Interactive comment on "Urban organic aerosol composition in Eastern China differs from North to South: Molecular insight from a liquid chromatography-Orbitrap mass spectrometry study" *by* Kai Wang et al.

Anonymous Referee #3

General Comments:

This manuscript presents the application of UHPLC/Orbitrap MS to three sets of samples collected in urban locations in China. The observed molecular formula are classified according to the components (CHO, CHON, CHOS, etc.) and overall trends in the abundance and characteristics of these molecular formulas are compared between the three samples. The paper is well written and the results are clearly communicated. I appreciate the careful and detailed break down of trends by compound class. The largest gap that I see is a lack of discussion/consideration for the information provided

by the chromatography itself. I recommend publication after the following concerns are addressed.

Specific comments:

1. Throughout the manuscript, a comparison is made between overlapping molecular formula. However, no reference to the LC data set is provided. The separation should provide some information on whether the observed molecular formulas actually correspond to the same chemical structure.

a. What fraction of the overlapping formulas in Figure 2 (and discussed throughout the document) have the same or very similar retention times in the column?

b. Possible identities for some compounds are given (e.g. page 10 lines 289-330 and page 11 lines 342-355). How do the retention times of these compounds compare to standards run with the same column settings?

2. How clean was the separation? Were multiple ions observed under each chromatography peak? Were there differences in this between the three sample locations? It looks like there are a lot more compounds in the Changchun samples as a whole. If there was more charge competition during ionization (especially in positive ion mode) in specific samples, it would be good to clarify this because you are using the signal to directly correlate to abundance.

3. The positive ion mode data for Guangzhou and Shanghai are interesting in that they appear to be dominated by (one?) very high abundance peak. The fact that amines can ionize very well in ESI+ is discussed in the manuscript. However, it would be helpful to provide more context/information on these samples. In particular, I suggest including (possibly in the supplemental) the reconstructed MS for Shanghai+ and Guangzhou+ with a zoom in on the lower abundance ions (indicate those CHN+ ions are off scale). I also recommend making comparisons in the CHN+ section both in terms of numbers of chemical formula and in terms of abundances (as is currently provided).

4. The conclusion that there is more photochemical aging in the southern locations compared to the more northern location is made throughout the manuscript. However, I would argue that the differences could also be due to more fresh emissions being sampled in the northern location compared to the southern ones. This is supported by, for example, the data in Figure S1 which I think shows similar higher O/C numbers across all three, but more total lower O/C and aromatic compounds for Changchun. I recognize that van Krevelen diagrams can hide some depth (different molecular formulas with the same O/C and H/C). If this is the case here, that should be clarified in the text.

What was the photoactive radiation level at each site (were any of them cloudy)? Is there any information on the age of the air mass (back trajectories, wind speed, wind direction, etc.)? The other possible explanations for this trend should be addressed: different source types and different overall age of the OA material (i.e. longer transport for the southern locations).

5. For the CHON and CHN compounds, how many contained two or three nitrogen atoms? I can see evidence for a few of them in Figures 4 and S2 but it is hard to quantify. This may be a useful piece of information to include in the supplemental.

6. Throughout the text, differences are stated to be significant but no significance tests are performed. I would recommend re-phrasing to avoid implying a level of analysis that was not carried out.