We thank the reviewers for their careful revisions of the manuscript which helped improving the clarity and quality of the text. Please find our point-by-point responses below.

I want to thank the authors for the taking the time to address all comments. I have, however, still some concerns, that should be addressed prior to publication in AMT. These concern mainly the decreasing signal intensities during a measurement sequence, that were addressed in previous comments Reviewer1-Comment2; Reviewer1-Comment2; Reviewer2-Comments10-13.

1. I still have a hard time judging whether this observation prevents a reliable quantification for samples with variable matrices or not. During consecutive measurements within a sequence, both analyte signals and total signals seem to decrease. The authors describe the effect and did characterization experiments to well characterize the decrease of the signal and correct for it, which is good. But the reproducibility is unclear to me. In Figure S8, the Lockmass signal decreases by a factor of <2 over the course of 19h, but from Figure S7, the same compound signal decreases by a factor of ~5 over 25h. On the other hand, on two different days, the signals seem very comparable, as shown in figure R3 in the authors response as well as in figure S5 of the SI. Are these experiments comparable and what might explain the difference?

To ensure reproducibility of the signal between sequences injected on different days, a source cleaning is performed systematically before each sequence. Source cleaning together with purge and equilibration steps contributes to the signal recovery and comparability as observed in Figure R3. The intensity of the signal is also checked systematically on our reference compound (also called “lockmass”), to ensure there is no issue with the mass spectrometer detection before sample injection. After purging and equilibration, the mass precision of the instrument is assessed via injection of a quality control mixture (SST solution, provided by Waters). Then, the experiments are comparable although of the instrumental signal decrease during the sequence as the signals were further normalized to the internal standard, which accounts for that lost (Figure S9). We believe that the difference highlighted by the reviewer is due to a sensitivity variation of the instrument over time. This is better observed on the following question.

Regarding the concern of the reviewer due to time variations in Figures S7 and S8, we would like to clarify that they belong to different tests with Leucine-Enkephalin. First, the Lockmass, which is a solution of pure Leucine-Enkephalin at 100 pg/µL in a mixture of Acetonitrile/Water/Formic acid 50/50/0.1%, is continuously infused in the source of the mass spectrometer, at 15µL/min, in parallel to the flow coming from the UPLC. Lockmass signal is recorded every 5 minutes for the full length of the analysis sequence and serves as mass calibration correction in real time. The second Leucine-Enkephalin test is part of the SST mixture, which serves as quality control. In this mixture, Leucine-Enkephalin is at 2.5 µg/mL in Water/Acetonitrile/Formic acid 95/5/0.1%. As explained in the main manuscript, Figure S7 represents the signal measure for the SST injected before and after a sequence of 34 samples (approx. 25 hours), while Figure S8 shows the Lockmass signal, which was directly infused in the mass spectrometer.

Lines 146-147 were added in the manuscript: Between sequences, a source cleaning step was performed to increase the instrument sensitivity (as detailed in Section 2.5).

2. It seems to me that the method itself causes the reduction of signal intensity. Have the authors tested the signal intensity with the cleaning step, that was mentioned in the response to Reviewer1-Comment1?

We discarded that the reduction in signal intensity over time was method dependent as such behavior was observed during samples analysis using the method described in the main manuscript, SST injected using a different elution method and the direct infusion of the Lockmass solution. It is an instrumental issue, specific to the Vion series.

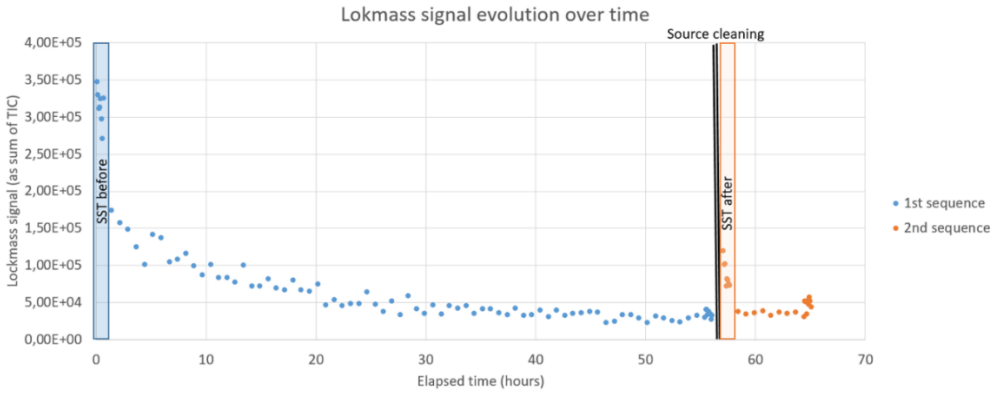

Figure R1. Lockmass responses as TIC during sequences injected at two different days (SST sequence then samples) and a source cleaning step in-between.

We have tested the signal variation between different sequences (injection days) for the stability of the Lockmass with a source cleaning step between them. As observed in Figure R1, the signal increases after the cleaning step, without necessarily recovering the same absolute intensity as before and decreases again during the sequence, following a similar behavior at the beginning of each sequence, independently of the total analysis method (extraction, elution, detection). To overcome this, the sequences analysis was selected in a period where we can ensure the instrument response and performed the time dependent calibrations. A new calibration is performed for each sequence.

3. More to that point, how did the authors exclude that the compound-specific decrease of the signal is matrix dependent?

The matrix effect is not excluded in this work, as it can explain for example, the behavior of nitrophenol compounds as highlighted in the manuscript. However, we considered that the signal decrease is not only matrix dependent as during the Lockmass and SST tests the signal decrease was also observed (Question 2). Compound-specific signal variability can also be explained by the differences in ionization efficiencies of different compounds in ESI.

5. Where all measurement sequences performed in the same way, with the samples at the same position in the sequence to apply the correction of the respective replicate or how exactly was the correction factor applied here?

Thank you for point this out. We would like to clarify that there is not a correction factor applied between sequences or samples. During each sequence, the calibration and samples were injected by triplicates. Each measurement was performed in the following order by triplicate: Calibration (from less concentrated to more concentrated) in the same positions and samples (in randomized order). The samples were randomized to take into accounts experimental variability bias.

6. The authors point out, that the drop of sensitivity is a known problem with waters systems, but the provided link refers to an actual technical issue with the instrument that can be addressed and fixed. Was that specific issue occurring with the instrument used in this study and has it been fixed?

Signal decrease over time is indeed a known instrumental issue in Waters MS instruments. A parameter is implemented to correct for this issue for some instruments of Waters (like Synapt, or Xevo, which work with MassLynx software). The link below we provide mentions a software-implemented solution with the addition of the “Automatic Detector Check” function. However we did not observe any improvement of the sensitivity loss when activating the automatic detector check, because the function does not work with our version of QTOF (Vion), as confirmed by Waters Service Engineers. Unfortunately, there is no possible software update that could solve the problem for our version of the QTOF.

https://support.waters.com/KB_Inst/Mass_Spectrometry/WKB99320_How_to_enable_or_disable_Automatic_Detector_Check_on_a_SYNAPT_G2-Si

Minor comments:

I would suggest to label plots in figures with multiple plots to better link the figure caption to the individual plot

Figures and captions were modified in the manuscript as suggested.

L.92 and some more, I would suggest “Time-of-Flight”

Modified as suggested by the reviewer

L.100 should read “first-time”

Corrected

L.333 should read: “Stability tests were performed”

Corrected