Response to reviewers

Reviewer comments are in black ***italic*** type. Author responses are indented and in normal font labeled with [R]. Line numbers in the responses correspond to the revised manuscript without track-changes. Modifications to the manuscript are in *italics*.

Reviewer #1

I thank the authors for the responses to my first set of comments and their edits to the manuscript. However, I do not feel like they have done enough to warrant publication, and I continue to have major issues with this manuscript. Overall comment: I think that the authors have collected an interesting dataset, and the manuscript gives some glimpses into interesting conclusions that might be reached from the data (e.g., I am particularly interested in the spatial variation in inorganic PM2.5). However, there does not seem to be a coherent story. I instead got the impression that there were a set of discrete explanations for each chunk of the data (e.g., high emitting vehicles or regional transport), even if that explanation did not hold up for another part of the dataset. I think the authors need to step back and tell a coherent story about the full dataset. If that is not possible (and it may not be!), they should be straightforward about the limitations of the dataset and the analyses presented in the paper. For example, the small number of haze days (there seem to be two, but the authors are not forthcoming about this) is a limitation of the dataset. That is fine if there are only two haze days, but right now I feel like some of the details are being downplayed, and that makes me wonder if other aspects of the data collection and analysis are not being shown. Hopefully the comments below help to flesh out the concerns listed in this overall comment.

[R0] We thank the reviewer for the valuable feedback and constructive suggestions. We have major changes to the manuscript to address the reviewer's comments. Detailed information about the data collection and analysis have also been added to the main text and the supplementary. We think the results are presented in a much better way in the revised manuscript and the story behind has been clarified. Detailed responses are given below.

Major comment 1: My first set of comments criticized the manuscript for relying on what seemed to be a single transit of the 4th-ring road on two separate days (one haze and one non-haze). I don't feel like that comment was adequately addressed. In their response, the authors state "Averaging the data for the whole measurement period or all clean days would smooth out the spatial variability." However, by relying on only one or a few sampling passes to make their point, the authors risk overdue influence by quasi-random events such as driving near high emitting vehicles. Spatial aggregation over multiple sampling drives is needed to remove the influence of these events and to reveal the longer-term spatial patterns. Averaging over multiple drives is critical if the authors want to draw general conclusions from the mobile sampling. The influence of quasi-random high emission events is shown graphically by Apte et al (2017). Other papers, including Gu et al (which the authors cite) address the issue of "how many" mobile sampling passes are needed to build robust spatial patterns with mobile sampling.

[R1] We agree with the reviewer that averaging is necessary to derive a longer-term

spatial pattern for general conclusions from the mobile sampling. Our original focus was mainly on the spatial variations, and therefore only one driving cycle of pollutant distributions is shown for non-haze vs haze conditions. The presented spatial variations might be biased by quasi-random emission events. In the revised manuscript, we have averaged all drives from 9 AM to 4 PM over 8 non-haze days to derive the non-haze spatial patterns. The haze-day case is limited to 1-day average of the data, which has been clarified in the main text about the data limitation. But we have discussed some of the key features of the data in Line 117-138. The haze day herein represents a typical winter-haze event in the later high relative humidity stage. The findings from the haze-day spatial patterns are confirmed by the analysis of another haze-day in the 2021 mobile campaign. The revised manuscript now focused on discussing the general spatial patterns. Sections are re-organized and figures are replaced.

Major comment 2: The authors added Figure 2 to try and address my comment about temporal or drive-to-drive variation. However, this figure generates more questions than answers for me. I don't understand what Figure 2 shows. There are box plots, but how are they constructed? Is there one CV calculated for each time around the ring? One for each sampling day? Additionally, why is the organic PM normalized to $PM_{2.5}$ mass, but the inorganic components are normalized to the sum of inorganics? Why not normalize everything to $PM_{2.5}$ mass?

[R2] The original Figure 2 shows the box plots of CV for each drive cycle on the 4th Ring Road for all the cycles in the 8 non-haze days and 2 haze days. Because the revised manuscript now focuses on the spatial pattern of mean concentrations, this figure is no longer necessary and has been deleted from the main text. We have now used the magnitude of concentration variation and the CV values of the spatial patterns of mean concentrations to discuss about the spatial variability. Tremendous work has been done in previous studies to investigate the temporal variations of common gas pollutants and aerosol species. Temporal variations of these pollutants are not our focus herein. We therefore only discussed about the temporal variations for VOC and OVOCs in Sect. 3.3 of the revised manuscript.

Major comment 3: I am still not convinced by the author's reasoning for a lack of spatial heterogeneity on the haze days. They try to explain this away with a hand-waving nod to "regional transport." However, on stagnant haze days, local plumes do not disperse, and their impacts should be larger. For example, Lines 136-138 attribute HOA hotspots on the clean day to high emitting vehicles. If occasional high emitting vehicles are truly the source of the hotspots, the authors should detect these (or similar) hotspots on the haze days. A lack of these hotspots would seem to undermine the conclusion that pollutants are more homogeneous on the haze day. Rather, it would mean that the mobile lab simply passed fewer high emitting vehicles on the haze day. Another example of stagnant plumes on haze days: Lines 217-233 discuss high on-road emissions and titration of O3 on the highway. This is evidence of a strong emission source and spatial gradients associated with that source. And on haze days the data should see vehicle plumes, unless traffic volumes are vastly different on haze and non-haze days (or perhaps bad luck passing high emitting vehicles on the haze days).

[R3] We agree with the reviewer that local plumes are less dispersed much under stagnant conditions. The revised manuscript presents the averaged spatial patterns, which are much clearer about the enhanced impacts of local sources. For example, in Line 231, we state that "*The mean mixing ratios of CO were however greater than the non-haze day ratios, indicating accumulated pollution"*. In Line 270-271, we state that "*Hot spots of HOA and COA became more evident, which is consistent with the less-dispersed primary emissions under stagnant conditions (Figure 4)"*. While the spatial variabilities for all pollutants were significant for all pollutants during the non-haze days, the spatial variability for secondary aerosol species (e.g., OOA, sulfate, nitrate, and ammonium) and OVOCs have been largely reduced. The haze in NCP is usually developed regionally, meaning that the polluted air mass travels and would become more polluted when it suspends in urban areas to accumulate local emissions and secondary production under stagnant conditions. During the haze event, polluted air mass arrives and leads to significantly greater urban background concentrations for both primary and secondary pollutants. Meanwhile, secondary formation can be enhanced because of the elevated precursor concentrations during the haze day and heterogeneous and aqueous pathways for aerosol species that occur during the high-RH haze stage. The two facts drives a rather homogeneous distribution of aerosol composition because secondary species dominate the mass during the haze day. This has been clarified in the new Sect. 3.2 - "Spatial distribution and variability during the haze day".

Major comment 4: Some information about sources or land use would be helpful. There are some general descriptions in the text, but a graphical representation would be better. Most readers are not familiar with the land use in Beijing. Linking the land uses to the observed spatial variations in a more concrete way would help drive home the conclusions of this manuscript.

[R4] We have added Figure 1 for information about land use and vehicle emissions in the revised manuscript. Indeed, the land use information is helpful. For example, in Line 79-80, we added the following "*The 4th Ring Road is a 65-km-long urban highway that passes through residential, commercial and services, park, and transportation areas in the megacity (Figure 1a)"*. In Line 163-165, We discussed as follows: "*Overall, the spatial pattern of NOx was consistent with the bottom-up emission inventory for (1) the nonuniform vehicle emissions on the 4th Ring Road and (2) high concentrations in the east segment of the 4th Ring Road where the traffic volume was high (Figure 1b and Figure S10 in the Supplement)"*. In Line 192-197, we discussed about local sources as follows: "*The 40-s $PM_{2.5}$ measurements by TOF-ACSM may roughly represent a maximum area of 0.16 km^2 (for a mean speed of 6 m s^{-1} and wind direction perpendicular to the mobile path) upwind when the mobile laboratory was run on the 4th Ring Road by cycles. The HOA hot spots are generally consistent with the locations where the traffic volume was high and the driving speed was relatively low (Figure S10 of the Supplement). The COA hot spots are consistent with the places where the 4th Ring Road passes through sparsely located residential areas (Figure 1a)"*.

Major comment 5: Fig 4 is hardly discussed in the text. Some of the OVOCs have high PFs on the non-haze days. What are possible sources? The text focuses on vehicles as the main source of spatial variation - do vehicles emit things like furoic acid?

[R5] We have revised this figure (now Figure 6) with common VOC species likely related to primary vehicle emissions and secondary production. We have added more detailed discussions for Figure 6 in Line 253-265 as follows: "*The calculated PF for VOCs ranged from 11-67% (median) (Figure 6). High PF values were found for hydrocarbons and some OVOCs (e.g., C8H8, C10H8 and C4H4O), indicating a major contribution of transient localized sources (e.g., traffic, industrial facilities) to these species. The time series of these so-called primary species showed low baselines and sharp peaks (Figure S6). By contrast, OVOCs (i.e., with 2 or more oxygen in their formulae) that were typically considered as secondary species had low PF values (median: 11-16%) and elevated baseline contribution from photochemistry. Significance tests indicate greater PF values for the primary species during the non-haze days, meaning that the localized sources contributed more to the measured concentrations during the non-haze days than during the haze day ($p < 0.001$). During the haze day, the localized emissions should be accumulated near the source under stagnant conditions. Indeed, the peak concentrations of primary VOC species were significantly greater (e.g., ~2-4× for C6H6 and C7H8) (Figure S6). The lower PF values (by 30% for C6H6 and C7H8) during the haze day were caused by much more elevated baselines (e.g., ~9× for C6H6 and C7H8) that represent urban background affected by polluted air mass from regional transport plus gradually mixed local emissions. The mean VOC concentrations at the PKU roof site increased for about 2 times during the haze day, which agrees with the elevated baselines (Table S3 and Figure S12 in the Supplement).*"

Major comment 6: In my first round of comments I questioned whether some of the OOA spatial variation could be the result of misapportionment. The authors responded in part with "OOA can be contributed by many precursors and processes. It is not surprised to see a great spatial variability." I disagree vigorously. Primary OA will be spatially variable because it is emitted by local sources. OOA, which requires chemical processing, would be expected to be more spatially homogeneous. (At least I would consider this the null hypothesis, and the authors would need to disprove the null, which they have not done.)

[R6] Yes, non-perfect separation of POA and OOA by the PMF analysis may lead to misplaced spatial variability in OOA (Section A3). We have added detailed descriptions about source apportionment of OA by PMF in Section A3 of the Supplement, Figures S2-S5, and Table S1. The uncertainty of the PMF analysis has been clarified. The signal of m/z 44 of the PMF factors is sensitive to the rotation choice, which may introduce some uncertainty of the PMF results. During the non-haze days, the mean spatial pattern of OOA still shows moderate spatial variability (now Figure 4). In Line 197-208, we explained this as follows: "*Moreover, the mass concentrations of the sum of OOAs varied from 0 to 15 $\mu g\ m^{-3}$. Local photochemical production of SOA is a significant source of OA in Beijing in winter, although the solar radiation is reduced (Duan et al., 2020; Lu et al., 2019a). The photochemical production depends on the distributions of SOA precursors and oxidants.*

In the northwest corner where hydrocarbons showed high concentrations, the OOA mass loadings were indeed high. Because the majority of the SOA precursors (i.e., intermediate volatility and semivolatile organic species from anthropogenic sources) were not measured by the PTR-Qi-ToF (Liao et al., 2021; Miao et al., 2021), it is difficult to investigate more about the OOA source. The measurements in Pittsburgh also showed a significant spatial heterogeneity of primary carbonaceous components such as HOA, COA, and BC (Gu et al., 2018). Less spatial variability presented for OOAs in the Pittsburgh study. The OA mass loadings in Pittsburgh were however much less than the loadings in Beijing. The SOA formation can be significantly more efficient and complicated under conditions of high oxidative capacity and abundant precursors in Beijing than in Pittsburgh (Lu et al., 2019a; Li et al., 2021; Yang et al., 2019). Non-perfect separation of POA and SOA by the PMF analysis may also lead to misplaced spatial variability in OOA (Section A3)".

Minor comments

Line 84-87 refer to traffic volumes and composition on the 4th ring road. This needs a reference.

[R7] We thank the reviewer for the suggestion, and have added references. The traffic volume information is updated to the mean daily volume in November 2018.

Paragraph starting on line 88 - the vehicle speed was typically 60 km/h and AMS sampling times were 40 s. This gives a spatial resolution of about 700 m for PM mass and composition, and should be stated.

[R8] We thank the reviewer for the suggestion and have stated the resolution in Line 93-94 as follows "*The time resolution was 40 s, corresponding to a spatial resolution of ~0.7 km for a driving speed of 60 km h-1*".

Line 82 notes that measurements were collected on Nov 7-21 and Jan 21. Was sampling conducted on all of the November days? How many haze and non-haze days were sampled?

[R9] Sampling was conducted almost every day excluding some days for instrument and OFR maintenance. The effective sampling covered only 8 non-haze days and 1 haze day in 2018, which has been clarified in Line 122.

Figure 3 would benefit from the Clean and Haze days using the same color scale for each pollutant.

[R10] We have revised the color scales (now Figure 2).

Line 164 - where does the value of 0.16 km^2 come from?

[R11] Assuming the wind is persistently perpendicular to the mobile path at a mean speed of 6 m s^{-1}, the maximum area that the mobile measurement could represent is 60 km h^{-1} × 6 m s^{-1} × 40 s × 40 s = 0.16 km^2. We have clarified this in Line 193.

Line 202 and 203 - concentrations for "non-haze" days are listed twice; clearly one should be haze days.

[R12] Yes. We have revised it as follows "*For CO, the mean mixing ratio of 1.5 ± 0.7 ppmv was about three times greater than the urban background level (0.5 ± 0.3 ppm at the PKU roof site), indicating significant contributions of localized sources during the non-haze days.*" The haze-day discussion has been moved to Sect. 3.2.

Line 287-288, when describing Fig 6, the authors state "Day-to-day variations are not included because of the possible change of sources." I do not understand what this means. Does it mean that Fig 6 is presented for a single haze and non-haze day?

[R13] Yes, Figure 6 (now Figure 8) is presented for the haze day and a single non-haze day. The averaged spatial patterns have been discussed in the revised Section 3.1 and 3.2. Herein, we selected one clean day to compare with the haze day for correlations of VOCs.

Figue S5 shows VOC baselines. How were these determined and were they used?

[R14] We have described how these baselines were determined in Line 106-108 as follows: "*Baseline concentrations for each 2-s point in the 20-s smoothed data that represents were calculated as the 5th percentile concentration within a rolling window of 60 (i.e., 120 s) to represent the well-mixed urban background conditions.*"

Not much discussion of Fig 4.

[R15] As replied in [R5], we have revised this figure (now Figure 6) and added more discussion about this figure in Line 255-267.

I still don't understand how stagnant conditions in the haze day lead to things being more regional. More photochemically active? But if weather is stagnant, you should see stronger plumes near sources.

[R16] As replied in [R3], we agree with the reviewer that the stagnant conditions may enhance the impact of local plumes. Indeed, we have seen it in the data. But the particle composition was dominated by OOA, sulfate, nitrate, and ammonium during the haze day. Those secondary species had reduced spatial variability compared to the non-haze case.