We thank the reviewers for taking time to read the article and their valuable comments on our paper. To facilitate the revision process we have copied the reviewer comments in black text. Our responses are in regular blue font. We have responded to all the reviewer comments and made alterations to our paper (**in bold text**)

Anonymous Referee #1

Received and published: 24 April 2016

This paper discusses the application of high resolution peak fitting commonplace in HR-AMS analysis to the TOF-ACSM. While this instrument is not designed to produce data over UMR (like the Q-ACSM), it may be possible to obtain some information and will be of much interest to those operating such instrument. The paper employs the current state-of-the-art regarding peak resolution within the AMS community and critically assesses its ability to deliver usable data. Recommendations for data analysis and future instrument modifications (e.g. addition of a m/z calibration source) are given. The paper is methodical and well written. I have a few comments, but these are mostly minor. This paper is well suited to AMT and will provide a useful reference for future data inversion. It may also have applications to other mass spectrometric techniques.

Comments:

In general, there are a lot of figures, possibly more than is needed to support the main article, so some could probably be moved into the supplement. I would also observe that only about half of the figures are referred to in the text, which might imply that some of them are not vital. However, some of the methodological figures (e.g. 1 and 2) should be retained, as these help to illustrate the specific steps described.

Following the reviewer's recommendation, we propose moving the following figures to the supplementary information:

Figure 3: ToF-ACSM vs high-resolution MS for W+ isotope pattern. Very minor edits to the text required.

Figure 5: Compensating position of calibration peaks. Reference to figure in text altered.

Figure 8: Fitting of calibration fit residuals. Since this chapter was removed (see later discussion), the figure is no longer in the manuscript.

Figure 17: Sensitivity of fitted As2+ intensities to mass calibration. Reference to figure in text altered. Figure numbers are updated and all figures are referred in text using new numbers.

Page 5 line 3: I am not sure what is meant by 'background ions' because ions are unlikely to persist in the electric fields of a time-of-flight mass spectrometer. If the authors are referring to scattered ions or electrons, they should specify this.

AR: The point here is to compare the signal of the tungsten isotopes with that of the surrounding ions. We clarify the text thus: **the surrounding peaks, which we will refer to here as "background", that is the signal at the mass to charge ratios surrounding the calibrant ion signal**

Page 5 line 10: More detail should be given here, specifically the low-pass filter algorithm and parameters and the actual width of the box used. I would note that step ii could leave a slightly positive offset in peak-free data depending on the amount of noise, so the authors should comment on whether this is an issue.

AR: Following text was added to the manuscript: **By default the width of the box is set to 40 mass spectral data points (50 ns). It is possible that, in regions adjacent to large peaks with significant gradients, the box method generates a baseline slightly lower than the optimum. However, the at the peak most significantly impacted by this effect, NO^+, the negative offset is observed to be at most 5% of peak height and typically lower. This offset is also removed during calculation of the difference spectrum, although it does contribute to a slightly higher imprecision. The low-pass filter (Zahradnik et al., 2004) used in step (i) removes by default frequencies higher than 16 GHz.**

For details on the low-pass filter algorithm utilised by the Igor environment in which the ToF-ACSM data analysis software runs, please see Zahradník, P., and M. Vlcek, Fast Analytical Design Algorithms for FIR Notch Filters, IEEE Trans. on Circuits and Systems, 51, 608 - 623, 2004.

Page 6, line 11: An example should be given for step iii

AR: The software automatically plots the peaks defined in the steps i-ii. The user can choose the representative peaks for the step iii. The plots of peak shape (below) in the SMEARIII campaign were added to the supplement (Fig S2).
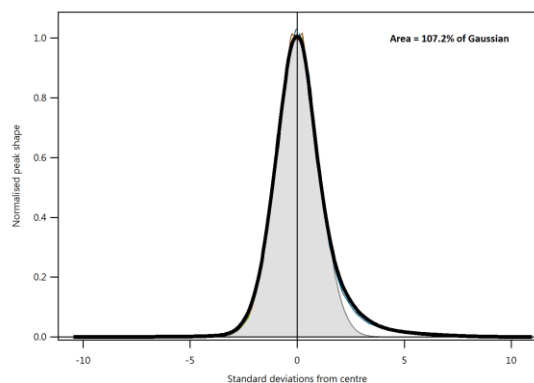
Fig S2: The observed non-Gaussian peak shape calculated using a representative mass spectrum from the SMEARIII campaign data.

Page 9, line 8: 'Mass defect' is probably more appropriate than 'mass excess', as these can be negative.

AR: We chose to show mass excess rather than defect because, as in Kendrick's original publication (Kendrick, E.: A mass scale based on CH_2 = 14.0000 for high resolution mass spectrometry of organic compounds. Anal. Chem. **35**: 2146–2154,1963) describing the well known "Kendrick plots", mass defect is rather confusingly negative for masses with a positive offset from the nominal value. We find mass excess thus more intuitive.

Page 16: The ability of the AMS to detect sea salt is not universally well established. The results here should be compared against Ovadnevaite et al. (2012, Doi10.1029/2011jd01737).

AR: Following text was added to the manuscript: **Ovadnevaite et al., 2012 demonstrated using laboratory testing and ambient data that seasalt can be quantitatively measured with HR-ToF-AMS. They demonstrated that altogether 36 high resolution ions contributed to the sea salt fragmentation pattern, but only 7 ions had a significant influence on the derived aerosol mass, namely $^{23}Na^+$, $^{35}Cl^+$,$H^{35}Cl^+$,$^{23}Na^{35}Cl^+$ and their isotopes $^{37}Cl^+$,$H^{37}Cl^+$, $^{23}Na^{37}Cl^+$. In clean Antarctic conditions with minimum interference from organic ions, the ToF-ACSM can detect $^{23}Na^+$, $^{35}Cl^+$,$H^{35}Cl^+$, $^{23}Na^{35}Cl^+$ ions and their isotopes.**

Page 20, line 5: There is no need to refer to 'poor woman's PMF' by name; the technique has had other nicknames over the years (e.g. 'poor person's PMF', 'marker method'), so it is probably best just referring to the paper.

AR: Agreed. The name "poor woman's PMF" is removed and instead a reference to corresponding paper is used.

Anonymous Referee #2
Received and published: 30 March 2016

The manuscript describes in detail the methodology used for separation of isobaric ions using constrained peak-fitting in data from the ToF-ACSM and thoroughly discusses good practices and limitations of the procedures applying them to three datasets recorded in different environments. The manuscript is well written and structured, and it falls within the journal's main subject area "validation of measurement techniques". I recommend the publication of the manuscript in AMT, after addressing the following issues:

GENERAL COMMENTS
The manuscript could be more specific on the advantages of the methodology with respect to using the unit mass data together with the fragmentation pattern in order to determine mass concentrations. The manuscript spends several pages discussing the application of the peak-fitting to ammonium, nitrate and sulphate ions, concluding that it is not possible to deconvolve the ammonium ions and that it is possible to do it with the sulphate and nitrate ions; the implications or importance of this conclusion are not explained. On the other hand, the cases for which the peak-fitting represents a real advantage (HOA/OOA constants and the oxidation state) are not as thoroughly explained and discussed.

AR: In order to extract reliable information from high-resolution peak-fitting, the results need to be verified through careful sensitivity analyses and comparisons to established methods such as the formulation of species intensities according to fragmentation patterns applied to peak integrations. The process of investigating peak-fitting uncertainties is, as this article shows, complex and requires significant care and attention. A large part of this article is thus dedicated to describing this process. In contrast, the calculation of the estimated oxidation state (based on Stark et al. 2015) and revised

HOA/OOA constants (based on Ng et al., 2011) is straightforward and therefore requires less discussion of the details, even if the results are important.

SPECIFIC COMMENTS

SECTION 3.2.4. The discussion in this section is very vague. The bias described in this section, as well as the methodology to compensate for it should be explained in more detail for the benefit of the reader. Several questions remain unanswered after reading the section. For example: a) "In certain cases, a bias exists in the mass calibration fit residuals..". Which cases? How often is this bias observed? Why is this effect observed? b) "..apparent polynomial shape in mass-to-charge space, as exemplified in Fig. 8." How many terms did the polynomial used have? How was the number of terms chosen? c) "A MS with known calibrant ions in the central mass range exhibiting strong mass spectral peaks can thus be used to minimize the residuals, and this subtraction applied to all calibrations in the entire dataset time-series". Which subtraction is this paragraph referring to? How is the subtraction performed?

AR: As the technique referred to in this section is only rarely appropriate for application to ToF-ACSM data, and given the reviewer's justified comments about the vague discussion, the whole section 3.2.4 was removed from the manuscript, concentrating the calibration discussion on the other, more relevant, issues.

SECTION 4.2 (Nitrate) Compare mass concentration results using the integrated data together with fragmentation table vs. the peak-fitting, as it was done for sulphate.

AR: The following figures showing integrated signal vs peak fitting results for nitrate were added to the supplement as suggested.
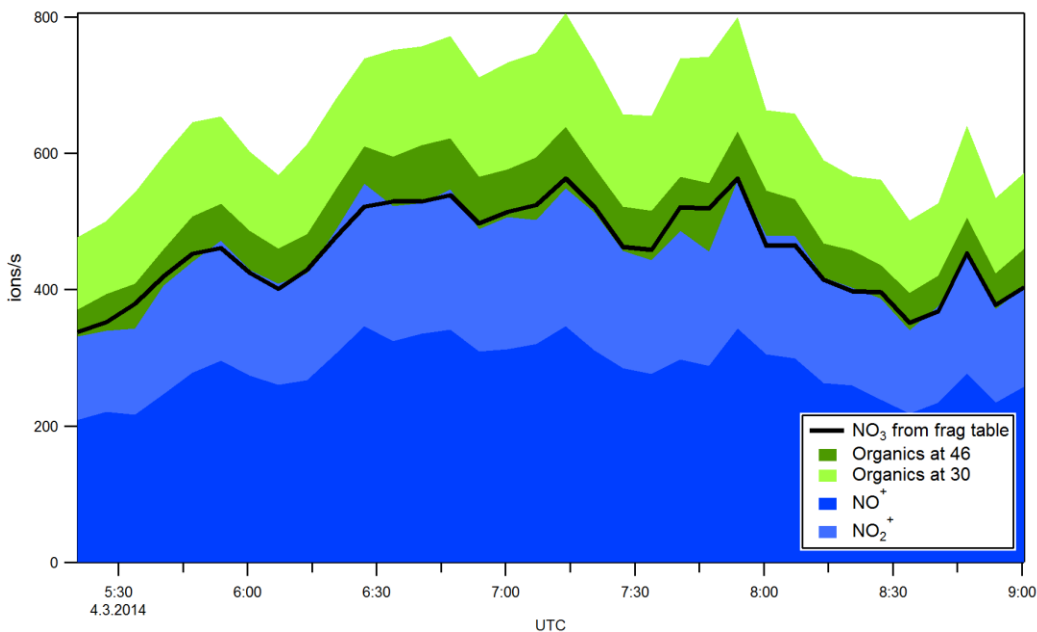
Fig_S7: Time-series of the nitrate intensity derived from application of the fragmentation table to integrated (UMR) data compared to the sum of the fitted intensities for the nitrate ions NO^+ and NO_2^+ and their isobaric organic interferences. On average, the sum of NO^+ and NO_2^+ represents 102.7% of the integrated signal.

PAGE 20, LINE 20. "Stark et al. used the 20 higher mass resolving power of their data together with comprehensive ion lists to analyse the mass excess of all the fitted ions using the constrained peak-fitting methods described in section 3. A mass resolving power of 500 precludes this approach, so instead, the simpler bulk mass spectrum method also proposed by Stark et al. was employed." It is not clear what is the difference between the method originally used by Stark et al., and the simpler method used for this work. Can the authors better explain it?

AR: In addition to utilising the fitted peak intensities and positions, Stark et al. showed that a weighted average of mass excess values of the mass spectral data points themselves was also a valid technique for estimating the oxidation state. We add a clarification to the text:

In this method, the average mass excess is calculated from the individual mass excesses of the mass spectral data points, weighted by their respective intensities.

The manuscript describes a way to extract the signal of selected ions from an Aerodyne TOF-ACSM whose signals are disturbed by isobaric ions. This is done by fitting peaks with a custom shape to the expected isobaric ions. Since the position and the shape of the peaks are known, the only parameters to be fitted are the heights of the peaks. The results are used to calculate refined time series of selected compounds.

An important feature of the publication is the discussion of the sensitivity of the results to uncertainties in the mass calibration, which covers the influence of counting statistics and isobaric ions. Furthermore the limitations of the method with respect to the compounds that can be analyzed are presented as well as techniques to evaluate the quality of the results. This is very important to allow other users to use the method appropriately.

As a technical paper in the field of aerosol science, the paper fits well in the journal Atmospheric Measurement Techniques and I recommend to accept the manuscript for publication after addressing the influence of the peak shape and after adding a comparison to a HR-TOF-AMS.

GENERAL COMMENTS

An important factor that is not covered is the influence of the peak shape on the results. The authors assume a fixed peak shape without showing that the peak shape is really constant over the whole spectrum, nor do they show that the influence of an instability of the peak shape is not relevant to the results. A possible analysis for the sensitivity to the peak shape would be to split the mass spectrum into two sections A and B and determine the peak shape for these sections separately. The different peak shapes could be compared to determine how stable the peak shape actually is. In a second step, the peak shape of section A could be used to perform the deconvolution for an ion signal in section B. The resulting time series could be compared to those using the peak shape of section B.

AR: The measured imperfect peak shape arises as a result of the energy distribution that the ions take in the extraction region of the mass analyser. As electrostatic fields are used to guide the ions, ions of equal energy/charge thus behave identically in the TOF and a consistent peak shape can be expected provided all ion species exhibit the same energy distribution. In certain cases, such as expansion through a nozzle under vacuum or delayed extraction (Peregudov et al., IJMS, 2010), one observes rather a constant velocity than constant energy but neither of these are relevant to the ToF-ACSM. Only the central section

of the ion plume at the source is extracted thus eliminating those ions with energies in the tails of the distribution. As a result, all ion species across the mass-to-charge range are generally observed to exhibit the same energy distribution in the extraction region - and hence a consistent peak shape. This is importantly not the case for ions whose origin does not lie within the source plume, for example surface ionisation of K+ on the ToF-ACSM heater.

The assumption of consistent peak shape has been widely employed in the atmospheric community (e.g. Cappellin et al., 2011); nonetheless this does not prove its validity. A comprehensive study of the impact of a varying or incorrectly-determined peak shape would be a useful addition to the literature but arguably something that ought be approached in the first instance using simulations. For this study, we rather prefer to answer the reviewer's points directly and demonstrate that the assumption of consistent instrumental peak shape is justified using the ToF-ACSM datasets studied in this work.

Peregudov, O. N., and Buhay, O. M.: The peak shape model for magnetic sector and time-of-flight mass spectrometers, International Journal of Mass Spectrometry, 295, 1-6, 10.1016/j.ijms.2010.06.009, 2010.

To support that "the peak shape is really constant over the whole spectrum":

Firstly, we propose the addition of the following figure to supplementary information. In order to be able to unequivocally state that the peak shapes studied represent isolated ions, we use a calibration mass spectrum where SF6 gas was injected at the ioniser and compare the peak shapes of the SFx fragment ions. No obvious trend moving from low to high mass-to-charge is visible:
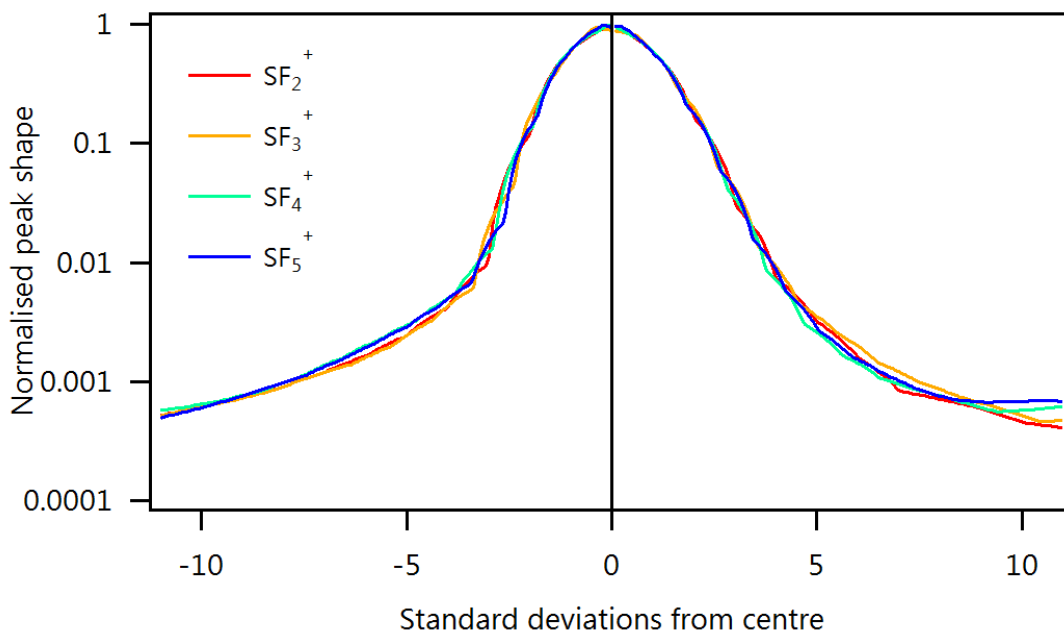
Figure S4: Normalised peak shape for various ions across a calibration MS with SF6 gas: peak shape is consistent for the different SF6 fragment ions and shows no obvious trend with mass-to-charge ratio.

Secondly, the further addition of the following figure would help support the claim that the peak shape is invariant with time (provided voltages are not adjusted). A comparison of the peak shape of the N2+ ion shows negligible change over nearly a month of measurements:
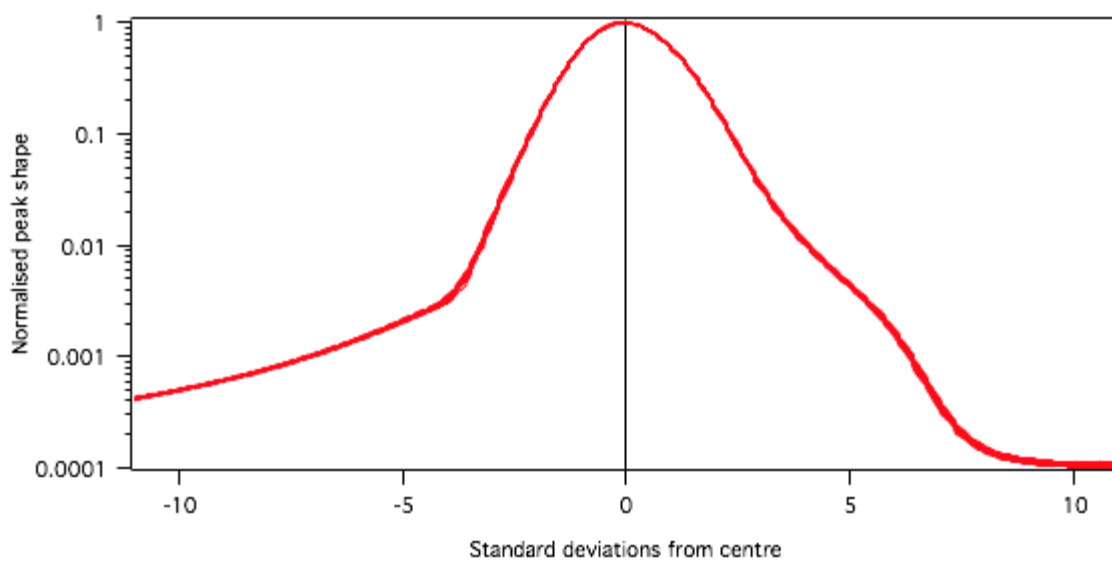
Figure S3: Normalised peak shape for the N2+ ion peak at 28 Th from 25 daily averages representing nearly a month of acquired data with consistent TOF tuning (voltages). No significant changes are observed over this time period, such that the individual traces are barely visible.

To quantitatively demonstrate that the "influence of an instability of the peak shape is not relevant to the results" is somewhat more challenging. Arguably an incorrectly-defined peak shape (i.e. user error) would lead to larger errors than any instrumental instabilities. The questions are posed: What should be changed in the peak shape and how should the impacts of those changes be quantified? For this reason we argue that a general (not ToF-ACSM specific) simulation approach is the preferred method to address this question. Nonetheless we propose that the following additions to this study to alert the reader to this issue:

Section 3.1.2.2. Peak Shape, add following text (relating to the figures above):

The assumption of consistent peak shape has been widely employed in the atmospheric community (e.g. Cappellin et al., 2011); nonetheless this does not prove its validity. It is importantly not the case for ions whose origin does not lie within the source plume, for example surface ionisation of K+ on the ToF-ACSM heater. User error - i.e. an incorrect definition of peak shape - is also a potential source of error and arguably of greater probability and magnitude than any potential instrumental artefacts. A comprehensive study of the impact of a varying or incorrectly-determined peak shape would be a useful addition to the literature. However, to highlight the validity of the invariant peak shape assumption for this study, we i) verified that the peak shape was invariant over a large mass range by comparing the shape for known isolated ions in a calibration measurement where SF6 gas was injected at the ioniser (see Figure S2), and ii) verified that the peak shape was invariant over time by comparing the shape of the N_2^+ peak for nearly a month of data (Figure S3).

Added to references:

Cappellin, L., Biasioli, F., Granitto, P. M., Schuhfried, E., Soukoulis, C., Costa, F., Mark, T. D., and Gasperi, F.: On data analysis in PTR-TOF-MS: From raw spectra to data mining, Sensors and Actuators B-Chemical, 155, 183-190, 10.1016/j.snb.2010.11.044, 2011.

Section 4.2. Nitrate, add following text as penultimate paragraph:

It is noted that this assessment of the uncertainties assumes an invariant and correctly-defined peak shape. Inaccuracies in the peak shape arising from either instrumental artefacts or user error in its definition would propagate into errors in the fitted intensities discussed above. Although a quantitative assessment of these effects is outside of the scope of this study, we illustrate the impact of an incorrect peak shape by comparing the fit results for NO+ and its interferences using the user-defined and Gaussian peak shapes. For a representative MS from the SMEAR dataset, we observe that 73% of the area is assigned to the NO^+ ion using peak-fitting with the user-defined shape. Using Gaussians, this increases to 87%, a relative change of +19%. Whilst such a large inaccuracy in peak shape would not be expected (and indeed the fit residuals are large and apparent using Gaussians), this highlights the importance of carefully defining the peak shape as described in section 3.1.2.2.

What is also missing is a real comparison of the results from the deconvolution of the low resolution spectra to the results from a HR-TOF-AMS. Such a comparison could be used to directly validate the results from the presented method. The presented examples, although very useful to understand how to properly apply the method, only show that the results are plausible. A comparison to high resolution data would directly reveal the quality of the deconvolution. (See specific comments)

AR: The authors agree that a comparison to HR-ToF-AMS could be an interesting addition to the article. However, the focus of this article is to study applicability, uncertainties and limitations of high resolution peak fitting to *ToF-ACSM* data. As such, peak-fitting results need to be verified through careful sensitivity analyses and comparisons to established methods such as the formulation of species intensities according to fragmentation patterns applied to peak integrations. We argue that an instrument intercomparison does not fit within the scope of this paper. It would be extremely hard to assess whether observed differences in fitted intensities between a ToF-ACSM and HR-ToF-AMS dataset arose from peak-fitting errors, instrumental artefacts, mass calibration imperfections, IE calibration, baseline errors, thresholding errors and so on. We therefore disagree that an intercomparison would necessarily reveal the quality of the results as the reviewer suggests. Nonetheless, we agree with the reviewer that the intercomparison of intensities calculated by integration and peak fitting from multiple co-located instruments should be a focus of analysis of future data analysis. For example, the recent ACTRIS campaign with six ToF-ACSMs and one HR-ToF-AMS could represent an excellent dataset for this

purpose. However, this undoubtedly lengthy discussion merits outright publication of its own paper and would detract from the central message - that one needs to take care in setting up peak fitting parameters and address potential uncertainties in the results - were it to be added to this (already fairly lengthy and technically detailed) manuscript.

Specific comments:

3.1.1 The baseline calculation assumes that the baseline is always reached between peaks. This is surely not the case at high m/z. Although this is not relevant in the mass to charge range covered in the manuscript, it might be worth noting when the method is described.

AR: This is a very valid point and the following text has been added to the manuscript: **We note that sometimes the baseline might not be reached between large peaks, especially at higher m/Q range. For ToF-ACSM mass range (up to m/Q ~200) this is however not an issue.**

3.1.2.2 This section completely relies on the peak shape being constant over the whole mass range. I doubt that this is the case, especially when looking at Fig. S11. This fit clearly shows that the peak shape does not well represent the peak shape of the signals at this mass. The left region of the fit is significantly different from the fitted curve.

AR: The authors accept that Fig. S11 is indeed an example of a poor quality peak fit. Subtraction of the filter MS, the large MS baseline present in both the ambient and filter measurements as well as smoothing of the signal-to-noise-limited data all contributed to the observed discrepancy. We propose to update the figure with a version that, whilst using a similar integration period, does not attempt to smooth the data and, rather than list the unknown presumed organic interferences (of which there are many possibilities, with exact masses up to ~ 79.15 Th), just fit a single unknown peak alongside the ion of interest, CH3SO2+:
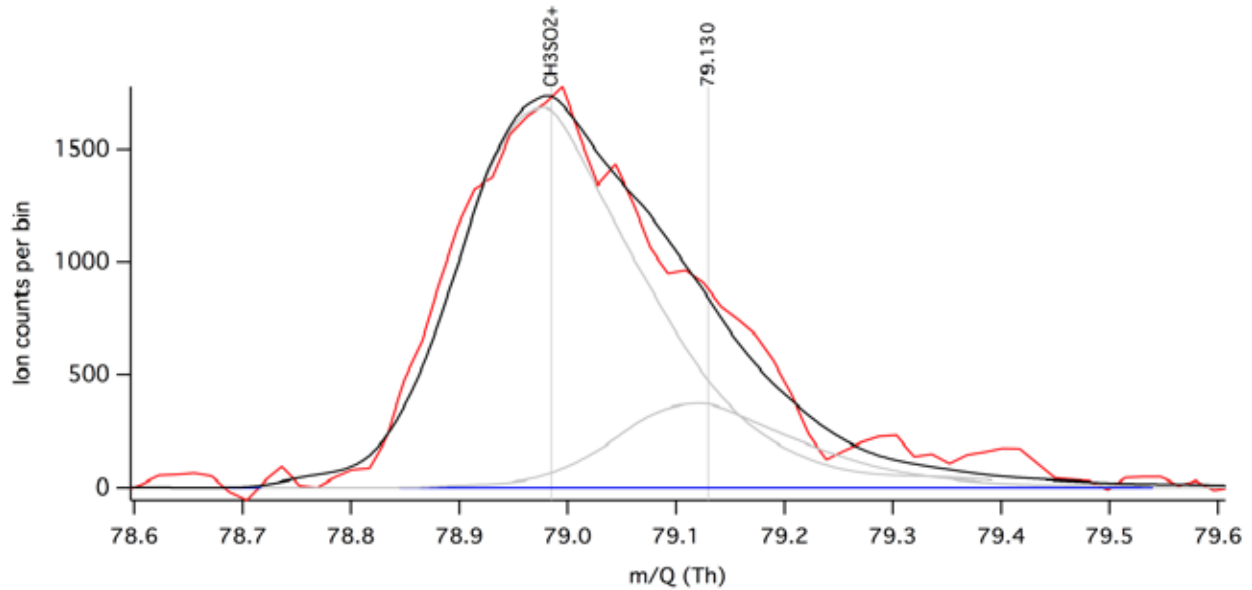
We refer to the discussion presented previously regarding the assumption of consistent peak shape.

3.2 A short comment on why the spectrometer resolving power, although being an intrinsic limitation of the hardware, does not need to be considered would be helpful. It would be beneficial to know to what extend this could influence the result, even if it possibly just means stating that it is irrelevant compared to other influences.

AR: Increased resolution means sharper peaks and therefore, for a given signal/noise ratio, better determination of peak position using peak fitting. However, during data analysis there is nothing that can be done to change the measured peak widths so for this discussion on analysis techniques, it is not a central discussion point. We alter the text for clarification thus:

Since i) is limited by the hardware, it may be considered an intrinsic limitation during data analysis. Whilst narrower peaks would improve the fitted peak positions, the measured peak width is of course fixed during analysis and thus plays no role in the parameters that a user may tune to optimise the mass calibration.

3.2.2.2 Please explain what is meant by "background peaks". It can only be deduced later from the text that it must be the signal on the mass to charge ratios surrounding the calibrant ion signal.

AR: Clarification for the background peaks "**signal on the mass to charge ratios surrounding the calibrant ion signal**" was added to the manuscript.

3.2.4 This is de facto an addition of more parameters to the mass calibration.

AR: The whole chapter was removed as explained in the response to reviewer 2.

3.3 The parameterization of Cubison and Jimenez (2015) assumes the ideal case of a constant peak shape. This is not the case here. "Imprecision in such a constrained fitting procedure may arise from.. (iv) the instrument transfer function ("peak shape")… Since the errors from iv are not included, our results represent a best-case scenario for the precision of the retrieved parameters." (Cubison and Jimenez (2015)) Please add the information that the parameterization represents a best case scenario.

AR: The following text was appended to the paragraph introducing this parameterisation:

It is noted that Cubison and Jimenez (2015) assumed a constant peak shape in deriving this parameterisation, which thus represents the best-case scenario and should be interpreted accordingly.

4.2 Figure 11 shows the time series of the results of the separated ion signals. On the left, the unperturbed fitted signals are compared to the full signals and it is noted that the total signal is well reproduced by the best fit. How does this look like for the perturbed case? Is the reproduction of the total signal by the separated signals much worse in this case? This would be a good indication that the comparison to the total signal is actually sensitive to poor deconvolution of the isobaric ion signals. Such a graph would be nice in the supplemental.

AR: We use SO^+ ion data measured at SMEARIII station as an example of the perturbed timeseries. Figure S14 represents the normal time series of SO^+ and time series using 20 ppm positive and negative perturbation. Figures S3-S7 show the difference between perturbed time series and fitted time series for all ions. For SO+ for the optimum case, the sum of the fitted traces was 3 % of the integrated signal, and then for the perturbed cases it was 3.9 % and 4% for positive and negative perturbations respectively. Figure below was added to the supplement and a reference to the text.
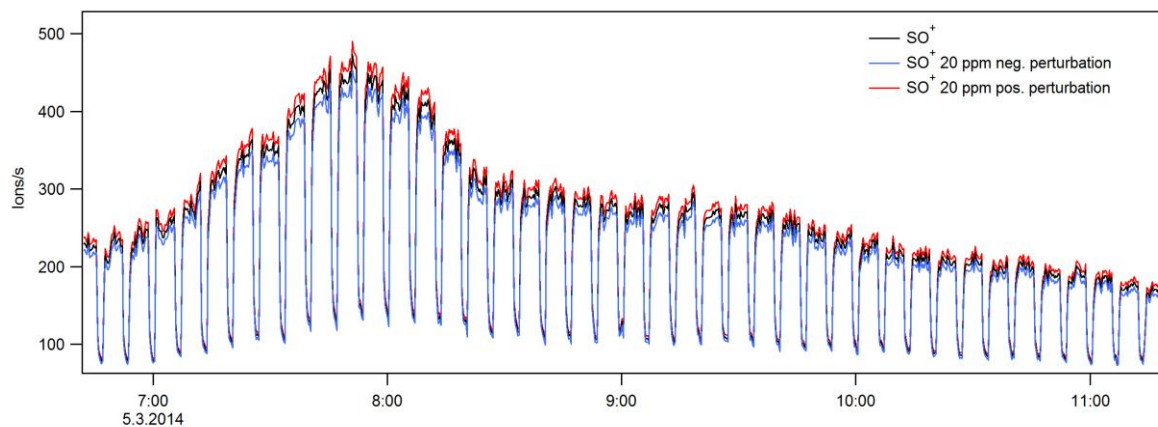
Figure S14: shows an example of positively and negatively perturbed timeseries for SO^+ -ion measured during the SMEARIII campaign. The difference between both negatively and positively perturbed and unperturbed timeseries was on average 4%.

4.8 Second paragraph: Using a transfer function to convert HR-TOF-AMS fitted spectra to ETOF spectra is not a valid verification of the method. The fixed transfer function simplicitly adds the assumption that the shape of the peaks is constant. However, this is an assumptions that needs to be verified.

AR: In the discussion on peak shape above, we demonstrate that the ion peak shape is indeed invariant over the mass spectrum to a level which renders any anomalies insignificant at the side of other sources of error such as mass calibration and counting statistics. However, even if this were not the case and changes in the peak shapes tails of a few percent were observed moving from low to higher masses, this would actually not significantly impact the oxidation state discussion presented in section 4.8. The mass excess values are weighted according to the intensity, so it is those data immediately surrounding the peak centre which contribute most to the average mass excess value. Clearly, if the shape near the peak ($<$ 2 std dev) significantly deviated from that of a Gaussian this could impact the calculation, but we argue that it is clear from the above discussion that this is not the case, nor are the authors aware of any data in the community which shows an invariant peak shape in this region. As such, we argue that the methodology presented in the paper to support the validation of the oxidation state estimation is valid.

Technical corrections:

p2 l2: Atmospheric aerosols

AR: corrected as suggested.

p2 l16: better: in literature p3

AR: corrected as suggested.

l22: Why did you assume a collection efficiency of 0.5?

Previous studies have shown that collection efficiency of 0.5 is very typical for the ambient urban aerosol (e.g. Aurela, M.: Chemical and Source Characterization of Submicron Particles at Residential and Traffic Sites in the Helsinki Metropolitan Area, Finland, doi: 10.4209/aaqr.2014.11.0279, 2015). The collection efficiency could be calculated based on aerosol composition and relative humidity (RH). However, we did not have RH results available. Since the aim of article is to explore peak fitting, the chosen collection efficiency does not influence the fitting or any other conclusions made from the data. Value 0.5 was chosen just to show approximate concentration level in urban background station during the measurements. Following sentence was added to the manuscript: **A collection efficiency of 0.5 was assumed based on previous studies conducted on SMEARIII station (e.g. Aurela et al., 2015)**

p6 l29: an imperfect mass calibration

AR: corrected as suggested.

Table 4: Sensitivity (without -)

AR: corrected as suggested.