Response to referee #2:

General comments

This measurement report is clearly presented and includes valuable observations and methods, important for monitoring and understanding aerosol sources and variability in urban and suburban air quality. Through regular tethered balloon measurements, a substantial number of vertical profiles of aerosol size distribution have been collected in the boundary layer, occasionally reaching into the free troposphere. Three distinct particle size modes are attributed to different sources: local fossil fuel combustion emissions and secondary aerosol formation (< 0.3 μm), aerosol produced in religious activities involving burning incense and other materials (0.3 – 2.5 μm), and coarse mode dust (>1 μm). Relative and absolute weighting of these size modes in the boundary layer were attributed to both meteorological (precipitation, RH, etc.) and anthropogenic causes (in particular, religious burning practices). Free tropospheric conditions appear to be decoupled from the boundary layer, with an occasional residual layer between the BL and FT.

The work presented here is appropriately submitted as a Measurement Report, though direct links for access to data (per ACP policy) appear to be missing at this time.

We highly appreciate the referee's valuable comments and instructive suggestions. We have addressed each comment as below and corresponding revisions have been made in the manuscript. We have also added a direct link to the data presented in this study and revised the section of "Data availability" as "The data in this study can be publicly accessed via https://doi.org/10.5281/zenodo.6374312 (Ran et al., 2022)."

Specific comments

Please address the following technical comments to help clarify and improve the manuscript:

- **The two POPS calibration materials (ammonium sulfate and PSL) have different refractive indices. What assumptions are being made about atmospheric aerosol optical properties for measurement with the POPS? Sizing on any such optical particle counter is sensitive to selection of refractive index. Non-monotonic Mie surfaces may influence reported size distributions in certain size regions. This can show up as relatively sharp features in size distributions like the peak seen at 230 nm or 500-900 nm in Figure 4. See Gao et al., 2016 (Figure 8 in particular) for potential influences of refractive index mismatch between calibration and measured material. High resolution binning can also introduce erroneous peaks**

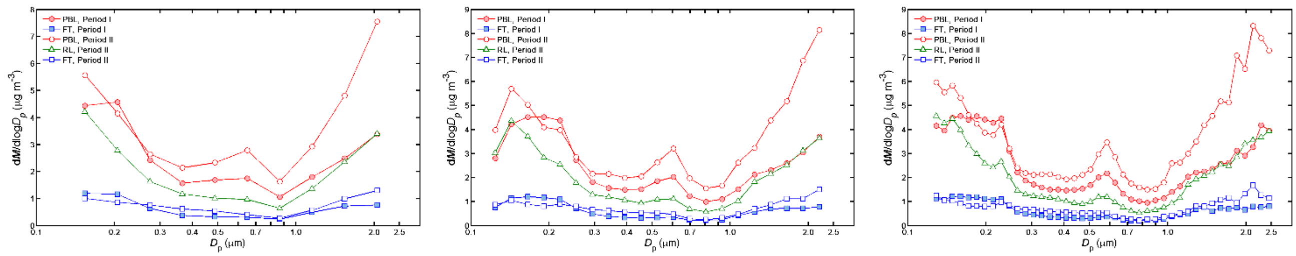
in distributions, again particularly in regions of Mie resonances or "flat" portions of the Mie curve. Would your analysis change if either/both the 230 nm and 500-700 nm peaks in Figure 4 were due to instrumental settings and assumptions, rather than real features?

We appreciate the referee's valuable comment. We agree with the referee that sizing is sensitive to the selection of the refractive index and using two calibration materials with different refractive index for particle sizing could introduce misleading interpretation of the data measured by an optical counter such as POPS. We intended to report ammonium sulfate-based PNSDs, as the refractive index of ambient dry particles is closer to that of ammonium sulfate (Shingler et al., 2016) rather than polystyrene latex sphere (PSL), though PSL is widely used in the calibration of OPCs. However, without available methods to classify super-micron ammonium sulfate particles, a calibration curve combined from experimental responses of ammonium sulfate particles at submicron sizes and PSL at super-micron sizes was adopted in the original manuscript. As found in Gao et al. (2016), there are strong oscillations in theoretical response curves for polystyrene latex sphere (PSL) and dioctyl sebacate (DOS) when particle diameters are larger than 600 nm, leading to unrealistic particle sizing. Similar oscillations were also found for ammonium sulfate particles based on our calculation. Therefore, the response curve for converting the optical signal to the particle size above 600 nm should be smoothed to be monotonically increasing. In the revised manuscript, we re-calibrated POPS data using a combined calibration curve from the experimental response of ammonium sulfate particles for diameters smaller than 600 nm and the smoothed theoretical response of ammonium sulfate particles for diameters larger than 600 nm. By using the new calibration method, PNSDs were obtained assuming only the refractive index of ammonium sulfate. Though new results were quantitatively different, but the changes in numbers and plots were rather small and qualitative conclusions remained the same. The manuscript has been revised accordingly. We have also revised the description of POPS calibration as below.

Page 6, Line 101: The POPS was calibrated by establishing a relationship between the scattering signal and the particle size before the campaign. Both polystyrene latex sphere (PSL) with known sizes and ammonium sulfate particles with sizes selected by a differential mobility analyzer were employed. Though the experimental responses of the two calibration materials generally agreed well with the simulated theoretical responses, both theoretical response curves were found highly oscillatory above the particle size of 600 nm (Gao et al., 2016). Considering that the refractive index of ammonium sulfate is closer to ambient dry aerosols than PSL (Shingler et al., 2016), a combined calibration curve from the experimental response of ammonium sulfate particles for diameters smaller than 600 nm and the smoothed theoretical response of ammonium sulfate particles for diameters larger than 600 nm was used to obtain 42 logarithmically equal size bins over the size range of 0.124~2.55 μm.

We agree with the referee that high solution binning might result in fake peaks and valleys in PNSDs due to the non-monotonic response. We performed a comparison among average PMSDs for different layers using POPS data generated with a resolution of 8, 16 and 32 bins per magnitude. The plots below showed a peak in the size range of 0.5~0.7 μm for all resolutions. However, the resolution of 8 bins per magnitude was too coarse to sufficiently display the characteristics of the PMSDs. As for the small peak around 230 nm indicated by the referee, it was indeed more pronounced in PMSDs based on a resolution of 32 bins. Since no solid evidence or physical/chemical reasons to prove such a peak was real, we speculated this might be resulted from uncertainties in the measurement or calibration. We added in the manuscript a discussion as below.

Page 18, Line 315: Plainly, a distinct mode below 0.3 μm existed for all layers during both periods... It was noteworthy that the small peak around 0.23 μm might be resulted from uncertainties in the measurement or calibration. Thus, caution should be taken when drawing specific conclusions.



Plots of average particle mass size distributions for different layers during the two periods using POPS data with a resolution of (left) 8 bins, (middle) 16 bins, (right) 32 bins per magnitude.

References:

Gao, R. S., Telg, H., McLaughlin, R. J., Ciciora, S. J., Watts, L. A., Richardson, M. S., Schwarz, J. P., Perring, A. E., Thornberry, T. D., Rollins, A. W., Markovic, M. Z., Bates, T. S., Johnson, J. E., and Fahey, D. W.: A light-weight, high-sensitivity particle spectrometer for PM2.5 aerosol measurements, Aerosol Sci. Technol., 50, 88-99, doi: 10.1080/02786826.2015.1131809, 2016.

Shingler, T., Crosbie, E., Ortega, A., Shiraiwa, M., Zuend, A., Beyersdorf, A., Ziemba, L., Anderson, B., Thornhill, L., Perring, A. E., Schwarz, J. P., Campazano-Jost, P., Day, D. A., Jimenez, J. L., Hair, J. W., Mikoviny, T., Wisthaler, A., and Sorooshian, A.: Airborne characterization of subsaturated aerosol hygroscopicity and dry refractive index from the surface to 6.5 km during the SEAC4RS campaign, J. Geophys. Res. Atmos., 121, 4188-4210, doi:10.1002/2015JD024498, 2016.

- **A figure illustrating the combined POPS + GRIMM 11-C distribution and the validity of the weighting and combination method of the individuation distributions would be fitting, perhaps in the SI.**

We have added in the supplement a figure to illustrate how the PNSD and the PMSD measured by POPS and GRIMM 11-C were combined using the weighting factors for each instrument. We have revised the manuscript accordingly.

Page 8, Line 2: An example of combining the PNSD from POPS and GRIMM 11-C and the PMSD from POPS and GRIMM 11-C using the weighting factors for each instrument was illustrated in Fig. S2.

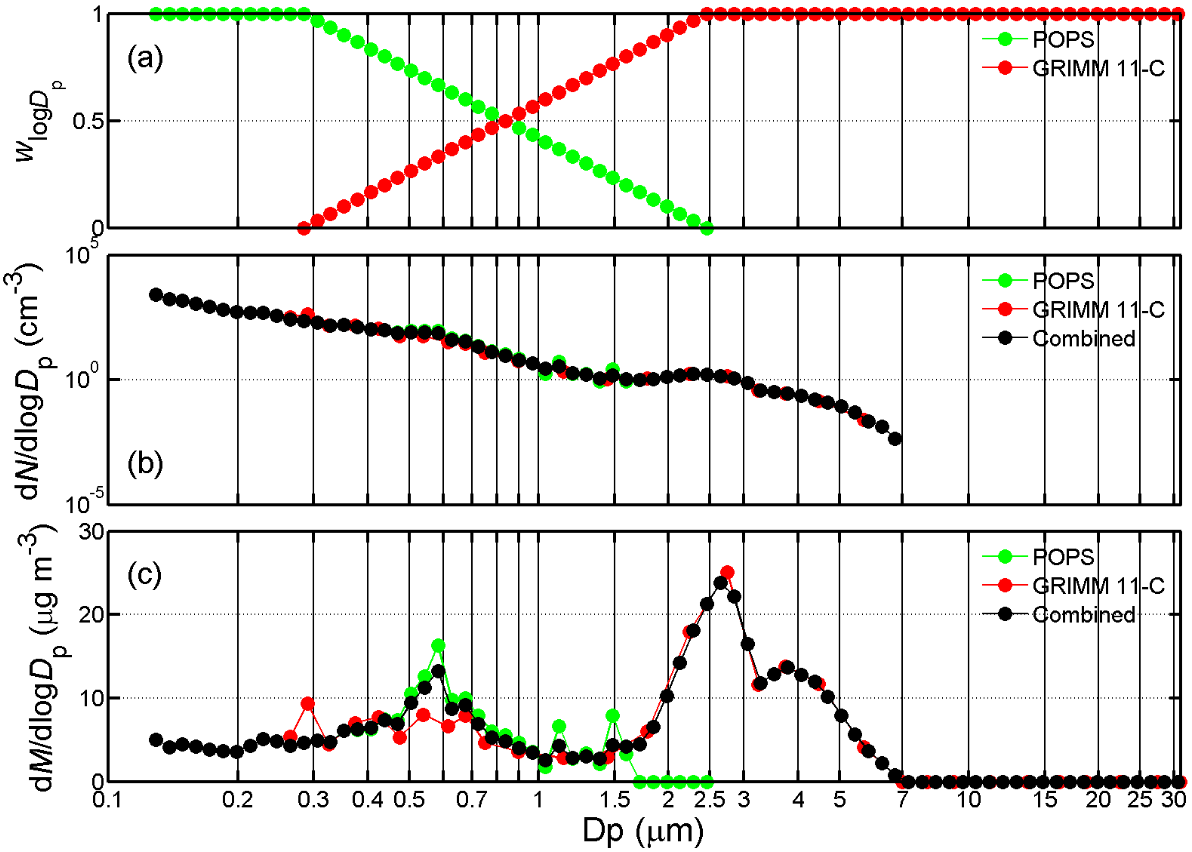


Figure S2 An example of combining the PNSD from POPS and GRIMM 11-C (b) and the PMSD from POPS and GRIMM 11-C (c) using the weighting factors ($w_{\log Dp}$) for each instrument (a).

- **Is there a reason residual layers were only seen in Period II, not in Period I? Also, how was the upper limit of the RL defined, to distinguish from the FT?**

The residual layer (RL) usually occurrs after the collapse of the planetary boundary layer (PBL) in the evening and might last until the early morning on the next day. The residues would usually be gradually mixed down into the PBL along with the evolution of the PBL after sunrise in the morning. As described in the manuscript, profiles in Period II were mainly collected in the early morning and at night, when the maximum height reached by the tethered balloon exceeded the top of the PBL and the RL could be observed if it existed. During Period I, all profiles were collected before the evening and a few in the early morning. The FT was identified for 14 profiles but no RL was identified. For four profiles collected in the morning, there was actually a layer above the PBL and beneath the FT with N_a to be a little higher than that in the FT but still very low (~100 cm^{-3}). Since it was insufficient for us to confidently decide whether it was the RL or still a part of the FT, we did not include them into either the RL or the FT. Another reason that we did not identify them as the RL was that the statistical results might be misleading if we included these cases, especially when the determination was not solid. We thank the referee for this question to make us realize that it would be better to give a clearer and more precise description of this issue. We have revised the manuscript as below.

Page 16, Line 286: In total, the RL was identified for 29 profiles in Period II, whereas none was identified in Period I. A layer above the PBL and beneath the FT was observed for four profiles collected in the morning in Period I, with N_a to be a little higher than that in the FT but still very low (~100 cm^{-3}). Since it was insufficient to decide whether the layer was the RL or still a part of the FT, also it would be misleading if they were used to represent the average condition of the RL during Period I, they were identified as neither the RL nor the FT.

When the maximum height reached by the tethered balloon exceeded the top of the identified RL, the upper limit of the RL was determined by subjectively analyzing vertical distributions of N_a and PMSDs. A sharp reduction in N_a could be found from the RL to the FT, where the average N_a was at least less than half of that in the RL. Both N_a and PMSDs greatly varied across the top of the RL and gradually became almost unchanged after entering into the FT.

Technical corrections

I believe the following grammatical changes will clarify the authors' intended meanings.

- **Line 20: "in consistence" change to "consistent"**

We have revised the manuscript accordingly.

- **Line 57: "religious activities involved incense burning…" change to "religious activities which involve incense burning"**

We have revised the manuscript accordingly.

- **Lines 159 & 160: "averagely" change to "average"**

We have revised the manuscript accordingly.

- **Line 177: "Averagely" change to "On average"**

We have revised the manuscript accordingly.

- **Line 185: "The last but not the least" change to "Finally" or "Last but not least"**

We have revised the manuscript accordingly.

- **Lines 193 & 194: "as being already pointed out by" change to "as has been pointed out by"**

We have revised the manuscript accordingly.

- **Line 259: "evolved to an enough height in the next day, or there could also be contributions" change to "had evolved to sufficient height the next day, or if there could also be contributions"**

We have revised the manuscript accordingly.

- **Line 263: “were almost less than” change to “were generally less than”**

We have revised the manuscript accordingly.

- **Line 366: “was” change to “were”**

We have revised the manuscript as “It was found that on non-holidays and holidays* contributions of potential sources were…”

- **Line 378: “entangle” change to “disentangle”**

We have revised the manuscript accordingly.