Responses to Comments on Manuscript amt-2020-267
(Orbitool: A software tool for analyzing online Orbitrap mass spectrometry data)

We thank the reviewer for the constructive comments that help to improve this manuscript. The manuscript was revised correspondingly. The reviewer's comments are addressed in the following paragraphs. The comments are shown as sans-serif dark red texts and our responses are shown as serif black texts. Changes are highlighted in the revised manuscript and shown as "quoted underlined texts" in the responses. References are given at the end of the responses.

Reviewer #2

Summary and recommendation:

In this study, Cai et al. describe the development of a new software tool for the analysis of Orbitrap data, focusing in particular on data treatment procedures that are common in atmospheric sciences. The authors demonstrate the usefulness of the developed procedures and explain how the algorithms process the data. Moreover, they show examples of laboratory and field data that were processed using the new software. The software is planned as an open source tool and freely available on the internet after user registration.

In my opinion, the authors did a very good job and created an impressive tool for Orbitrap data analysis, which follows routines that are familiar to the CIMS community. I agree with the authors that Orbitrap mass spectrometers will play a more and more important role in atmospheric sciences in the future. Therefore, a software tool tailored to the needs of online Orbitrap MS data is highly desirable and will certainly help CITOFMS users to transition to CI-Orbitrap MS.

The manuscript is well-written and explains in a reader-friendly manner all modules of the software. Therefore, I have only some minor comments, which should be addressed before final publication in AMT.

Response: We thank the reviewer for these positive comments. To facilitate atmospheric researches with Orbitrap mass spectrometers with various needs, Orbitool will be updated continuously and a forum will be shortly open to collect user's demands.

Minor comments:

1) P6L17f: I cannot follow the authors explanation on the averaging weight for the spectra. Could you explain in more detail why the scan number is more important than the duration of a single scan?

Response: This sentence was revised as: "The data recorded in the same file is averaged using an averaging function in the Thermo RawFileReader library and the average is calculated with respect to scan number."

The Thermo RawFileReader library calculates the average spectrum with respect to scan number. Xcalibur^{TM} uses the same averaging function. The duration of a single scan does not vary much or the total ion current is relatively stable. However, an implicit assumption is that the parameters cannot be changed during the time interval of an averaged spectrum. We added the following sentence in the revised manuscript: "Hence, it is assumed that the parameters determining the duration of every single scan (e.g., injection time and the number of micro scans) are kept constant during each time interval."

2) P7L5: How do the authors define background here? Is it only electronic background?

As far as I know there is already a built-in background subtraction of signals by the acquisition software from the manufacturer.

Response: Thanks for this comment. We added the following text to better explain how the background is treated. "When converting electronic signals into a spectrum via the Fourier transform, the instrument removes signals below the noise threshold defined by the Orbitrap. Signals exceeding this noise threshold are converted into peaks recorded in the data file. Since this procedure does not remove all the noise signals, some noise peaks are also recorded in the data file. In this study, "noise" refers to these noise peaks." To avoid confusion, we removed "background" when referring to noise.

3) P7L9f: For readers less familiar with the common procedures in the CIMS community, it might be helpful to explain shortly why the noise estimation is conducted in the mass defect range of 0.5-0.8 Da.

Response: There was a short explanation below this sentence: "Such a mass defect range is chosen because most of the observed compounds in atmospheric measurements are located outside of this mass range."

4) P8L13f: The description of the peak width normalization should be revised. Currently, it remains unclear why the peaks are normalized to 200 Da after dividing them by the square root of m.

Response: We revised this sentence as follow: "According to this relationship, the measured widths of these selected peaks are normalized to widths at 200 Da by dividing them by $\sqrt{m/200\ \mathrm{Da}}$."

We also added the following information: "200 Da is chosen as the reference value because the instrumental resolution of the Orbitrap is usually reported at this mass."

5) P9L13: As far as I know, the default setting of XCalibur is to remove lock masses and reference signals from the mass spectra. Could the authors check again?

Response: We agree with the reviewer and rewrite this sentence to avoid confusion.
The revised sentence is: "Instead of excluding the lock masses from the spectra (i.e., the default setting in Xcalibur), Orbitool treats the lock masses as normal signals since the abundance of the reagent ion is usually a critical parameter to characterize the performance of a CI-mass spectrometer."

6) P10L4ff: It remains unclear whether isotopic patterns of candidate formulae are considered in the formula assignment procedure. If they are not considered so far, I would suggest to include this in future versions of the software. Even if the isotopic signals might deviate at low concentrations from the expected intensity, it would still be reasonable to check for the presence of such signals. Could the authors comment on this?

Response: Orbitool has an isotope list for less abundant isotopes and the isotope pattern can be used for peak determination.

To clarify this, we added a new paragraph: "Orbitool also supports the identification of isotopes. The users can adjust the possible isotopes in an isotope list and all the isotopes in this list will be considered. Using the default algorithm for peak determination, the chemical formulae containing a less abundant isotope can be determined only when its corresponding formula containing the abundant isotope is found in the spectrum. In addition, this default algorithm also checks whether the abundance ratio of these two species is consistent with their natural abundances within the uncertainty range. To facilitate studies such as isotope labeling experiments, Orbitool also provides an algorithm that does not restrict isotope abundances during peak determination. However, it is important to mention that the Orbitrap provides non-linear responses when the concentrations of the analytes are very low (i.e., < 1 × 10^6 molecules cm^{-3}, at a 10-minute integration time). As a result, the calculated isotope abundance may deviate from its true value (Riva et al., 2020)."

7) P12L19: Over which time span were the spectra averaged that are shown in Fig. 4?

Response: We added "measured in Shanghai (31/10/2019 between 14:00-14:30)" to the caption of Figure 4.

8) P13L10f: Could the authors provide a list of observed signals and assigned molecular formulas in the supporting material? Moreover, do the authors suggest the presence of gas-phase organosulfates and nitrooxy organosulfates which are among the CHOS and CHONS compounds?

Response: We do not claim that we observed organosulfates and nitrooxy organosulfates. What the molecular assignment is showing is the presence of sulfur-containing compounds (CHOS/CHONS) in the gas phase. Such compounds might be clusters of carboxylic acid, sulfuric acid and amine (e.g., $(C_xH_yO_z)(H_2SO_4)NO_3^-$ or $(C_xH_yO_z)(H_2SO_4)(C_mH_nN)NO_3^-$) as such clusters are expected to exist in polluted environments with high concentrations of precursors (Lin et al., 2019). Earlier studies have also reported the presence of organosulfur compounds in the gas phase (e.g., Ehn et al., 2010). However more work is needed to determine the sources as well as the chemical processes forming such species, but such investigation is beyond the scope of this study. We are currently working on separate studies to characterize the composition and sources of the gaseous species observed during this field campaign.

9) P14L15ff: Several authors are missing in the list of "Author Contributions" (i.e., C.G.,M.E., C.H., and P.Y.)

Response: Thanks. We added their contributions in the revised manuscript.

Technical comments:

1) Abstract / L2: replace "wildly-used" with "widely-used"

2) P3L1: replace "produce" with "emit" (or a similar verb)

3) P6L15: replace "within each time bin" with "for each time bin"

4) P6L23: better use "ion polarity" instead of "charge polarity"

5) P7L1: replace "positively charged spectra" with "spectra from positive mode" (and accordingly for "negatively charged spectra")

6) P8L22: word missing after "this"

7) P10L6: better use "ion polarity" instead of "charge polarity"

8) P10L20 and P11L11: replace "exported into file" by "exported into other file formats"

9) P11L8: replace "These" by "The"

10) P12L14: remove "50th"

11) P13L1: replace "then" by "the"

12) Figure 3: It is difficult to distinguish the colors of the different percentiles

13) Figure 4: typos in the legend ("denoized" should read as "denoised")

Response: We thank the reviewer for these suggestions. The manuscript was corrected accordingly.