Response to Referee #1 (ACPD-15-C11505-2016)

The manuscript of Tang et al. elucidates potentials and limits of the Ensemble Kalman filter (EnKF) for chemical data assimilation (DA) and cross-correction of reactive gases and emissions (O_3 and NOx) in the framework of air-quality forecasts. The first part of the paper provides an extended validation of the previous study of Tang et al. (2011) with a focus on NO_2 forecasts. The observed degradation of NO_2 forecasts at some locations motivates the authors to examine the behavior of EnKF in a simplified model setting. DA experiments in such a controlled environment permit to identify the likely cause of the degradation, i.e. strong non-linearities between the controlled NOx emissions and the observed/assimilated O_3 concentration.

First, I appreciated the fact that the authors further validated their previous study and published these new results, even if this partially question the method that was employed in Tang et al. (2011). The EnKF is a powerful and flexible DA algorithm but requires particular care when applied to correct unobserved variables or parameters in complex models. Studies that use EnKF to cross-correct unobserved variables or model parameters should more often try to provide in-depth validation of assimilation results, as the authors did here.

Second, I liked the methodology that was used by the authors, i.e. reproduce the observed behavior within a simplified model. This allowed a reasonable scientific explanation for the NO_2 degradation and, more in general, permitted to highlight the effect of strong non-linearities in chemical DA. As the authors also stated, this topic is often not well discussed in the chemistry DA literature and deserves further research. It would have been nicer if the authors could propose an algorithm to automatically detect strong non-linear regimes and at least avoid the analysis degradation within the EnKF. This limits a bit the impact of the study for the air-quality DA community.

The manuscript is concise and well structured, although multiple sentences should be rewritten in a better English. Hence, I recommend publication in ACP as a companion paper of Tang et al. (2011), after the following comments are considered.

Response: Great thanks to the reviewer for the valuable comments. Accordingly, the manuscript has been revised with improvement of the language. A point-by-point response to the review's comments is given as follows.

Specific comments:

1) Page 35694, line 27: 'the fast variability of the relationship between ozone concentrations and NOx emissions' is not very clear. The O_3-NOx emissions 'relationship' is a result of complex chemical reactions involving other species, radiation, temperature etc. Therefore, the 'relationship' is by definition not unique and saying that it varies 'fast' has not a precise scientific meaning. I suggest the authors to either remove this sentence or rephrase to make it scientifically sound.

Response: We agree. We have revised this sentence in the revised manuscript. "*The mixed effects observed in the cross-variable DA, i.e., positive DA impacts on NO_2 forecast over some urban sites, negative DA impacts over the other urban sites and weak DA impacts over suburban sites, were found to be strongly associated with the limitations of the EnKF in a strong nonlinear system.*"

2) Page 35695, line 14-15: '*... the divergence of the influences of the initial condition optimization ...*' is not clear. Do the authors mean that the initial condition has a weak influence on chemical forecasts?

Please rephrase. It is also worth reminding that chemical species have a large range of life-times and can depend on different processes (emissions, photolysis etc.). This implies that this statement is not very informative without saying to which species and which forecast's duration we refer to.

Response: We agree. This sentence has been corrected in the revised manuscript: "*One of the major challenges in CDA is that the impact of the initial conditions on the forecast of air pollutants such as ozone, decreases with simulation time (Gaubert et al., 2014; Jimenez et al., 2006).*"

3) Page 35695, line 24: I could not find demonstrations of improvements of ozone forecasts in Hanea et al (2004). Please remove the reference if not pertinent to the text.

Response: Thanks for this comment. We have removed this reference in the revised manuscript as suggested. "*… their applications have provided with significant improvement of ozone forecasts (Tang et al., 2011).*"

4) Page 35698, line 8-9: '*fully supports nonlinear evolution of a model...*' might lead to a wrong interpretation since the EnKF is based on Gaussian hypothesis and, as the authors show, it fails when non-linearities become too prominent. I guess the authors mean that EnKF can be implemented quite easily because the full non-linear model is employed during the ensemble forecast step. Please rephrase.

Response: Thanks for this comment and suggestion. We have rewritten this sentence in the revised manuscript as suggested. "*EnKF can directly calculate the background error covariance from the ensemble forecasts of the highly nonlinear model, which is very suitable for data assimilation in complex high-dimensional models (Carmichael et al., 2008).*"

5) Page 35698, line 16: see comment 3 for Hanea et al. 2004, Lin et al. 2008 is missing in the list of references and van Loon et al. 2000 does not demonstrate improved forecast skills for ozone (this concerns also page 35709, line 1). I suggest the authors to provide a more complete list of references that demonstrate the successful improvement of reactive gases forecasts through DA. Otherwise the authors should acknowledge that more research is needed in this regard.

Response: We have provided two new references to support the statement for improving forecasts through DA in the revised manuscript. The reference for Lin et al. (2008) is also added to the list of references. "*Further applications of the EnKF in improving forecast skills of dust and ozone through emission optimization have been reported (Constantinescu et al., 2007; Eben et al., 2005; Lin et al., 2008; Tang et al., 2011).*"

6) Page 35699, line 9-10: are the samples extracted from a normal distribution? Can the authors also precise the criteria that have been used to choose an ensemble of 50 members. How were the assimilation performances evaluated?

Response: Thanks for these comments! The samples are extracted from a normal distribution using the method proposed by Evensen (1994). The ensemble size is chosen after several sensitivity tests for the O_3 data assimilation (DA). Figure 1 displays the root mean square errors (RMSEs) of analyzed O_3 concentrations in the O_3 DA experiments with the EnKF under different ensemble members. The model domains and observation network is the same as in this study. As can be seen, the RMSEs in the tests with the ensemble size less than 30 are significantly higher than those in the other tests, which may

be related to the spurious correlation induced by the small ensemble size. The RMSEs decreased with the increase of the ensemble size. However, due to the linear increase of the computational cost with the ensemble member, we took 50 members as a relatively good balance between computational efficiency and assimilation performance of the O_3 analysis. Furthermore, previous studies (e.g., Carmichael et al., 2008; Constantinescu et al., 2007) applying EnKF in chemical transport model took this ensemble size for ozone data assimilation. Due to space limit in the Journal, the sensitivity result presented in Fig.1 is not showed in the revised manuscript. However, we have clarified these issues in the revised manuscript.

"*... random samples are extracted from a normal distribution using the method proposed by Evensen (1994). The ensemble size (set as 50) is chosen based on several sensitivity experiments of ozone data assimilation. The experiments are performed with the same model domains and observation network as those employed in this study. The results suggest that an ensemble of 50 members keeps good balance between computational efficiency and assimilation performance of ozone analysis. Due to space limit in the Journal, this result is not presented in the revised manuscript, but provided in the supplement material.*"

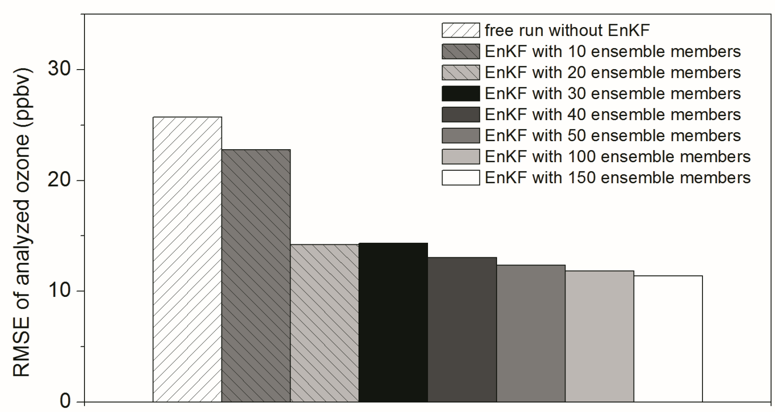

Figure 1. Root mean square errors (RMSEs) of the analyzed ozone concentrations over Beijing and its surrounding areas in the ozone data assimilation experiments that are conducted with ensemble Kalman filter (EnKF) for different ensemble members.

7) Page 35701, Sec. Data assimilation algorithm: Are the authors using some inflation and/or localization technique for the EnKF? If yes please describe it briefly in the text.

Response: Thanks. We have added some sentences to clarify this issue. "*In order to avoid filter divergence, the NO_2 photolysis rate and vertical diffusion coefficient are perturbed by Gaussian distributed random noise, and the NOx emissions (to be updated by the EnKF) are perturbed by a time-correlated Gaussian distributed random noise. A detailed description for these perturbations has been given by Tang et al. (2011). Moreover, to reduce the spurious impact caused by the finite ensemble size, localization is performed for analysis and only observations within a localization scale are used to update the NOx emissions at a model grid. The localization scale is set as 45km following the configuration of Tang et al. (2011).*"

8) Page 35702, Sec. Surface observation network: the authors should report some information about the measurement method and instrumental uncertainties of the employed in-situ NO_2 measurements. The issue of representativity of NO_2 measurements for the model grid should also be briefly discussed. Compared to O_3, NO_2 measurements in urban environment can be largely affected by local pollution and be not representative of a 10km model pixel. For example, are some of the used NO_2 sites exposed to heavy road traffic?

Response: Thanks! We have added some sentences in the revised manuscript with regard to this issue. "*The measurements of NO_2 and O_3 are observed by online instruments (Model 42C& 42I NO-NO_2-NOx Analyzer and Model 49C&49I O_3 Analyzer from Thermo Scientific). The direct comparison between the simulated and observed NO_2 data often suffered from the representativeness errors of the NO_2 measurements. In this study, the stations close to the main roads with heavy traffic are not included in order to reduce the influence of the representativeness errors of the NO_2 measurements. Nevertheless, under certain resolutions (9km for example), the representativeness errors still persist in NO_2 measurements over urban areas. Increasing model resolution therefore reduces the uncertainties.*"

9) Page 35703, lines 10-13: It is not very clear to me why small emissions of NOx cannot undergo *'significant'* changes with DA. If the variance of the ensemble is set as a percentage of the NOx emissions themselves, the DA correction is expected to be also proportional to the emissions and, therefore, locally significant. This should be the case unless the O_3 is not sensitive to NOx in low NOx regimes. Can the authors provide more insights on this? Looking at the corresponding O_3 ensemble spread and EnKF correction at suburban sites could also help.

Response: Thanks for raising this issue. According to your comment, Fig. 2 shows the hourly NO_2 concentrations from the observation, the simulation without DA and the simulation with DA at the suburban site (Yongledian as an example). Figure 3 displays the ensemble spread of the hourly NO_2 forecasts at YLD in the data assimilation experiment using the EnKF. As can be seen in Fig. 2, the simulation without DA significantly underestimated the NO_2 concentrations at YLD, which is probably caused by the very low emission rates of NOx in the model. Under this situation, the perturbations on the NOx emissions still resulted in a relative small ensemble spread (shown in Fig. 3) in the DA using the EnKF, and the ensemble spread is significantly smaller than the errors in the real case. This would lead to weak corrections to the NOx emission over the suburban areas. On the other hand, the DA brought out significant errors of the NO_2 forecast at YLD during some period (especially on August 10 and 16), which may be induced by some wrong adjustments of the NOx emission over urban areas. Therefore, the minor changes of the RMSEs after DA are mainly caused by the above two reasons. We have clarified this issue in the revised manuscript. "*Over the suburban sites, the DA shows minor influence on NO_2 forecasts and has no statistically significant impacts on the RMSEs over 5 of the 6 suburban sites. There are mainly two reasons for the minor DA impacts over the suburban sites.*"

Firstly, the emission rates of NOx in the model were very low over suburban regions and the simulation without DA significantly underestimated the NO_2 concentrations. Even with the perturbations on the NOx emission the ensemble spread is significantly smaller than the errors in the real case, which would weak the DA impacts of the EnKF. On the other hand, within the influences of the air pollutants transported from urban regions, the wrong adjustments of the NOx emission observed over some urban areas may induce significant errors into the NO_2 forecasts over some periods."

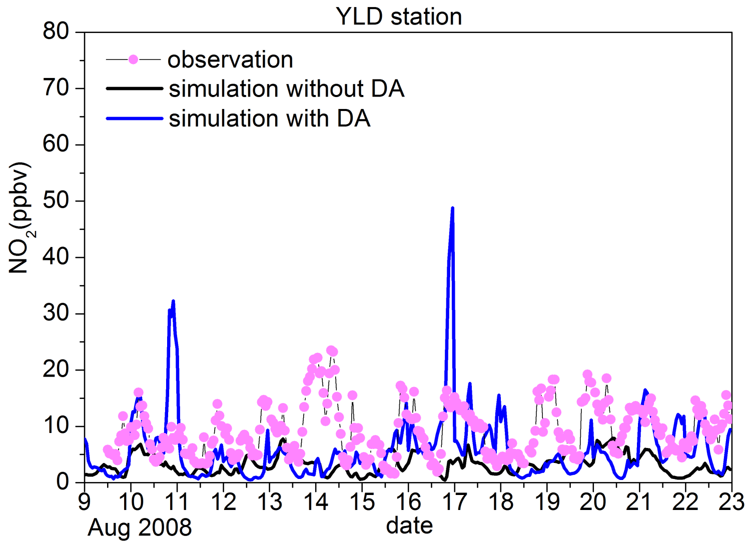

Figure 2. Time series of the hourly NO_2 concentrations obtained from the observation (magenta dots), the simulation without data assimilation (DA) (black line) and the simulation with DA (blue line) at the suburban site of Yongledian (YLD).

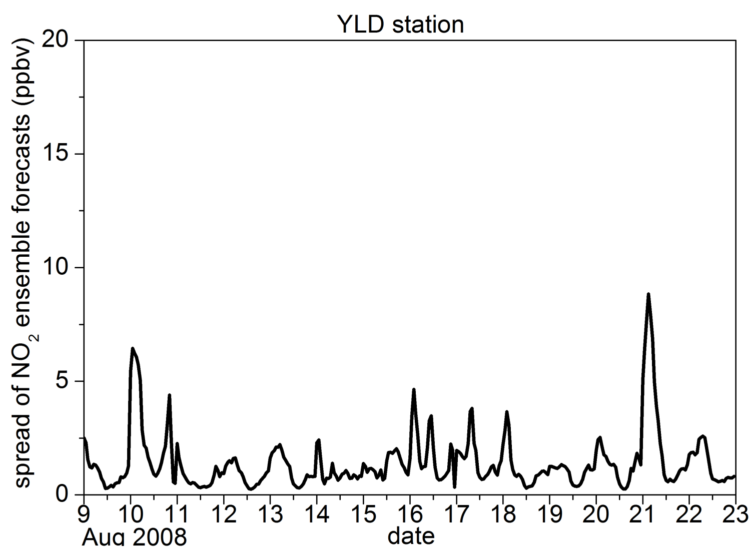

Figure 3. Ensemble spread of the hourly NO_2 ensemble forecasts at the suburban station of Yongledian (YLD) in the data assimilation with the EnKF.

10) Page 35704, lines 23-24: larger errors of modeled NO_2 in ppb units can also just be related to larger values of NO_2 concentration, which normally occurs in early morning and late evening, when NO_2 photo dissociation is not active and the boundary layer is shallow. Is the percentage error showing the same behavior?

Response: Thanks for your comments. According to your comment, we provide Fig. 4 showing daily variation of the root mean square errors (RMSEs) and the relative errors of the NO_2 forecast in the free run of model over the urban stations (BY, CP, IAP, TJ and YF) with negative DA impacts. The relative errors present a similar daily variation as the RMSEs. The relative errors of the NO_2 forecasts in night and morning are also much higher than those during the daytime.

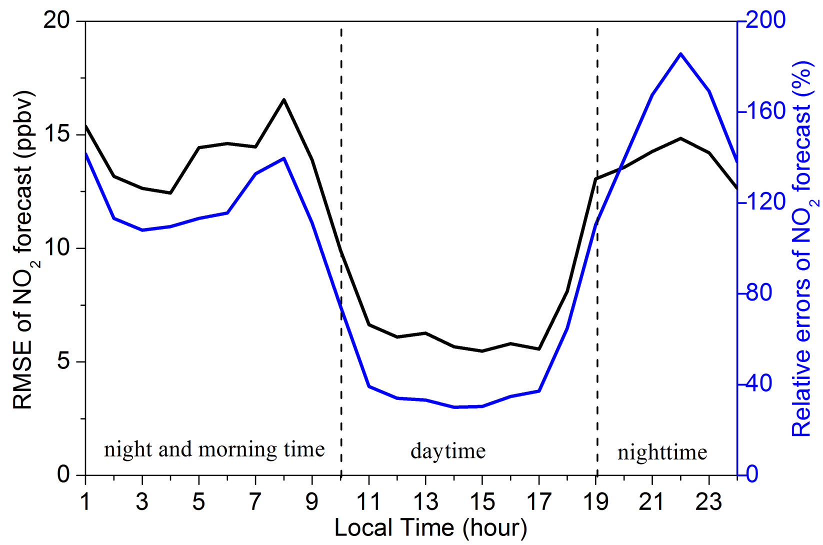

Figure 4. Daily variation of the NO_2 forecast errors in the free run of model at the urban stations (BY, CP, IAP, TJ and YF) with negative DA impacts. The black line represents the root mean square errors (RMSEs) and the blue line is the relative errors (percentage error).

11) Page 35708, lines 6-7: ' ... *except for dealing with the non-linear relationship* ...'. this part of the sentence is not clear, please clarify what you mean by 'except' and rephrase in case

Response: Thanks. We have rewritten this part in the revised manuscript. "*In the real case, model errors exist, and the DA scheme needs simultaneously to properly quantify model uncertainties and deal with the nonlinear problem between assimilated observations and adjusted variables.*"

12) Page 35708, line 23: *'rapid variations'* see comment n. 1

Response: We have revised this sentence in the manuscript. "*Through the idealized DA experiments, the mixed effect was found to be strongly associated with the difficulty in dealing with the highly nonlinear DA problem especially under the presence of large model biases.*"

13) Page 35709, lines 17-20: The largest non-linearities arise from the chemical mechanism. Please explain why changing the model resolution would affect the non-linear behavior of the system and therefore the results of DA.

Response: Thanks for raising this issue. Thunis et al. (2015) reported some (minor) impacts of the spatial model resolution on the non-linearity behavior of the regional air quality modeling.

However, the affect is still not very clear, and we have removed this part in the revised manuscript.

14) Page 35709, lines 19-20: *'Except for inversely estimating emissions* ... ' I cannot understand the exception. Doesn't this study show that the estimation of NOx emissions assimilating O_3 observation deals with chemical non-linearities? Please clarify this sentence.
Response: We have removed this sentence in the revised manuscript.

Technical corrections:
Please consider proof-reading the manuscript by an English native speaker. I provide here some suggestions for some sentences that should be ameliorated.

1) Page 35694, lines 2-3 '... *that has been validated as an efficient approach for improving ozone forecast*' -> ' that has been used in the companion study to improve ozone forecasts over Bejing and surrounding areas'

Response: We have revised this in the revised manuscript as suggested. "*... that has been used in the companion study to improve ozone forecasts over Beijing and surrounding areas.*"

2) page 35694, line 16: remove *'as a further investigation'*

Response: We have removed this in the revised manuscript as suggested.

3) page 35695, line 7: '... *that closely integrates ... is recognized* ... ' > ' ... integrates ... and is recognized ...'

Response: We have revised this as suggested. "*Chemical data assimilation (CDA) integrates models and observations to better represent the chemical state of atmosphere and is recognized as a technique for improving the simulations and forecasts of air pollutants such as ozone and aerosols*"

4) page 35700, lines 8-9: remove *'provide various ... initial estimations) and* '

Response: We have removed this in the revised manuscript as suggested.

5) page 35704, line 8: ' *varies from the day to the night and the morning*' > ' is different between day-time, night-time and morning hours'

Response: We have revised this as suggested. "*Over the urban sites with negative DA impacts, the performance of the data assimilation is different between day-time, night-time and morning hours.*"

6) page 35706, lines 11-13: '... *are combined by EnKF to produce linear correlations between them during the calculation of* ...' does not sound very well in English, please rephrase

Response: We have revised this sentence in the revised manuscript. "*At the analysis step, samples of the O_3 concentrations and NOx emissions are integrated into the EnKF to calculate the background error covariance in Eq. (5).*"

7) page 35706, lines 24-28: same as above
Response: We have revised this sentence in the revised manuscript. "*As can be seen from the results in Fig. 4(a-c), the most plausible cause of the negative DA impact on the NOx emission*

estimation is the linearizing analysis of the EnKF in dealing with the cross-variable (O_3 to NOx emission) DA problem of a highly nonlinearly chemical system."

References

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

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

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

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

Gaubert, B., Coman, A., Foret, G., Meleux, F., Ung, A., Rouil, L., Ionescu, A., Candau, Y., and Beekmann, M.: Regional scale ozone data assimilation using an ensemble Kalman filter and the CHIMERE chemical transport model, Geosci. Model Dev., 7, 283–302, doi:10.5194/gmd-7-283-2014, 2014.

Lin, C., J. Zhu, and Z. Wang: Model bias correction for dust storm forecast using ensemble Kalman filter, J. Geophys. Res., 113, D14306, doi:10.1029/2007JD009498, 2008.

Jimenez, P., Parra, R., and Baldasano, J. M.: Influence of initial and boundary conditions for ozone modeling in very complex terrains: A case study in the northeastern Iberian Peninsula, Environmental Modelling & Software, 22, 1294-1306, 2007.

Tang, X., Zhu, J., Wang, Z.F., 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, 2011.

Thunis, P., Clappier, A., Pisoni, E., Degraeuwe, B.: Quantification of non-linearities as a function of time averaging in regional air quality modeling applications, Atmos. Environ., 103, 263-275, 2015.