RESPONSE TO REVIEWERS

Ms. Ref. No.: Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2021-428.

Title: Assessment of strict autumn-winter emission controls on air quality in the Beijing-Tianjin-Hebei region

Journal: Atmos. Chem. Phys. Discuss.

Reviewer comments are in blue. Responses are in black and include line numbers consistent with the updated manuscript with changes tracked.

We wish to thank the reviewers for their thoughtful and extensive assessment of our paper. Their comments and suggestions have ultimately improved the quality and rigour of the underlying science. We also make very minor editorial changes to the manuscript. These changes are also tracked.

Response to RC#1:

The manuscript evaluated emission controls on air quality in the BTH region with observation products and CTM simulations. The topic is of great importance to environmental policymakers. However, several issues should be addressed properly in the manuscript for publication in ACP.

General comments for the modification

1. The objectives and motivations do not seem clear in the manuscript (e.g., Estimating the emissions reduction for mitigation measures, reproducing the AW2017 case, or determining contributions of parameters to the air quality). The authors had better make your objectives and motivations articulate and explicit in the manuscript. The manuscript is also lacking in the implication of this work to report to the readers or scientific community. Accordingly, the results should be are congruent with the objectives and the implication of the study.

We have reworded text in the Abstract, Introduction and Conclusions to better articulate the motivation and objectives, and implications of this work (lines 32-38, lines 45-52, lines 116-121, and lines 524-526).

2. Second, validation is crucial for evaluating the effects of emission controls on air quality. The manuscript did not discuss the validation of NO_x, SO_2, CO precursors for the AW2017 simulation, although there are some comparisons of particulate matters. Thus, the authors need to compare the simulated gaseous species with observations (e.g., in-situ ground or satellite data).

We agree with the reviewer that validation is crucial. This is why we evaluate the surface monitoring network before using these data. We also evaluate model performance of gaseous air pollutants, not just $PM_{2.5}$ for AW2016 (lines 340-378). We already use the surface observations of trace gases to derive scale factors for emissions in AW2017 (lines 319-337), so it would not be an independent validation if we again compared the model to these same network observations.

3. Lastly, the authors had better reorganize the manuscript to strengthen the methodology (i.e., Adding a method section).

Thank you for your suggestion. We have reorganized the text to include a dedicated Data and methods section (lines 123-192).

Specific comments for the modification

1. Lines 142-144: The manuscript did not mention what was utilized for NO_2 observation during the APHH campaign. Was it different with the Chemiluminescence detection system? If both measurements are not based on the same principle, the differences can be caused by instrumental sensitivity (as the authors mentioned). However, the different local sources at both network sites are also an important issue that cannot be ignored. There is ~3 km distance between them. The authors need to discuss it.

We now state that NO_2 was measured using Teledyne T500~U CAPS analyser during APHH campaign (lines 149-150). This type of instrument provides a direct NO_2 measurement, which does not involve catalytic conversion or reagents which introduce measurement artifacts inherent in traditional chemiluminescence instruments. We agree that the different local sources at both network sites are also an important issue, but the difference between the two is small (~5%). Regardless, we reword the text to mention these two issues (lines 212-221).

2. Lines 210-213: It is an important part of the methodology. The authors used scale factors of 1.5 for NO_x, 2.4 for CO, and 2.1-6.8 for SO_2 to conduct the CTM simulation for the AW2016 case. Were the spatially same factor applied? Also, no matter which (ground or satellite) observation data is used for the emission estimations, there are two crucial issues of i) nonlinearity between emissions and concentration of a species (e.g., NO_2) and ii) transfer between adjacent grid cells in the calculation. The authors need to clarify how the scale factors are derived (i.e., procedure). Furthermore, in particular, for the scale factor of NO_x, the authors need to explain how to treat the relation between observed NO_2 and the NO_x emissions (usually emitted as NO).

We now elaborate on our description of the approach we use to scale MEIC for AW2016 so that our approach and intention are clearer. That is, that we apply a single scaling factor to the whole domain NO_x and CO emissions and spatially variable scaling factors to just seven SO_2 emissions grids to address biases in the MEIC emissions for AW2016 (lines 280-285, lines 293-303). There are many papers that use NO_2 observations to derive NO_x emissions with the model providing the conversion factor, at least as early as Martin et al. (2003). We now quantify the non-linearities that result from using observed surface concentrations to address biases in emissions (lines 355-358, 365-367, 371-374, Figure S2) and assess the size of these against the size of the biases in the MEIC inventory (lines 359, 367-369, 374-375, Figure S1).

3. Lines 219-222: It is well known that CO is a final product of NMVOC oxidations in many textbooks. So, it is not easy to agree that modeled CO is relatively unaffected by NMVOC emissions. The authors need to explain some reasons in the manuscript in terms of the lifetime of NMVOCs and their chemical evolution during long-range transport. The enhanced levels of CO would occur in other remote areas other than BTH regions through long-range transport.

We have added text to discuss this result further by also acknowledging that the limited impact of scaling NMVOCs emissions on CO could in part be because of model errors in the oxidation of NMVOCs leading to formation of CO (lines 307-312).

4. Lines 231-235: I think there is a more important reason for the inconsistency. That is interference (e.g., HNO_3 and PANs) in the NO_2 chemiluminescence detection instrument equipped with a molybdenum converter, which converts NO_2 to NO. Here, the molybdenum

converter also oxides NO_z (≈ HNO_3 + PANs) to NO under typically operational temperature 300 – 350 °C (refer to Winer et al., 1974 and Dunlea et al., 2007). Dunlea et al. reported the interference in the chemiluminescence detection accounting for up to 50% of ambient NO_2 Considering this issue, the correlation between the simulated and observed NO_2 would be better. In other words, the data points of NO_2 in Fig. 4 would shift to the left, and the intercept would decrease. The authors had better discuss and/or reanalyze it.

We have added discussion about the role of this interference in affecting the agreement between the model and observations. We also use the model results to calculate NO_2 concentrations that include contributions from NO_x reservoir compounds known to thermally decompose to NO_2 and assess the influence this has on comparison of the model to the measurements (lines 348-352).

5. Lines 242-243 & Figure 4: Although the scale factors of 2-7 were applied to grid cells somewhere (which was not specified in the manuscript, but probably around Shanxi province) in the MEIC SO_2 emission, the SO_2 concentrations were still significantly under-predicted. The under-predicted SO_2 concentrations can influence SO_2 and PM_5 in the BTH areas via the atmospheric chemical and physical processes (e.g., secondary aerosol formation and the transport to the BTH) because SO_2 has ~ 5 days lifetime. Accordingly, the estimation of the emission changes for the AW2017 simulation is probably hampered by low simulated SO_2. The authors had better discuss how to treat this issue in your estimate. Also, the authors need to present the results for the AW2017 case, similar to Fig. 4.

We have reworded the text to clarify the approach we use to address biases in MEIC emissions of SO_2 (lines 301-303). We have also included Figure S1 that compares the modelled SO_2 using the original MEIC emissions to the observations to illustrate that there are a few points that make the largest contribution to the model underestimate. The grids that are scaled to address the emissions underestimate are also indicated.

6. Lines 270-272 and Fig. 5: It is not easy to agree that the errors in the boundary layer dynamics are related to the overestimation of nitrate alone. The issue should also apply to sulfate and others. Therefore, the errors in the boundary layer would not be the main reason for the overestimation. It is reasonable to discuss the overestimation of nitrate in terms of understanding like a relationship between SO_2 and sulfate (as the authors mentioned). However, as shown in Figs. 4 and 5, the modeled NO_2 concentration (a precursor of nitrate) is underestimated while nitrate is overestimated. It is a logical contradiction. Thus, the authors need to re-examine the overestimation of nitrate, considering the 4th comment pointed out by this reviewer.

We have edited the text to better reflect the possible causes in the model overestimate of aerosol nitrate identified by Miao et al. (2020) as due to uncertainties in formation and processing of aerosol nitrate (lines 401-404).

7. Lines 285-298: The authors need to discuss a clear description of how to estimate the emissions fluxes for AW2017. It is also required to explain how to treat the nonlinearity between emissions and concentration in the estimation.

We now reword the text so that it is clear how we estimated AW2017 emissions fluxes (lines 319-337). We quantify and discuss the size of the non-linear responses (lines 355-358, 365-367, 371-374, Figure S2) and how these compare to the size of discrepancies between the model with the original MEIC emissions and the observations (lines 359, 367-369, 374-375, Figure S1).

8. Lines 330-340: Zhang et al. (2010) mentioned "NH_3 emission varied greatly from city to city from HS1617 (AW2016 in this study) to HS1718 (AW2017). In some cities, NH_3 emissions were largely reduced, such as in Beijing (6.4%), Taiyuan (33%), and Zhengzhou (19.6%), while the NH_3 emissions showed increases in some other cities, such as Tianjin (5.0%), Shijiazhuang (0.2%) and Jinan (35.2%)". These variations are not marginal. Also, some studies reported that the SO_2 and NO_2 emissions have a decreasing trend while atmospheric NH_3 experienced a significant increasing trend (Xia et al., 2016; Ge et al., 2019). If NH_3 emissions increase in your simulation for the AW2017 case, what change would be expected in the concentration of $PM_{2.5}$?

The increase in NH_3 abundances in Xia et al. (2016) and Ge et al. (2019) is due to decline in acidic aerosols resulting from decline in precursor emissions of SO_2 and NO_x and so cannot be used to assess changes in emissions. We now report the increased modelled NH_3 concentrations in BTH, and discuss in the manuscript the challenges in relating NH_3 concentrations to emissions, due to this dependence on abundance of acidic aerosols (lines 464-468). We now also reference Xia et al. (2016) and Ge et al. (2019) in this discussion (line 466). We identified and corrected a minor mistake in quoting the numbers from Zhang et al. (2021) (line 472).

Minor comments for the modification

1. 1: Provide information on the number of data in Figure 1.

Added to Figures 1 and 2.

2. Line 142: "<10%". Clarify it, as for example, 0-10%, ~10%, or ~%.

Updated to ~5% (line 212)

3. The authors mentioned several grid points, for example, "seven grid squares" (Lines 213), "2 grid points" (Line 242), "13 grids" (Line 305), and "14 model grids" (Line 318). Clarify or leave out because readers cannot find out such information in the manuscript.

We have added Figure S1 to show the locations of the "seven grid squares" and the "two grid points". We have also updated Figure 7 (lines 820-827) to show the "thirteen Beijing grids" and the "thirteen [14 was incorrect] model grids" where $PM_{2.5}$ declines in model, but increases in observations. We also change values, such as "2", to "two" throughout to ensure consistency.

Response to RC#2:

This manuscript evaluates the impact of emission reduction policies in China on changes in winter PM2.5 concentrations. In particular, the authors evaluate the emission reduction and PM2.5 concentration change in the BTH area during 2017AW using the chemical transport model and various observational data. I expect that such an attempt will greatly help policymakers assess the impact of policy implementation not only in China but also in many polluted regions. However, in order for this manuscript to be published on ACP, various issues must be resolved. Some detailed comments are below:

- Lines 27-28: The observed PM2.5 concentration was significantly reduced compared to the expected target values (15%). It seems better to emphasize this part in comparison with observations rather than models.

Agreed. We have updated the text (lines 28-30).

- Lines 30-31: It is difficult to easily determine the effect of the emission reduction policy and the inter-annual variability of the meteorological field, respectively. Among the PM2.5 concentration reductions derived from the model (20%), if the effect of the emission reduction policy is 8%, does the remaining 12% mean the effect of the meteorological field?

We have reworded the text in the Abstract (lines 33-36) and provided additional explanatory text to our approach and a new Figure 9 so that it is clearer how we distinguish the effect of emission reduction policy and interannual variability in meteorology (lines 481-487 and line 833).

- Lines 65-76: If MEE implemented a strong emission reduction policy during 2017AW, is there any data on the amount of emission reduction estimated by MEE? If any, it should be compared and discussed with the emission reductions assumed in this study.

Despite our best efforts, we were not able to find public data from the MEE or other government departments or local authorities on their estimate of emissions reductions. This is why we used changes in observed surface air pollution to estimate changes in emissions. We now compare relative emission reductions in our study with those provided by Zhang et al. (2021) (in their Figure 6) using a regional emission inventory that is not publicly available (lines 459-461).

- Lines 89-90: Authors should address the biases of bottom-up emissions inventories in more detail in the introduction part.

We now reference the emission inventory uncertainty estimates summarized in the review by Li et al. (2017) (lines 112-114).

- Lines 108-120: It is recommended to organize the station information in a table and mention only those that require detailed explanation.

Done (Table S1, lines 134-135, lines 145-146 and lines 149-153).

- Line 168: Different from the values shown in Figure 3. Are the emission control and target areas different?

Thanks for pointing out this inconsistency. The values in Figure 3 reflect an earlier version of this figure and have been updated. The emission control and target areas are the same.

- Line 199: Is emission reduction necessary during the two-month spin-up period before the reduction policy is implemented? Emissions during the spin-up period should also be mentioned.

We do not include emissions reductions during the two-month spin-up period, as there was no action plan in place during these two months. We now clarify this (lines 293-295).

- Lines 211-213: The authors scaled up MEIC emissions from observations and models for the 2016 AW period. However, the authors do not specifically mention the criteria for increasing NOx and CO emissions by 1.5 and 2.4 times, respectively (Line 203 does not provide such information). Although the authors uniformly increased NOx and CO, do the differences in model and observation appear uniformly across the entire domain? Limitations on this should be mentioned. Moreover, although SO2 concentrations would be underestimated in most regions, the increase in emissions was applied to only 7 model grids. Also, how is the number 2.1-6.8 times calculated?

We elaborate on our approach of scaling MEIC for AW2016 to enhance clarity (lines 280-283, lines 293-303). We have also added Figure S1 that compares modelled surface concentrations obtained with the original MEIC inventory for AW2016 to the observations. Our use of a uniform scaling factor for NO_x and CO is based on the reasonable spatial correlation between the model and observations, whereas the underestimate in SO_2 is mainly driven by a few points (Figure S1).

- Lines 236, 244: The model underestimates despite the increase in emissions. The authors mention the positive bias of the monitoring network as one of the causes. Although there is a bias of two points compared to APHH in Fig. 2, it is not clear whether the value can represent the bias in the entire domain. Even SO2 is inconsistent. In addition, APHH and both observation points are located in urban, so they are greatly affected by mobile sources. Therefore, we cannot be sure that the difference between APHH and the two points represents the bias of CNEMN and BJENM.

We agree with the reviewer. We now acknowledge that the difference between APHH and the two points may not be representative of a bias in CNEMN and BJENM measurements across the whole domain (line 371).

- Lines 285-298: Are 2017AW emissions scaled up in the same way as 2016AW? It is not clearly described in the manuscript.

Emissions in AW2016 and AW2017 are scaled using the same monitoring network, but following slightly different approaches. We now include greater detail on how we scale the original MEIC emissions to address biases in emissions in AW2016 (lines 280-283, line 293-303) and how we scale the resultant AW2016 emissions to get AW2017 emissions using the observed relative changes in air pollutants formed from these precursor emissions (lines 319-337).

- Line 289 and Lines 315-319: The authors scaled up emissions outside of BTH, but did not change emissions outside of the area shown in Figure 3. Is there any clear reason for that? Are MEIC emissions underestimated only in BTH and its vicinity but are assumed to be similar elsewhere? As the authors mentioned in the manuscript, PM2.5 affects different regions through long-range transport, so the impact of fixed emissions outside the domain should be mentioned.

We do scale emissions beyond BTH to account for the possible influence of uncertainties in neighbouring emissions on modelling air pollution in BTH. We already state in the manuscript "The scale factors we apply to the whole domain in Fig. 3 are uniform values of 1.5 for NO_x emissions and 2.4 for CO emissions" (lines 295-298). We do not scale emissions beyond the domain shown in Figure 3, as this domain is already considerably larger than BTH. The impacts of fixed emissions outside the wider domain should be small (lines 326-328).

- Line 340: Since the influence of the meteorological field is also an important part of this study, it is recommended to present the changed PM2.5 concentration field as a 2-D map due to the interannual variability of the meteorological field. This will allow us to evaluate in more detail the impact of interannual variability in regional meteorological fields.

We have added a Figure 9 (line 833) and accompanying text (lines 483-487).

- Lines 346-348: The authors should highlight the significant differences between Zhang et al. (2021) and this study.

Updated (lines 494-496).

- Line 345: Zhang et al. (2019) evaluated the difference between the meteorological fields in December 2017 and December 2016. Therefore, the authors should also compare for the same period (December). This is because, even in winter, the meteorological field can have large fluctuations from month to month.

We now ensure this comparison is consistent by also providing values from the model for the same time periods (Decembers 2016 and 2017) as the Zhang et al. (2019) paper (lines 490-492).

- Line 364: "PM2.5 in BTH decreased by 28% from 103 μg m-3 to 75 μg m-3 in the control period relative to the previous year". However, in the abstract, it is described as "PM2.5 in BTH in autumn-winter 2017/2018 relative to the previous year is 27%, declining from 103 to 75 μg m-3". The same value should be used for the same content.

Apologies. This is a typo. We have corrected this to 27% throughout (lines 29, 436, 513).

- Lines 368-370: The authors used observations and models to evaluate emission reductions in the BTH region. However, there may be large errors in the scale-up assumed by the authors. Although there are difficulties to be derived only with limited data, the authors should evaluate the reliability of the presented numbers and make important comments about the effect of uncertainty.

We agree that the scaling has uncertainties. We acknowledge the limitations of our approach (lines 319-321). We now also discuss that this is a substantial improvement on just using the default bottom-up emission inventories without any constraints from surface measurements (Figure S1). We now quantify the non-linearities that result from using observed surface concentrations to address biases in emissions (lines 355-358, 365-367, 371-374, Figure S2) and assess the size of these against the size of the biases in the MEIC inventory (lines 359, 367-369, 374-375, Figure S1).

- Line 371: How 8% is derived should be presented in more detail in Session 4.

We now provide details on this (lines 481-490), we also corrected the minor mistake of "8%" to "9%".

- Line 737: Does "n" mean NMB?

Thank you for spotting this omission. We have updated the caption to state that “n” is the number of points in the comparison (lines 831).

References:

Ge, B. Z., Xu, X. B., Ma, Z. Q., Pan, X. L., Wang, Z., Lin, W. L., Ouyang, B., Xu, D. H., Lee, J., Zheng, M., Ji, D. S., Sun, Y. L., Dong, H. B., Squires, F. A., Fu, P. Q., and Wang, Z. F.: Role of Ammonia on the Feedback Between AWC and Inorganic Aerosol Formation During Heavy Pollution in the North China Plain, Earth and Space Science, 6, 1675-1693, 10.1029/2019ea000799, 2019.

Li, M., Liu, H., Geng, G. N., Hong, C. P., Liu, F., Song, Y., Tong, D., Zheng, B., Cui, H. Y., Man, H. Y., Zhang, Q., and He, K. B.: Anthropogenic emission inventories in China: a review, Natl. Sci. Rev., 4, 834-866, https://doi.org/10.1093/nsr/nwx150, 2017.

Martin, R. V., Jacob, D. J., Chance, K., Kurosu, T. P., Palmer, P. I., and Evans, M. J.: Global inventory of nitrogen oxide emissions constrained by space-based observations of NO_2 columns, Journal of Geophysical Research-Atmospheres, 108, 10.1029/2003jd003453, 2003.

Xia, Y. M., Zhao, Y., and Nielsen, C. P.: Benefits of of China's efforts in gaseous pollutant control indicated by the bottom-up emissions and satellite observations 2000-2014, Atmospheric Environment, 136, 43-53, 10.1016/j.atmosenv.2016.04.013, 2016.

Zhang, Y., Chen, X., Yu, S., Wang, L., Li, Z., Li, M., Liu, W., Li, P., Rosenfeld, D., and Seinfeld, J. H.: City-level air quality improvement in the Beijing-Tianjin-Hebei region from 2016/17 to 2017/18 heating seasons: Attributions and process analysis, Environ. Pollut., 274, 116523, https://doi.org/10.1016/j.envpol.2021.116523, 2021.