We are thankful to the two reviewers for their thoughtful comments and suggestions. The manuscript was revised accordingly. Listed below are our point-by-point responses in blue to each referee's comments that are repeated in italic.

Response to Reviewer #1

General Comments:

This paper reports real-time aerosol mass spectrometer measurement results at a high altitude level (260m) in Beijing. The aerosol composition, variation, sources, as well as the influence of meteorology were detailedly discussed. As the authors stated, although the ground aerosol measurements were performed a lot before, those at a higher altitude have been rare, and need to be explored to help reveal the formation mechanisms of air pollution in Beijing. Generally, I think this paper provides an interesting dataset and some valuable results for the severe air pollution in China, and could be accepted by ACP after carefully considering the following revisions.

We thank the reviewer's positive comments.

Specific Comments:

The title. Since 260m is still in the range of near ground boundary layer, it is not strict to state that the measurements in this study were "above the urban canopy ".

Urban canopy layer is the layer of air closest to the surface in cities, which is approximately the mean building height (Rotach, 1999). The average building height is approximately 50 m to the south of the sampling site and approximately 20 m for other directions (Song et al., 2013). Considering that the sampling site in this study is located at 260 m, which is well above the urban canopy layer, we used "above the urban canopy" in the title.

Abstract. The description of the vertical differences, which should be the most important feature of this paper, were too vague.

Thank the referee's comments. This paper presented a detailed characterization of aerosol particle composition and sources above the urban canopy in Beijing before and during APEC. In addition to the response of aerosol chemistry to emission controls, the vertical differences are also important and interesting. Following the reviewer's suggestions, we expanded the descriptions of vertical differences in the revised manuscript.

Two different types of aerosol instruments were used simultaneously at two altitudes, including an ACSM and an AMS. Although the two instruments are both based on mass spectrometry, their structures are significantly different. Some comparisons between them ever found that they may have larger difference for high particle concentrations. Were the two instruments systematically compared on the ground before this campaign? This is important to state how much difference between the ground and the 260 m level was from the instruments.

Yes, we did compare the measurements from the two instruments side by side before

the campaign. The HR-AMS and ACSM were deployed for two weeks at the same site with the same inlet. All submicron aerosol species measured by the ACSM were highly correlated with those by the HR-AMS ($r^2 > 0.97$). Although the total NR-PM_1 mass measured by the ACSM agreed well with that by HR-AMS ($r^2 = 0.99$, slope = 0.99), the regression slopes of ACSM against HR-AMS varied from 0.61–1.24 for different aerosol species. Because ACSM was found to have a larger uncertainty in quantification of submicron aerosol species, particularly in determination of relative ionization efficiency, the mass concentrations of aerosol species measured by the ACSM at 260 m were further corrected using the regression slopes of ACSM/HR-AMS obtained from the inter-comparison study. As a result, the comparisons of vertical differences in this study have considered the uncertainties of the two instruments. For clarification, the inter-comparisons between the two instruments were added in the revised manuscript.

Page 22904, Line 17. "Considering that the peak time corresponds to lunch time, we concluded that it was attributed mainly to local cooking sources." This conclusion is too arbitrary. Have the mass spectra at noon been checked for features of cooking emissions?

We thank the referee's comments. As shown in Fig.4, the noon peak of HOA was corresponding well to the lunch time when cooking emissions were significant. To further investigate the influences of cooking sources on organic aerosol (OA) at 260 m, the organic spectra at noon time (10:00-14:00) were extracted and used for PMF analysis. Again, two factors, i.e., HOA and OOA, were identified by PMF analysis. The mass spectra of the two OA factors, which were similar to those from PMF analysis of the entire dataset, are shown in Fig. R1. It is clear that the mass spectrum of Factor1, i.e., HOA, was characterized by a high ratio of *m/z* 55/57 (1.6), which is consistent with the spectral characteristics of fresh cooking aerosols (Mohr et al., 2009;He et al., 2010). In contrast, the traffic-related HOA is generally characterized by comparable *m/z* 55 and *m/z* 57 (Ng et al., 2011). In addition, the diurnal profile of traffic related HOA is often characterized by two peaks at morning and evening rush traffic hours with the lowest concentration occurring at noon time. Therefore, based on the spectral characteristics and diurnal profile of HOA, we concluded that HOA was likely mixed with cooking organic aerosol. It is possible that cooking emissions can be mixed to the height of 260 m due to the high temperature and vertical turbulence exchange in the daytime.

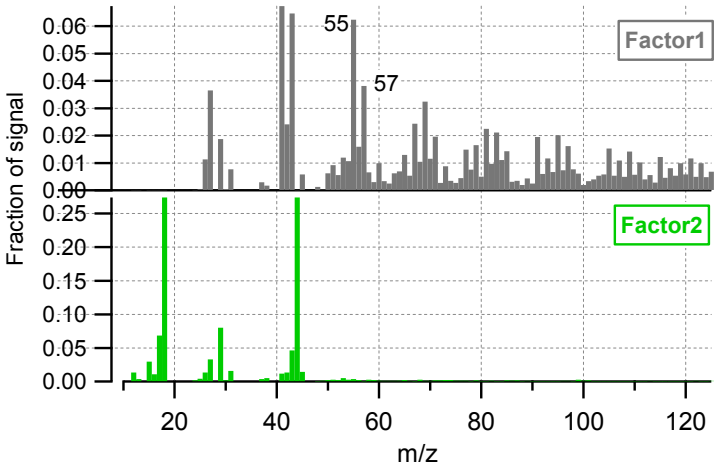

Figure R1. Mass spectra of two factors from PMF analysis of organic aerosol spectra during noon time period (10:00-14:00).

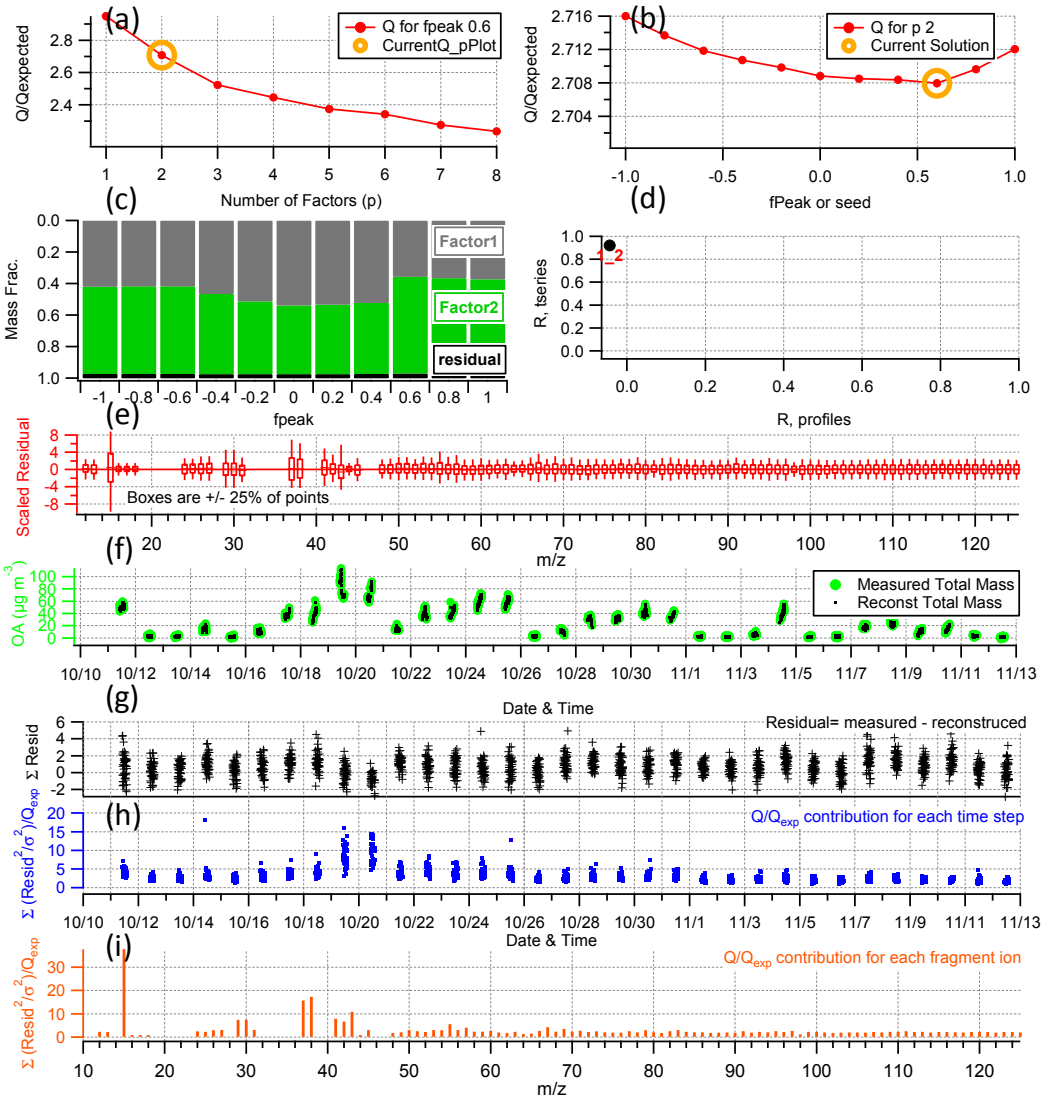

Figure R2. Summary of key diagnostic plots of the noon PMF results for 2-factor solution (fpeak=0.6): (a) Q/Q_{exp} as a function of number of factors (P) selected for PMF modeling. For the 2-factor solution, (b) Q/Q_{exp} as a function of fPeak;(c) fractions of OA factors vs. fPeak, (d) correlations among PMF factors, (e) the box and whiskers plot showing the distributions of scaled residuals for each *m/z*, (f) time series of the measured organic mass and the reconstructed organic mass (two factors), (g) variations of the residual (= measured − reconstructed) of the fit, (h) the Q/Q_{exp} for each point in time, and (i) the Q/Q_{exp} values for each *m/z*.

Fig. 13. This figure may be very misleading if without the comparison between the two instruments, as mentioned above.
Thank the reviewer's comment. The HR-AMS and ACSM were indeed compared before the campaign. The inter-comparisons between the two instruments were now added in Section 2.2 in the revised manuscript (also see our response above).

Page22913, Line 1. A typo of double "cause".
Corrected.

Response to Reviewer #2

General Comments:

This paper presents a case study of pollution and meteorology during the APEC summit in Beijing, with a specific focus on ACSM measurements on the Beijing Meteorological Tower. While the techniques and the processes under investigation are by no means cutting-edge, the facility in unique in its capacity to study pollution and dynamics in a megacity environment. Furthermore, the APEC case study presents a very interesting case that will allow new insights into air quality control strategies and source apportionment to be made. As such, I find this very relevant to ACP. It is worth noting that another paper from this platform and study period, Xu et al., (2015), is also currently under discussion, however having read both papers, I am satisfied that there is not too much overlap because that mainly focuses on the detailed measurements of the HR-AMS. However, while this paper is well-written, I do find that some of the interpretation needs revision and there needs to be more accountability on the PMF and clustering analyses, so therefore I recommend publication subject to the following comments.

We thank the referee's comments.

Specific Comments:

The authors need to include more evidence and reasoning in the supplementary information about their choice of PMF outputs used in the analysis. Specifically, why a 2 factor solution was considered the most reliable and why a nonzero value of fpeak was used.

Thank the referee's comments. PMF analysis of ACSM organic aerosol spectra in this study may not yield higher order factorization as easily due to lower signal-to-noise ratio and the fact that the primary emissions will be homogenized to an extent at the high altitude. As shown in Fig. R3, the time series of the third factor from 3-factor solution is much nosier than the other two factors, and the mass spectrum showed unrealistically high *m/z* 12 and *m/z* 15. Therefore, 2-factor solution was chosen in this study.

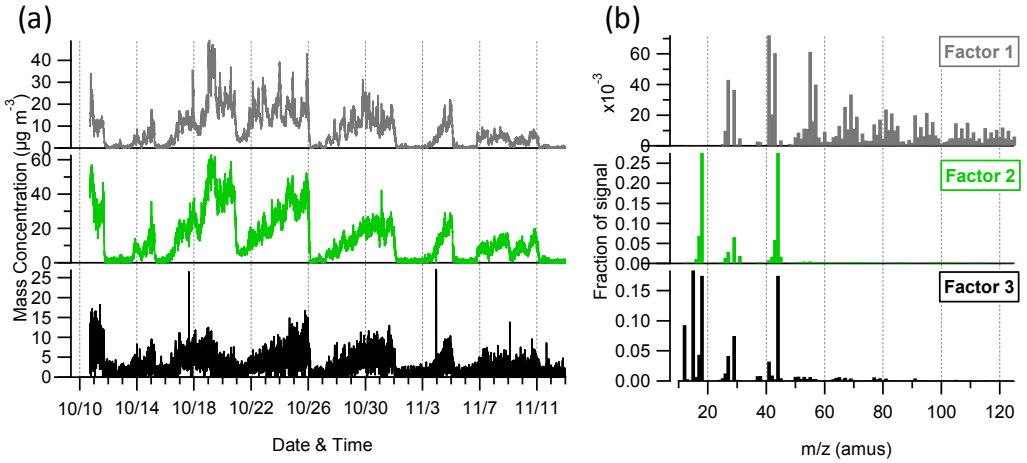

Figure R3. (a) Time series and (b) mass spectra of three OA factors from 3-factor

solution (fpeak=0.4)

We also evaluated the PMF solution as a function of fpeak. While the HOA contribution was fairly stable across different fpeak values (average ±1σ: 39%±4%), the mass spectra were quite different, particularly *m/z* 44. By comparing with the standard spectrum of HOA (Ng et al., 2010) at different fpeaks, we found that the spectrum of HOA at fpeak = 0.4 presented the best correlation (Table S1). The HOA spectrum at fpeak = 0 showed unexpectedly high *m/z* 44 which is a typical characteristics of secondary organic aerosol. Therefore, we chose the solution at fpeak = 0.4 rather than fpeak = 0. Such information was now added in supplementary and also partly in section 2.2 in the revised manuscript.

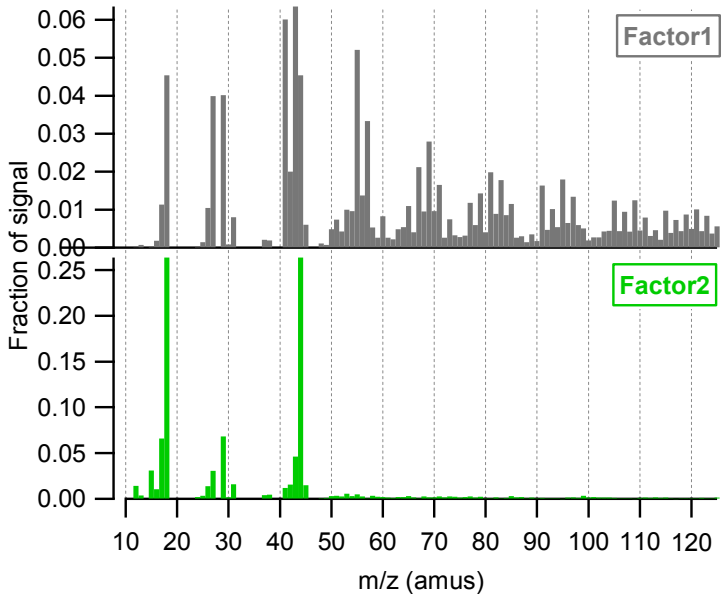

Figure R4. Mass spectra of two OA factors from 2-factor solution at fpeak=0.0

The fact that the authors conclude that the cooking and biomass burning were contained in the HOA factor but could not be resolved begs the question of whether this would be possible using the ME2 algorithm and lead to an improvement in the quality and depth of the science. Have the authors tried doing this?
The individual cooking and biomass burning aerosols were not resolved due to the lower signal-to-noise ratio of the ACSM and also the fact that the primary emissions can be mixed to an extent at a high altitude, e.g., 260 m in this study. The reviewer has a good point that the more advanced analysis with ME-2 algorithm might work for this dataset. Considering this treatment by applying the HR-AMS factor spectra as target profiles for ME-2 analysis would open a whole new avenue of enquiry and introduce new ambiguities to the manuscript, we didn't try ME-2 in this study yet. Nevertheless, we really appreciate the reviewer's suggestions, and will try the ME-2 analysis in the near future.

I do not agree with the conclusions reached regarding aerosol acidity. The correlations on figure 8 are very good and large quantities of nitrate were measured. This to me implies that the aerosol was consistently pH neutral, because an acidic aerosol would not be able to support nitrate in the particle phase. I think it is far more likely that one or more of the inorganic calibration values was wrong. What RIE values did the authors use and how were these determined?

Thank the reviewer's comments. The RIE of ammonium was determined from the pure ammonium nitrate during IE calibration. The default RIEs were used for other species, i.e., 1.4 for organics, 1.1 for nitrate, 1.2 for sulfate, and 1.3 for chloride (Canagaratna et al., 2007). It should be noted that the ACSM concentration was further corrected based on the inter-comparisons with the HR-AMS measurements, particularly for sulfate. By comparing the sulfate concentrations measured by the ACSM and the HR-AMS, a RIE of 0.98 was used for the ACSM sulfate.

We are also surprised the high nitrate content in such acidic aerosols. Similar high concentration of nitrate in acidic aerosols was also observed in previous AMS studies in Beijing (Zhang et al., 2014). Thus, we believe it might be true. There are several reasons which might explain this phenomenon. Firstly, biomass burning aerosol might be an important fraction of HOA in this study as indicated by the prominent *m/z* 60 in the HOA spectrum although it was not resolved. Biomass burning could emit a considerable amount of chloride in the form of KCl rather than NH_4Cl. Therefore, we might overestimate the predicted NH_4^+ when chloride was all counted as NH_4Cl, and hence overestimate the particle acidity. In fact, after excluding chloride, sulfate and nitrate were neutralized for most of the time periods, particularly the low mass loading periods (Fig. R5). These results suggest that chloride in this study might mainly exist in the form of KCl rather than NH_4Cl. In addition, higher acidity was observed at higher mass loading periods as indicated in Fig. R5. This result was consistent with previous findings in Beijing (Zhang et al., 2014). One of the reasons was due to the high humidity during the high mass loading periods. High humidity would facilitate the transformation of acidic gases, e.g., HNO_3 into liquid phase particles. We agree with the reviewer that such high nitrate concentration in acidic environment needs to be interpreted carefully. In the revised manuscript, we expanded the discussions and also the uncertainties on aerosol particle acidity.

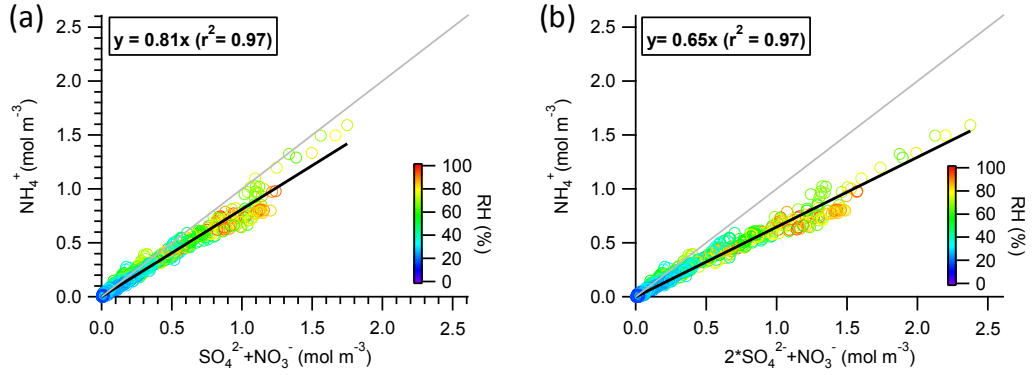

Figure R5. Scatter plots of molar ammonium concentration versus the sum of molar concentrations of sulfate and nitrate (a) $SO_4^{2-}+NO_3^-$, (b) $2*SO_4^{2-}+NO_3^-$. The data were color coded by RH.

The concluding line on page 22906 seems a little out-of-place considering the discussion that follows it concerning back trajectories and sources to the south. Would it not make more sense to talk about the regulatory implications later in the manuscript?

Thank the referee's comments. As suggested, the concluding description was moved to Section 3.2.5.

Insufficient information is given regarding the clustering process applied to the back trajectories. What clustering algorithm was applied? How did the authors determine the optimum number of clusters? The authors should also report the number of trajectories that contribute to each cluster, as well as the percentages, so that their significance can be assessed.

In this study, we calculated the three-day (72 h) back trajectories every hour at 500 m height using Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT, NOAA) 4.9 model (Draxler and Hess, 1997;Li et al., 2015). The clustering of trajectories is based on the total spatial variance (TSV) method (Draxler et al., 2012). This method minimizes the inter-cluster differences among trajectories while maximizing the inter-cluster differences, which has been widely used in previous studies (Sun et al., 2014;Zhang et al., 2014;Y. J. Li, 2015). Such information was now added in the revised manuscript as "Section 2.3 Air mass trajectory analyses."

Figure R6 shows the change of TSV as function of number of clusters before and during APEC. The changes in TSV decreased substantially from 2 and 3 before and during APEC, respectively. Therefore, the solutions with 3 and 4 clusters before and during APEC were used for further evaluation. We also evaluated 4-cluster solution before APEC. Figure R7 shows the back trajectories results for the 3-cluster solution before APEC. The contribution of cluster1 (C1), cluster 2 (C2), and cluster3 (C3) was 54%, 21%, and 25%, respectively. As shown in Fig. R6a and Fig. 11a, extending the 3-cluster solution to the 4-cluster solution will split C3 into two clusters i.e., C3 and C4. Considering that the APEC period has similar C3 and C4 clusters and also the chemical composition of C3 and C4 in 4-cluster solution were quite different, we chose four clusters before and during APEC for a better comparison of chemical composition and total aerosol loading from the same or similar trajectories before and during APEC.

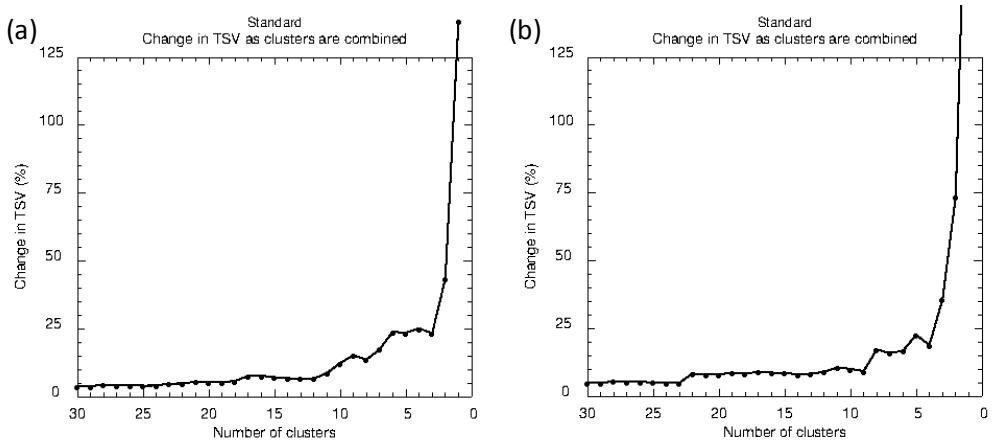

Figure R6. The variation of total spatial variance (TSV) as function of number of clusters **(a)** before and **(b)** during APEC.

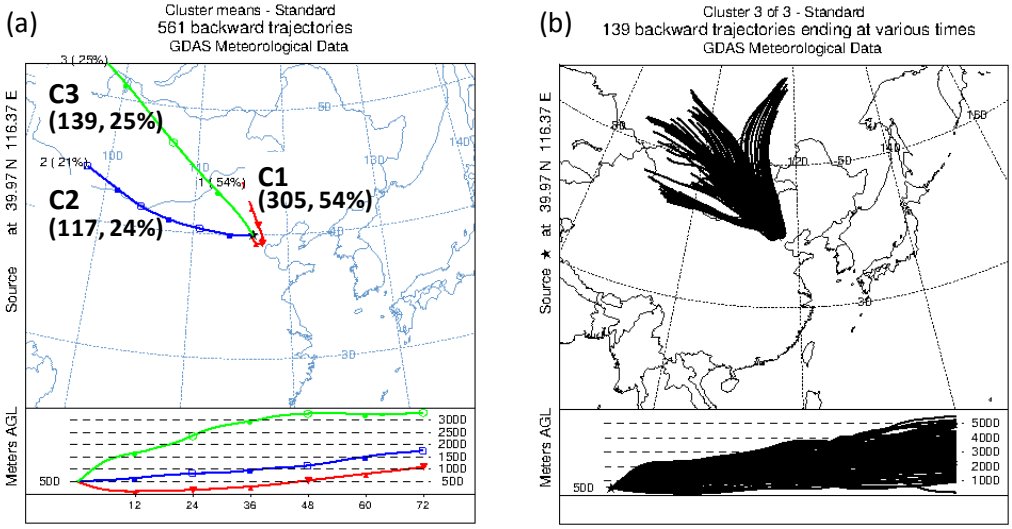

Figure R7. (a) Back trajectory results of 3-cluster solution, and (b) all trajectories of the cluster 3 (C3) before APEC. The number of trajectories and its percentage to the total number of trajectories for each cluster are also shown.

As suggested, the number of trajectories that contribute to each cluster was added on Fig.11 in the revised manuscript.

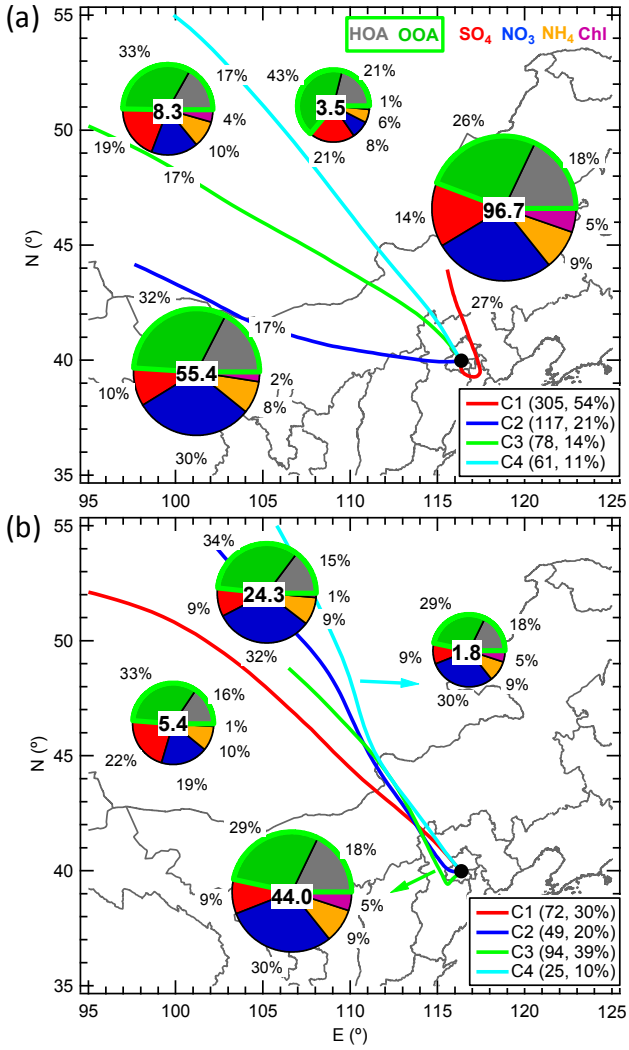

Figure 11. The average NR-PM_1 composition for each cluster **(a)** before and **(b)** during APEC. The numbers on the pie charts refer to the average total NR-PM_1 mass for each cluster. In addition, the number of trajectories and its percentage to the total trajectories are also shown in the legends.

Figures: With the vertical profiles of wind speed and direction, it is not completely clear whether this is from the LIDAR or the in situ measurements. This should be made clearer. Also, the white areas on the plots should be explained.

The vertical profiles of wind speed and wind direction in Fig.1 and Fig. 14 were from the measurements by the Doppler wind lidar. The white areas in the Figure indicate that the data were not available. Other meteorological parameters in the Figures were all from the tower measurements. Following the reviewer's suggestions, we clarified this in the revised manuscript.

References

Canagaratna, M., Jayne, J., Jimenez, J. L., Allan, J. A., Alfarra, R., Zhang, Q., Onasch,

T., Drewnick, F., Coe, H., Middlebrook, A., Delia, A., Williams, L., Trimborn, A., Northway, M., Kolb, C., Davidovits, P., and Worsnop, D.: Chemical and microphysical characterization of aerosols via Aerosol Mass Spectrometry, Mass Spectrom. Rev., 26, 185-222, 2007.

Draxler, R., Stunder, B., Rolph, G., Stein, A., and Taylor, A.: HYSPLIT4 user's guide, version 4, report, NOAA, Silver Spring, Md, in, 2012.

Draxler, R. R., and Hess, G.: Description of the HYSPLIT4 modeling system, Air Resources Laboratory, Silver Spring, Maryland, 1997.

He, L. Y., Lin, Y., Huang, X. F., Guo, S., Xue, L., Su, Q., Hu, M., Luan, S. J., and Zhang, Y. H.: Characterization of high-resolution aerosol mass spectra of primary organic aerosol emissions from Chinese cooking and biomass burning, Atmos. Chem. Phys., 10, 11535-11543, 10.5194/acp-10-11535-2010, 2010.

Li, Y., Lee, B., Su, L., Fung, J., and Chan, C.: Seasonal characteristics of fine particulate matter (PM) based on high-resolution time-of-flight aerosol mass spectrometric (HR-ToF-AMS) measurements at the HKUST Supersite in Hong Kong, Atmos. Chem. Phys., 15, 37-53, 2015.

Mohr, C., Huffman, J. A., Cubison, M. J., Aiken, A. C., Docherty, K. S., Kimmel, J. R., Ulbrich, I. M., Hannigan, M., and Jimenez, J. L.: Characterization of primary organic aerosol emissions from meat cooking, trash burning, and motor vehicles with high-resolution aerosol mass spectrometry and comparison with ambient and chamber observations, Environ. Sci. Technol., 43, 2443-2449, 2009.

Ng, N., Canagaratna, M., Jimenez, J., Zhang, Q., Ulbrich, I., and Worsnop, D.: Real-time methods for estimating organic component mass concentrations from aerosol mass spectrometer data, Environ. Sci. Technol., 45, 910-916, 2010.

Ng, N. L., Canagaratna, M. R., Jimenez, J. L., Zhang, Q., Ulbrich, I. M., and Worsnop, D. R.: Real-Time Methods for Estimating Organic Component Mass Concentrations from Aerosol Mass Spectrometer Data, Environ. Sci. Technol., 45, 910-916, 10.1021/es102951k, 2011.

Rotach, M. W.: On the influence of the urban roughness sublayer on turbulence and dispersion, Atmos. Environ., 33, 4001-4008, http://dx.doi.org/10.1016/S1352-2310(99)00141-7, 1999.

Song, T., Sun, Y., and Wang, Y.: Multilevel measurements of fluxes and turbulence over an urban landscape in Beijing, Tellus B, 65, 2013.

Sun, Y., Jiang, Q., Wang, Z., Fu, P., Li, J., Yang, T., and Yin, Y.: Investigation of the sources and evolution processes of severe haze pollution in Beijing in January 2013, Journal of Geophysical Research: Atmospheres, 119, 4380-4398, 2014.

Y. J. Li, B. P. L., L. Su, J. C. H. Fung, and C.K. Chan: Seasonal characteristics of fine particulate matter (PM) based on high-resolution time-of-flight aerosol mass spectrometric (HR-ToF-AMS) measurements at the HKUST Supersite in Hong Kong, 10.5194/acp-15-37-2015, 2015.

Zhang, J., Sun, Y., Liu, Z., Ji, D., Hu, B., Liu, Q., and Wang, Y.: Characterization of submicron aerosols during a month of serious pollution in Beijing, 2013, Atmos. Chem. Phys., 14, 2887-2903, 2014.