Response to Reviewer #1's comments on He et al. 2019 Atmospheric Physics and Chemistry manuscript

We thank the anonymous reviewer for thoroughly reading our manuscript and providing helpful comments and suggestions, which lead a significant improvement of the manuscript. The detailed responses to comments are listed below (text in italic is the reviewer's comments, and the blue text highlighted is our response):
**All line numbers are from the clean version of revised manuscript.*
***Please see the revised manuscript in the attachment.*

Summary

This paper discusses aircraft and surface observations of important air pollutant species over China. Such measurements are critically important for understanding the poor air quality in China, which adversely affects the health of 100's of millions of people. However, this paper does not adequately describe the measurements, and the results are poorly discussed, in some cases reaching contradictory conclusions. In at least one instance, the results are compromised by a fundamental math error. The results of the photochemical grid modeling appear to be extremely poor. I recommend that the paper be rejected, and the authors encouraged to resubmit the paper, if the issues detailed below can be adequately addressed.

Response: We appreciate the comments from the anonymous reviewer. We added more information about the measurements to the supplementary material and detailed discussion on the results. The manuscript has been revised according to these comments as discussed below.

1) Much more detailed information regarding the aircraft measurements must be given, perhaps in the Supplementary Material. The single paragraph in Section 2.1 is inadequate. Questions that immediately occurred to me include: Were the instruments zeroed and calibrated in flight? How can ozone be measured at 1 Hz frequency when the specifications of the ozone instrument state that the response time is 20 seconds with a 10 second lag time? How can formaldehyde be measured at 1 Hz frequency when the specifications of that instrument state that the time resolution is 90 sec with a delay time ~300 sec. How were the lag times of the various instruments synchronized with the GPS system? What was the aircraft air speed (i.e., 1 Hz measurement frequency corresponds to what spatial resolution)? What was the rate of ascent and descent (this is important given the evident time resolution of some of the instruments)? Given the disparate time resolution of the instruments, how were comparable spatial average concentrations calculated? What are the accuracies and precisions of the 1 Hz measurements? (Please provide a table listing these instruments parameters and explanations or references of how these parameters were determined.) Can any evidence be provided to demonstrate these tabulated accuracy and precision parameters are realistic (e.g., the results of in-flight instrument comparisons with another aircraft)? A full description of these and all such instrument issues must be provided for the interested reader.

Response: As suggested by the reviewer, we have included the Table S1 in the supplementary material showing the aircraft instruments and their sample frequencies (average time), precisions and accuracies. In the last sentence of the second paragraph of Section 2.1 in the revised

manuscript, we have clarified that while the data acquisition system was logging the data at a frequency of 1 Hz, trace gas analyzers had their sample average times longer than 1 second as shown in Table S1.

Table S1. UMD Aircraft Instrumentation

Variable	**Method**	**Sample Frequency**	**Precision/Accuracy***
Position	GPS	1 s	Horizontal: ~1 / ±2.5 m Vertical: ~1 / ±3.75 m
Meteorology (T, RH, P, 2-D Wind)	Cloud water inertial probe (CWIP): Hotwire, advanced heading reference system, 5-hole gust probe	1 s	T: 0.2 / ±0.5 ℃ P: 2.6 hPa / ±0.25% of FS RH: 1 / ±2% WS: 0.5 / ±1.0 m/s WD: 5 / ±10 °
Greenhouse gas CO_2/CH_4/CO/H_2O	Cavity Ring Down Spectroscopy Picarro Model G2401-m	2 s	CO_2: 0.02 / ±0.1 ppm CH_4: 0.2 / ±1 ppb CO: 4.2 / ±10 ppb
Ozone (O_3)	UV Absorption	10 s	1 ppb / ±1%
Sulfur Dioxide (SO_2)	Pulsed Fluorescence	10 s	0.1 ppb / ±3%
Nitrogen Dioxide (NO_2)	Cavity enhanced absorption spectroscopy, Los Gatos	1 s	0.05 ppb / ±5%
Reactive Nitrogen (NO-NOy)	Chemiluminescence	10 s	0.05 ppb / ±3%
Aerosol Scattering, b_{scat} (450, 550, 700 nm)	Nephelometer	1 s	±5x10^{-7} m^{-1} / ±5%
Aerosol Absorption, b_{abs} (565 nm)	Particle Soot Absorption Photometer (PSAP)	1 min	±5x10^{-7} m^{-1} / ±5%
Black Carbon (370, 470, 520, 590, 660, 880, 950 nm)	Aethalometer	2 min	0.05 µg/m^3 / ±5%
VOCs	Grab Canisters/GC-FID	5-6 / flight	Species dependent
Formaldehyde	Wet chemistry and fluorescence detection	90 seconds	0.1 ppb / ±5%

**The precisions are from the instrument specifications provided by the manufacturers while the accuracies are estimated from the uncertainties in calibration standards and mass flow controllers that are used to control flow rates of zero air and calibration gas.*

Depending on wind speed and direction (i.e., head wind or tail wind), the Y12 aircraft travels at a ground speed of 50-80 m/s. So for an instrument with a 10-second average time, the spatial resolution would be 500-800 m. The average rate of ascent and descent was about 2.5-3 m/s.

All the instruments used in this study had been tested and used for airborne measurements during the past two decades in the United States and China (Dickerson and Delany, 1988; Doddridge et al., 1998; Marufu et al., 2004; Taubman et al., 2004a; Taubman et al., 2004b; Taubman et al.,

2006; Dickerson et al., 2007; Hains et al., 2008; Krotkov et al., 2008; He et al., 2012; Li et al., 2012; Brent et al., 2013; He et al., 2013; He et al., 2014; He et al., 2016; Ren et al., 2018; Salmon et al., 2018). Due to the limited space and power in the small research aircraft employed in both the United States and China, we did not use onboard calibration system for the trace-gas instruments except calibrating CO_2 and CH_4 for Picarro during flights in the United States (Ren et al., 2018). Some instruments such as SO_2, NO/NO_y, and NO_2 analyzers have internal 'zero' and we conducted 'chemical zero' frequently during the flight (He et al., 2012; Li et al., 2012; Brent et al., 2013; He et al., 2016). All the instruments were serviced and calibrated before and after the airborne campaign. Based on our experiences, these instruments are usually stable for 1-2 month-long campaigns. A side-by-side flight with the NASA P3B aircraft measurements during the 2011 DISCOVER-AQ campaign (Brent et al., 2013) and analysis of UMD and NASA aircraft measurements showed the our measurements are comparable with NASA instruments (He et al., 2014). In the original manuscript, we provided several key references which include the details of instruments to keep this paper concise.

We have revised this as: "*Measurements of ambient air pollutants were logged at 1 Hz frequency but the average times for different instruments are different as shown in Table S1. All measurements were synchronized based on the Picaroo measurements of CO_2, and CO with time, geolocation and altitude from the Global Position System (GPS).*" In the Supplementary Material, we also state: "*Table S1 shows the aircraft instrumentation on the Y12 aircraft during ARIAs. For synchronization of aircraft measurements, we first synchronized Picarro measurements of CO_2, CO and CH_4, which have a 6-second lag time that takes for a plume to transport from the aircraft sample inlet to the Picarro cavity to be detected the analyzer. The lag time was determined by introducing a pulse of CO_2/CO/CH_4 calibration standard into the sample inlet and then measure the time it takes for the Picarro detects the pulsed signals. All other measurements were then synchronized based on concurrent peak appearance of these measurements and the Picarro measurements.*"

For model comparison, since CMAQ generated hourly outputs, we used 10-min averaged data to alleviate the impacts caused by the relatively coarse temporal resolution (see discussion in Goldberg et al., 2016). The formaldehyde instrument was only used on the ground with 1-min averaged data of the data logger, although it has an average time of 90 seconds and a lag time of 300 seconds. Then we used hourly averaged HCHO concentrations as compared with CMAQ simulations. Our goal is not to measure high-frequency characteristics of the atmosphere such as turbulence, and the 1-Hz sampling rate was used to synchronize all the measurements. We added the following sentences "*The delay and lag time of each instrument was considered during the post-processing of observation data and averaged to 1-min record for further analysis and model evaluation.*" in Line 171 to explain it clearly.

2) Lines 264-276: This paragraph suggests that underestimation of ozone precursors in CMAQ could lead to the poor model performance. Poor simulation of the boundary layer depth also could lead to the poor model performance; a balanced discussion of both of these issues should be given.

Response: We agree that the PBL dynamics could play an important role in accurately simulating ozone concentrations. The A^2BC campaign had measurements of PBL heights from LIDAR and other remote sensing instruments, however improving the PBL simulation is beyond the scope of

this study. We added one sentence "*It is worth noting that PBL dynamics could also play an important role in accurately simulating the concentrations of air pollutants, especially with the complex terrain at the Xingtai supersite (Figure S4 in the supplementary material). However, evaluation of the PBL simulations and advections in CMAQ is beyond the scope of this study, and we focus on the photochemistry of ozone here*" in Line 294 to explain the potential impacts from PBL height.

3) Figure S2d seems to show a comparison of measured versus modeled NO and NOy, yet the measurement of these two species is not described in Section 2.1. What is going on here? Similarly Figure 4c gives NO measurements. If NOy was measured, why is its model measurement comparison not given in Figure 4? Clearly the description of the aircraft measurements must be improved as noted in Comment 1 above.

Response: Thanks for pointing out the missing information. We did have NO and NO_y measurements onboard of the Y12 aircraft. NO was measured using a commercially available Thermo Environmental NO/NO_y analyzer, and we used a hot molybdenum converter (at 375 ^oC) next to the gas inlet to convert all reactive nitrogen to NO_y. However, due to high electricity demand from the converter and limited power supply on the airplane, we only turned on this instrument for few research flights such as 06/11/2016. Because of limited observations we did not compare airborne NO_y measurements vs. CMAQ simulations. We added the introduction about the NO/NO_y instrument as "*Nitrogen oxide (NO) and reactive nitrogen compounds (NO_y) concentrations were analyzed using a commercial available NO analyzer (Model 42C, Thermo Environmental Instruments) with a hot molybdenum convertor working at 375 ^oC. Ambient gas input was switched with and without the convertor frequently to measure NO and NO_y simultaneous. However, due to high power demand of the instrument and convertor, NO and NO_y were only measured during some research flights.*" in Line 154 of the revised manuscript.

4) Figure 4 compares 10-minute averages of measured concentrations with model results. A discussion of this averaging must be given, as the aircraft covers a significant distance (~30 km?) and can cover a significant altitude range in 10 minutes. Is this really a reasonable comparison?

Response: Our CMAQ simulations had hourly 3-D outputs, so the comparison of hourly CMAQ outputs with 1-min averaged aircraft measurements could introduce large uncertainty into the analysis. Goldberg et al. (2016) developed the approach which used 10-min mean measurements to better evaluate the model performance. To explain it, we added following sentences as "*Since CMAQ generated hourly outputs, to alleviate the uncertainty of comparing 1-min aircraft data and hourly model simulations, we used 10-min averaged aircraft measurements which were matched to the closest hourly model output following the approach described in Goldberg et al. (2016)*" in Line 302 of the revised manuscript.

5) Lines 264-276: This paragraph is confused, contradictory and inaccurate. Figure 4 shows that NO is neither under- nor over-estimated by CMAQ. Figure S2 indicates that CMAQ under-estimates (not over-estimates) NOy. This description requires rethinking and rewriting

Response: Figure 4 summarized measurements from all 11 flights which stand for the regional characteristics of air pollution over the NCP during the one and half month campaign in 2016.

This paragraph described the surface measurements and CMAQ performance. Figure S5 in the revised manuscript only showed one example of model evaluation (1-min mean) on 06/11/2016, indicating the NO_y is underestimated by CMAQ (similar results in other research flights). These results suggest that significant portion of nitrogen compounds could exist in the format of organic nitrates such as PAN, and also indicate that aged air masses with lots of NO_y were measurements over the NCP. In Line 264-276 of the original manuscript, this paragraph discussed the comparison of CMAQ and surface observations at Xingtai supersite. Unlike the airborne measurements, surface observations at Xingtai site were significantly impacted by nearby emissions sources especially for the primary pollutants including CO and NO_x. The 12-km CMAQ cannot resolve the local high values like 6~7 ppmv CO and ~100 ppbv NO_x. Thus we disagree that results from Xingtai super site and Y-12 aircraft are contradictory. The former consisted of local fresh released emissions, while the latter represented the regional footprint and aged air mass.

6) It is not possible for me to assess the validity of the analysis in Section 3.2 because 1) only a very brief discussion of the approach and results are given, 2) the reference for the method (Halliday et al., 2018) is "to be submitted" and is thus not available, and 3) no illustrations of the EF calculations are shown in the paper or in the Supplementary Material. The description of this analysis must be improved.

Response: The Halliday et al. paper has been published recently which contains all the detailed information about this technique. We updated this reference in the revised manuscript and the citation information as:

Halliday, H. S., DiGangi, J. P., Choi, Y., Diskin, G. S., Pusede, S. E., Rana, M., Nowak, J. B., Knote, C., Ren, X., He, H., Dickerson, R. R., and Li, Z.: Using Short-Term CO/CO2 Ratios to Assess Air Mass Differences over the Korean Peninsula during KORUS-AQ, Journal of Geophysical Research: Atmospheres, 0, 10.1029/2018jd029697, 2019.

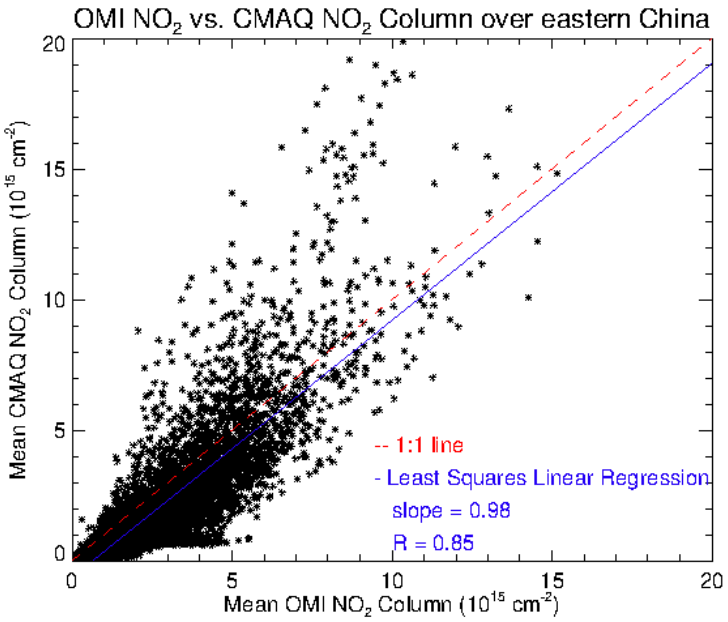
To better explain this technique, we added the following sentences "*To evaluate the emission inventory data in the NCP, we used a plume recognition method to calculate the emission enhancements (EEs) from Y12 observations. We first selected 60 1-s aircraft measurements with a 60-s moving window. Then we conducted linear regression of observed air pollutant (CO, NO_x, etc.) concentrations vs. CO_2 concentrations in each 60-s window and calculated the slope (i.e. $\Delta CO/\Delta CO_2$ and $\Delta NO_x/\Delta CO_2$) and correlation (R). The slope is defined as EEs in each window, standing for a 'plume' tested in the 60-s window. Lastly, we only selected EEs that are within the PBL (below 1.5 km AGL in this study) and statistically significant ($R^2 > 0.6$), so the values of these selected EEs can act as a proxy of EFs in the air mass observed.*" in Line 326 of the revised manuscript.

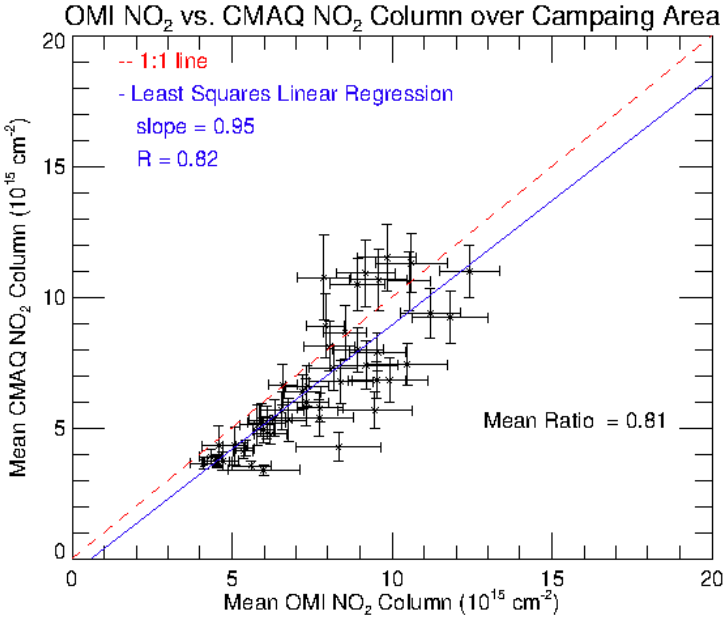
7) Lines 307-328: In these paragraphs the authors apparently make an elementary math error. They appear to be comparing the average of many, independently determined emission ratios (EFs) based on measurements with the ratio of the total EDGAR inventory emissions. Such a comparison is not valid because in general the arithmetic average of a distribution of ratios is not equal to the ratio of the means of the numerators and denominators of the ratios. The underlying numbers must be corrected and the discussion modified accordingly.

Response: We thank the reviewer for raising this question. Here we are comparing two different types of values: the emission enhancements (EEs) that were derived from the aircraft measurements near the sources and can act as a proxy of averaged emission factors (EFs), and the EFs calculated from the EDGAR emission inventory data. The EDGAR inventory has emissions for 4 sectors: Energy, Industry, Transportation, and Residential. The averaged EF value here is calculated as the mean of ratios from 4 different sectors. Thus, we believe these two numbers are 'comparable', which are both mean of several EEs or EFs. We added one sentence as "*The 2010 EDGAR inventory has emissions for 4 sectors: Energy, Industry, Transportation, and Residential. We calculated the CO/CO_2, NO_x/CO_2, and NO_x/CO ratios through averaging the EFs from these 4 sectors (Fig. 6).*" in Line 323 to better explain it.

8) Section 3.3 compares satellite measurements with CMAQ model results for NO2, CH2O and CO. The discussion proceeds with no considerations of systematic uncertainties in either the satellite measurements or the model results. The authors note that over the aircraft campaign area CMAQ predicts 81% of NO2 satellite column measurement and has good agreement with CH2O (<20% underestimation). To me this seems to be excellent agreement, and that it is not legitimate to interpret <20% differences as indications of emission uncertainties. This section is not acceptable without robust quantitative discussion of the systematic uncertainties in both the satellite measurements and the model results.

Response: Model uncertainty is difficult to quantify and we usually treat them as the 'true' results from model outputs. On the other hand, the satellite products could have some uncertainties during the 2-month campaign. Canty et al. (2015) qualitatively analyzed the uncertainty of satellite NO_2 products over Maryland during the 2010 NASA DISCOVER-AQ campaign which use >25% overestimation as compared with CMAQ results. He further improved the CMAQ chemistry (CB05 chemical mechanism) to obtain better modeling results. To show the uncertainties from OMI observation and CMAQ simulation, we added two figures showing the comparison of OMI and CMAQ NO_2 columns over eastern China and the campaign area (Fig. 8 in the revised manuscript) and related discussion as "*We plotted OMI and CMAQ*





NO_2 columns over eastern China and the campaign area (Fig. 8). Generally, NO_2 columns from OMI and CMAQ agreed well over the eastern China (Fig. 8a) but large discrepancies with both overestimation and underestimation existed. For the aircraft campaign area, CMAQ underestimates NO_2 columns (slope = 0.95 and mean ratio = 0.81, i.e., only predicts 81% of OMI NO_2 column) with uncertainties relatively smaller within the 2-month period (Fig. 8b)" in Line 380 of the revised manuscript.

9) Sections 3.1 and 3.3 seem to reach inconsistent conclusions. Section 3.1 suggests that CMAQ generally underestimates observed concentrations of major air pollutants, often by large factors

(Line 273: factors of 2 to 4 for all air pollutants; Line 287: a factor of 5 for VOCs). Yet as noted in the previous comment, Section 3.3 finds agreement within 20%. Section 3.4 goes back to the idea that the CMAQ run substantially underestimates the concentrations of ozone and its major precursors in the NCP. Such inconsistencies must be fully and quantitatively addressed before emissions within the CMAQ modeling can be objectively adjusted.

Response: Section 3.1 and 3.3 described the model evaluation with different observations. The largest underestimation is found when comparing hourly CMAQ simulations with surface observations from a single monitoring site. The factors of 2 to 4 indicated the existence of nearby sources, which cannot be resolved by the 12-km CMAQ as discussed in the manuscript. When evaluating the model performance with aircraft measurements, the airborne observations cover a larger area within one flight day and represent the regional mean of air pollutant distribution, i.e., averaging spatially, so we found better model performance. Section 3.4 compared the monthly mean column contents from CMAQ and satellites, i.e., we averaged the measurements both temporally and spatially, thus better model performance was anticipated. In this study, we adjusted the emissions based on the satellite products, in Section 3.5 (previously Section 3.4 in the original manuscript) only surface observations and aircraft measurements which are independent measurements of satellite observations were used to evaluate the model performance.

10) Section 3.4 is not satisfactory. Figures 10 and S3 present time series of observations and model results, but the agreement is quite poor regardless of the model run. These comparisons should be based on an objective measure of overall model performance so that the reader can appreciate how well or poorly each of the model simulations actually reproduced the observations.

Response: We conducted quantitative evaluation of all 11 research flights versus all sensitivity CMAQ experiments like Figure 4. To shorten the length of the manuscript, we summarized the statistics of linear regression analysis in Table 2. We added Figure 4 type results for each sensitivity experiments as Figure S10 in the supplementary materials and one sentence as "*Table 2 summarizes the model performance of CMAQ as compared with aircraft measurements and scatter plots for each CMAQ sensitivity experiment are showed in Figure S10 of the supplementary material*" in Line 485 of the revised manuscript.

11) Figure 11 compares observations and model result for a selected flight segment. The agreement appears to be extremely poor. Again, these comparisons should be based on an objective measure of overall model performance (for all flights) so that the reader can appreciate how well or poorly the model simulation actually reproduced the observations.

Response: Please see our response to Comment 10 above.

12) Much of the Conclusions and Discussion section is speculative and/or not quantitatively supported by the results discussed previously.

Response: We appreciate the reviewer to point it out. We revised the manuscript thoroughly to add more quantitative analysis and discussion. Please see these modifications in the track-and-change version of the manuscript.

Minor issues:

1) Line 122: The country of the Environmental Protection Agency (EPA) should be indicated. China (where the research is located) has an EPA, but I assume that this sentence refers to the U.S. EPA.

Response: We added 'U.S' here to clarify that the CMAQ model is developed by the U.S. EPA.

2) Figure 1 should be improved. It is not possible for a reader unfamiliar with Chinese geography to easily understand the region of China actually covered by the flight tracks.

Response: We revised the Figure 1b, and adding the proximity of flight region in it.

3) Lines 245-248: The statement on these lines is not accurate. The description is not of the generally observed concentrations, but rather reflects the maximum concentrations observed.

Response: The review is correct that we discussed the maximum concentrations observed here. To make it clear, we revised this sentence as ". *Generally, we observed high concentrations of air pollutants, with maximum values as >100 part per billion by volume (ppbv) of O_3, >20 ppbv of NO_2, >500 ppbv of CO, and >450 part per million by volume (ppmv) of CO_2, in the aircraft campaign area (defined as 36.5-38.5 °N, 114.0-115.5 °E hereafter)*" in Line 260 of the revised manuscript.

4) The Sections are not properly numbered; two are labeled Section 3.3.

Response: We appreciate the review pointed it out. The section number has been corrected.

Reference:

Brent, L. C., Thorn, W. J., Gupta, M., Leen, B., Stehr, J. W., He, H., Arkinson, H. L., Weinheimer, A., Garland, C., Pusede, S. E., Wooldridge, P. J., Cohen, R. C., and Dickerson, R. R.: Evaluation of the use of a commercially available cavity ringdown absorption spectrometer for measuring NO2 in flight, and observations over the Mid-Atlantic States, during DISCOVER-AQ, Journal of Atmospheric Chemistry, 1-19, 10.1007/s10874-013-9265-6, 2013.

Dickerson, R. R., and Delany, A. C.: Modification of a commercial gas filter correlation CO detector for enhanced sensitivity, Journal of Atmospheric and Oceanic Technology, 5, 424-431, 1988.

Dickerson, R. R., Li, C., Li, Z., Marufu, L. T., Stehr, J. W., McClure, B., Krotkov, N., Chen, H., Wang, P., Xia, X., Ban, X., Gong, F., Yuan, J., and Yang, J.: Aircraft observations of dust and pollutants over northeast China: Insight into the meteorological mechanisms of transport, Journal of Geophysical Research-Atmospheres, 112, D24s90, 10.1029/2007jd008999, 2007.

Doddridge, B. G., Morales-Morales, R., Rhoads, K. P., Merrill, J. T., Novelli, P. C., Dickerson, R. R., Connors, V. S., and Reichle, H. G.: Ground-based and airborne observations of

carbon monoxide during NASA measurements of air pollution from satellite (MAPS) missions SRL-1 and SRL-2, Journal of Geophysical Research-Atmospheres, 103, 19305-19316, 1998.

Goldberg, D. L., Vinciguerra, T. P., Anderson, D. C., Hembeck, L., Canty, T. P., Ehrman, S. H., Martins, D. K., Stauffer, R. M., Thompson, A. M., Salawitch, R. J., and Dickerson, R. R.: CAMx Ozone Source Attribution in the Eastern United States using Guidance from Observations during DISCOVER-AQ Maryland, Geophysical Research Letters, 2015GL067332, 10.1002/2015gl067332, 2016.

Hains, J. C., Taubman, B. F., Thompson, A. M., Stehr, J. W., Marufu, L. T., Doddridge, B. G., and Dickerson, R. R.: Origins of chemical pollution derived from Mid-Atlantic aircraft profiles using a clustering technique, Atmospheric Environment, 42, 1727-1741, 10.1016/j.atmosenv.2007.11.052, 2008.

He, H., Li, C., Loughner, C. P., Li, Z., Krotkov, N. A., Yang, K., Wang, L., Zheng, Y., Bao, X., Zhao, G., and Dickerson, R. R.: SO2 over central China: Measurements, numerical simulations and the tropospheric sulfur budget, Journal of Geophysical Research: Atmospheres, 117, doi:10.1029/2011JD016473, 2012.

He, H., Stehr, J. W., Hains, J. C., Krask, D. J., Doddridge, B. G., Vinnikov, K. Y., Canty, T. P., Hosley, K. M., Salawitch, R. J., Worden, H. M., and Dickerson, R. R.: Trends in emissions and concentrations of air pollutants in the lower troposphere in the Baltimore/Washington airshed from 1997 to 2011, Atmospheric Chemistry and Physics, 13, 7859-7874, 10.5194/acp-13-7859-2013, 2013.

He, H., Loughner, C. P., Stehr, J. W., Arkinson, H. L., Brent, L. C., Follette-Cook, M. B., Tzortziou, M. A., Pickering, K. E., Thompson, A. M., Martins, D. K., Diskin, G. S., Anderson, B. E., Crawford, J. H., Weinheimer, A. J., Lee, P., Hains, J. C., and Dickerson, R. R.: An elevated reservoir of air pollutants over the Mid-Atlantic States during the 2011 DISCOVER-AQ campaign: Airborne measurements and numerical simulations, Atmospheric Environment, 85, 18-30, 10.1016/j.atmosenv.2013.11.039, 2014.

He, H., Vinnikov, K. Y., Li, C., Krotkov, N. A., Jongeward, A. R., Li, Z. Q., Stehr, J. W., Hains, J. C., and Dickerson, R. R.: Response of SO2 and particulate air pollution to local and regional emission controls: A case study in Maryland, Earth Future, 4, 94-109, 10.1002/2015ef000330, 2016.

Krotkov, N. A., McClure, B., Dickerson, R. R., Carn, S. A., Li, C., Bhartia, P. K., Yang, K., Krueger, A. J., Li, Z. Q., Levelt, P. F., Chen, H. B., Wang, P. C., and Lu, D. R.: Validation of SO2 retrievals from the Ozone Monitoring Instrument over NE China, Journal of Geophysical Research-Atmospheres, 113, D16s40 10.1029/2007jd008818, 2008.

Li, C., Stehr, J. W., Marufu, L. T., Li, Z. Q., and Dickerson, R. R.: Aircraft measurements of SO2 and aerosols over northeastern China: Vertical profiles and the influence of weather on air quality, Atmospheric Environment, 62, 492-501, 10.1016/j.atmosenv.2012.07.076, 2012.

Marufu, L. T., Taubman, B. F., Bloomer, B., Piety, C. A., Doddridge, B. G., Stehr, J. W., and Dickerson, R. R.: The 2003 North American electrical blackout: An accidental experiment in atmospheric chemistry, Geophysical Research Letters, 31, L13106 10.1029/2004gl019771, 2004.

Ren, X., Salmon, O. E., Hansford, J. R., Ahn, D., Hall, D., Benish, S. E., Stratton, P. R., He, H., Sahu, S., Grimes, C., Heimburger, A. M. F., Martin, C. R., Cohen, M. D., Stunder, B.,

Salawitch, R. J., Ehrman, S. H., Shepson, P. B., and Dickerson, R. R.: Methane Emissions From the Baltimore-Washington Area Based on Airborne Observations: Comparison to Emissions Inventories, Journal of Geophysical Research: Atmospheres, 0, doi:10.1029/2018JD028851, 2018.

Salmon, O. E., Shepson, P. B., Ren, X., He, H., Hall, D. L., Dickerson, R. R., Stirm, B. H., Brown, S. S., Fibiger, D. L., McDuffie, E. E., Campos, T. L., Gurney, K. R., and Thornton, J. A.: Top-Down Estimates of NOx and CO Emissions From Washington, D.C.-Baltimore During the WINTER Campaign, Journal of Geophysical Research: Atmospheres, 123, 7705-7724, doi:10.1029/2018JD028539, 2018.

Taubman, B. F., Marufu, L. T., Piety, C. A., Doddridge, B. G., Stehr, J. W., and Dickerson, R. R.: Airborne characterization of the chemical, optical, and meteorological properties, and origins of a combined ozone-haze episode over the eastern United States, Journal of the Atmospheric Sciences, 61, 1781-1793, 2004a.

Taubman, B. F., Marufu, L. T., Vant-Hull, B. L., Piety, C. A., Doddridge, B. G., Dickerson, R. R., and Li, Z. Q.: Smoke over haze: Aircraft observations of chemical and optical properties and the effects on heating rates and stability, Journal of Geophysical Research-Atmospheres, 109, D02206

10.1029/2003jd003898, 2004b.

Taubman, B. F., Hains, J. C., Thompson, A. M., Marufu, L. T., Doddridge, B. G., Stehr, J. W., Piety, C. A., and Dickerson, R. R.: Aircraft vertical profiles of trace gas and aerosol pollution over the mid-Atlantic United States: Statistics and meteorological cluster analysis, Journal of Geophysical Research-Atmospheres, 111, D10s07 10.1029/2005jd006196, 2006.