts were always larger than those of the control run.

The O_3 DA results were dependent on the NO_2 DA results in the daytime, due to chemical transformation. Both the biases and the RMSEs were larger, as compared with those of the control run (see Figures 5–7). However, at nighttime, when there was significant systematic underestimation in the control run, the biases in fcICsEs had very similar values to those of the analysis. Also, the biases in fcICs ranged between the analysis and the control run; and the RMSEs of both sets of forecasting experiments were about 10% smaller than those of the control run. All these results indicate that the DA system performed well at night.

5.3 Emission optimization results

Besides improved pollutant forecasts, improved estimates of emissions were expected from the joint DA procedure. The MEIC-2010 was constructed on the basis of annual statistical books in which the data were often 2–3 years older than the actual year (Chen et al., 2016). However, consistent efforts aimed at reducing and managing anthropogenic emissions have been made over the past decade to mitigate air pollution. Thus, there was a large difference between the emission year and our simulation year.

Besides, the spatial allocations of these emissions over small spatial scales, and the monthly allocations, will also lead to some uncertainties. Lastly, the emissions inventory cannot fully capture the day-to-day variability or the actual daily variations, though its differentiation in terms of working days and weekend days, plus the daily variations, can be taken into account in practical applications. However, in this assimilation procedure, the differentiation in terms of working days and weekend days, plus the daily variations, was ignored. Therefore, the prescribed anthropogenic emissions were subject to large uncertainties.

Figures 8 and 9 display the spatial distribution of the prescribed emission rates and the differences between the analysis and the prescribed emission rates of $PM_{2.5}$, PM_{10}, NH_3, SO_2, NO and CO averaged over all hours from 6 to 16 October 2014 in the NCP region. The assimilated emission rates of $PM_{2.5}$, SO_2, NO and CO were lower than the prescribed emissions on the whole. In the BTH region especially, the differences reached -0.02 $\mu g \cdot m^{-2} \cdot s^{-1}$, -2.9, -8.8 and -24.65 $mol \cdot km^{-2} \cdot hr^{-1}$, which was a reduction of about 10%–20% of the prescribed emissions. For PM_{10} emissions, the assimilated values were very close to the prescribed ones, indicating that the prescribed PM_{10} emissions had small uncertainties for the NCP region. For NH_3 emissions, the assimilated values were a little larger than the prescribed emissions in large industrial cities like Beijing, Tianjin, Baoding, Xingtai, Handan, and Taiyuan. However, they were smaller than the prescribed emissions in agricultural regions, especially in Shandong Province and Henan Province. However, in the BTH region, the assimilated NH_3 emissions were very close to the prescribed emissions on the whole.

Figure 10 shows the time series of the emission scaling factors and the emissions. As concluded in Peng et al. (2017), the forecast emission scaling factors changed with the analyzed emission scaling factors due to the use of the time smoothing operator. Besides, although the prescribed emissions were constant when designing the assimilation experiment, the analyzed emission scaling factors showed obvious variation with time, as did the analyzed emissions. For the assimilated SO_2 and NO emissions in particular, the diurnal variations were perfect. In addition, the difference between the assimilated emissions and the prescribed emissions were consistent with

those in Figures 8 and 9. The assimilated emissions of $PM_{2.5}$, SO_2, NO and CO were apparently lower than the corresponding prescribed emissions. Whereas, the values of the assimilated emissions of PM_{10} and NH_3 were very close to their corresponding prescribed emissions.

In order to investigate the impact of optimized emissions on chemical simulations, a simulation (fcEs) using the optimized emissions were performed from 5 to 16 October 2014. Same as the control run, the ICs were the ensemble mean of the spin-up forecasts at 00:00 UTC on 5 October 2014. Thus the difference between the fcEs and the control run is the anthropogenic emissions. The results showed that the fcEs performed very similar to the control run in the whole in the BTH region. For $PM_{2.5}$, PM_{10} and CO, the values of the fcEs were a little smaller than those of the control run, which were consistent with the difference of the anthropogenic emissions. For SO_2 and NO_2, fcEs performed much better than the control run in most time though significant systematic overestimation still existed during the nighttime. For O_3, miner improvements were also gained due to the better simulation in fcEs for NO_2.

5.4 Discussion

From the results presented above, it is clear that improvements were achieved for almost all the 72-h verification forecasts using the optimized ICs and emissions for $PM_{2.5}$, PM_{10}, SO_2 and CO concentrations in the BTH region. However, the 72-h NO_2 verification forecasts performed much worse than the control run, due to the assimilation. Plus, the 72-h O_3 verification forecasts performed worse than the control run during the daytime, due to the worse performance of the NO_2 forecasts, although they did perform better at night. However, relatively favorable NO_2 and O_3 forecast results were gained for the Yangtze River delta and Pearl River delta (PRD) regions (see Figure 11). In the PRD region, during the daytime, the three NO_2 forecasts (i.e., the control run, the fcICsEs, and the fcICs) performed similarly, and had relatively small biases and RMSEs. At nighttime, when there was significant systematic overestimation in the control run, the biases and the RMSEs in fcICsEs were much smaller than those in the control run. For the O_3 72-h verification forecasts, fcICsEs

performed much better than the control run, except for the first 8 h. Also, fcICs improved the O_3 forecasts to some extent from the 9- to 72-h forecast range. These results indicate that DA is still an effective way to improve NO_2 and O_3 forecasts.

Regarding the failure to improve the NO_2 and O_3 forecasts in the BTH region, there are three likely factors. And certainly, NO_2 and O_3 forecasts in other areas are also facing similar challenges.

Firstly, there are still some limitations for the EnKF method. EnKF assimilation is influenced greatly by model errors and observation errors. There are many sources of uncertainties in air-quality forecast that were not directly considered in this study (such as chemical schemes and parameterizations, meteorology, and emissions). And it is very difficult to accurately evaluate the uncertainties of models, though the covariance inflation technique was simply applied for all state variables to roughly compensate for model errors. Therefore, we can only obtain suboptimal results through EnKF assimilation. Furthermore, for short-lived chemical reactive species, such as NO_2 and O_3, they undergo highly complex nonlinear photochemical reactions, even on timescales of hours, such that the forecast accuracy is largely dependent on the chemical process as well as the physical transportation process, the ICs, and the emissions. However, those complex photochemical reaction processes are not precisely described in current chemical mechanisms, e.g., heterogeneous reactions (Yang et al., 2015), the photolysis of nitrous acid and $ClNO_2$ during daytime (Zhang et al., 2017), and so on. Therefore, on the one hand, there are still large uncertainties for NO_2 and O_3 forecasts; whilst on the other hand, it is very difficult for NO_2 and O_3 DA to accurately estimate the model errors with a limited ensemble size. Thus, NO_2 and O_3 assimilations do not perform well (Elbern et al., 2007; Tang et al., 2016). However, for SO_2 and CO, which are representative of long-lived chemical reactive species, the chemical reaction process does not work on timescales of hours, meaning that to some extent hourly chemical DA has the potential to improve their forecasts. For CO in particular, due to its inertness, we might be able to obtain high-quality ICs and emissions through DA. The primary sources of aerosol are the dominant part of the atmospheric aerosol concentration. So, 72-h aerosol forecasts may perform similarly to CO, albeit there are

large uncertainties in the chemical model.

Secondly, as stated in the above paragraph, the analysis ICs and emissions are only a mathematical optimum under the existing conditions. In addition, only part of the chemical ICs and emissions are involved in the DA experiment; and VOC ICs and emissions, which may greatly influence the NO_2 and O_3 forecasts, were not included here because of the absence of VOC measurements. Although we carried out two DA sensitivity experiments to adjust the VOC ICs and emissions through the use of NO_2 or O_3 measurements, we were still unable to gain improved NO_2 and O_3 forecasts in the BTH region in both DA experiments. VOC measurements are needed to reduce uncertainties of VOC ICs and emissions. In addition, almost all available data were observed in cities, and no observation stations located in rural. Thus, the atmospheric environmental monitoring system was still spatially heterogeneous.

Another important point is that there are still limitations to the current chemical mechanisms used in our model, such as the treatment of model error. NO is the primary species of NO_x emissions in city areas, and reacts directly with O_3 to form NO_2 ($NO+O_3 \rightarrow NO_2+O_2$). Thus, O_3 concentrations may inversely correlate with NO_2 concentrations at night. Consequently, air quality models may systematically underestimate O_3 concentrations. Currently, DA can only revise the ICs and the emissions in this work. It cannot change the model performance, especially when there are certain uncertainties for the meteorological simulation.

6. Summary

In this study, we developed an EnKF system to simultaneously assimilate surface measurements of PM_{10}, $PM_{2.5}$, SO_2, NO_2, O_3 and CO via the joint adjustment of ICs and source emissions. This system was applied to assimilate hourly pollution data while modeling an extreme haze event that occurred in early October 2014 over North China. In order to evaluate the impact of DA, two sets of 72-h verification forecasts were performed. One was conducted with the optimized ICs and emissions, and the other with only optimized ICs and the prescribed emissions. A control experiment without DA was also performed for comparison.

The results showed that both verification forecasts performed much better than the control simulations for $PM_{2.5}$, PM_{10} and CO. Obvious improvements were achieved for almost the entire 72-h forecast range. For the first-day forecast especially, near perfect forecasts results were achieved. However, with longer-range forecasts, the impact of DA quickly decayed. In addition, the forecasts with only optimized ICs and the prescribed emissions performed similarly to that with the optimized ICs and emissions, indicating that ICs play key roles in $PM_{2.5}$, PM_{10} and CO forecasts during severe haze episodes.

Also, large improvements were achieved for SO_2 forecasts with both the optimized ICs and emissions for the whole 72-h forecast range. However, similar improvements were achieved for SO_2 forecasts with the optimized ICs only for just the first 3 h, and then the impact of the ICs decayed quickly to zero. This demonstrates that the joint adjustment of SO_2 ICs and emissions is an efficient way to improve SO_2 forecasts.

Even though we failed to improve the NO_2 and O_3 forecasts in the BTH region, relatively favorable NO_2 and O_3 forecast results were gained in other areas. Also, the forecasts with both the optimized ICs and emissions performed much better than the forecasts with only optimized ICs and the prescribed emissions. These results indicate that there is still potential to improve NO_2 and O_3 forecasts via the joint adjustment of SO_2 ICs and emissions.

However, only a case was investigated in this work. Thus it is uncertain if the conclusions about different performance of forecasts for various species would hold in a general. Therefore, more case studies are needed to obtain general conclusions in future works.

References

Barbu, A. L., Segers, A. J., Schaap, M., Heemink, A.W., and Builtjes, P. J. H.: A multi-component data assimilation experiment directed to sulphur dioxide and sulphate over Europe, Atmos. Environ., 43, 1622–1631, 2009.

Carmichael, G. R., Daescu, D. N., Sandu, A., and Chai, T.: Computational aspects of chemical data assimilation into atmospheric models, in Science Computational ICCS 2003. Lecture Notes in Computer Science, IV, 269–278, Springer, Berlin, 2003.

Carmichael, G. R., Sandu, A., Chai, T., Daescu, D. N., Constantinescu, E. M., and Tang, Y.: Predicting air quality: improvements through advanced methods to integrate models and measurements, J. Comput. Phys., 227, 3540–

3571, 2008a.
Carmichael, G. R., Sakuraib, T., Streetsc, D., Hozumib, Y., Uedab, H., Parkd, S. U., Funge, C., Hanb, Z., Kajinof, M., Engardtg, M., Bennetg, C., Hayamih, H., Sarteleti, K., Hollowayj, T., Wangk, Z., Kannaril, A., Fum, J., Matsudan, K., Thongboonchooa, N., and Amanno, M.: MICS-ASIA II: the model intercomaprison study for Asia phase II methodology and overview of findings, Atmos. Environ., 42, 3468–3490, 2008b.
Chai, T., Carmichael, G. R., Tang, Y., Sandu, A., Hardesty, M., Pilewskie, P., Whitlow, S., Browell, E. V., Avery, M. A., Nedelec, P., Merrill, J. T., Thompson, A. M., and Williams, E.: Four dimensional data assimilation experiments with International Consortium for Atmospheric Research on Transport and Transformation ozone measurements, J. Geophys. Res., 112, D12S15, doi:10.1029/2006JD007763, 2007.
Chen, D., Liu, Z., Fast, J., and Ban, J.: Simulations of sulfate – nitrate – ammonium (SNA) aerosols during the extreme haze events over northern China in October 2014, Atmos. Chem.Phys., 16, 10707 – 10724, doi:10.5194/acp-16-10707-2016, 2016.
Chen, F. and Dudhia, J.: Coupling an advanced land surfacehydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity, Mon. Weather Rev., 129, 569 – 585, doi:10.1175/1520-0493(2001)129<0569:Caalsh>2.0.Co;2, 2001.
Chin, M., Rood, R. B., Lin, S. J., Muller, J. F., and Thompson,A. M.: Atmospheric sulfur cycle simulated in the global model GOCART: Model description and global properties, J. Geophys. Res.-Atmos., 105, 24671 – 24687, 2000.
Chin, M., Ginoux, P., Kinne, S., Torres, O., Holben, B. N., Duncan, B. N., Martin, R. V., Logan, J. A., Higurashi, A., and Nakajima, J.: Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and Sun photometer measurements, J. Atmos. Sci., 59, 461 – 483, 2002.
Chou, M.-D. and Suarez, M. J.: An efficient thermal infrared radiation parameterization for use in general circulation models, NASA Tech. Memo., TM 104606, vol. 3, 25 pp., NASA Goddard Space Flight Cent., Greenbelt, MD, USA, 1994.
Dai, T., Schutgens, N. A. J., Goto, D., Shi, G. Y., and Nakajima, T.: Improvement of aerosol optical properties modeling over Eastern Asia with MODIS AOD assimilation in a global non-hydrostatic icosahedral aerosol transport model, Environ. Pollut., 195, 319–329, 2014.
Ding, A. J., Huang, X., Nie, W., Sun, J., Kerminen, V. M., Petaja, T., Su, H. L., Cheng, Y. F., Yang, X. Q., and Wang, M.: Enhanced haze pollution by black carbon in megacities in China, Geophys. Res. Lett., 2873-2879, doi:10.1002/2016GL067745, 2016.
Ding, J., van der A, R. J., Mijling, B., Levelt, P. F., and Hao, N.: NO_x emission estimates during the 2014 Youth Olympic Games in Nanjing, Atmos. Chem. Phys., 15, 9399–9412, doi:10.5194/acp-15-9399-2015, 2015.
Elbern, H., Schmidt, H., and Ebel, A.: Variational data assimilation for tropospheric chemistry modelling, J. Geophys. Res., 102, 15967–15985, 1997.
Elbern, H. and Schmidt, H.: A 4D-Var chemistry data assimilation scheme for Eulerian chemistry transport modelling, J. Geophys. Res., 104, 18583–18598, 1999.
Elbern, H., Schmidt, H., Talagrand, O., and Ebel, A.: 4D-variational data assimilation with an adjoint air quality model for emission analysis, Environ, Model. Softw., 15, 539–548, 2000.
Elbern, H. and Schmidt, H.: Ozone episode analysis by four dimensional variational chemistry data assimilation, J. Geophys. Res., 106, 3569–3590, 2001.
Elbern, H., Strunk, A., Schmidt, H., and Talagrand, O.: Emission rate and chemical state estimation by 4-dimensional variational inversion, Atmos. Chem. Phys., 7, 3749–3769, doi:10.5194/acp-7-3749-2007, 2007.
Evensen, G.: Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistcs, J. Geophys. Res., 99, 10143 – 10162, 1994.

606 Gaspari, G. and Cohn S. E.: Construction of correlation functions in two and three dimensions, Quart. J. R. Meteorol.
607 Soc. 125 (1999), 723–757.
608 Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O., and Lin, S.-J.: Sources and distributions
609 of dust aerosols simulated with the GOCART model, J. Geophys. Res., 106, 20255 – 20273,
610 doi:10.1029/2000JD000053, 2001.
611 Grell, G., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled
612 "online" chemistry within the WRF model, Atmos. Environ., 39, 6957 – 6975,
613 doi:10.1016/j.atmosenv.2005.04.027, 2005.
614 Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P., Klinger, L., Lerdau, M.,
615 McKay, W., Pierce, T., Scholes, B., Steinbrecher, R., Tallamraju, R., Taylor, J., and Zimmerman, P.: A global
616 model of natural volatile organic compound emissions, J. Geophys. Res., 100, 8873 – 8892,
617 doi:10.1029/94JD02950, 1995.
618 Guerrette, J. J. and Henze, D. K.: Development and application of the WRFPLUS-Chem online chemistry adjoint
619 and WRFDA-Chem assimilation system, Geosci. Model Dev., 8, 1857–1876, doi:10.5194/gmd-8-1857-2015,
620 2015.
621 Hakami, A., Henze, D. K., Seinfeld, J. H., Chai, T., Tang, Y., Carmichael, G. R., and Sandu, A.: Adjoint inverse
622 modeling of black carbon during the Asian Pacific Regional Aerosol Characterization Experiment, J. Geophys.
623 Res.-Atmos., 110, D14301, doi:10.1029/2004JD005671, 2005.
624 Heemink, A.W. and Segers, A. J.: Modeling and prediction of environmental data in space and time using Kalman
625 filtering, Stoch. Environ. Res. Risk A., 16, 225–240, 2002.
626 Henze, D. K., Hakami, A., and Seinfeld, J. H.: Development of the adjoint of GEOS-Chem, Atmos. Chem. Phys., 7,
627 2413–2433, doi:10.5194/acp-7-2413-2007, 2007.
628 Henze, D. K., Seinfeld, J. H., and Shindell, D. T.: Inverse modeling and mapping US air quality influences of
629 inorganic $PM_{2.5}$ precursor emissions using the adjoint of GEOS-Chem, Atmos. Chem. Phys., 9, 5877–5903,
630 doi:10.5194/acp-9-5877-2009, 2009.
631 Hong, S. Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment
632 processes, Mon. Weather Rev., 134, 2318–2341, doi:10.1175/Mwr3199.1, 2006.
633 Houtekamer, P. L., Mitchell, H. L., Pellerin, G., Buehner, M., Charron, M., Spacek, L., and Hansen, B.: Atmospheric
634 data assimilation with an ensemble Kalman filter: Results with real observations, Mon. Weather Rev., 133,
635 604 – 620, 2005.
636 Huang, X., Song, Y., Zhao, C., Li, M., Zhu, T., Zhang Q., and Zhang, X.Y.: Pathways of sulfate enhancement by
637 natural and anthropogenic mineral aerosols in China, J. Geophys. Res. – Atmos., 119, 24, 14165-14179, 2014.
638 Jiang, Z., Liu, Z., Wang, T., Schwartz, C. S., Lin, H.-C., and Jiang, F.: Probing into the impact of 3DVAR
639 assimilation of surface PM_{10} observations over China using process analysis, J. Geophys. Res.-Atmos., 118,
640 6738–6749, doi:10.1002/jgrd.50495,2013.
641 Li, Z., Zang, Z., Li, Q. B., Chao, Y., Chen, D., Ye, Z., Liu, Y., and Liou, K. N.: A three-dimensional variational data
642 assimilation system for multiple aerosol species with WRF/Chem and an application to $PM_{2.5}$ prediction, Atmos.
643 Chem. Phys., 13, 4265–4278, doi:10.5194/acp-13-4265-2013, 2013.
644 Li, M., Zhang, Q., Streets, D. G., He, K. B., Cheng, Y. F., Emmons, L. K., Huo, H., Kang, S. C., Lu, Z., Shao, M.,
645 Su, H., Yu, X., and Zhang, Y.: Mapping Asian anthropogenic emissions of nonmethane volatile organic
646 compounds to multiple chemical mechanisms, Atmos. Chem. Phys., 14, 5617 – 5638, doi:10.5194/acp-14-
647 5617-2014, 2014
648 Liu, Z., Liu, Q., Lin, H. C., Schwartz, C. S., Lee, Y. H., and Wang, T.: Three-dimensional variational assimilation
649 of MODIS aerosol optical depth: implementation and application to a dust storm over East Asia, J. Geophys.

Res., 116, D23206, doi:10.1029/2011JD016159, 2011.
McLinden, C.A., Fioletov, V., Shephard, M.W., Krotkov, N., Li, C., Martin, R.V., Moran, M.D., and J. Joiner,: Space-based detection of missing sulfur dioxide sources of global air pollution, Nat. Geosci., 9, 496–500, doi:10.1038/ngeo2724, 2016.
Mijling, B., van der A, R. J., Boersma, K. F., Van Roozendael, M., De Smedt, I., and Kelder, H. M.: Reduction of NO_2 detected from space during the 2008 Beijing Olympic Games, Geophys. Res. Lett., 36, L13801, doi:10.1029/2009GL038943, 2009.
Mijling, B. and van der A, R. J.: Using daily satellite observations to estimate emissions of short-lived air pollutants on a mesoscopic scale, J. Geophys. Res., 117, D17302, doi:10.1029/2012JD017817, 2012.
Mijling, B., van der A, R. J., and Zhang, Q.: Regional nitrogen oxides emission trends in East Asia observed from space, Atmos. Chem. Phys., 13, 12003–12012, doi:10.5194/acp-13-12003-2013, 2013.
Miyazaki, K., Eskes, H. J., Sudo, K., Takigawa, M., van Weele, M., and Boersma, K. F.: Simultaneous assimilation of satellite NO_2, O_3, CO, and HNO_3 data for the analysis of tropospheric chemical composition and emissions, Atmos. Chem. Phys., 12, 9545– 9579, doi:10.5194/acp-12-9545-2012, 2012.
Miyazaki, K. and Eskes, H.: Constraints on surface NO_x emissions by assimilating satellite observations of multiple species, Geophys. Res. Lett., 40, 4745–4750, doi:10.1002/grl.50894, 2013.
Miyazaki, K., Eskes, H. J., Sudo, K., and Zhang, C.: Global lightning NOx production estimated by an assimilation of multiple satellite data sets, Atmos. Chem. Phys., 14, 3277–3305, doi:10.5194/acp-14-3277-2014, 2014.
Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, J. Geophys. Res.-Atmos., 102, 16663–16682, doi:10.1029/97jd00237, 1997.
Nie, W., Ding, A., Wang, T., Kerminen, V.-M., George, C., Xue, L., Wang, W., Zhang, Q., Petäjä, T., Qi, X., Gao, X., Wang, X., Yang, X., Fu, C., and Kulmala, M.: Polluted dust promotes new particle formation and growth, Sci. Rept., 4, 6634, 2014.
Pagowski, M., Grell, G. A., McKeen, S. A., Peckham, S. E., and Devenyi, D.: Three-dimensional variational data assimilation of ozone and fine particulate matter observations: some results using the Weather Research and Forecasting – Chemistry model and Grid-point Statistical Interpolation, Q. J. Roy. Meteor. Soc., 136, 2013–2024, doi:10.1002/qj.700, 2010.
Pagowski, M., and Grell, G. A.: Experiments with the assimilation of fine aerosols using an ensemble Kalman filter, J. Geophys. Res.-Atmos., 117, D21302, doi:10.1029/2012jd018333, 2012.
Pagowski, M., Liu, Z., Grell, G. A., Hu, M., Lin, H.-C., and Schwartz, C. S.: Implementation of aerosol assimilation in Gridpoint Statistical Interpolation (v. 3.2) and WRF-Chem (v.3.4.1), Geosci. Model Dev., 7, 1621-1627, https://doi.org/10.5194/gmd-7-1621-2014, 2014.
Parrish, D. D., M. Trainer, M. P. Buhr, B. A. Watkins, and F. C. Fehsenfeld, Carbon monoxide concentrations and their relation to concentrations of total reactive oxidized nitrogen at two rural U.S. sites, J. Geophys. Res., 96, 9309– 9320, 1991.
Parrish, D. D., and Zhu, T.: Clean Air for Megacities, Science, 326, 674-675, 408 doi:10.1126/science.1176064, 2009.
Peng, Z., Zhang, M., Kou, X., Tian, X., and Ma, X.: A regional carbon data assimilation system and its preliminary evaluation in East Asia, Atmos. Chem. Phys., 15, 1087-1104, doi:10.5194/acp-15-1087-2015, 2015.
Peng, Z., Liu, Z., Chen, D., and Ban, J.: Improving PM2. 5 forecast over China by the joint adjustment of initial conditions and source emissions with an ensemble Kalman filter, Atmos. Chem. Phys., 17, 4837-4855, https://doi.org/10.5194/acp-17-4837-2017, 2017.
Pope, C. A.: Review: Epidemiological basis for particulate air pollution health standards, Aerosol Sci. Tech., 32, 4–

14, 2000.

Pope, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., and Thurston, G. D.: Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution, J. Am. Med. Assoc., 287, 1132–1141, 2002.

Sandu, A., Daescu, D., Carmichael, G. R., and Chai, T.: Adjoint sensitivity analysis of regional air quality models, J. Comput. Phys., 204, 222–252, 2005.

Schutgens, N. A. J., Miyoshi, T., Takemura, T., and Nakajima, T.: Sensitivity tests for an ensemble Kalman filter for aerosol assimilation, Atmos. Chem. Phys., 10, 6583–6600, doi:10.5194/acp-10-6583-2010, 2010a.

Schutgens, N. A. J., Miyoshi, T., Takemura, T., and Nakajima, T.: Applying an ensemble Kalman filter to the assimilation of AERONET observations in a global aerosol transport model, Atmos. Chem. Phys., 10, 2561–2576, doi:10.5194/acp-10-2561-2010, 2010b.

Schutgens, N., Nakata, M., and Nakajima, T.: Estimating Aerosol Emissions by Assimilating Remote Sensing Observations into a Global Transport Model, Remote Sens., 4, 3528–3543, 2012.

Schwartz, C. S., Liu, Z., Lin, H. C., and McKeen, S. A.: Simultaneous three-dimensional variational assimilation of surface fine particulate matter and MODIS aerosol optical depth, J. Geophys. Res., 117, D13202, doi:10.1029/2011JD017383, 2012.

Schwartz, C. S., Liu, Z., Lin, H.-C., and Cetola, J. D.: Assimilating aerosol observations with a "hybrid" variational-ensemble data assimilation system, J. Geophys. Res.-Atmos., 119, 4043–4069, doi:10.1002/2013JD020937, 2013.

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.

Stockwell, W. R., Kirchner, F., Kuhn, M., and Seefeld, S: A new mechanism for regional atmospheric chemistry modeling, J. Geophys. Res., 102(D22), 25,847–25,879, doi:10.1029/97JD00849, 1997.

Sun, Y. L., Wang, Z. F., Du, W., Zhang, Q., Wang, Q. Q., Fu, P. Q., Pan, X. L., Li, J., Jayne, J., and Worsnop, D. R.: Longterm real-time measurements of aerosol particle composition in Beijing, China: seasonal variations, meteorological effects, and source analysis, Atmos. Chem. Phys., 15, 10149 – 10165, doi:10.5194/acp-15-10149-2015, 2015.

Tang, X., Wang, Z. F., Zhu, J., Gbaguidi, A., Wu, Q. Z., Li, J., and Zhu, T.: Sensitivity of ozone to precursor emissions in urban Beijing with a Monte Carlo scheme, Atmos. Environ., 44, 3833–3842, 2010.

Tang, X., Zhu, J., Wang, Z. F., and Gbaguidi, A.: Improvement of ozone forecast over Beijing based on ensemble Kalman filter with simultaneous adjustment of initial conditions and emissions, Atmos. Chem. Phys., 11, 12901–12916, doi:10.5194/acp-11-12901-2011, 2011.

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.

Tang, X., Zhu, J., Wang, Z., Gbaguidi, A., Lin, C., Xin, J., Song, T., and Hu, B.: Limitations of ozone data assimilation with adjustment of NO_x emissions: mixed effects on NO_2 forecasts over Beijing and surrounding areas, Atmos. Chem. Phys., 16, 6395-6405, https://doi.org/10.5194/acp-16-6395-2016, 2016.

van der A, R. J., Peters, D. H. M. U., Eskes, H., Boersma, K. F., Van Roozendael, M., De Smedt, I., and Kelder, H. M.: Detection of the trend and seasonal variation in tropospheric NO_2 over China, J. Geophys. Res., 111, D12317, doi:10.1029/2005JD006594, 2006.

van der A, R. J., Mijling, B., Ding, J., Koukouli, M. E., Liu, F., Li, Q., Mao, H., and Theys, N.: Cleaning up the air: effectiveness of air quality policy for SO_2 and NO_x emissions in China, Atmos. Chem. Phys., 17, 1775-1789, https://doi.org/10.5194/acp-17-1775-2017, 2017.

van Loon, M., Builtjes, P. J. H., and Segers, A. J.: Data assimilation of ozone in the atmospheric transport chemistry model LOTOS, Environ. Model. Softw., 15, 603–609, 2000.

Wang, Z., Li, J., Wang, Z., Yang, W., Tang, X., Ge, B., Yan, P., Zhu, L., Chen, X., and Chen, H.: Modeling study of regional severe hazes over mid-eastern China in January 2013 and its implications on pollution prevention and control, Sci.China-Earth Sci., 57, 3–13, 2014.

Whitaker, J. S., and Hamill, T. M.: Ensemble data assimilation without perturbed observations, Mon. Weather Rev., 130, 1913–1924, 2002.

Wild, O., Zhu, X., and Prather, M. J.: Fast-j: Accurate simulation of in- and below-cloud photolysis in tropospheric chemical models, J. Atmos. Chem., 37, 245–282, doi:10.1023/A:1006415919030, 2000.

Xie, Y., Ding, A., Nie, W., Mao, H., Qi, X., Huang, X., Xu, Z., Kerminen, V.-M., Petäjä, T., Chi, X., Virkkula, A., Boy, M., Xue, L., Guo, J., Sun, J., Yang, X., Kulmala, M., and Fu, C.: Enhanced sulfate formation by nitrogen dioxide: Implications from in situ observations at the SORPES station, J. Geophys. Res. – Atmos., 120, 24, 12679-12694, 2015.

Yang, Y. R., Liu, X. G., Qu, Y., An, J. L., Jiang, R., Zhang, Y. H., Sun, Y. L., Wu, Z. J., Zhang, F., Xu, W. Q., and Ma, Q. X.: Characteristics and formation mechanism of continuous hazes in China: a case study during the autumn of 2014 in the North China Plain, Atmos. Chem. Phys., 15, 8165–8178, doi:10.5194/acp-15-8165-2015, 2015.

Yumimoto, K., Uno, I., Sugimoto, N., Shimizu, A., Liu, Z., and Winker, D. M.: Adjoint inversion modeling of Asian dust emission using lidar observations, Atmos. Chem. Phys., 8, 2869-2884, doi:10.5194/acp-8-2869-2008, 2008.

Yumimoto, K., Uno, I., Sugimoto, N., Shimizu, A., and Satake, S.: Adjoint inverse modeling of dust emission and transport over East Asia, Geophys. Res. Lett., 34, L00806, doi:10.029/2006GL028551, 2007.

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, R., Wang, G., Guo, S., Zamora, M. L., Ying, Q., Lin, Y., Wang, W., Hu, M., and Wang, Y.: Formation of Urban Fine Particulate Matter, Chem. Rev., 115, 3803–3855, doi:10.1021/acs.chemrev.5b00067, 2015a.

Zhang, L., Shao, J. Y., Lu, X., Zhao, Y. H., Hu, Y. Y., Henze, D. K., et al.: Sources and processes affecting fine particulate matter pollution over North China: An adjoint analysis of the Beijing APEC period. Environmental Science & Technology, 50(16), 8731–8740. https://doi.org/10.1021/acs.est.6b03010, 2016.

Zhang, L., Li, Q., Wang, T., Ahmadov, R., Zhang, Q., Li, M., and Lv, M.: Combined impacts of nitrous acid and nitryl chloride on lower-tropospheric ozone: new module development in WRF-Chem and application to China, Atmos. Chem. Phys., 17, 9733-9750, https://doi.org/10.5194/acp-17-9733-2017, 2017.

Zhao, X. J., Zhao, P. S., Xu, J., Meng,,W., Pu,W.W., Dong, F., He, D., and Shi, Q. F.: Analysis of a winter regional haze event and its formation mechanism in the North China Plain, Atmos. Chem. Phys., 13, 5685–5696, doi:10.5194/acp-13-5685-2013, 2013.

Zheng, B., Zhang, Q., Zhang, Y., He, K. B., Wang, K., Zheng, G. J., Duan, F. K., Ma, Y. L., and Kimoto, T.: Heterogeneous chemistry: a mechanism missing in current models to explain secondary inorganic aerosol formation during the January 2013 haze episode in North China, Atmos.Chem.Phys., 15, 2031-2049, 10.5194/acp-15-2031-2015, 2015.

List of Figures and Tables

Table 1. WRF-Chem model configurations in this study.

780 Table 2. State vectors in the data assimilation system.

781 Figure 1. The model domain (left) and the North China Plain (right). Black dots are the
782 observational sites used for assimilation, and red stars are the observational sites used
783 for validation. The green frame marks the Beijing–Tianjin–Hebei region.

784 Figure 2. Time series of prior ensemble mean RMSE (blue line) and total spread (black
785 line) for $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO and O_3 concentrations aggregated over all
786 observations over the Beijing–Tianjin–Hebei region. Units for all these variables are
787 $\mu g\ m^{-3}$.

788 Figure 3. Time series of the area-averaged ensemble spread for the emission scaling
789 factors over the Beijing–Tianjin–Hebei region.

790 Figure 4. Time series of the hourly pollutant concentrations in the Beijing–Tianjin–
791 Hebei (BTH) region obtained from observations (red line), the control run (black line),
792 the analysis (pink line), the first-day forecast from fcICsEs (fcICsEs24, blue line), and
793 the first-day forecast from fcICs (fcICs24, blue line). The observations were obtained
794 from the 47 independent sites in the BTH region. The modelled time series were
795 interpolated to the 47 independent sites using the spatial bilinear interpolator method.
796 The shaded backgrounds indicate the distribution of the observations, where the top
797 edge represented the 90^{th} percentile and the bottom edge the 10^{th} percentile. Units:
798 $\mu g\ m^{-3}$.

799 Table 3. Comparison with observations of the surface $PM_{2.5}$ mass concentrations in the
800 Beijing–Tianjin–Hebei region from the control experiment, the assimilation experiment,
801 and the first-day forecast, over all analysis times from 6 to 16 October 2014. Units:
802 $\mu g\ m^{-3}$.

803 Figure 5. Bias of surface $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO and O_3 as a function of forecast
804 range calculated against all the independent observations over the Beijing–Tianjin–
805 Hebei region shown in Figure 1. The 72-h forecasts were performed at each 0000 UTC
806 from 6 to 14 October 2014 and the statistics were computed from 6 to 14 October. Units:
807 $\mu g\ m^{-3}$.

808 Figure 6. As in Figure 5 but for RMSE. Units: $\mu g\ m^{-3}$.

809 Figure 7. Normalized RMSE (assimilation divided by control) for fcICsEs and fcICs
810 for $PM_{2.5}$, PM_{10}, SO_2 and CO.

811 Figure 8. Spatial distribution of the prescribed emissions (top panels) of $PM_{2.5}$ (left), PM_{10} (middle),
812 and NH_3 (right) and the corresponding time-averaged differences between the ensemble mean
813 analysis and the prescribed values at the lowest model level averaged over all hours from 6 to

16 October 2014 in the NCP region. Units for $PM_{2.5}$ and PM_{10} emissions: $\mu g \cdot m^{-2} \cdot s^{-1}$; and for NH_3 emissions: $mol \cdot km^{-2} \cdot hr^{-1}$.

Figure 9. As in Figure 8 but for SO_2 (left), NO (middle), and CO (right). Units for SO_2, NO and CO emissions: $mol \cdot km^{-2} \cdot hr^{-1}$.

Figure 10. Hourly area-averaged time series extracted from the analyzed emission scaling factors (black line), the forecast emission scaling factors (green dashed line), the analyzed emissions (blue line), and the prescribed emissions (blue dashed line) in the Beijing–Tianjin–Hebei region. Units for $PM_{2.5}$ and PM_{10} emissions: $\mu g \cdot m^{-2} \cdot s^{-1}$; and for NH_3, SO_2, NO and CO emissions: $mol \cdot km^{-2} \cdot hr^{-1}$.

824

825 Table 1. WRF-Chem model configurations in this study.

Parameterization	WRF-Chem Option
Aerosol scheme	Goddard Chemistry Aerosol Radiation and Transport (Chin et al., 2000, 2002)
Photolysis scheme	Fast-J (Wild et al., 2000)
Gas-phase chemistry	Regional Atmospheric Chemistry Mechanism (Stockwell et al., 1997)
Microphysics	the WRF single-moment 5 class scheme
Longwave radiation	Rapid Radiative Transfer Model longwave scheme (Mlawer et al., 1997)
shortwave radiation	Goddard shortwave radiation scheme (Chou and Suarez, 1994)
Planetary boundary layer	Yonsei University boundary layer scheme (Hong et al., 2006)
cumulus parameterization	Grell-3D scheme
Land-surface model	NOAH (Chen and Dudhia, 2001)
Dust and sea salt emissions	Goddard Chemistry Aerosol Radiation and Transport (Chin et al., 2002)

826

827

828

829 Table 2. State vectors in the data assimilation system.

Observations	$PM_{2.5}$	$PM_{10\text{-}2.5}$	SO_2	NO_2	CO	O_3
Mass concentration	P_{25}, S, OC_1, OC_2 BC_1, BC_2, D_1, D_2, S_1, S_2	P_{10}, D_3, D_4, D_5 S_3, S_4,	SO_2	NO, NO_2	CO	O_3
Scaling factors	$\boldsymbol{\lambda}_{PM2.5}$, $\boldsymbol{\lambda}_{NH3}$	$\boldsymbol{\lambda}_{PM10}$	$\boldsymbol{\lambda}_{SO2}$	$\boldsymbol{\lambda}_{NO}$	$\boldsymbol{\lambda}_{CO}$	—

830

831

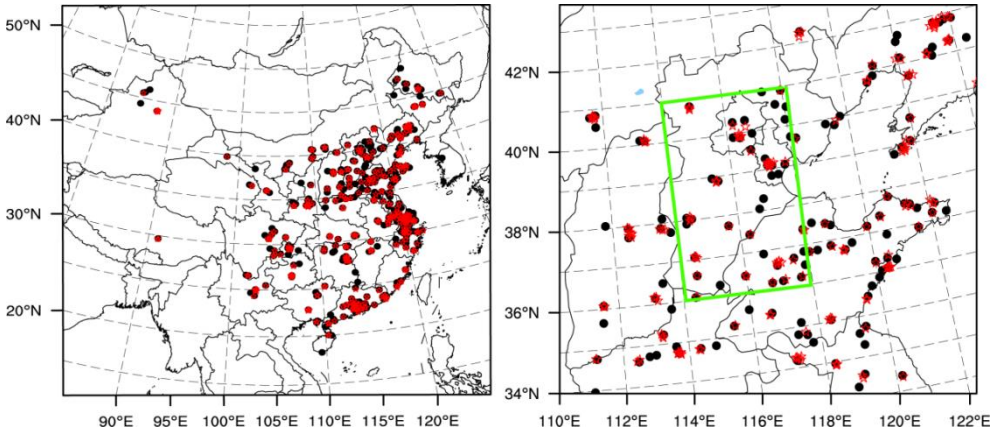


Figure 1. The model domain (left) and the North China Plain (right). Black dots are the observational sites used for assimilation, and red stars are the observational sites used for validation. The green frame marks the Beijing–Tianjin–Hebei region.

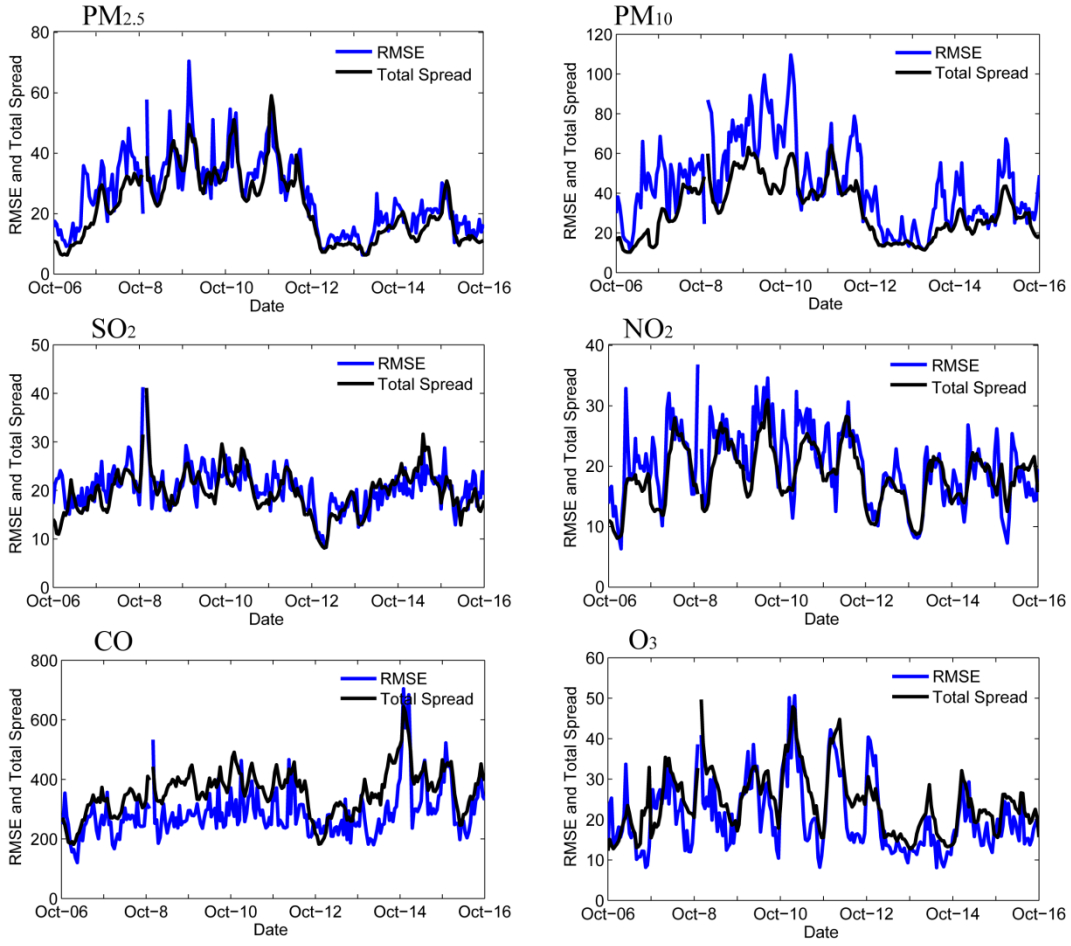


Figure 2. Time series of prior ensemble mean RMSE (blue line) and total spread (black line) for $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO and O_3 concentrations aggregated over all observations over the Beijing–Tianjin–Hebei region. Units for all these variables are $\mu g\ m^{-3}$.

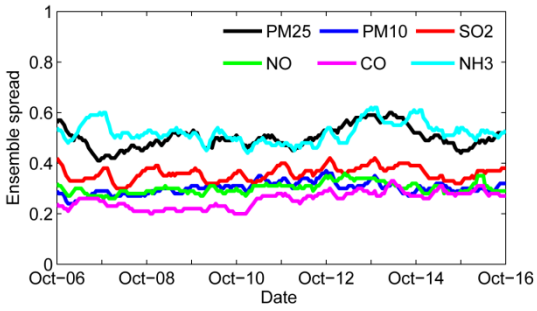


Figure 3. Time series of the area-averaged ensemble spread for the emission scaling factors over the Beijing–Tianjin–Hebei region.

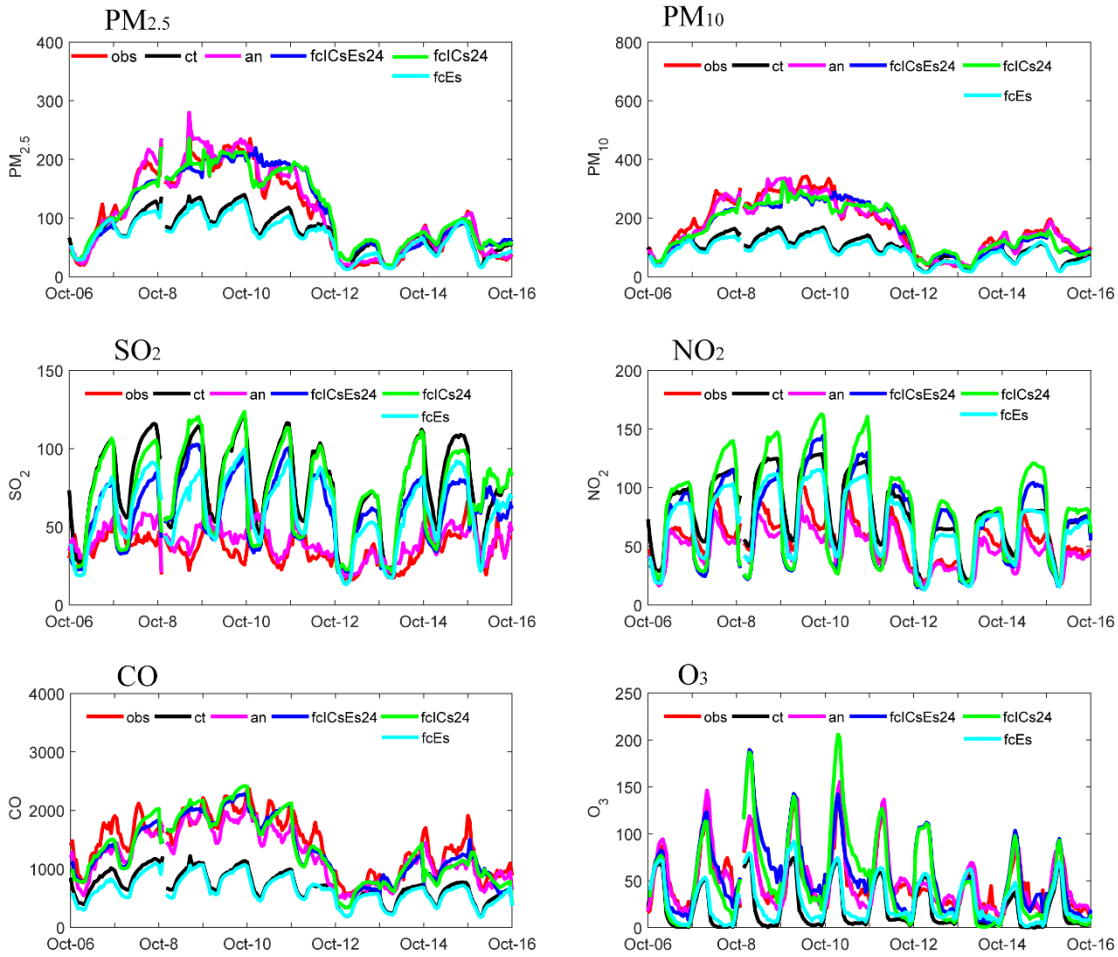


Figure 4. Time series of the hourly pollutant concentrations in the Beijing–Tianjin–Hebei (BTH) region obtained from observations (referred to as "obs", red line), the control run (referred to as "ct", black line), the analysis (referred to as "an", pink line), the first-day forecast from fcICsEs (referred to as "fcICsEs24", blue line), the first-day forecast from fcICs (referred to as "fcICs24", blue line), and the simulation only using the optimized emissions (referred to as "fcEs", cyan line). The observations were obtained from the 47 independent sites in the BTH region. The modelled time series were interpolated to the 47 independent sites using the spatial bilinear interpolator method. Units: μg m^{-3}.

Table 3. Comparison with observations of the surface $PM_{2.5}$ mass concentrations in the Beijing–Tianjin–Hebei region from the control experiment, the assimilation experiment, and the first-day forecast, over all analysis times from 6 to 16 October 2014. Units: μg m^{-3}.

Species	Experiment	Mean observed value	Mean simulated value	BIAS	RMSE	CORR
$PM_{2.5}$	Control	114.8	80.7	−34.1	92.1	0.740
	Analysis		119.9	5.1	51.5	0.891
	fcICsEs24		121.2	6.5	77.8	0.736
	fcICs24		123.1	8.3	75.1	0.748
PM_{10}	Control	174.6	96.9	−77.7	134.6	0.691
	Analysis		169.0	−5.6	63.4	0.890
	fcICsEs24		162.7	−11.9	98.7	0.716
	fcICs24		164.3	−10.3	95.9	0.726
SO_2	Control	33.0	81.1	48.1	66.6	0.088
	Analysis		41.1	8.1	27.9	0.540
	fcICsEs24		62.0	29.0	51.2	0.120
	fcICs24		75.7	42.7	65.8	0.038
NO_2	Control	56.4	78.8	22.4	39.7	0.545
	Analysis		48.0	−8.3	31.7	0.557
	fcICsEs24		71.8	15.4	46.2	0.408
	fcICs24		82.8	26.4	55.5	0.414
CO	Control	1318.0	752.3	−565.7	962.7	0.354
	Analysis		1157.5	−160.4	618.9	0.705
	fcICsEs24		1418.4	100.4	805.1	0.476
	fcICs24		1448.2	130.2	838.2	0.439
O_3	Control	57.5	26.5	−31.0	50.8	0.463
	Analysis		59.6	2.1	31.1	0.753
	fcICsEs24		63.5	6.0	49.0	0.460

	fcICs24	58.98	1.5	50.5	0.478

867
868

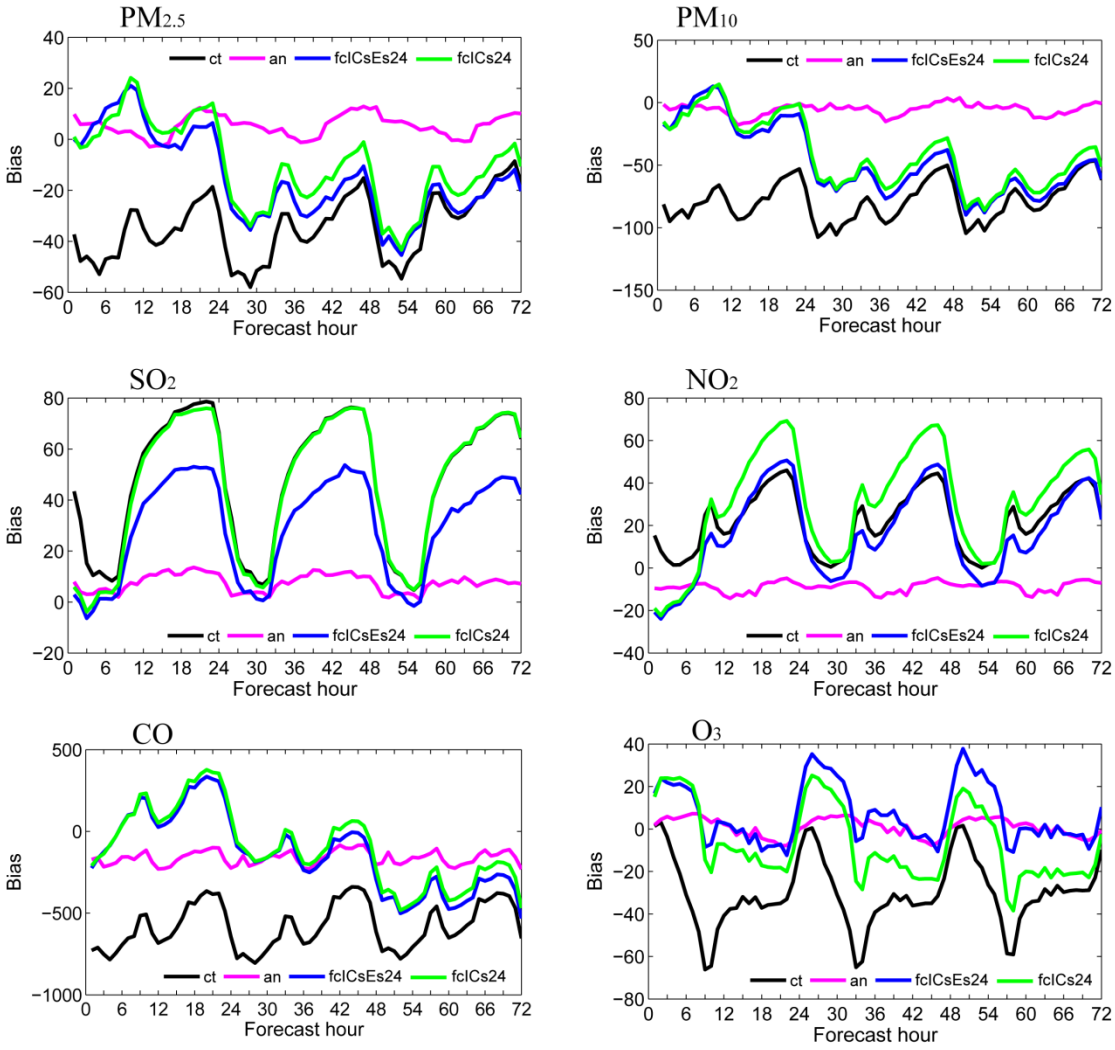


Figure 5. Bias of surface $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO and O_3 as a function of forecast range calculated against all the independent observations over the Beijing–Tianjin–Hebei region shown in Figure 1. The 72-h forecasts were performed at each 0000 UTC from 6 to 14 October 2014 and the statistics were computed from 6 to 14 October. Units: μg m^{-3}.

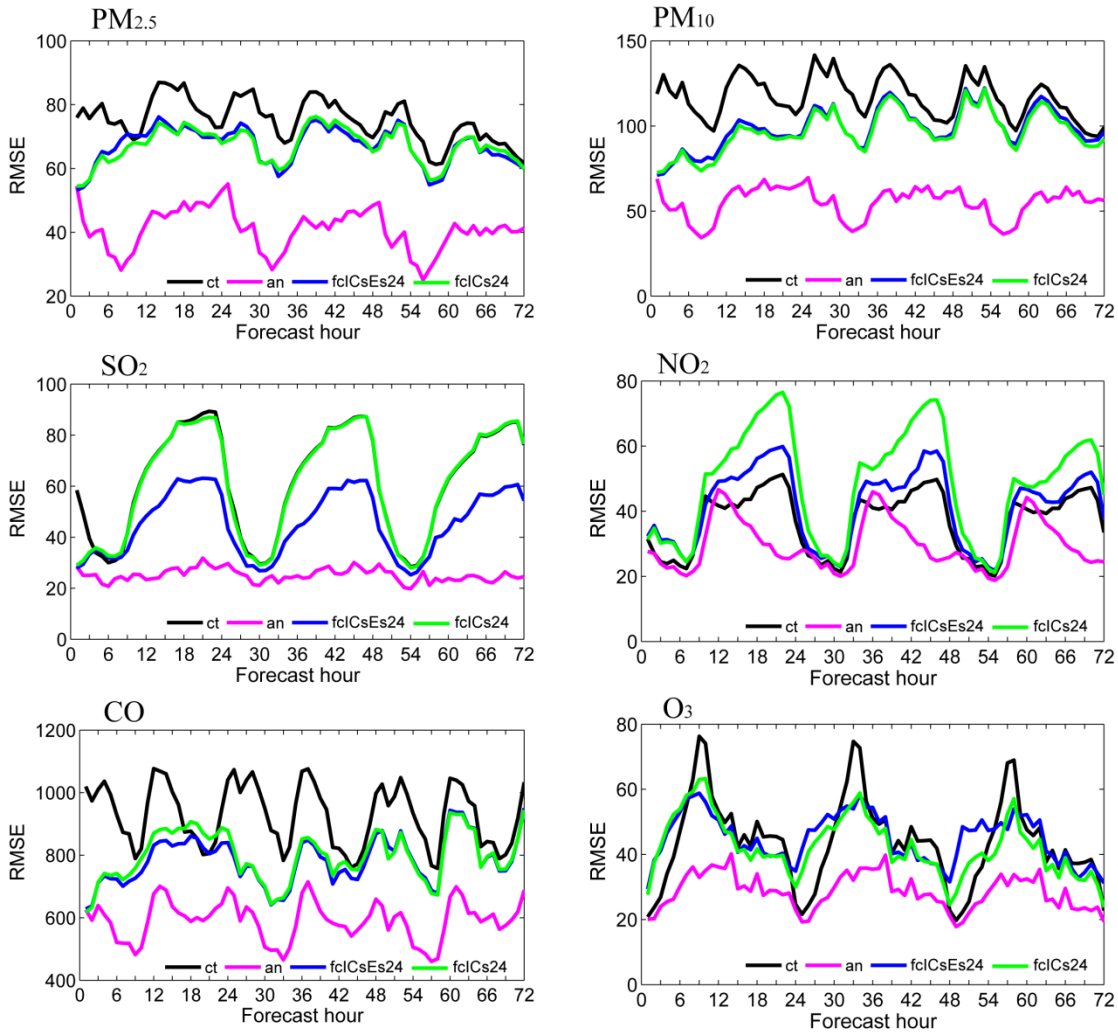


Figure 6. As in Figure 5 but for RMSE. Units: μg m^{-3}.

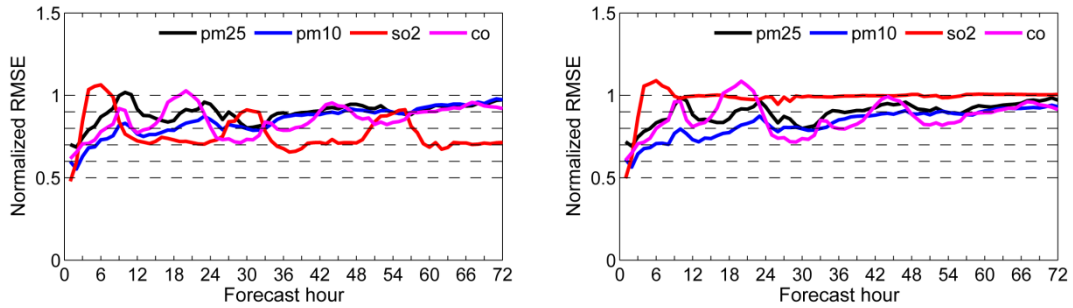


Figure 7. Normalized RMSE (assimilation divided by control) for fcICsEs and fcICs for $PM_{2.5}$, PM_{10}, SO_2 and CO.

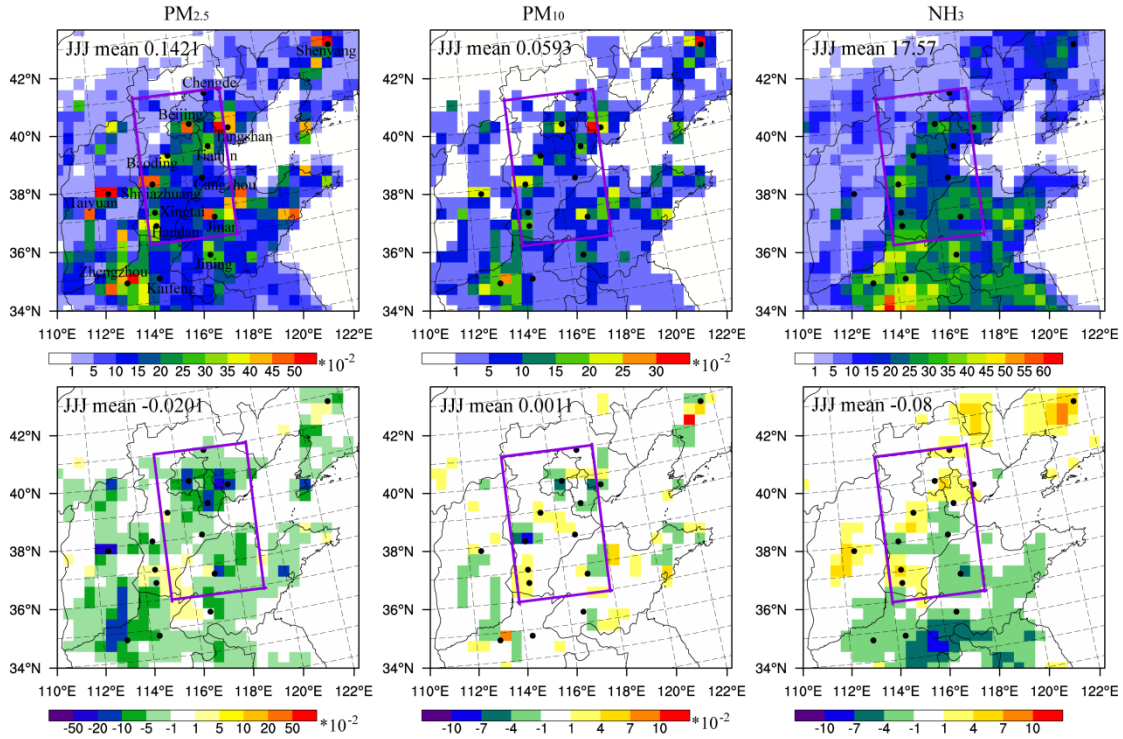


Figure 8. Spatial distribution of the prescribed emissions (top panels) of $PM_{2.5}$ (left), PM_{10} (middle), and NH_3 (right) and the corresponding time-averaged differences between the ensemble mean analysis and the prescribed values at the lowest model level averaged over all hours from 6 to 16 October 2014 in the NCP region. Units for $PM_{2.5}$ and PM_{10} emissions: $\mu g \cdot m^{-2} \cdot s^{-1}$; and for NH_3 emissions: $mol \cdot km^{-2} \cdot hr^{-1}$.

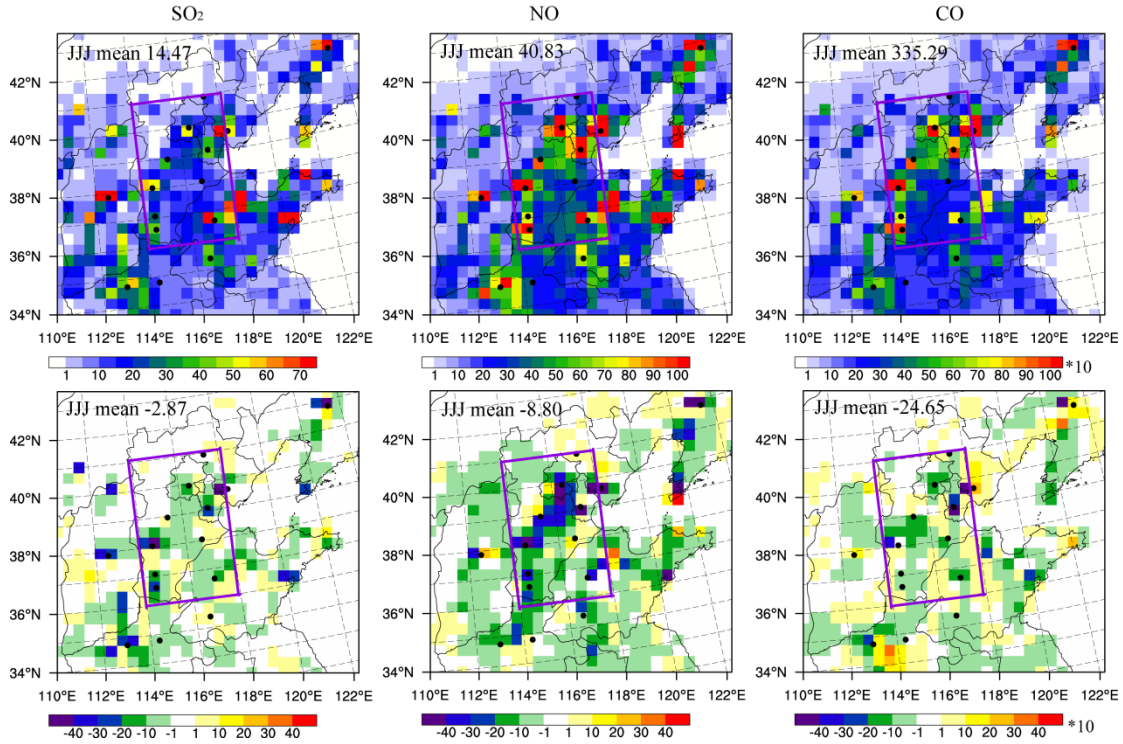


Figure 9. As in Figure 8 but for SO_2 (left), NO (middle), and CO (right). Units for SO_2, NO and CO emissions: $mol \cdot km^{-2} \cdot hr^{-1}$.

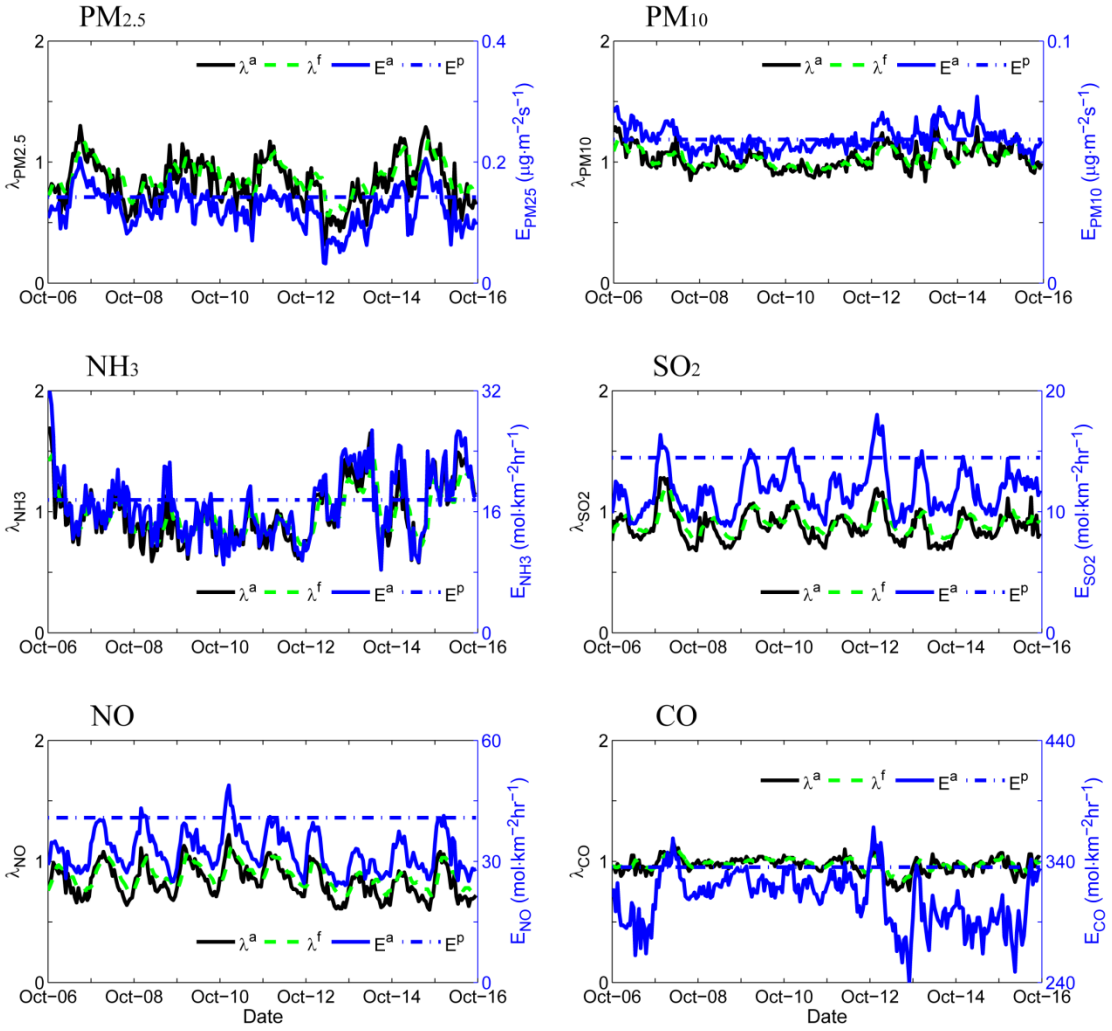


Figure 10. Hourly area-averaged time series extracted from the analyzed emission scaling factors (black line), the forecast emission scaling factors (green dashed line), the analyzed emissions (blue line), and the prescribed emissions (blue dashed line) in the Beijing–Tianjin–Hebei region. Units for $PM_{2.5}$ and PM_{10} emissions: $\mu g \cdot m^{-2} \cdot s^{-1}$; and for NH_3, SO_2, NO and CO emissions: $mol \cdot km^{-2} \cdot hr^{-1}$.

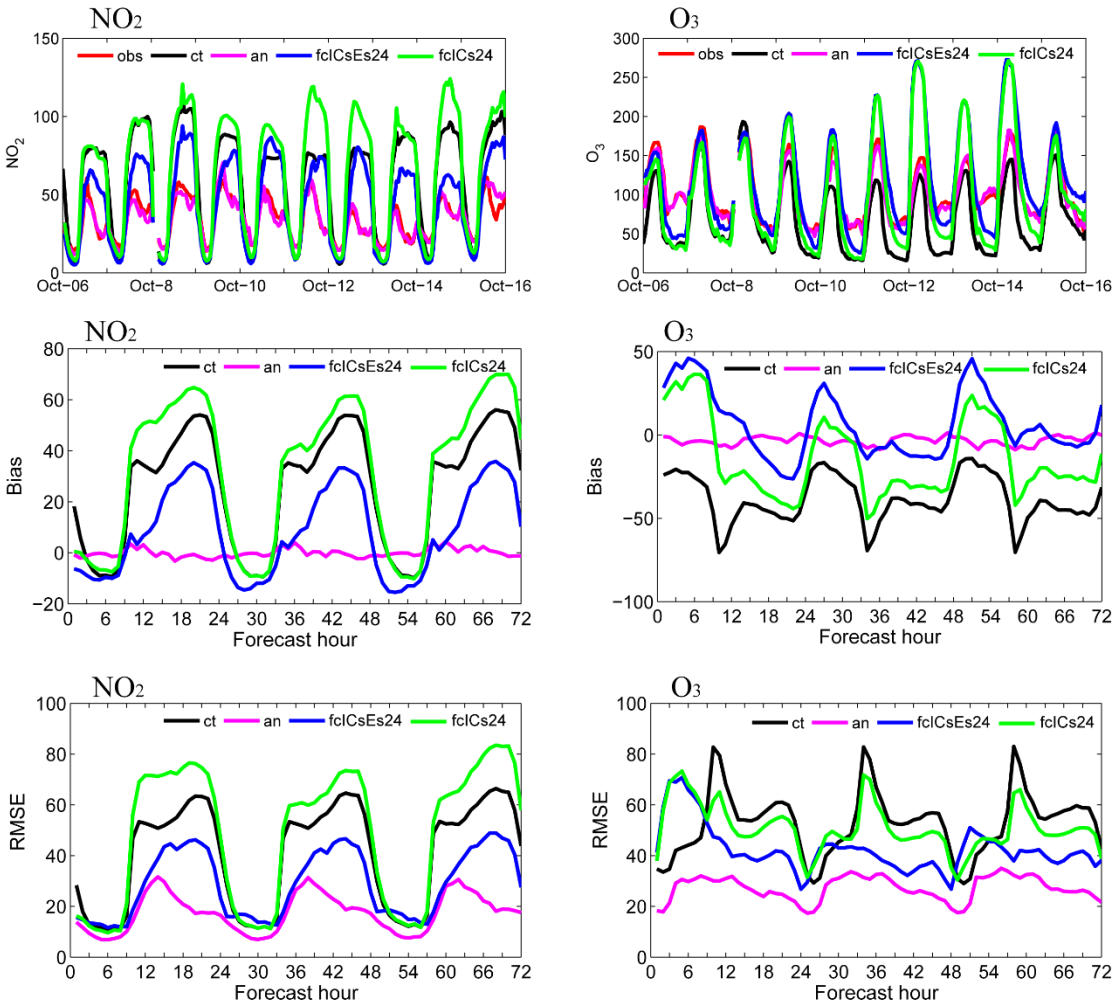


Figure 11. NO_2 and O_3 time series of the hourly pollutant concentrations in the Pearl River Delta region (PRD, 21 °–24 °N, 112.5 °–115 °E) obtained from observations (referred to as "obs", red line), the control run (referred to as "ct", black line), the analysis (referred to as "an", pink line), the first-day forecast from fcICsEs (referred to as "fcICs24", blue line), and the first-day forecast from fcICs (referred to as "fcICs24", blue line). The bias and RMSEs of surface NO_2 and O_3 as a function of forecast range calculated against all the independent observations (34 sites) over the PRD region. Units: μg m^{-3}.