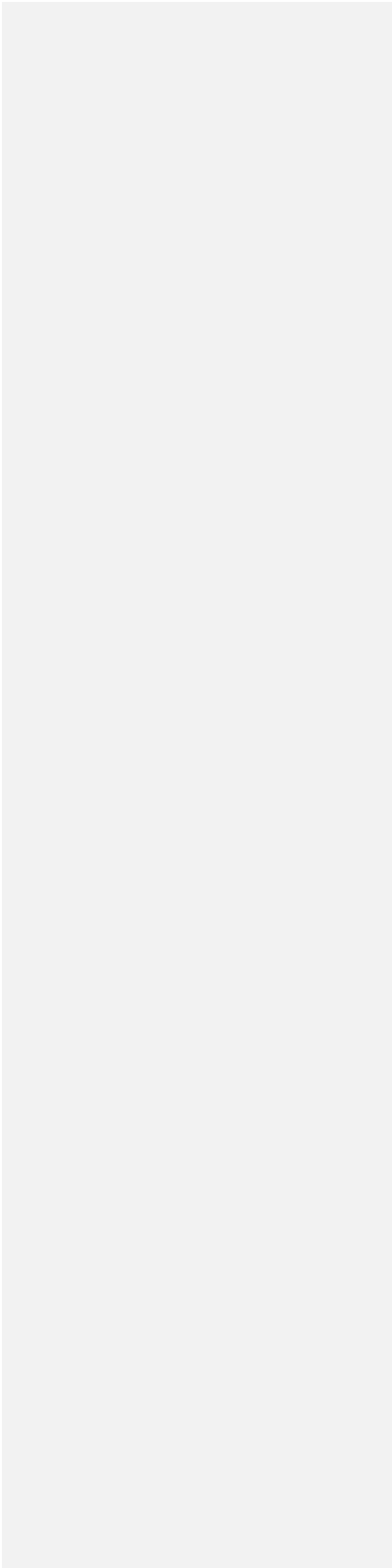


Response to the editor's comments

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

We would like to thank the editor for the valuable comments. We have revised the manuscript as suggested. We provide a marked-up manuscript version showing the changes made by us in the following part.



1 Limitations of ozone data assimilation with adjustment of NOx emissions: mixed 2 effects on NO₂ forecasts over Beijing and surrounding areas

3 Xiao Tang¹, Jiang Zhu¹, Zifa Wang¹, Alex Gbaguidi², Caiyan Lin³, Jinyuan Xin¹, Tao Song¹, Bo
4 Hu¹

5 ¹LAPC, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

6 ²AECOM Asia, Hong Kong, China

7 ³Aviation Meteorological Center, Air Traffic Management Bureau, Civil Aviation Administration of
8 China, Beijing, China

9 Correspondence to: Xiao Tang (tangxiao@mail.iap.ac.cn)

10 Abstract

11 This study investigates a cross-variable ozone data assimilation (DA) method based on an ensemble
12 Kalman filter (EnKF) that has been used in the companion study to improve ozone forecasts over
13 Beijing and surrounding areas. The main purpose is to delve into the impacts of the cross-variable
14 adjustment of nitrogen oxides (NOx) emissions on the nitrogen dioxide (NO₂) forecasts over this
15 region during the 2008 Beijing Olympic Games. A mixed effect on the NO₂ forecasts was observed
16 through application of the cross-variable assimilation approach in the real-data assimilation (RDA)
17 experiments. The method improved the NO₂ forecasts over almost half of the urban sites with
18 reductions of the root mean square errors (RMSEs) by 15%~36% in contrast to big increases of the
19 RMSEs over other urban stations by 56%~239%. Over the urban stations with negative DA impacts,
20 improvement of the NO₂ forecasts (with 7% reduction of the RMSEs) was noticed in the night and
21 morning versus significant deterioration in daytime (with 190% increase of the RMSEs), suggesting
22 that the negative DA impacts mainly occurred during daytime. Ideal data assimilation (IDA)
23 experiments with a box model and the same cross-variable assimilation method confirmed the mixed
24 effects found in the RDA experiments. In the same way, NOx emission estimation was improved in
25 the night and morning even under large biases in the prior emission, while it deteriorated in daytime
26 (except for the case of minor errors in the prior emission). The mixed effects observed in the
27 cross-variable DA, i.e., positive DA impacts on NO₂ forecasts over some urban sites, negative DA
28 impacts over the other urban sites and weak DA impacts over suburban sites, highlighted the
29 limitations of the EnKF under strong nonlinear relationships between chemical variables. Under
30 strong nonlinearity between daytime ozone concentrations and NOx emissions uncertainties (with
31 large biases in the a prior emission), the EnKF may come up with inefficient or wrong adjustments to
32 NOx emissions. The present findings reveal that bias correction is essential for the application of the

apple 16/5/6 3:27 PM
已删除: .

apple 16/5/6 3:27 PM
已删除: .

apple 16/5/6 3:27 PM
已删除: .

apple 16/5/6 3:28 PM
已删除: .

apple 16/5/6 3:28 PM
已删除: .

apple 16/5/6 3:28 PM
已删除: .

apple 16/5/6 3:28 PM
已删除: .

apple 16/5/6 3:28 PM
已删除: .

apple 16/5/6 3:28 PM
已删除: .

apple 16/5/6 3:28 PM
已删除: .

apple 16/5/6 3:28 PM
已删除: .

apple 16/5/6 3:28 PM
已删除: .

apple 16/5/6 3:28 PM
已删除: .

apple 16/5/6 3:50 PM
已删除: .

apple 16/5/6 9:16 AM
已删除: tendency

1 | EnKF in dealing with the DA problem over strong nonlinear system.

apple 16/5/6 9:19 AM
已删除: inconsistency

2 | 1. Introduction

3 | Chemical data assimilation (CDA) integrates models and observations to better represent the
4 | chemical state of the atmosphere and is recognized as a technique for improving the simulations and
5 | forecasts of air pollutants such as ozone and aerosols (Carmichael et al., 2008; Sandu et al., 2011;
6 | Zhang et al., 2012). The role of CDA in optimizing initial and boundary conditions has been explored
7 | in several applications to improve forecasts of ozone and aerosol (Gaubert et al., 2014; Pagowski et
8 | al., 2014). Nevertheless, significant challenges persist in CDA.

9 | One of the major challenges in CDA is that the impact of the initial conditions on the forecast of
10 | air pollutants such as ozone decreases with simulation time (Gaubert et al., 2014; Jimenez et al.,
11 | 2006). To overcome such an obstacle, emissions with large uncertainties and strong impacts on air
12 | quality modeling, identified as the crucial sources of uncertainties and considered to be the key
13 | control variables (Beekmann and Derognat, 2003; Hanna et al., 2001), have been integrated into the
14 | CDA. The importance of emissions as control variables in the CDA has also been documented
15 | recently (Carmichael et al., 2008; Koohkan et al., 2013; Zhang et al., 2012). Accordingly, advanced
16 | CDA techniques that enable inverse or cross-variable adjustments of emissions have been established
17 | and their applications have provided significant improvement of ozone forecasts (e.g., Tang et al.,
18 | 2011).

19 | However, the performances of such advanced CDA on the forecasts of other pollutants related to
20 | tropospheric ozone are rarely reported and have not aroused enough attention. In this field, few
21 | studies stand out (Elbern et al., 2007; van Loon et al., 2000). Elbern et al., (2007) carried out two sets
22 | of data assimilation experiments with a four dimensional variational inversion method: (1)
23 | assimilation of ozone (O₃) and nitrogen dioxide (NO₂) observations simultaneously, and (2)
24 | assimilation of only O₃ observations. Both experiments resulted in reductions of nitrogen oxides
25 | (NO_x) emissions after data assimilation in most cases even if the model underestimated the NO_x
26 | concentrations before data assimilation. Similar results were reported by van Loon et al. (2000)
27 | through the assimilation of O₃ observations and adjustments of sulfur oxides (SO_x) emissions using
28 | an ensemble Kalman filter. The method enhanced the emission rates of SO_x when significant
29 | over-prediction of SO₂ concentrations occurred. Such inconsistencies, i.e., the emissions enhanced

apple 16/5/6 9:25 AM
已删除: subsist

1 under the overestimation of concentrations or the emissions reduced under the underestimation of
2 concentrations, reveal some gaps between ozone forecast improvement and precursor emission
3 optimization and call for a comprehensive evaluation of the cross-variable chemical data assimilation
4 techniques.

5 Tang et al. (2011) employed a high horizontal resolution (9km) model to perform the
6 assimilation of O₃ observations with the ensemble Kalman filter and the adjustment of NO_x
7 emissions for O₃ forecast improvement over Beijing and its surrounding areas. However, the impact
8 of ozone assimilation on the precursor (NO₂ and, volatile organic compounds) uncertainty was not
9 elucidated. This paper (as an extension of Tang et al (2011)), based on the assimilation experiments
10 performed by Tang et al., (2011), attempts to analyze in detail the impacts of the cross-variable ozone
11 data assimilation on NO₂ forecasts over Beijing and surrounding areas during the 2008 Beijing
12 Olympic Games. Both real O₃ data assimilation (with a 3-dimensional chemical transport model) and
13 ideal O₃ data assimilation experiments (with a box model) are performed to investigate the state of
14 NO₂ and NO_x emissions during assimilation processes in order to provide further insights into the
15 scientific potential of the assimilation method.

16 Section 2 describes the chemical transport model employed, the data assimilation algorithm and
17 the surface observation network used for the data assimilation. Results from the real data assimilation
18 experiments and the ideal data assimilation experiments are presented in Sect. 3. Section 4 presents
19 conclusions and discussion.

20 **2. Methodology**

21 **(1) Chemical transport model**

22 The chemical transport model used for O₃ simulations was the Nested Air Quality Prediction
23 Modeling System (NAQPMS) (Wang et al., 2001). Several applications of NAQPMS have been
24 reported for simulating the chemical processes and transports of ozone, modeling the processes of
25 aerosol and acid rain, and providing operational air quality forecasts in megacities such as Beijing
26 and Shanghai (Wang et al., 2006). It contains modules for modeling the processes of emissions,
27 advection, diffusion, dry and wet deposition, gaseous phase, aqueous phase, heterogeneous and
28 aerosol chemical reactions. The gas-chemistry processes were simulated by the Carbon-Bond
29 Mechanism Z (CBM-Z) which includes 133 reactions for 53 species (Zaveri and Peter, 1999). The

apple 16/5/6 9:26 AM

已删除: &

1 dry deposition modeling followed the scheme of Wesely (1999). The vertical eddy diffusivity was
2 parameterized based on a scheme by Byun and Dennis (1995). The O₃ simulations were configured
3 with three nested domains and the horizontal resolutions were 81km, 27km and 9km respectively.
4 The first domain covered East Asia with a 81km resolution and the second domain contained North
5 China with a 27km resolution. The third domain displayed in Fig. 1 covered Beijing and its
6 surrounding areas with 9km resolution. Vertically, the model was set as twenty terrain-following
7 layers, nine of which were within the lowest 2 km of the atmosphere and the height of the first layer
8 near the surface was 50 m. The Fifth-Generation National Center for Atmospheric Research
9 (NCAR)/Penn State Mesoscale Model (MM5; Grell et al., 1994) was employed to provide the hourly
10 meteorological inputs for NAQPMS. The regional emission data of the Intercontinental Chemical
11 Transport Experiment-Phase B (INTEX-B) Asia inventory for 2006 with 0.5° × 0.5° resolution
12 (Zhang et al., 2009) and the local high-resolution emission inventory were combined to provide the
13 emission data for NAQPMS (Tang et al., 2011).

14 **(2) Data assimilation algorithm**

15 The assimilation algorithm employed was the ensemble Kalman filter (EnKF) proposed by
16 Evensen (1994). The main feature of this method consists of a series of ensemble samples generally
17 produced via ensemble forecasts to calculate the background error covariance of state variables. It
18 serves as an approximate version of the Kalman filter (Kalman, 1960). [The](#) EnKF can directly
19 calculate the background error covariance from the ensemble forecasts of the highly nonlinear model,
20 which is very suitable for data assimilation in complex high-dimensional models (Carmichael et al.,
21 2008). Its implementation is very simple and does not require an adjoint model which is a very
22 cumbersome task for complex high-dimensional model. It can be used for combined state and
23 parameter estimation (Evensen, 2009). In the field of air pollution, the EnKF has been shown to be an
24 efficient method in optimizing concentrations. Further applications of the EnKF in improving dust
25 and ozone forecast skills through emission optimization have been reported (e.g., Constantinescu et
26 al., 2007; Eben et al., 2005; Lin et al., 2008; Tang et al., 2011).

27 In the present study, the EnKF was employed to assimilate ozone observations for the
28 corrections of NO_x emissions. The main purpose is to elucidate the performances of that method
29 during the cross-variable assimilation of O₃ observations. The sequential algorithm proposed by
30 Houtekamer and Mitchell (2001), as a variant of [the](#) EnKF, was adopted for its efficiency in

1 computation. The first step of the implementation was to perturb ozone concentrations, NOx
2 emissions and other key uncertainty sources of ozone modeling, i.e., photolysis rates and vertical
3 diffusion coefficients, as described by the following equations:

$$4 \quad \mathbf{x}'(\mathbf{i}) = \mathbf{x}^b + \boldsymbol{\zeta}(\mathbf{i}), \quad i = 1, 2, \dots, N \quad (1)$$

$$5 \quad \mathbf{e}'(\mathbf{i}) = \mathbf{e}^b + \boldsymbol{\varepsilon}(\mathbf{i}), \quad i = 1, 2, \dots, N \quad (2)$$

$$6 \quad \mathbf{q}'(\mathbf{i}) = \mathbf{q}^b + \boldsymbol{\phi}(\mathbf{i}), \quad i = 1, 2, \dots, N \quad (3)$$

7 where \mathbf{x} , \mathbf{e} , and \mathbf{q} are ozone concentrations, emissions, and other parameters (NO₂ photolysis rates
8 and vertical diffusion coefficients) respectively, and the superscript b represents their background
9 values in the model. The superscript ' represents the ensemble samples of these variables after
10 perturbing the background values by random samples of $\boldsymbol{\zeta}$, $\boldsymbol{\varepsilon}$, and $\boldsymbol{\phi}$. ~~$\boldsymbol{\zeta}(\mathbf{i})$, $\boldsymbol{\varepsilon}(\mathbf{i})$, and $\boldsymbol{\phi}(\mathbf{i})$ are the~~
11 ~~random samples~~, extracted from a normal distribution using the method proposed by Evensen
12 (1994). N is the ensemble size. The ensemble size (set as 50) was chosen based on several
13 sensitivity experiments of ozone data assimilation. The experiments were performed with the
14 same model domains and observation network as those employed in this study. The results
15 suggest that an ensemble of 50 members keeps good balance between computational efficiency
16 and assimilation performance of ~~the~~ ozone analysis.

17 In order to avoid filter divergence, the NO₂ photolysis rate and vertical diffusion coefficient
18 were perturbed by Gaussian distributed random noise, and the NOx emissions (to be updated
19 by the EnKF) were perturbed by a time-correlated Gaussian distributed random noise.
20 Estimating the uncertainty of the NOx emissions used for the modeling during the Beijing Olympic
21 Games was a hard task. The INTEX-B Asia inventory (Zhang et al., 2009) was estimated to contain
22 31% uncertainty in ~~the~~ NOx emission estimation. But the base year of this inventory is 2006. Another
23 key factor affecting the emission uncertainty is the temporary air pollution control measures during
24 the Beijing Olympic Games. The control measures were estimated to reduce the NOx emissions by
25 36% to 47% (Wang et al., 2009; 2010). This would induce large biases into the emission inventory
26 and lead to significant increase of the uncertainties of the emission inventory. Therefore, we
27 estimated the uncertainty of the NOx emissions to be 60 % of the first guess emission rates, about
28 twice the uncertainty in the INTEX-B Asia inventory. The uncertainties of vertical diffusion
29 coefficients in ozone modeling have been estimated by Beekmann and Derognat (2003), Hanna et al.

apple 16/5/6 9:32 AM

已删除: T

apple 16/5/6 9:32 AM

已删除: were

(1998) and Moore et al. (2001) and rang from 25% to 50%. We estimated the uncertainty of vertical diffusion coefficients to be 35% of the first guess values which are close to the average estimates of the above three estimates of Beekmann and Derognat (2003), Hanna et al. (1998) and Moore et al. (2001). Also with reference to the studies of Hanna et al. (1998) and Moore et al. (2001), the uncertainty of the modeled photolysis rates was estimated to be 30%. The uncertainty of the modeled O₃ concentrations at the initial time was estimated to be 50% after comparing the modeled O₃ concentrations with the O₃ observations. Based on the method suggested by Evensen (1994), the perturbations of the variables in three dimensions were implemented through adding a pseudo smooth random field. The random samples were Gaussian distributed with zero mean. The horizontal and vertical scales of initial error correlations could be effectively controlled using this method. The scales were set as 54 km in the horizontal and 3 model grids in the vertical (approximately 200 m) as in Tang et al. (2011).

Ensemble samples of the emissions, the vertical diffusion coefficients, the photolysis rates and the O₃ concentrations were used to derive ensemble forecasts of ozone. In order to achieve cross-variable adjustment for NO_x emissions, an extended state variable was defined as:

$$\mathbf{U}'(i) = \begin{bmatrix} \mathbf{x}'(i) \\ \mathbf{e}'(i) \end{bmatrix}, i = 1, 2, \dots, N \quad (4)$$

where $\mathbf{x}'(i)$ and $\mathbf{e}'(i)$ represent the ozone concentrations and the emissions after perturbations as in Eq. (1). Through the ensemble forecast $\mathbf{x}'(i)$ is strongly dependent on $\mathbf{e}'(i)$, which makes it convenient for estimating the correlation between x and e and for cross-variable adjustment of NO_x emissions. The background error covariance of the extended variable could be directly calculated from the ensemble forecast results during the simulation period:

$$\mathbf{P} = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{U}'(i) - \bar{\mathbf{U}}') (\mathbf{U}'(i) - \bar{\mathbf{U}}')^T \quad (5)$$

where $\bar{\mathbf{U}}'$ is the mean of the ensemble samples of the extended state variable and N is the ensemble size.

This algorithm treats the observations as random variables and perturbs them (Houtekamer and Mitchell, 1998). The perturbations on the O₃ observations and the perturbations on the emissions, the vertical diffusion coefficients, the photolysis rates and the O₃ concentrations described above would be helpful to prevent filter divergence of the EnKF in our data assimilation experiments. When ozone observations are available, they were perturbed according to the observation errors (Gaussian with

apple 16/5/6 9:34 AM
已删除: .

apple 16/5/6 9:34 AM
已删除: ing

apple 16/5/6 9:34 AM
已删除: ion

apple 16/5/6 9:35 AM
已删除: ion

apple 16/5/6 9:37 AM
已设置格式: 下标

apple 16/5/6 1:13 PM
已删除: to prevent filter divergence of the EnKF

apple 16/5/6 1:18 PM
已删除: .

1 mean zero and covariance \mathbf{R} , including both measurement errors and representativeness errors):

$$2 \mathbf{y}'(\mathbf{i}) = \mathbf{y} + \mathbf{Y}(\mathbf{i}), i = 1, 2, \dots, N \quad (6)$$

$$3 \mathbf{Y} \in N(0, \mathbf{R}).$$

4 As suggested by von Loon et al. (2000), the observation errors were assumed to be within 10% of the
5 original observation value and uncorrelated in time and space. It is worth noting that some other
6 variants of the EnKF (e.g., the ensemble square root filter (EnSRF) proposed by Whitaker and Hamill,
7 2002) do not need the perturbations on observations but can also provide accurate analyses.

8 Then the ensemble samples of the extended variables from the ensemble forecasts could be
9 updated through assimilating the ozone observations:

$$10 \mathbf{U}^a(\mathbf{i}) = \mathbf{U}'(\mathbf{i}) + \mathbf{K}(\mathbf{y}'(\mathbf{i}) - \mathbf{H}\mathbf{U}'(\mathbf{i})), i = 1, 2, \dots, N \quad (7)$$

$$11 \mathbf{K} = \mathbf{P}\mathbf{H}^T(\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R})^{-1} \quad (8)$$

12 where \mathbf{H} represents a linear operator mapping the extended state variable from model space to
13 observational space, and \mathbf{K} is the Kalman weight calculated based on the background error
14 covariance and the observation error covariance. $\mathbf{U}^a(\mathbf{i})$ is the updated ensemble sample of the
15 extended state variable and was used for the sequential ozone forecast. The updating of the ensemble
16 ensembles of the extended variables was conducted one time every 1 hour (1h), and the updated NOx
17 emissions were then used for the NO₂ forecast of the next hour. The ensemble mean of $\mathbf{U}^a(\mathbf{i})$ was
18 taken as the best estimate, after assimilating observations and was used as the output analysis state
19 (e.g. the blue dots in Figures 4 and 5) for comparisons with the observation and the simulation. To
20 reduce the spurious impact caused by the finite ensemble size, localization was performed for
21 analysis and only observations within a localization scale were used to update the NOx
22 emissions at a model grid. The localization scale was set as 45km following the configuration of
23 Tang et al. (2011).

24 (3) Surface observation network

25 We employed a regional surface air quality network over Beijing and its surrounding areas
26 during the 2008 Beijing Olympic Games including 17 stations established by the Beijing
27 Environment Monitoring Center and Chinese Academy of Science (Xin et al., 2010). Figure 1
28 displays the distributions of these stations and the non-industrial NOx emission rates of the
29 observation regions in the innermost third model domain with 9km horizontal resolution. As can be

apple 16/5/6 9:53 AM

已删除: ion

apple 16/5/6 9:57 AM

已删除: (e.g. the blue dots in Figures 4 and 5)

1 seen, 11 urban stations (CP, PEK, BY, IAP, YF, BD, CZ, QHD, SJZ, TS, TJ) are located in the urban
2 areas with high non-industrial NO_x emission rates, and the other 6 (LF, XH, XL, YJ, YuF, YLD) are
3 in the suburban areas with relatively low non-industrial NO_x emission rates. The network provides
4 observations of O₃ and NO₂ at the same temporal resolution as the model (i.e., 1h). The
5 measurements of NO₂ and O₃ were observed by online instruments (Model 42C& 42I NO-NO₂-NO_x
6 Analyzer and Model 49C&49I O₃ Analyzer from Thermo Scientific). The O₃ observations were
7 assimilated hourly into the model to adjust NO_x emissions. The direct comparison between the
8 simulated and observed NO₂ data often suffered from the representativeness errors of the NO₂
9 measurements. In this study, the stations close to the main roads with heavy traffic were not included
10 in order to reduce the influence of the representativeness errors of the NO₂ measurements.
11 Nevertheless, under certain resolutions (9km for example), the representativeness errors still persisted
12 in NO₂ measurements over urban areas. In order to independently validate the assimilation results,
13 three of the observation stations were withdrawn from the assimilation and were used for the
14 validation. NO₂ observations not used in the assimilation were also used to assess the impacts of the
15 cross-variable assimilation on the NO₂ forecasts.

16 3. Results

17 3.1 Real data assimilation experiment

18 The real data assimilation (RDA) experiment assimilated the surface ozone observations over Beijing
19 and surrounding areas to adjust the NO_x emissions over these areas in the NAQPMS. The experiment
20 was based on the study of Tang et al. (2011) in which the assimilation of real O₃ observations with
21 the EnKF was performed to correct NO_x emissions. The experiment focused on a two-week period
22 from 00:00 local time (LT) 9 August to 00:00 LT 23 August in 2008. The initial conditions of the
23 simulation were from a two-week spin-up model run. The initial conditions of ozone, NO_x emissions
24 and vertical diffusion parameters were perturbed at 19:00 LT on 8 August 2008 according to the
25 equations (1), (2) and (3) and were used to derive ensemble runs of NAQPMS. After 5h free
26 ensemble runs, the observed ozone data starting at 00:00 LT on 9 August was assimilated hourly into
27 the third model domain (displayed in Fig. 1) of NAQPMS to adjust the NO_x emissions. Adjusted
28 factors of the NO_x emissions were then used for the NO₂ forecast of the next hour. Both daytime and
29 nighttime observations were assimilated. We only adjusted the variables in the first three vertical

apple 16/5/6 10:07 AM

已删除: ed

apple 16/5/6 10:07 AM

已删除: to be

apple 16/5/6 1:28 PM

已删除: By considering possible large errors
in the modeling of vertical profiles of air
pollutants, w

1 layers near the surface, which could reduce the influence of the modeling errors of vertical mixing on
2 data assimilation. A free run of NAQPMS without data assimilation (NonDA) was also performed as
3 a reference run to validate the assimilation results of the RDA experiment.

4 Figure 2 compares the root mean square errors (RMSEs) of the 1 h ensemble mean forecast of
5 NO₂ at the 17 stations in the RDA experiment with the RMSEs in the NonDA experiment. The
6 RMSE of each site was calculated based on the hourly differences between NO₂ observation and the
7 ensemble mean forecast of NO₂ from 00:00 LT 9 August to 00:00 LT 23 August in 2008. The
8 number of valid observations used for each station is listed in Figure 2. The differences of the
9 RMSEs before and after DA were statistically significant over 11 stations (TJ, BY, YF, IAP, CP, XH,
10 CZ, PEK, QHD, SJZ and TS) at the 95% level of the t-test, while there were no statistically
11 significant differences of the RMSEs before and after DA over 6 stations (XL, YuF, YJ, YLD, LF
12 and BD). The RMSEs of the NO₂ forecasts in the free run of the model were dominated by the biases
13 which accounted for 55~90% (biases of the NO₂ forecasts divided by the RMSEs of the NO₂
14 forecasts) of the RMSEs. Biases noticed in simulations performed over urban sites are relatively
15 larger than those over the suburban ones. The free model run overestimated NO₂ concentrations at
16 most of the urban stations, while underestimated them at most of the suburban ones. The DA impacts
17 on the NO₂ forecast varied substantially from the suburban to the urban stations. At urban station
18 such as BD, PEK, CZ, QHD, SJZ, and TS, the RMSEs were reduced by 15%~36% after DA,
19 resulting in improvement of NO₂ forecasts in contrast to large increases, ranging from 56~239% of
20 the RMSEs at CP, BY, IAP, YF and TJ. At the suburban sites, the DA showed minor influence on
21 NO₂ forecasts and had no statistically significant impacts on the RMSEs over 5 of the 6 suburban
22 sites. Such minor DA impacts over the suburban sites could be explained firstly, by the fact that
23 emission rates of NO_x in the model were very low over suburban regions and that the simulation
24 without DA significantly underestimated the NO₂ concentrations. Even with the perturbations on the
25 NO_x emission, the ensemble spread was significantly weaker than the errors in the real case, and
26 thereby reduced the DA impacts of the EnKF. Secondly, in regards to the influences of the air
27 pollutants transport from urban regions, observed negative DA impacts over some urban areas may
28 have induced significant errors into the NO₂ forecasts. The above results suggest the adjustment of
29 the NO_x emissions by the ozone data assimilation has a mixed effect on the NO₂ forecast (i.e., weak
30 DA impacts over suburban sites, positive DA impacts over some urban sites and negative DA

apple 16/5/6 1:44 PM

已删除: (Bias/RMSE)

apple 16/5/6 10:11 AM

已删除: it

apple 16/5/6 10:12 AM

已删除: On the other hand

1 impacts over others). Nevertheless, the assimilation produced significant improvement of ozone
2 forecasts over all these sites, consistent with Tang et al. (2011).

3 Further investigations were conducted on the variation of such mixed effects of the data
4 assimilation on NO₂ forecasts over both the first week (from 00:00 LT 9 August to 00:00 LT 16
5 August in 2008) and the second week (from 00:00 LT 16 August to 00:00 LT 23 August in 2008). As
6 a result, the DA mixed effects were relatively stable during the Beijing Olympic Games. Figures 3
7 (a-c) display daily variation of the 1h NO₂ forecast RMSEs in the RDA experiment and the NonDA
8 experiment over the urban stations with positive DA impacts (CZ, PEK, QHD, SJZ, and TS), those
9 with negative DA impacts (BY, CP, IAP, TJ and YF) and the suburban stations (LF, XH, YLD, YJ
10 and YuF with weak DA impacts). At the suburban stations, the cross-variable DA also showed very
11 weak impacts on the NO₂ forecast in both the daytime and nighttime. At the urban stations with
12 positive DA impacts, the cross-variable assimilation presented consistent positive DA impacts in
13 daytime, nighttime and morning, with a 23% reduction of the RMSE_v during daytime and a 21%
14 reduction in night and morning.

15 At the urban sites with negative DA impacts, the performance of the DA was different between
16 daytime, nighttime and morning hours. Adjusting NO_x emissions improves the forecasts of NO₂
17 concentrations during most of the night and the morning time by reducing 7% of the RMSE_v in
18 contrast to the deterioration of the forecast in the daytime with 190% increase of the RMSE_v. This
19 finding suggests that the impacts of the cross-variable assimilation on the NO₂ forecast during
20 daytime are opposite to those in night and morning at these urban sites. Negative DA impacts mainly
21 occur in the daytime. As described by Tang et al. (2010b), daytime ozone is strongly nonlinearly
22 related to high NO_x emissions over urban areas (in particular over central Beijing), whereas
23 nighttime ozone is mainly controlled by the titration reaction of O₃-NO and the relationship between
24 the nighttime ozone and the NO_x emissions has weak nonlinearity. Due to the obvious discrepancy
25 between daytime ozone and nighttime ozone chemistry, further experiments were carried out to
26 elucidate the impact of the chemistry on the cross-variable assimilation. We describe these
27 experiments in the following section.

28 Another phenomenon observed in Figs. 3(a-b) is that the errors in NO₂ forecasts with the free
29 model run in night and morning were much higher than those in daytime. This might be due to the

apple 16/5/6 10:13 AM
已删除: as reported by

apple 16/5/6 2:24 PM
已删除: s

apple 16/5/6 10:15 AM
已删除: s

apple 16/5/6 10:15 AM
已删除: s

apple 16/5/6 10:15 AM
已删除: In clear, n

apple 16/5/6 10:17 AM
已删除: with

apple 16/5/6 10:20 AM
已删除: in following section

1 large uncertainties in modeling of nighttime boundary layer over urban regions (Kleczek et al., 2014).
2 Although the modeling of vertical diffusion was taken as a key uncertainty source in our data
3 assimilation, its uncertainty was not constrained by the data assimilation. Therefore, high errors still
4 occurred in the nighttime NO₂ forecasts after data assimilation, as shown in Figs. 3(a-b).

apple 16/5/6 10:21 AM

已删除: subsist

5 3.2 Ideal data assimilation experiment

6 An ideal experiment with a known true state provided a simple way to investigate the potential
7 consequences of some key inspected factors in a highly complex system. The true states are normally
8 the simulated observations generated by a model run or a data assimilation system (Timmermans et
9 al., 2015). In order to investigate the possible cause of observed mixed effects in RDA experiment,
10 this study employed a simplified box model including the main chemical processes of NAQPMS
11 (Xiang et al., 2010). Within the ideal data assimilation (IDA) experiments, the true state of ozone
12 concentrations and NO_x emissions were assumed to be known and were generated from the box
13 model run. The main purpose is to closely monitor the impacts of ozone chemistry on the
14 cross-variable assimilation method experimented in the RDA. However, this investigation did not
15 take into account complex transport processes and the removal processes were simulated by
16 multiplying the concentrations by removal coefficients. The experiments with the box model were
17 conducted on the IAP station where negative impact on NO₂ forecasts is observed in the RDA
18 experiment. Emission rates and meteorological parameters are from the inputs used by NAQPMS.

apple 16/5/6 10:25 AM

已删除: conducted

19 The IDA experiments focused on the negative DA impacts on the daytime NO₂ forecasts. The a
20 priori emission rates from NAQPMS and their corresponding O₃ concentrations modeled with the box
21 model were assumed to be the true state and were used for validation of the optimized emissions from
22 DA. Ensemble runs of the box model were initialized by the ensemble forecasts of the chemical
23 species of NAQPMS at 19:00 LT on 11 August 2008; NO_x emissions were perturbed to provide
24 ensemble samples of emissions during the following ensemble runs of the model. At 12:00 LT on 12
25 August 2008, the artificial O₃ observation was assimilated into the box model to adjust the NO_x
26 emissions. Artificial O₃ observations were generated through adding slight random errors to the true
27 state of O₃ concentrations. To be consistent with the RDA experiment, the random errors for perturbing
28 observations were also assumed to be within 10% of the true value. Three error scenarios for NO_x
29 emissions (10%, 30% and 50% underestimates) were assumed and separately applied to simulations of

apple 16/5/6 10:25 AM

已删除: Firstly, t

apple 16/5/6 2:26 PM

已删除: ions

1 the box model. In order to avoid dealing with complex model errors, the errors in NOx emissions were
2 assumed to be the only error sources of ozone modeling. For each error scenario, cross-variable
3 adjustment of the NOx emissions through assimilating the artificial O₃ observations with the EnKF
4 was conducted. Figures 4(a-c) show the O₃ concentrations and NOx emissions before and after DA,
5 with their ensemble samples before DA at 12:00 August 12, 2008.

6 Figure 4a presents the results under the first scenario with 10% underestimation of NOx
7 emissions (S1). The analyzed O₃ concentration and NOx emission after DA were close to their true
8 state, suggesting an improvement of the NOx emission estimation from the cross-variable assimilation.

9 Figure 4b shows the results under the second scenario with 30% underestimation of NOx emissions
10 (S2). The DA inefficiently reduced the error in NOx emission, since large errors (about 20%) still
11 persisted in the optimized NOx emission. Ensemble samples of O₃ concentrations shown in Fig.4b
12 were obtained from the ensemble runs of the box model that were derived from the ensemble samples
13 of NOx emissions (also shown in Fig.4b). The ensemble forecasts of O₃ concentrations presented high
14 nonlinear responses to the perturbations of NOx emissions. This suggests that the EnKF with Monte
15 Carlo simulations can predict the nonlinear evolutions of error statistics of the O₃ modeling. At the
16 analysis step, the ensemble samples of O₃ concentrations and NOx emissions were integrated into the
17 EnKF to calculate the background error covariance in Eq. (5). The linearized relationship between the
18 O₃ concentrations and the NOx emissions is presented in Fig. 4b. Noticeable discrepancies appear
19 between the nonlinear relationship denoted by the ensemble samples and the linearized relationship at
20 the analysis step. This significantly weakens the performance of the EnKF in the cross-variable
21 adjustment.

22 In the third scenario (S3) with NOx emissions underestimated by 50%, enhanced deterioration of
23 the NOx emission estimations was observed (Fig. 4c). The DA closely adjusted the simulated O₃
24 concentration to the true state, but induced an additional bias to the previously underestimated NOx
25 emission. Such negative DA impact on NOx emission estimation was similar to the phenomenon
26 observed in the daytime NO₂ forecast over some urban stations in the RDA experiment. From the
27 results in Fig. 4(a-c), the most plausible cause of the negative DA impact on NOx emission estimation
28 is the linearizing analysis of the EnKF used to deal with the cross-variable (O₃ to NOx emission) DA
29 problem of a highly nonlinearly chemical system. With a large bias in the a priori estimation of NOx
30 emissions, the cross-variable assimilation may induce enhancement of the bias in NOx emissions. The

apple 16/5/6 10:26 AM

已删除: Obviously, t

apple 16/5/6 10:26 AM

已删除: properly

apple 16/5/6 10:34 AM

已删除: o

apple 16/5/6 10:35 AM

已删除: in

apple 16/5/6 10:35 AM

已删除: ing

1 results of the three IDA experiments (i.e., positive DA impact under the first and second scenarios and
2 negative impact under the third scenario) confirm the mixed effects of the cross-variable assimilations
3 observed in the RDA experiments, and suggest a strong link between the mixed effects and the
4 linearization process at the analysis step of the EnKF applied to a strongly nonlinear chemical system.

5 In order to consider error scenarios with overestimations of NO_x emission, four idealized DA
6 experiments in which NO_x emission was assumed to be overestimated by 10%, 30%, 50% and 100%
7 respectively were performed. The results are shown in Fig. 5(a-d). In the first three experiments with
8 10%, 30% and 50% overestimates of the a priori NO_x emission, the DA worked well and significantly
9 reduced the biases of the emission. In the fourth experiment with the largest bias in the a priori
10 emission estimation, the DA enhanced the bias of the emission estimate in daytime. These mixed DA
11 effects under different biases of the a priori emission estimation are similar to those observed in the
12 previous idealized experiments conducted with underestimate scenarios in Fig. 4(a-c). Both
13 underestimate and overestimate scenarios confirm the mixed effects of the DA.

14 Note that above IDA experiments do not consider the complex model errors (e.g., errors in
15 boundary layer or transport modeling). In the real case, model errors exist, and the DA scheme needs
16 to properly quantify model uncertainties and deal with the nonlinearity between assimilated
17 observations and adjusted variables simultaneously. Model errors may affect the results of the real DA.
18 Thus, in order to investigate the DA performance of adjusting NO_x emissions under the presence of
19 biases on other factors (e.g., boundary layer or chemical reaction modeling), we assumed that the NO₂
20 photolysis rate was overestimated by 20% in the idealized box modeling, since the errors of the NO₂
21 photolysis rates were found to be among the top five uncertainty sources of ozone modeling over
22 Beijing and surrounding areas during the Beijing Olympic Games (Tang et al., 2010a).

23 In order to investigate the performance of the DA method when the bias of the NO₂ photolysis
24 rate was not considered in the DA, we were blind to the bias of the simulated NO₂ photolysis rate and
25 no perturbation was operated on it in the first DA experiment. The NO_x emission was adjusted in the
26 same way as the idealized experiments described above. Figure 6a displays the results of the DA
27 experiment under the error scenario of a 30% overestimate in the a priori NO_x emission. The DA
28 corrected the NO_x emission, but led to an underestimate of the emission. This over-correction of the
29 NO_x emission by the DA could be associated with the bias in the simulated NO₂ photolysis rate.
30 Therefore, in the second experiment (Fig. 6b), we considered the uncertainty of the simulated NO₂

apple 16/5/6 10:42 AM
已删除: over

apple 16/5/6 10:42 AM
已删除: ing

apple 16/5/6 10:42 AM
已删除: ion

apple 16/5/6 2:28 PM
已删除: ion

apple 16/5/6 10:49 AM
已删除: clearly

apple 16/5/6 11:19 AM
已删除: Firstly,

apple 16/5/6 11:23 AM
已删除:

apple 16/5/6 11:24 AM
已删除: ,

apple 16/5/6 11:24 AM
已删除:

apple 16/5/6 11:24 AM
已删除: so that n

apple 16/5/6 12:20 PM
已删除: above-

apple 16/5/6 12:20 PM
已删除: .

apple 16/5/6 12:21 PM
已删除: ion

apple 16/5/6 12:22 PM
已删除: ion

1 photolysis rate and perturbed the NO₂ photolysis rate in the DA. The error scenario was the same as in
2 the first experiment. Under that condition, the DA performed better than for the first experiment,
3 without over-correction of the NO_x emission. The results of above experiments suggest that
4 considering the model errors is crucial for the assimilation performance; otherwise the DA leads to
5 over-correction of the state variables. In order to deal with this issue, simulated NO₂ photolysis rates
6 and vertical diffusion coefficients (considered as the key uncertainty sources of the O₃ modeling) were
7 perturbed to account for their uncertainties in the real DA experiment. The third DA experiment was
8 quite similar to the second one, but we increased the bias of the a priori NO_x emission to a 100%
9 overestimate. The results are shown in Fig. 6c. Under the large bias in the a priori NO_x emission, the
10 DA deteriorated the NO_x emission estimate. In short, despite considering the influence of the model
11 errors, the limitations of the DA method in dealing with the large bias of a highly nonlinear system are
12 still persistent.

13 To investigate the DA impacts on the NO_x emissions in night and morning, variations of O₃
14 concentrations and NO_x emissions before and after DA and their ensemble samples before DA at 8:00
15 August 13, 2008 (morning time) are shown in Figs. 7(a-c). Similar results (not shown here) were
16 obtained for other night and morning times. In Figs. 7(a-c), different level errors (10%, 30% and 50%
17 underestimates) in NO_x emissions were significantly reduced through the cross-variable assimilation
18 with the EnKF. The ensemble forecasts of morning O₃ concentrations show near-linear responses to
19 the uncertainties (or perturbations) of NO_x emissions; the linearization of the EnKF at the analysis step
20 worked to correct the biases in NO_x emissions. The positive DA impacts on the NO_x emission
21 estimate in the IDA experiments in night and morning were consistent with the improvement of the
22 NO₂ forecasts after data assimilation in the RDA experiment. In comparison with the mixed effects of
23 the DA in daytime, the positive DA impacts in night and morning in both RDA and IDA experiments
24 indicate that the assimilation of O₃ observations with the EnKF might be useful in optimizing NO_x
25 emissions and NO₂ forecasts in night and morning. Furthermore, the ensemble forecasts of O₃
26 concentrations show strong nonlinear responses to the perturbations of NO_x emissions during daytime
27 in Figs. 4(a-c) but present near-linear responses in night and morning in Figs. 7(a-c). This suggests the
28 variability of nonlinearity in the chemical system leads to different DA impacts during different
29 periods of the day.

30 4. Conclusion and discussion

apple 16/5/6 12:23 PM
已删除: that of

apple 16/5/6 12:23 PM
已删除: to

apple 16/5/6 12:25 PM
已删除: to

apple 16/5/6 12:25 PM
已删除: is

apple 16/5/6 12:25 PM
已删除: ion

apple 16/5/6 12:26 PM
已删除: ion

apple 16/5/6 12:27 PM
已删除: in sight of

apple 16/5/6 12:28 PM
已删除: trends

apple 16/5/6 12:28 PM
已删除: ion

apple 16/5/6 12:29 PM
已删除: properly

apple 16/5/6 12:29 PM
已删除: ion

apple 16/5/6 12:30 PM
已删除: of

1 The impacts of cross-variable adjustment of NO_x emissions on NO₂ forecasts were investigated
2 through assimilating O₃ observations with a variant of the EnKF (proposed by Houtekamer and
3 Mitchell, 2001) over Beijing and surrounding areas during the 2008 Beijing Olympic Games. Both
4 real DA experiments with a 3-dimensional chemical transport model and ideal DA experiments with
5 a simplified box chemical model were performed.

6 The results of the data assimilation experiments revealed mixed effects of the cross-variable
7 assimilation with the EnKF. The DA worked in improving the NO₂ forecasts and optimizing the NO_x
8 emissions in night and morning when the uncertainties of O₃ concentrations were almost linearized to
9 those of the NO_x emissions. During daytime, the data assimilation resulted in positive DA impacts on
10 NO₂ forecasts over some urban sites, negative over other urban sites and weak impacts over suburban
11 sites. Through idealized DA experiments, the mixed effects were found to be strongly associated with
12 the difficulty in dealing with a highly nonlinear DA problem, especially when there are large model
13 biases. The results highlighted a critical limitation of the EnKF for chemical DA despite its strong
14 performance for improving tropospheric ozone forecasts (e.g., Tang et al., 2011).

15 The results suggest that bias correction is crucial for the application of the EnKF in highly
16 nonlinear chemical DA problems. Alternatively, avoiding the cross-variable DA between two
17 strong-nonlinearly related variables such as NO_x emissions and O₃ is also a possible way to
18 overcome this issue. For example, assimilating NO₂ observations directly to optimize NO_x emissions
19 might produce a better result than assimilating O₃ observations to improve the NO₂ forecasts and
20 NO_x emission estimates. Nevertheless, the strong nonlinearity issue remains a critical challenge in
21 chemical DA. To summarize, DA approaches that enable dealing with high nonlinearity in both
22 model evolution and analysis step are needed. Particle filters such as the nonlinear filter method (e.g.,
23 Moral et al., 1996; van Leeuwen, 2009; 2010) might have potential in this field if their limitation for
24 high dimensional system application (Stordal et al., 2011) can be overcome.

25 Acknowledgements

26 This study was funded by the CAS Strategic Priority Research Program (Grant No. XDB05030200),
27 the National Natural Science Foundation (Grant No. 41205091 and No. 41305111) and Commonwealth
28 Project in Ministry of Environmental Protection (Grant No. 201309011). We would like to thank the
29 two anonymous reviewers and the editor for their valuable comments that help improve this paper.

apple 16/5/6 12:30 PM

已删除: properly

apple 16/5/6 12:31 PM

已删除: under

apple 16/5/6 12:31 PM

已删除: the

apple 16/5/6 12:32 PM

已删除: ion

apple 16/5/6 12:33 PM

已删除: the

apple 16/5/6 12:33 PM

已删除: In sum

apple 16/5/6 12:36 PM

已删除: its

apple 16/5/6 12:53 PM

已删除: and

1 **References**

2 Beekmann, M., Derognat, C.: Monte Carlo uncertainty analysis of a regional-scale transport
3 chemistry model constrained by measurements from the Atmospheric Pollution Over the Paris
4 Area (ESQUIF) campaign, *J. Geophys. Res.* 108, doi:10.1029/2003JD003391, 2003.

5 Byun, D.W., Dennis, R.: Design artifacts in Eulerian air quality models: evaluation of the effects of
6 layer thickness and vertical profile correction on surface ozone concentrations, *Atmos. Environ.*
7 29, 105–126, 1995.

8 Carmichael, G., Chai, T., Sandu, A., Constantinescu, E., Daescu, D.: Predicting air quality:
9 Improvements through advanced methods to integrate models and measurements, *J. Comput. Phys.*
10 227, 3540–3571, 2008.

11 Constantinescu, E.M., Sandu, A., Chai, T.F. and Carmichael, G.R.: Ensemble-based chemical data
12 assimilation. II: Covariance localization, *Q. J. R. Meteorol. Soc.* 133: 1245–1256, 2007.

13 Eben, K., Jurus, P., Resler, J., Belda, M., Pelikan, E., Kruger, B. C., and Keder, J.: An ensemble
14 Kalman filter for short-term forecasting of tropospheric ozone concentrations, *Q. J. R. Meteorol.*
15 *Soc.* (2005), 131, 3313–3322, 2005.

16 Elbern, H., Strunk, A., Schmidt, H., Talagrand, O.: Emission rate and chemical state estimation by
17 4-dimensional variational inversion, *Atmos. Chem. Phys.* 7, 3749–3769, 2007.

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

20 Gaubert, B., Coman, A., Foret, G., Meleux, F., Ung, A., Rouil, L., Ionescu, A., Candau, Y.,
21 Beekmann, M.: Regional scale ozone data assimilation using an ensemble Kalman filter and the
22 CHIMERE chemical transport model, *Geosci. Model. Dev.* 7, 283–302, 2014.

23 Grell, G.A., Dudhia, J., Stauffer, D.R.: A description of the fifth-generation Penn State/NCAR
24 mesoscale model (MM5), NCAR Technical Note NCAR/TN-398+STR, 117 pp, 1994.

25 Hanea, R.G., Velders, G., Heemink, A.: Data assimilation of ground-level ozone in Europe with a
26 Kalman filter and chemistry transport model, *J. Geophys. Res.* 109, doi:10.1029/2003JD004283,
27 2004.

28 Hanna, S.R., Chang, J.C., Fernau, M.E.: Monte Carlo estimates of uncertainties in predictions by a
29 photochemical grid model (UAM-IV) due to uncertainties in input variables, *Atmos. Environ.*,
30 32(21), 3619-3628, 1998.

31 Hanna, S.R., Lu, Z.G., Frey, H.C., Wheeler, N., Vukovich, J., Arunachalam, S., Fernau, M., Hansen,
32 D.A.: Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V
33 photochemical grid model applied to the July 1995 OTAG domain, *Atmos. Environ.* 35, 891-903,
34 2001.

35 Houtekamer, P.L., Mitchell, H.L.: Data assimilation using an ensemble Kalman filter technique, *Mon.*
36 *Wea. Rev.* 126, 796–811, 1998.

37 Houtekamer, P.L., Mitchell, H.L.: A sequential ensemble Kalman filter for atmospheric data
38 assimilation, *Mon. Wea. Rev.* 129, 123–137, 2001.

- 1 Jimenez, P., Parra, R., and Baldasano, J. M.: Influence of initial and boundary conditions for ozone
2 modeling in very complex terrains: A case study in the northeastern Iberian Peninsula,
3 *Environmental Modelling & Software*, 22, 1294-1306, 2007.
- 4 Kleczek, M.A., Steeneveld, G., Holtslag, A.A.M.: Evaluation of the Weather Research and
5 Forecasting mesoscale model for GABLS3: impact of boundary-layer schemes, boundary
6 conditions and spin-up, *Boundary-Layer Meteorol.* 152, 213–243, 2014.
- 7 Koohkan, M.R., Bocquet, M., Roustan, Y., Kim, Y., Seigneur, C.: Estimation of volatile organic
8 compound emissions for Europe using data assimilation, *Atmos. Chem. Phys.* 13, 5887–5905,
9 2013.
- 10 Lin, C., J. Zhu, and Z. Wang: Model bias correction for dust storm forecast using ensemble Kalman
11 filter, *J. Geophys. Res.*, 113, D14306, doi:10.1029/2007JD009498, 2008.
- 12 Moore, G.E., Londergan, R.J.: Sampled Monte Carlo uncertainty analysis for photochemical grid
13 models, *Atmos. Environ.*, 35(28), 4863-4876, 2001.
- 14 Moral, P.D.: Nonlinear filtering: interacting particle solution, *Markov Processes and Related Fields*, 2
15 (4), 555–580, 1996.
- 16 Pagowski, M., Liu, Z., Grell, G.A., Hu, M., Lin, H.-C., Schwartz, C.S.: Implementation of aerosol
17 assimilation in Gridpoint Statistical Interpolation (v.3.2) and WRF-Chem (v.3.4.1), *Geosci. Model*
18 *Dev.* 7, 1621–1627, 2014.
- 19 Sandu, A., Chai, T.: Chemical Data Assimilation—An Overview, *Atmosphere* 2, 426-463, 2011.
- 20 Stordal, A.S., Karlsen, H.A., Nævdal, G., Skaug, H.J., Vallès, B.: Bridging the ensemble Kalman
21 filter and particle filters: the adaptive Gaussian mixture filter, *Comput. Geosci.* 15, 293–305,
22 doi:10.1007/s10596-010-9207-1, 2011.
- 23 Tang, X., Wang, Z.F., Zhu, J., Wu, Q.Z., Gbaguidi, A.: Preliminary application of Monte Carlo
24 uncertainty analysis in ozone simulation, *Clim. Environ. Res.* (in Chinese) 15 (5), 541-550, 2010a.
- 25 Tang, X., Wang, Z.F., Zhu, J., Gbaguidi, A., Wu, Q.Z., Li, J., Zhu, T.: Sensitivity of ozone to
26 precursor emissions in urban Beijing with a Monte Carlo scheme, *Atmos. Environ.* 44, 3833-3842,
27 2010b.
- 28 Tang, X., Zhu, J., Wang, Z.F., Gbaguidi, A.: Improvement of ozone forecast over Beijing based on
29 ensemble Kalman filter with simultaneous adjustment of initial conditions and emissions, *Atmos.*
30 *Chem. Phys.* 11, 12901-12916, 2011.
- 31 Tang, X., Zhu, J., Wang, Z.F., Wang, M., Gbaguidi, A., Li, J., Shao, M., Tang, G.Q., Ji, D.S.:
32 Inversion of CO emissions over Beijing and its surrounding areas with ensemble Kalman filter,
33 *Atmos. Environ.* 81, 676-686, 2013.
- 34 [Timmermans, R.M.A., Lahoz, W.A., Attie, J.-L., Peuch, V.-H., Curier, R.L., Edwards, D.P., Eskes,](#)
35 [H.J., Builtjes, P.J.H.: Observing System Simulation Experiments for air quality, *Atmos. Environ.*](#)
36 [115, 199-213, 2015.](#)
- 37 van Leeuwen P.J.: Particle filtering in geophysical systems, *Mon. Wea. Rev.* 137, 4089–4114, 2009.

1 van Leeuwen P.J.: Nonlinear data assimilation in geosciences: an extremely efficient particle filter, Q.
2 J. R. Meteorol. Soc. 136, 1991–1999, doi:10.1002/qj.699, 2010.

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

5 Wang, S.X., Zhao, M., Xing, J., Wu, Y., Zhou, Y., Lei, Y., He, K.B., Fu, L.X., Hao, J.M.:
6 Quantifying the Air Pollutants Emission Reduction during the 2008 Olympic Games in Beijing,
7 Environ. Sci. Technol., 44 (7), 2490–2496, 2010.

8 Wang, Y., Hao, J., McElroy, M. B., Munger, J. W., Ma, H., Chen, D., Nielsen, C. P.: Ozone air
9 quality during the 2008 Beijing Olympics: effectiveness of emission restrictions, Atmos. Chem.
10 Phys., 9, 5237–5251, 2009.

11 Wang, Z., Maeda, T., Hayashi, M., Hsiao, L.F., Liu, K.Y.: A nested air quality prediction modeling
12 system for urban and regional scales: application for high-ozone episode in Taiwan, Water Air and
13 Soil Pollution 130, 391-396, 2001.

14 Wang, Z.F., Xie, F.Y., Wang, X.Q., An, J.L., Zhu, J.: Development and application of Nested Air
15 Quality Prediction Modeling System, Chinese Journal of Atmospheric Sciences (in Chinese) 30,
16 778–790, 2006.

17 Wesely, M.L.: Parameterization of surface resistances to gaseous dry deposition in regional-scale
18 numerical models, Atmos. Environ. 23, 1293-1304, 1999.

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

21 Wu, L., Mallet, V., Bocquet, M., Sportisse, B.: A comparison study of data assimilation algorithms
22 for ozone forecasts, J. Geophys. Res. 113, D20310, doi:10.1029/2008JD009991, 2008.

23 Xiang, W.L., An, J.L., Wang, Z.F., Wu, Q.Z., Tang, X.: Application of CBM-Z chemical mechanism
24 during Beijing Olympics, Clim. Environ. Res. (in Chinese) 15 (5), 551-559, 2010.

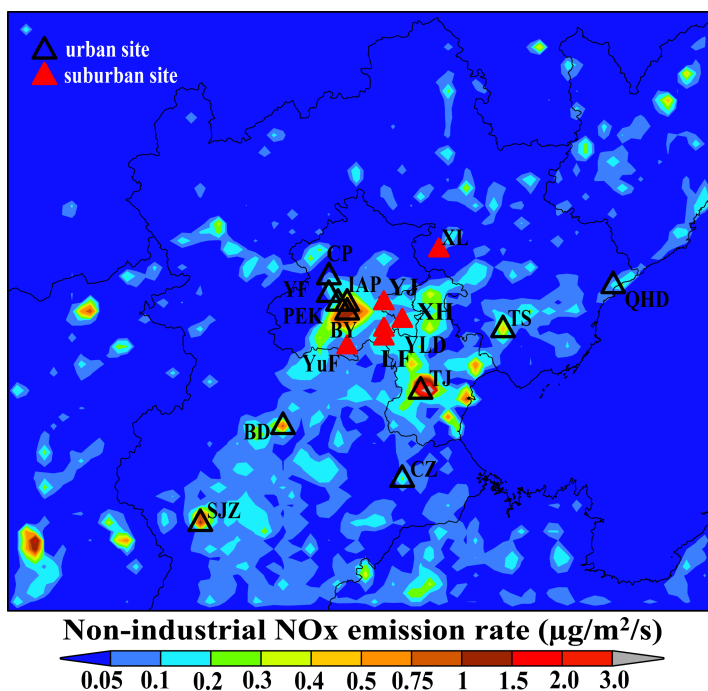
25 Xin, J.Y., Wang, Y.S., Tang, G.Q., Wang, L.L., Sun, Y., Wang, Y.H., Hu, B., Song, T., Ji, D.S.,
26 Wang, W.F., Li, L., Liu, G.R.: Variability and reduction of atmospheric pollutants in Beijing and
27 its surrounding area during the Beijing 2008 Olympic Games, Chinese Sci. Bull. (in Chinese) 55,
28 1937–1944, 2010.

29 Zaveri, R.A., Peters, L.K.: A new lumped structure photochemical mechanism for large-scale
30 applications, J. Geophys. Res. 104, 30387-30415, 1999.

31 Zhang, Q., Streets, D.G., Carmichael, G.R., He, K.B., Huo, H., Kannari, A., Klimont, Z., Park, I.S.,
32 Reddy, S., Fu, J.S., Chen, D., Duan, L., Lei, Y., Wang, L.T., Yao, Z.L.: Asian emissions in 2006
33 for the NASA INTEX-B mission, Atmos. Chem. Phys. 9, 5131-5153, 2009.

34 Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., Baklanov, A.: Real-time air quality forecasting,
35 part II: State of the science, current research needs, and future prospects, Atmos. Environ. 60,
36 656-676, 2012.

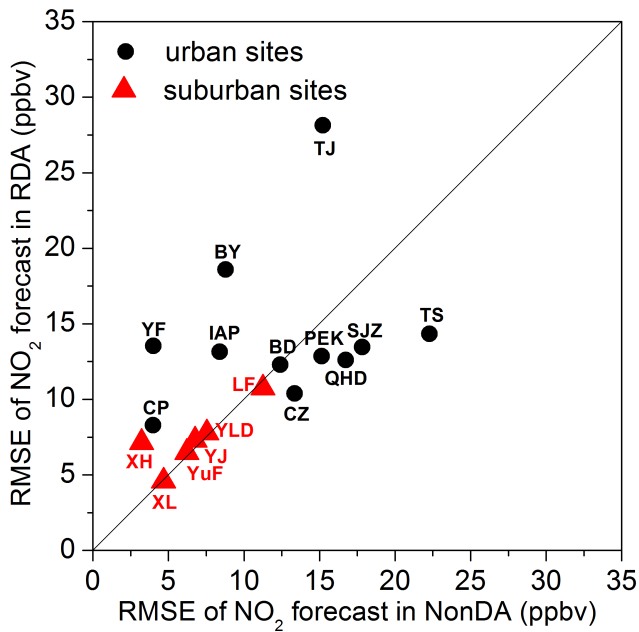
37



2

3 **Figure 1** Distribution of the observation stations and non-industrial NO_x emission rates in the third
 4 model domain (9km resolution) that covers Beijing and its surrounding areas. The non-industrial
 5 NO_x emission rates ($\mu\text{g}/\text{m}^2/\text{s}$) are divided into different bins (<0.05; 0.01-0.1; 0.1-0.2; 0.2-0.3;
 6 0.3-0.4; 0.4-0.5; 0.5-0.75; 0.75-1.0; 1.0-1.5; 1.5-2.0; 2.0-3.0) and represented by different shaded
 7 colors. The urban areas with high non-industrial NO_x emission rates are marked by the brown and
 8 red colors, and the suburban or rural areas with low non-industrial NO_x emission rates are marked by
 9 the green or blue colors. The 11 urban sites are denoted by the black triangles, and the 6 suburban
 10 stations are represented by the red triangles. The abbreviations of the station names are displayed
 11 close to the marks.

12

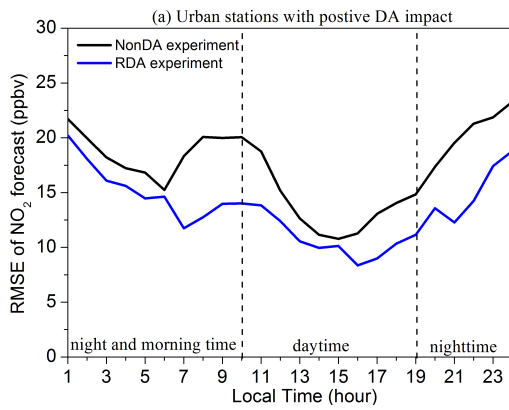


1

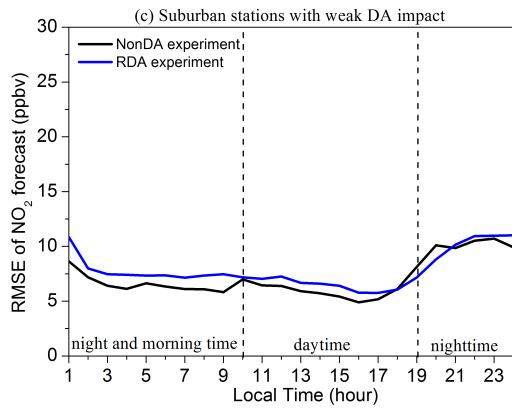
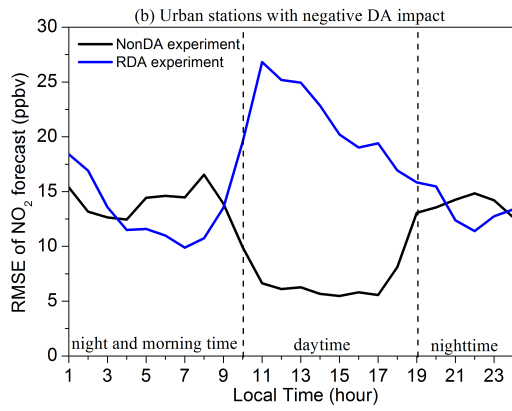
2 **Figure 2** Comparison of the root mean square error (RMSE) (ppbv) of 1h NO₂ forecasts at the 17
 3 stations of Beijing and its surrounding areas during the period of 00:00 local time (LT) 9 August to
 4 00:00 LT 23 August in 2008 in the real data assimilation (RDA) experiments and those in the
 5 reference (NonDA) experiment with a free run of the model. The comparisons at urban sites are
 6 denoted by the dots and those over suburban stations are represented by the triangles. The
 7 abbreviations of the station names are displayed close to the marks. The number of the valid
 8 observations used for the calculation is 336 at QHD, SJZ, TS, IAP, LF, YF and XH, and the
 9 numbers are 292, 226, 326, 317, 326, 320, 333, 321, 311, 323 at BD, PEK, BY, CZ, CP, TJ, XL, YJ,
 10 YLD and YuF respectively.

11

apple 16/5/6 2:49 PM
 已删除: s
 apple 16/5/6 12:37 PM
 已删除: s



12



1

2

3

4

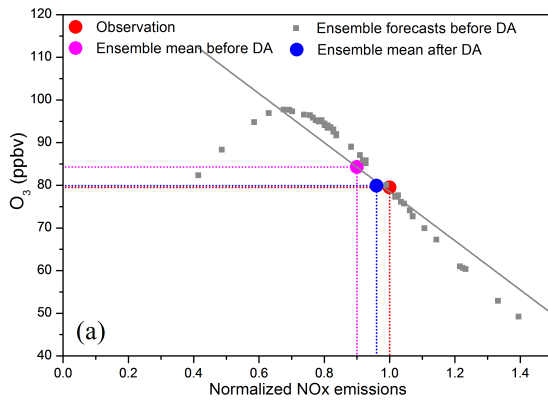
5

6

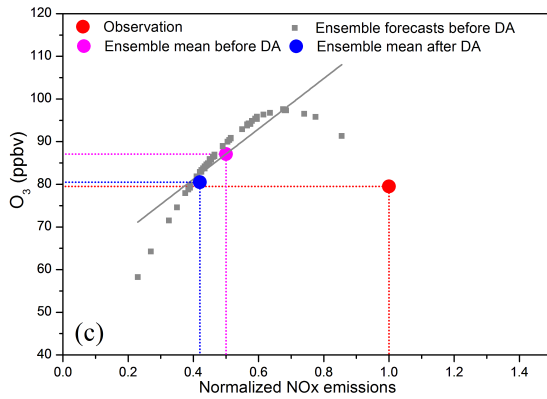
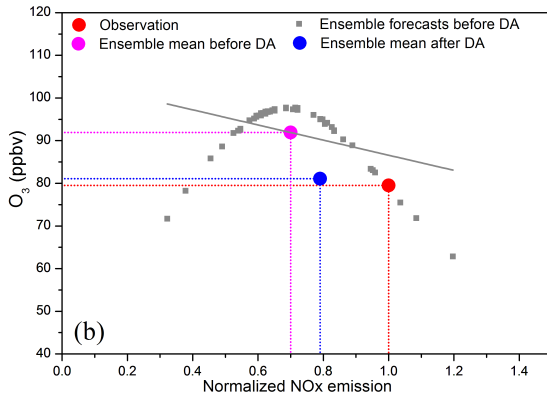
7

Figure 3 Daily variation of the 1h NO_2 forecast RMSE (ppbv) in the real data assimilation (RDA) experiments (blue line) and the reference (NonDA) experiment with a free run of the model (black line) over: (a) urban stations (CZ, PEK, QHD, SJZ, and TS) with positive DA impacts; (b) urban sites (BY, CP, IAP, TJ and YF) with negative DA impacts; (c) suburban stations (LF, XH, YLD, YJ and YuF) with weak DA impacts.

apple 16/5/6 12:37 PM
已删除: s



8



1

2

3

4

5

6

7

8

9

10

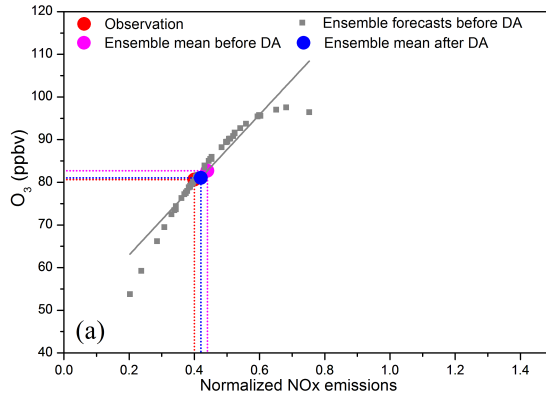
11

12

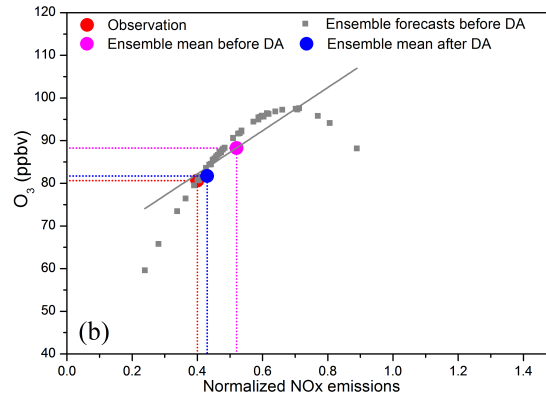
13

Figure 4 (a-c) O₃ concentrations (ppbv) and NO_x emissions (no unit, normalized by the true NO_x emission) before and after data assimilation (DA) and their ensemble samples before DA at 12:00 local time, on August 12, 2008 in the three ideal ozone data assimilation experiments with the prior NO_x emissions underestimated by 10% (a), 30% (b) and 50% (c) respectively. The grey squares denote the ensemble forecast O₃ concentrations corresponding to the perturbations of the NO_x emissions (ensemble forecasts before DA), and the magenta dot represents the result of the ensemble mean of the grey squares (ensemble mean before DA). The gray line represents a linear relationship calculated from the ensemble samples of O₃ concentrations and NO_x emissions. The red dot represents the true state of the NO_x emission and the observed O₃ concentration. The analyzed O₃ concentration and NO_x emission are denoted by the blue dot.

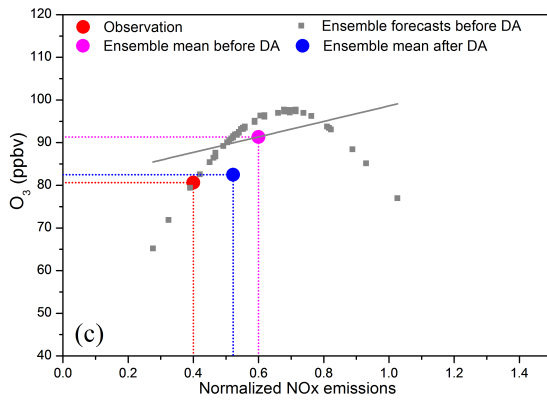
apple 16/5/6 12:46 PM
已删除: LT



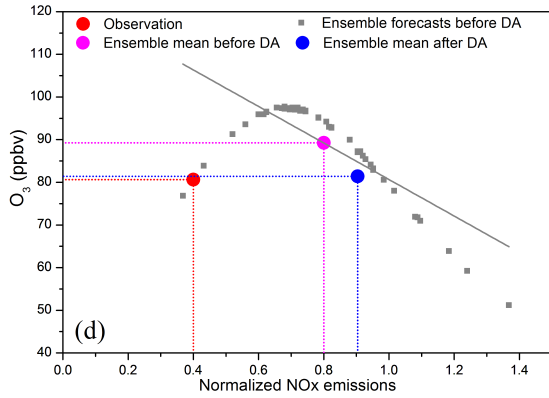
1



2

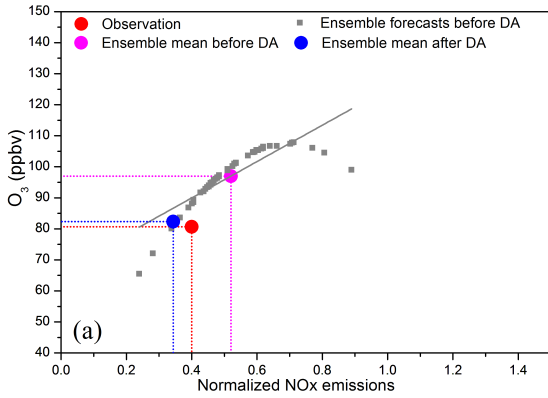


3



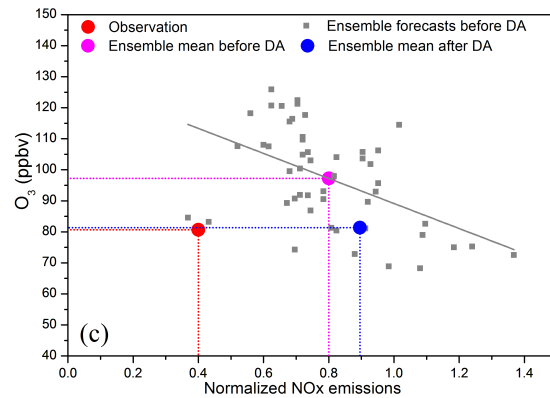
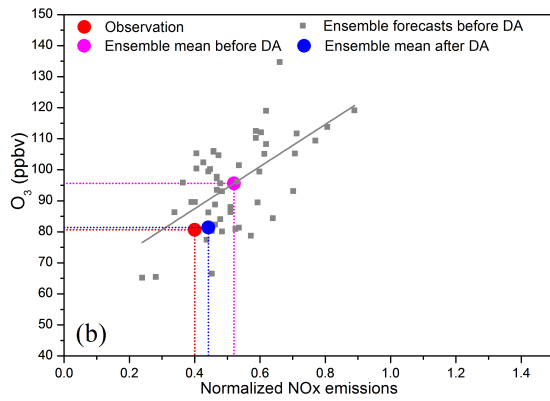
1
2
3
4
5
6
7
8
9

Figure 5 (a-d) O₃ concentrations (ppbv) and NO_x emissions (no unit, normalized by the true NO_x emission) before and after data assimilation (DA) and their ensemble samples before DA at 12:00 local time, on August 12, 2008 in the four idealized DA experiments. (a) DA experiment with 10% overestimate, in the a priori NO_x emission; (b) DA experiment with 30% overestimate, in the a priori NO_x emission; (c) DA experiment with 50% overestimate, in the a priori NO_x emission; (d) DA experiment with 100% overestimate, in the a priori NO_x emission. The magenta dot, the gray squares, the gray line, the red dot and the blue dot represent the same information as in Fig. 4.



10

- apple 16/5/6 3:41 PM
已设置格式: 字体:加粗
- apple 16/5/6 12:46 PM
已删除: LT
- apple 16/5/6 12:38 PM
已删除: ion
- apple 16/5/6 12:41 PM
已删除: estimation
- apple 16/5/6 12:39 PM
已删除: ion
- apple 16/5/6 12:41 PM
已删除: estimation
- apple 16/5/6 12:39 PM
已删除: ion
- apple 16/5/6 12:39 PM
已删除: ion



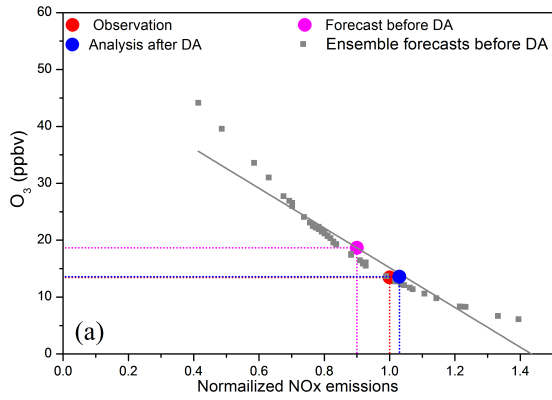
1

2

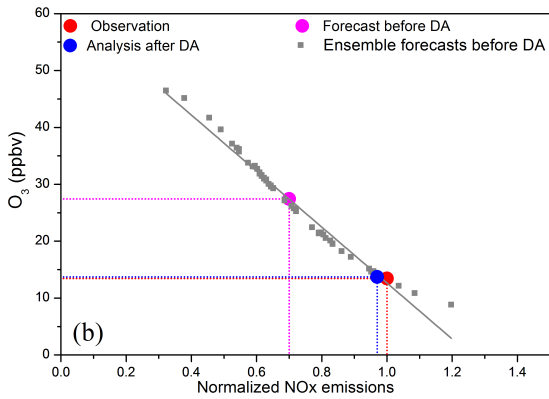
3 **Figure 6 (a-c)** O_3 concentrations (ppbv) and NO_x emissions (no unit, normalized by the true NO_x
 4 emission) before and after data assimilation (DA) and their ensemble samples before DA at 12:00
 5 local time, on August 12, 2008 in the three ideal DA experiments. The NO_2 photolysis rate is assumed
 6 to be overestimated by 20%. (a) The prior NO_x emission is overestimated by 30% and adjusted by
 7 the DA, but the uncertainty of the NO_2 photolysis rate is not included (there are no perturbations of
 8 the NO_2 photolysis rate) in the DA. (b) The same as the DA experiment in (a), but the uncertainty of
 9 the NO_2 photolysis rate is taken into account by perturbing it. (c) The same as the DA experiment in
 10 (b), but the bias in the prior NO_x emission is increased to 100%. The magenta dot, the gray squares,
 11 the gray line, the red dot and the blue dot represent the same information as in Fig. 4.

12

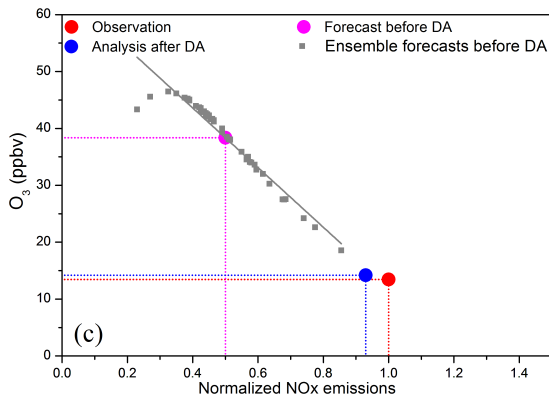
- apple 16/5/6 3:41 PM
已设置格式: 字体:加粗
- apple 16/5/6 12:42 PM
已删除: LT
- apple 16/5/6 12:42 PM
已删除: ing
- apple 16/5/6 12:43 PM
已删除:
- apple 16/5/6 12:43 PM
已删除: missed
- apple 16/5/6 12:43 PM
已删除: without perturbations on
- apple 16/5/6 12:43 PM
已删除: through



1



2



3

4 **Figure 7 (a-c)** O_3 concentrations (ppbv) and NOx emissions (no unit, normalized by the true NOx
 5 emission) before and after data assimilation (DA) and their ensemble samples before DA 08:00 [local](#)
 6 [time](#), on August 12, 2008 in the three ideal ozone data assimilation experiments with the prior NOx
 7 emissions underestimated by 10% (a), 30% (b) and 50% (c) respectively. The magenta dot, the gray
 8 squares, the gray line, the red dot and the blue dot represent the same information as [in Fig. 4](#).

apple 16/5/6 12:45 PM
 已删除: LT
 apple 16/5/6 12:45 PM
 已删除: s