

Reply to anonymous referee #1:

The referees effort to provide this report is appreciated. The review contains fundamental statements, some genuine concerns, and many minor remarks. As the genuine concerns are most important, I will address them first.

Reply to genuine concerns:

The reviewer is concerned that PTRwid may not be useful to process data recorded at 5-10 Hz, which is typically needed for Eddy-covariance analysis. I greatly appreciate that this point has been brought up because PTRwid can process such datasets without problems. Clearly this has to be unequivocally stated in the revised version. For example, in Park et al. (Atmos. Chem. Phys., 13, 1439–1456, 2013) we used 5 Hz data and saved a new file every 3 minutes. In order to ensure optimum performance SumSpectra of neighboring files are added until the specified maximum signal is accumulated (see second paragraph of section 2.1.6).

Library: As specified in Appendix D, the ion library contains more than 40000 ions in the extended mode ($C_{1-40}H_{0-82}O_{0-16}N_{0-2}$), which includes virtually all possible clusters with water molecules (e.g. the acetone-water cluster as $C_3H_7O_2H^+$ ion). Fragmentation is not considered at all since PTRwid does not aim at attributing chemical compounds. PTRwid attributes empirical ion formulas to the monitored peaks. For example, $C_5H_8H^+$ is attributed to m/z 69.070. Whether this signal originates from protonated isoprene or has to be attributed to a fragment of MBO is irrelevant for the attribution. In the revised version the library will be consistently referred to as ion-library and not as compound-library in order to avoid such misunderstandings.

Reply to fundamental statements: (reviewer comments in blue to provide the context)

Actually, the approach presented here to analyze PTR-TOFMS raw data is not unique as it follows another approach that has been already published (Müller et al. 2013)... Moreover, it is not clear what is the real advantage to use the approach presented here rather than the preexisting one of Müller et al (2013), and why there was the real need to develop another software tool instead of implementing the one already in use.

PTRwid is unique. The uniqueness is illustrated with several examples in the introduction (page 2, line 24 to page 3, line 17). PTRwid has not been developed in response but parallel to the other tool. The key components were already operational and used in our 2010 publication (Holzinger et al., Atmos. Chem. Phys., 10, 2257–2267, doi:10.5194/acp-10-2257-2010, 2010).

I suggest the author to consider publication in a more specialist journal that takes into account the development of software tools as a subject area of interest, after addressing the concerns risen above and the points listed below.

The 'aims and scope' page of AMT specifically lists 'techniques of data processing' as main subject area. So this work is clearly within the scope of AMT and in my opinion no other journal needs to be considered.

As a consequence, it should have been very interesting (and very useful for the scientific community) to compare the performances of the preexisting approach developed by Müller et al (2013) with those of the approach presented here by Holzinger, especially when these two different approaches have to post-process the same raw dataset recorded by same PTR-TOF-MS.

This will be a useful exercise, but I feel that this is beyond the scope of the current work.

Reply to minor remarks: (reviewer comments in **bold blue** to provide the context)

- **The manuscript includes too many figures (10 !!!); as some of those appear more like 'screenshots', they could be either merged in a few figures or moved to supplementary information (where there is already a Figure B1, oddly without a Figure A or B2. . .); by the way in Fig. 3 panels are missing.**

I see no arguments why any of these figures should be redundant. All the figures contain very specific information.

- **Line 10, page 2: it is written: "attributed peak", but attributed to what?**

Continue reading in line 11: "attributed peak mass"

- **Lines 15-19, page 2: I would simply say: "PTR-TOF-MS technology records full mass spectra with high mass resolution that allow differentiating between ions having the same nominal but different exact mass weight".**

Yes, this is point '(i)' made in the paragraph. But it should be realized that point '(ii)' (lines 16-18) is equally important.

- **Lines 4-7, page 3: actually the approach developed by Müller et al. (2013) shape the peaks on the basis of certain 'ions' that are always produced by the PTR-TOF-MS, as the primary ions m/z 21.022 (H₁₈₃₀⁺) and m/z 39.033 (H₂₀-H₁₈₃₀⁺) that do not saturate the detector and result isolated from other peaks (and thus easy to define and model).**

Exactly what I state in lines 4-6 on page 3: Mueller et al. retrieves the peak shape from selected peaks that can be entirely attributed to ions of the same composition...

- **Line 4, page 4: the author should consider that "single lab measurements" may run for time spans < than hours! (Brilli et al. 2011).**

See reply to genuine concerns.

- **Line 10, page 5: What "drifts in these parameter" the author does refer to?**

Most important are the parameters for the mass scale calibration (will be specified in the revised version).

- Line 14-15, page 5: The author should demonstrate/show to what extent the code presented here can give good performances even though analyzing files recorded for shorter time periods that 30 minutes.

See reply to genuine concerns.

- Line 17, page 5 and line 15 page 6: which are the “data containing parameter”?

Language will be clarified. PTRwid creates files containing, for example, the parameters for the mass scale calibration.

- Line 2, page 6: the author should specify what are the “scans” for?

‘scan’ is a verb in this context. The routine scans through the SumSpectrum. Why? The section is about detecting peaks.

- Line 8, page 6, line 19-20, page 6: the author speaks about “6 times”, “range of 20-10000”, “8 times”, but how all those thresholds have been defined and optimized?

These thresholds have been optimized to minimize false peak detections while keeping the number of overlooked peaks as low as possible.

- Lines 10-13, page 7: the author should make clear what the parameters “a, t0, m, x” refer to. In particular, the parameter “x” has been reported as “EX” from here to the end of the manuscript! The author should carefully double-check the manuscript to avoid this and other “typing” mistakes occurred throughout the manuscript.

All parameters are explained in the specified section. Also, there is no typo misspelling ex for x. So, this is a curious comment.

- Lines 23-24, page 8: please rephrase “. . .the corresponding TOF-bins of the corresponding. . .” as it sounds very bad.

Will be corrected.

- Lines 12-13, page 9: why and how the values of “90 nanoseconds” and “7 times” have been chosen?

These values work best to reliably retrieve the baseline.

- Line 8 and line 17, page 9: consistency is required as first it is written “Figure 4” and then “Fig. 4”.

Rather a (minor) typesetting issue...

- Line 7, page 10: something is missing within the sentence “. . .we consider all peaks the maximum signal. . .”

Continue reading in line 8 will complete the sentence. The missing part is a comma before 'we'.

- Lines 9-10, page 10: again, the author should explain how/why values of "800 counts" and "10 times" have been chosen.

The requested explanations are in lines 11-13, page 10: "The lower boundary is set to avoid distortion of the peak shape due to poor counting statistics, and the upper boundary is set to avoid distortion by the largest signals that cause saturation in the counting electronics of the mass spectrometer." The factor of 10 implies that signals 10 times the minimum signal do not exhibit saturation effects.

- Lines 5-19, page 11: this part is puzzling and difficult to follow and should be rephrased (avoiding such inappropriate wording as "routine executes all routines", in line 13).

Will be considered.

- Line 21, page 11: the abbreviation "ppm" refers here to "TOF bin width" but it confounds with another very similar abbreviation indicating concentrations (Part Per Million); I suggest the author to use a surrogate for such index so to avoid misunderstandings.

The use of ppm is common also in this context.

- Line 28, page; line 26 page 14. . . : the author mentions here and many more times throughout the manuscript about a Quadruple-PTR (namely PTR-MS or TD-PTR-MS), that do not require a mandatory post-processing step to elaborate the data as the PTR-TOF-MS. So why has the PTR-MS or TD-PTR-MS has been mentioned?

There is no reference to a quadrupole PTR-MS. In the revised version we will consistently use PTR-TOF-MS. TD-PTR-MS refers to our thermal desorption setup (also with a TOF-MS!) and is correctly referenced.

- Lines 15-19, page 14: what exactly are the numbers "487" and "1350" used by the author as constant multipliers? (nevertheless, in English language is "multiplying by" and not "multiply with").

These numbers result from the ratio between the stable oxygen isotopes ^{16}O , ^{17}O , and ^{18}O . $^{16}\text{O}/^{18}\text{O} \approx 487$ and $0.5 \times ^{16}\text{O}/^{17}\text{O} \approx 1350$.

- Line 16, page 14: it written "our off-line TD-PTR-MS method", but who is the other since Holzinger is the only author of this manuscript?

'Our' refers to the reference Timkovsky et al. that is given in the same line (line 26!)

- Line 20, page 15: the author refers to "engineering data", without providing any explanation of what kind of "engineering" has been done.

The term 'engineering data' refers to all data that are saved along with the actual TOF-spectra. I.e. voltages, pressures, temperatures, valve positions etc..

- Lines 3-26, page 17: the author should clarify the differences between the “attribute formulas” presented here and the one already (published) available on the web (<http://www.chemcalc.org/>)

I am not familiar with all features of chemcalc.org. The main feature of the attribute formulas’ tool is to check for consistency between the provided dataset and the natural abundance of the two stable carbon and nitrogen isotopes.

- Line 21, page 17: it is not clear the connection between “the measured signal” and the “attribute formulas” described in this subsection.

I do not get the reviewers point..

- Lines 19-21, page 18: this confirms my concerns about the application of this software tool in long-term (many months) eddy covariance flux measurements of VOC because the time consumed to open the big files will add to that required to post-process the same file thus further slowing down the data analysis and make it unsustainable.

See reply to genuine concerns.

- Lines 14 and 19, page 21: what is it “menue”?

‘Pulldown menue’ is a widget element of PTRwid. By clicking on the element an option can be selected from a list of options.

- Line 18, page 23: “the Utrecht PTR-TOF-MS” is different form those manufactured by Ionicon in Innsbruck (Austria)?

The Utrecht instrument is an Ionicon instrument. But flexibility in adjusting some parameters is not only useful for different instruments, but may be also required to account, for example, for instrumental conditions such as MCP ageing or TOF tuning, or to optimize for different types of measurements.