International Journal of Mass Spectrometry 306 (2011) 1–8

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

journal homepage: www.elsevier.com/locate/ijms

Enhanced spectral analysis of C-TOF Aerosol Mass Spectrometer data: Iterative residual analysis and cumulative peak fitting

Markus Müller[1], Christian George, Barbara D'Anna*

Université de Lyon, Lyon, F-69626, France; université Lyon 1, Lyon, F-69626, France; CNRS, UMR5256, IRCELYON, Institut de recherches sur la catalyse et l'environnement de Lyon, Villeurbanne, F-69626, France

A R T I C L E I N F O

Article history:
Received 23 February 2011
Received in revised form 26 April 2011
Accepted 26 April 2011
Available online 13 June 2011

Keywords:
Aerosol Mass Spectrometer (AMS)
C-TOF AMS
Data analysis
Cumulative peak fitting
Iterative residual analysis

A B S T R A C T

We present a method for an improved spectral analysis of datasets measured by a C-TOF Aerosol Mass Spectrometer. Using cumulative peak fitting and iterative residual analysis, multiple isobaric peaks can be separated. These algorithms are described and are tested in various simulations, which have been performed to determine the benefits and constraints of this method. A smog chamber example is presented to demonstrate the performance and constraints for real data, with an example focusing on m/z 43. In this case, two isobaric peaks can be separated (m/z 43.015 and m/z 43.043) and identified as $C_2H_3O^+$ and $C_2H_5N^+$ respectively. Indistinguishable isobaric interferences slightly shift the exact m/z ratios resulting in a worse mass accuracy as expected from the simulations. Nevertheless, a 50 ppm mass accuracy is achieved, which is an impressive value for the identification of two isobaric peaks separated by 0.65 full width half maximum for an instrument with a mass resolving power of only $R = 1000$. According to the simulations, a mass resolving power of at least 4000 would be necessary to quantitatively separate the expected isobaric interferences that are separated by a m/z difference of about 0.004 at m/z 43.

1. Introduction

Particulate matter (PM) plays an important role in the atmosphere, affecting the climate [1], human health [2,3] and the ecosystem [4,5]. The properties and effects of PM are strongly dependent on concentration, size, shape (or morphology) and chemical composition. Their measurement in real time is difficult due to the inherent complexity of aerosols, their low mass concentrations (typically few micrograms per cubic meter), and the large variability of their properties in space and time. The presence of hundreds or more of organic and inorganic compounds complicates their chemical analysis, requiring specific tools [6,7].

At present, the Aerosol Mass Spectrometer (AMS) is the most commonly used instrument for real time characterization of mass, size and chemical composition of aerosols [8–11]. For chemical composition analysis, the AMS utilizes thermal flash evaporation followed by electron impact ionization and analysis of the ions with a mass spectrometer (MS). Three types of MS are currently available: a Quadrupol MS (Q-AMS) [8], a compact Time-of-Flight MS (C-TOF AMS, $R \sim 1000$) [12] and a high resolution Time-of-Flight MS (HR-TOF AMS) either operated in V-mode ($R \sim 2000$) or in W-mode ($R \sim 4000$) [13]. For the Q-AMS and the C-TOF AMS unity mass information is processed by the SQUIRREL data analysis package [12]; for the HR-TOF AMS the high resolution data evaluation extension PIKA is available, based on the methods described by DeCarlo et al. [13].

Challenges for the TOF AMS data processing and reduction are an accurate deconvolution of multiple, overlapping peaks around a nominal m/z ratio and a precise mass scale calibration for an accurate exact mass to chemical composition assignment. Using a C-TOF AMS, a low mass resolving power R and a low number of peak contributing data bins make the definition of initial fitting conditions and fitting constraints difficult. To overcome this problem the concept of cumulative fitting has been applied. Instead of fitting probability density functions (PDF) to certain intervals of the spectrum [13–15], the cumulative sum is calculated within these intervals and fitted by a cumulative distribution function (CDF) [16]. When applying a CDF, it is easy to calculate an upper constraint of a single peak area that corresponds to the sum of all data bins within a certain interval around a nominal m/z ratio. A small systematic error is added to the data by assuming the AMS to directly count single ions, instead of sampling analog signals and calibrating it to the single ion intensity afterwards. Nevertheless, the CDF method is consistent with the standard SQUIRREL data analysis and the ionization efficiency (IE) calibration based on the Brute Force Single Particle (BFSP) mode method [8,12,17]. Both are based on counting data bins within certain intervals around a nominal m/z ratio.

---

* Corresponding author. Tel.: +33 472 448 397; fax: +33 472 448 438.
E-mail address: barbara.danna@ircelyon.univ-lyon1.fr (B. D'Anna).

[1] Institut für Ionenphysik, Universität Innsbruck, Technikerstrasse 25, A-6020 Innsbruck, Austria.

1387-3806/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2011.04.007

In the following sections a detailed overview about the improved data reduction routines is given. Simulations have been performed to validate their performance. A data example highlighting the deconvolution of two isobaric peaks with the new procedure is presented around m/z 43.

2. Instrumentation

The C-TOF AMS is based on the successful design of the aerosol sizing and ionization system of the Q-AMS [8,12]. Air is introduced through a critical orifice into an aerodynamic lens. This aerodynamic lens efficiently separates the gas phase from the particles (~40–1500 nm diameter) and focuses the latter into a narrow beam. By a supersonic expansion, the particle beam is extracted into a sizing unit. A chopper is utilized in front of a high vacuum region to periodically block the particle beam. Evaluation of the particle time of flight (P-TOF) reveals information about their size distribution. Afterwards, the particles are flash vaporized at about 870 K on a hot surface and subsequently ionized by electron impact (routinely at 70 eV). Finally, an ion guide transfers the positively charged ions into a compact orthogonal extraction Time-of-Flight MS (CTOF platform, Tofwerk AG, Thun, Switzerland).

Three data acquisition modes are available: the mass spectrum (MS) mode, the Particle Time-of-Flight (P-TOF) mode and the Brute Force Single Particle (BFSP) mode [12]. In MS mode the build in chopper is used for internal background analysis. The chopper sequentially opens (MS-open) or blocks (MS-closed) the particle beam. When the beam is blocked, only interfering instrumental background compounds are measured. Subtraction of a MS-closed spectrum from the following MS-open spectrum results in an instrumental background corrected difference spectrum (MS-diff). This MS-diff spectrum is used for peak quantification and chemical species characterization.

A major challenge for C-TOF AMS is the low mass resolving power which varies between 700 and 1200, at m/z 28 and m/z 184 respectively. These values are slightly better compared to DeCarlo et al. [13], who report maximum C-TOF AMS mass resolving powers of about 1000. Another limitation is the low number of data acquisition bins per single peak that complicates the analysis. The analog-to-digital converter (ADC; AP240, Acqiris, Geneva, Switzerland) of the C-TOF AMS allows a data acquisition with a 1 ns data bin width. This results in only 10,000 data bins per single spectrum measured from m/z 4 to m/z 300. For example, a single peak at m/z 28 consists of 6 data bins within an interval of 2 times the full width at half maximum (FWHM) and at m/z 43 of about 7.5 data bins.

3. Data processing and data reduction

To improve the C-TOF performances custom data reduction algorithms, developed for trace gas analysis by means of the high resolution Proton Transfer Reaction Time of Flight (PTR-TOF) instrument [15,18], have been adapted and improved. These routines rely on a peak shape detection, a Time-of-Flight to m/z calibration using several selected "*mass scale calibration m/z*", an iterative dynamic peak detection producing a mass list, and the time-series analysis. Matlab (MATLAB R2009A, The MathWorks, Inc.) was utilized for programming; with adequate programming effort the developed algorithms can certainly be ported to many other programming languages like Igor (Igor Pro, WaveMetrics, Inc., USA). In the following paragraphs the associated principles are described, highlighting changes made to the algorithm used by Müller et al. [15].

3.1. Peak shape determination

A new approach for the peak shape detection was tested. Instead of fitting a complicated Gaussian shaped function [13,15], a true reference peak (e.g., O_2^+) is calculated from every single MS-open spectrum by analyzing the cumulative sum of a baseline corrected "*model peak*". This cumulative "*model peak*" is interpolated applying a cubic interpolation and normalized to the peak area A. Finally the peak center t is shifted to zero and the relative Time-of-Flight $\mathrm{TOF_{old}}$ is normalized to the peak width Δ as shown in Eq. (1).

$$\mathrm{TOF}_{\Delta} = \frac{\mathrm{TOF_{old}} - t}{\Delta} \tag{1}$$

3.2. Peak fit

For peak fitting, a nonlinear least-square problem is solved by applying a Trust-Region-Reflective Optimization algorithm [19]. Instead of fitting a function, the cumulative "*model peak*" is fitted to a cumulative signal of a peak ($peak_i$). The nonlinear least-squares problem is solved and fitted to $peak_i$ by stretching the normalized relative Time-of-Flight TOF_{Δ} of the "*model peak*" with the width Δ_i, moving it to the peak center t_i and multiplying it by the area A_i of $peak_i$. This is done by minimizing the difference of the interpolation (Matlab command *interp*1) shown in Eq. (2) and its corresponding cumulative signal using a nonlinear least-square problem solver (Matlab command *lsqnonlin*).

$$peak_i(\Delta_{i,}, t_i, A_i) = interp1(\mathrm{TOF}_{\Delta} \cdot \Delta_i + t_i, A_i \cdot model\,peak,\ \mathrm{TOF}_i) \tag{2}$$

As a result, the peak area A_i, peak width Δ_i and peak center t_i of $peak_i$ as well as the fitting residual and the coefficient of determination are recorded.

3.3. Mass scale calibration

The C-ToF-AMS is mainly a field instrument that can operate continuously for days and even weeks. A robust procedure to perform automated mass calibration of the instrument without the use of external standards is needed. For this reason, Time-of-Flight to m/z calibration is accomplished using known peaks of the instrument background spectra, which are highly reproducible in time and across instruments owing to the least isobaric interference as possible. According to the "*mass scale calibration m/z*" and the ideal quadratic relationship between the m/z ratio and the Time-of-Flight t of a TOF-MS (Eq. (3), e.g., [20]) the mass scale is calibrated by fitting the coefficients α and β for every single MS-open spectrum.

$$\frac{m}{z} = \left(\frac{t-\beta}{\alpha}\right)^2 \tag{3}$$

Fig. 1 shows the Time-of-Flight t to m/z calibration for the C-TOF AMS. In this example five "*mass scale calibration m/z*" (C^+, $C_2H_3^+$, O_2^+, CO_2^+ and $C_8H_5O_3^+$ at m/z 12, 27, 32, 44 and 149 respectively) have been used by the fitting algorithm to calculate the exact peak center. It has to be mentioned, that the selection of the "*mass scale calibration m/z*" is dependent on the application and related isobaric interferences.

For very low m/z ratios, however, a systematic m/z shift is observed. For C^+ a m/z shift by 0.019 is observable and cannot be explained yet. This unexpected systematic mass deviation for low m/z ratios has also been recognized by the PTR-TOF instrument. The power law fit reported by DeCarlo et al. [13], which allows a slight deviation from the ideal quadratic relationship (Eq. (3)), could solve this problem, but we identified it as a source for additional systematic error of the mass scale calibration at higher m/z ratios. Nevertheless, the systematic deviation was identified to be stable and therefore can be corrected by post analysis.

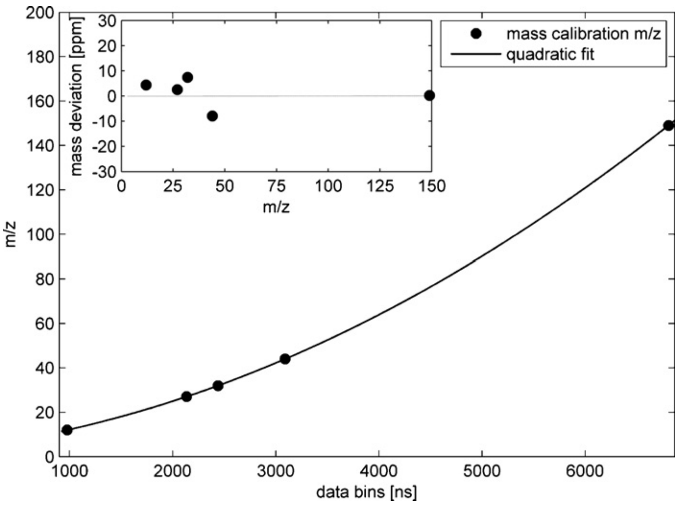

Fig. 1. Time-of-Flight to m/z calibration using five selected mass scale calibration m/z. The inset depicts the relative mass deviation from the expected ionic m/z ratio in parts per million (ppm).

The inset of Fig. 1 depicts the relative mass deviation of the five "*mass scale calibration m/z*". Mass scale deviations from the ideal quadratic relationship between 0.24 and 7.9 ppm are measured. The biggest challenge for an accurate m/z calibration using the C-TOF AMS is the identification of pure peaks without interferences. And it is important to mention, that the mass deviation for other m/z might be worse due to inseparable peaks. Small isobaric interferences from isotopes or other isobars will shift the mass center towards the interfering peak. Therefore for the C-TOF AMS, the application of the fitting routine for peak separation and exact mass analysis is limited to a certain m/z ratio. For high m/z ratios too many inseparable interfering peaks have to be expected. For this reason standard high m/z limit was empirically selected to be m/z 60, but the routines were tested up to m/z 100 during the analysis of environmental chamber experiments.

3.4. *Iterative peak identification*

The detection of multiple peaks is performed by using an iterative peak identification routine. The peak maxima that extend a user defined peak amplitude, around every nominal m/z_i (typically $10\Delta_i$) of every single MS-diff spectrum of a time series can be identified. All these peak maxima positions are merged together and filtered to eliminate possible double assignments. Only the bin numbers of well separated peaks (at least 1000 ppm) are stored into a preliminary mass list and later used as starting values for the peak assignment.

For the de-convolution of multiple isobaric peaks e.g., peaks covered in a shoulder of a larger peak, the routines described by Müller et al. [15] have been adopted for the cumulative peak fitting approach described above. For an iterative peak identification based on residual analysis a good knowledge about the expected peaks is necessary. Regarding Eq. (2), initial fitting conditions and respective fitting constraints for Δ_i, A_i, and t_i have to be calculated, as summarized in the following three paragraphs:

- According to Coles and Guilhaus [21], the expected coherence of the Time-of-Flight t_i and the width Δ_i of a $peak_i$ is described by Eq. (4).

$$\Delta_i = a\left(\frac{t_i^2}{b^2}+1\right)^{1/2} \tag{4}$$

  The coefficients a and b are solved by fitting the function to the predefined "*mass scale calibration m/z*" for every MS-open spectrum. Consequently the calculated values for Δ_i are used as the initial conditions for the peak fitting. The fitting constraints for Δ_i are empirically set to ±6%. Δ_i is set to be constant for all isobaric peaks.
- The initial condition for the peak area A_i is defined by the sum of all data bins within an interval of $10\Delta_i$ multiplied by the respective ratio of the amplitudes of all isobaric peaks. The upper fitting constraint is set to be the sum of all data bins within an interval of $10\Delta_i$.
- For the peak center t_i the preliminary mass list is applied. Fitting boundaries are empirically set to $\pm(\Delta_i/5)$ of the initial value.

With respect to the preliminary mass list, the cumulative signal within an interval of $10\Delta_i$ around the expected peak center is fitted. In an iterative approach, the fitting residual is analyzed, the position of a hidden peak is added to the preliminary mass list and a superposition of the corresponding number j of "*model peaks*" is fitted ($\Sigma_{i=1}^{j} peak_i$). This procedure is repeated until a maximum of 3 iterations is reached or the intensity of the residual is lower than an expected limit. This m/z dependent limit is calculated utilizing the residuals of the "*mass scale calibration m/z*" and therefore it is a measure how well the model peak shape suites to a certain m/z ratio. A new peak is added to the mass list if the intensity is higher than 8 times the standard deviation of the residual and the peak centers are better separated than $\Delta_i/3$. Finally, all fitted peak centers t_i are transformed into m/z ratios and saved into the mass list.

3.5. *Time series evolution*

At each acquisition time, the MS-diff spectrum is baseline corrected and analyzed using the routines described above. Fitting intervals are calculated; the lowest and the highest m/z ratios within each nominal m/z ratio are extended by $\pm5\Delta_i$ to get the fitting intervals, $[t_i - 5\Delta_i, t_{i+j} + 5\Delta_i]$ respectively. A superposition of the "*model peaks*" is fitted within the tight fitting constraints described above. Finally, peak areas are TOF-MS duty corrected and saved together with additional peak information to a file using the hdf5 file format (http://www.hdfgroup.org/HDF5/).

4. Simulations

To determine the quality of the fitting routines, 43,200 independent fits have been performed during some numerical testing. For this purpose, a theoretical double peak system (as defined below) has been chosen to assess the impacts of relative and absolute signal intensity, peak distance, noise, mass scale perturbation and temperature shifts. A "*model peak*" was calculated from several real spectra and adopted to correspond to a peak around m/z 43 and a mass resolving power of $R = 1000$. This resulted in a discrete peak with a FWHM of about 3.5 data bins.

During the simulations, two of these "*model peaks*" have been superposed at a peak center distance of 0.4, 0.8 and 1.2 FWHM (at m/z 43 these distances correspond to m/z differences of 0.018, 0.035 and 0.053 respectively). For simplicity, the ratio of the left peak (child) to the right peak (parent) was varied between 0.001 and 1. The intensity of the parent peak was chosen to be 100, 1000, 10,000 and 100,000 counts. Counts have been selected as the unit of the peak intensity to be independent of the integration time. To simulate the effects of the inaccuracy of the mass scale calibration a perturbation up to 10 parts per million (ppm) of the mass scale was applied. According to the mass scale deviation of the "*mass scale calibration m/z*" depicted in Fig. 1, 10 ppm mass scale inaccuracy has to be expected. Effects of a temperature shift were simulated by

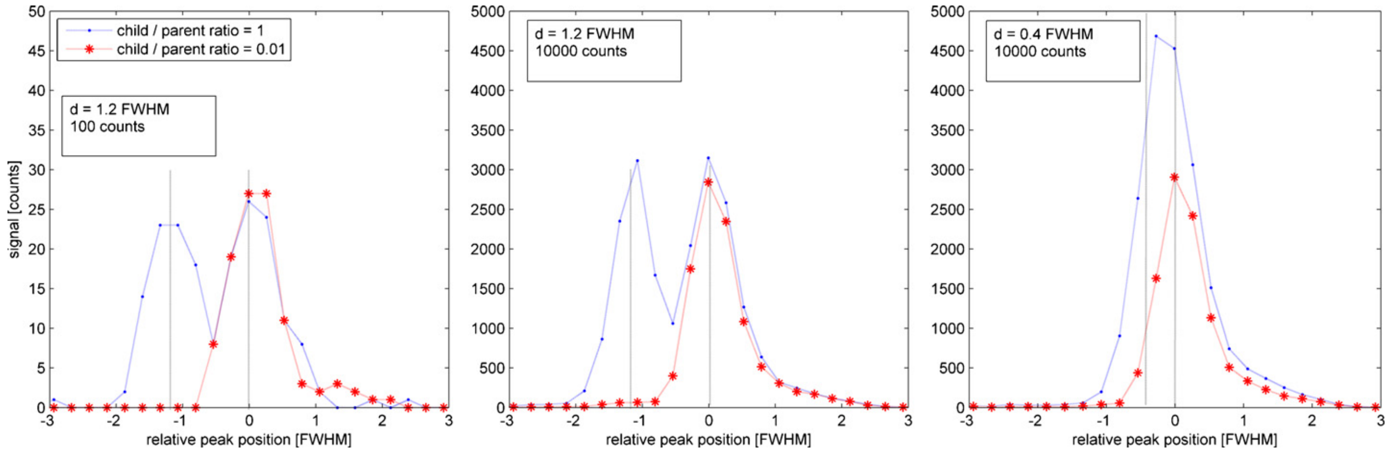

Fig. 2. Illustration of 6 simulated double peak examples around m/z 43. Peak distances of 0.4 and 1.2 FWHM, parent peak intensities of 100 and 10,000 counts and child / parent peak ratios of 0.01 and 1 are shown. Poisson noise was added to simulate ion counting statistics.Peak centers are indicated by vertical black lines.

shifting the data bins up to 100 ppm. For each of these initial conditions, MATLAB R2009a was used to add random Poisson noise to the calculated spectrum to simulate the AMS's counting error; and to achieve good statistics this procedure was performed 100 times per single initial condition. Fig. 2 depicts 6 examples of the two superposed peaks at peak distances of 0.4 and 1.2 FWHM, parent peak intensities of 100 and 10,000 counts and child/parent peak ratios of 0.01 and 1. It is visible, that at $d = 1.2$ FWHM the peaks are very well separated, whereas at $d = 0.4$ FWHM no individual peaks are visible. All resulting 43,200 spectra were analyzed by the fitting routines described above. Initial values and fitted results of the area and position of the parent and child peaks are compared.

4.1. Signal interference analysis

Fig. 3 (top panel) indicates a slight distance and signal dependent influence of the child peak to the area of the parent peak. At 0.4 FWHM distance, the deviation of the measured values to the input values vary between an underestimation of 7% at 100 counts parent peak area and a peak ratio of 0.001 to 0.8% at 100,000 counts and a peak ratio of 1. For peak distances $d \geq 0.8$ FWHM an overall accuracy better than 1.5% can be achieved. As expected, due to fitting a certain fraction of the parent peak's signal is transferred to the child peak. Fig. 3 (bottom panel) depicts the overlap of the dominating parent peak to the child peak. The biggest interference is visible for low child/parent ratios and low peak distances. At $d = 0.4$ and a child/parent ratio of 0.001 a signal between 7% and 3.5% of the parent peak (100 counts and 10,000 counts respectively) is contributed to the child. At $d = 1.2$ FWHM between 1% and 0.2% parent peak contribution is measured at a ratio of 0.001. By increasing child/parent ratio an increasing accuracy can be noticed. For a child/parent ratio of 0.1 an overestimation of the expected child area of 30%, 2% and 0.7% is measured at distances of $d = 0.4$, 0.8 and 1.2 FWHM respectively. Finally, this very well illustrates that a precise analysis can only be performed for reasonably well separated peaks.

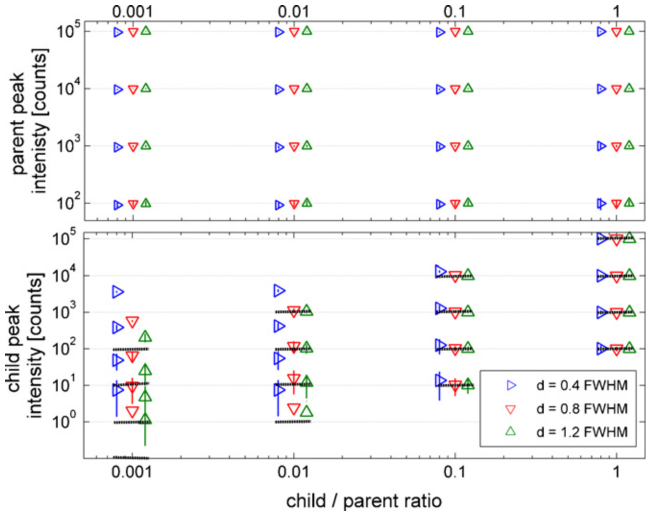

Fig. 3. Simulated co-contributions of the parent and child peak. 4 different child / parent peak ratios (0.001, 0.01, 0.1 and 1) and 4 different initial parent peak intensities (100, 1,000, 10,000 and 100,000 counts) are simulated and the effects for peak quantification for the parent (top panel) and the corresponding child (bottom panel) peak are plotted. Parent to child peak distances varied between d = 0.4 and d = 1.2 FWHM. Dashed black lines indicate the expected child peak intensities. For simplicity, all mass scale perturbation and temperature shift simulations have been merged together.

4.2. Error analysis

Fig. 4 shows the relations between the measured standard deviation and the calculated Poisson error for the parent and the child peak. The analysis shows a good agreement between measured and calculated error for distances $d \geq 0.8$ FWHM. While for $d = 0.4$ FWHM the measured standard deviation is 5 times higher than the

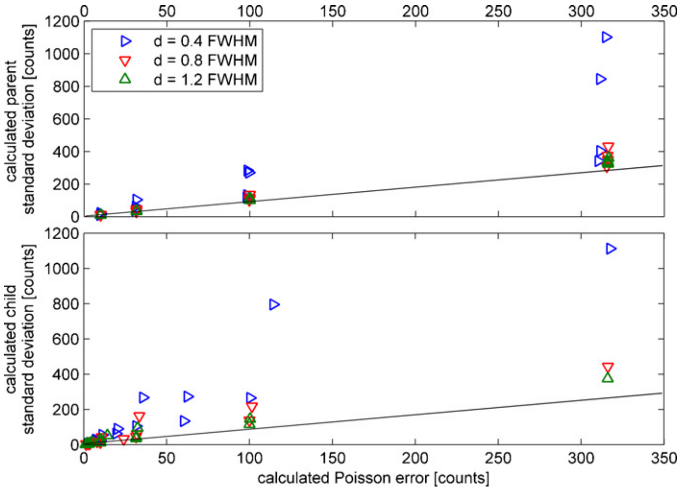

Fig. 4. Standard deviation analysis of the parent (top panel) and the child (bottom panel) peak. Comparison of the expected Poisson error with the simulated standard deviation of the area. The black lines depict the 1:1 relationships.

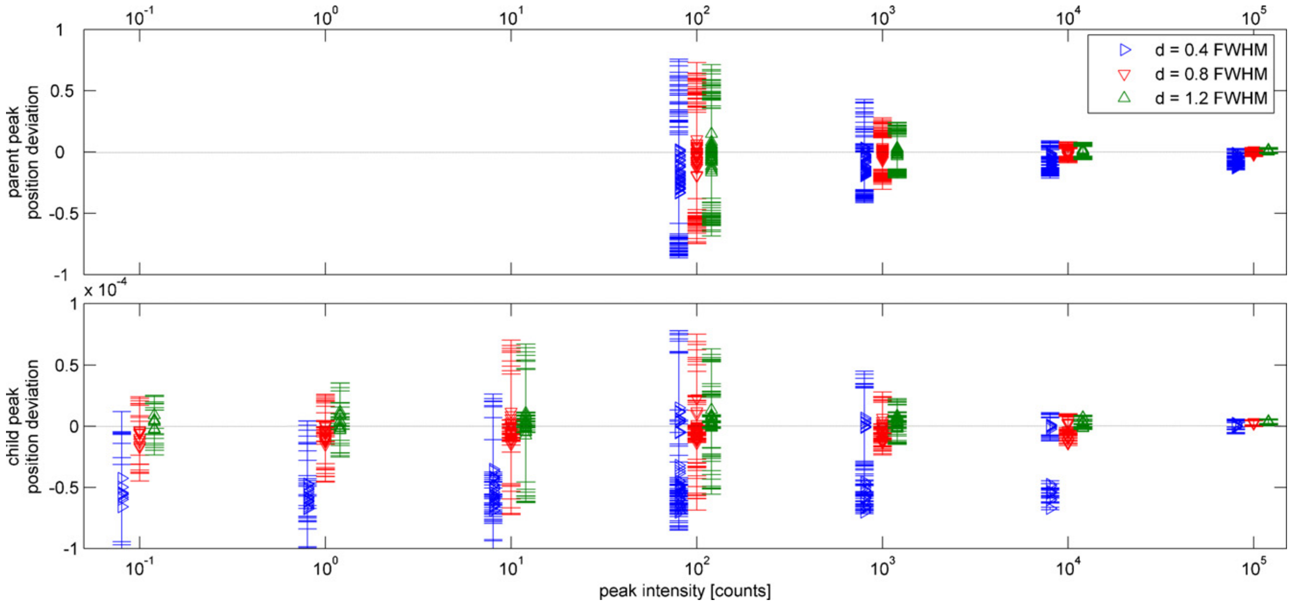

Fig. 5. Mass deviation analysis of the simulations. Peak intensity dependent exact peak center deviation for the parent (top panel) and the child (bottom panel) peak assuming a double peak system around m/z 43.

one expected by Poisson error. This can be explained either by the influence of the higher noise of the parent to the lower Poisson noise of the child peak or by adding additional noise due to insufficient separation by fitting both peaks. A noise analysis for single peaks of intensity between 100 and 100,000 counts has been performed by a PDF Gaussian shape fit [15]. Results show a 23–32% evaluated noise, at 100,000 and 100 count peak intensity respectively, as expected by pure Poisson counting noise. A single peak approximated by the CDF peak shape perfectly reflects the Poisson counting error. These differences between the PDF and CDF approach might be explained by the better fitting constrain of the CDF approach.

4.3. Mass deviation analysis

In addition to the peak area, the precise peak center and therefore the exact mass information is important for post analysis. The more accurate the mass center can be determined the easier is the assignment of an exact mass to a chemical composition as needed for elemental analysis of the aerosol [22,23] and the recently presented carbon oxidation state approach [24]. Fig. 5 depicts the relative peak center deviation according to the peak area for the parent and the child peak. For visualization, mean values and standard deviations of all 432 simulated initial conditions are plotted.

For the parent peak (Fig. 5, top panel) a maximum $1-\sigma$ mass scale deviation of 93 ppm and a minimum $1-\sigma$ mass scale deviation of 3 ppm is achieved, at $d = 0.4$ FWHM and 100 counts and $d = 1.2$ FWHM and 100,000 count signal intensity respectively. For the parent peak the mass scale deviation slightly decreases with increasing peak distance and significantly increases with increasing peak intensity. Therefore, at 10,000 count signal intensity, a maximum $1-\sigma$ mass scale deviation of 20 ppm at $d = 0.4$ FWHM and 6 ppm at $d = 1.2$ FWHM can be achieved. No influence of the mass scale perturbation and the temperature shift simulation on the parent peak position is visible.

For the child peak (Fig. 5, bottom panel), comparable results are achieved for distances $d \geq 0.8$ FWHM. For smaller peak distances ($d = 0.4$ FWHM), only the 1:1 child/parent peak ratios show results similar to the parent peak. At $d = 0.4$ FWHM all other child/parent peak ratios show a systematic offset scattering around −50 ppm, slightly dependent on the simulated mass scale perturbation and the temperature shift. This systematic deviation indicates problems of the fitting routine as the peak separation is not sufficient enough. This behavior is in agreement with the low accuracy of the area determination at $d = 0.4$ FWHM.

The simulated mass deviations are in agreement with the mass deviation results shown in Fig. 1. The slightly higher values shown in Fig. 1 might be explained by inseparable isobaric interferences that shift the "*mass scale calibration m/z*" positions, by unconsidered instrumental noise and by an imprecise baseline correction.

5. Real data example

Real data concerning the photo-degradation of amines which are precursor of secondary organic aerosol in the atmosphere [25,26] were used to illustrate the algorithm performance. These experiments were performed at the EUPHORE photo reactor facility in Valencia [27]. Fig. 6 shows the analysis steps at m/z 43. The photo-oxidation experiment started at about 07:15 UTC by the injection of gas precursors into the chamber. The chamber roof was opened at 07:51 UTC and closed at 10:53 UTC. Data acquisition stopped at 11:50 UTC. In addition, a multistep temperature thermogram (299–423 K) was performed between 10:15 UTC and 10:45 UTC to get additional volatility information of the aerosol.

The top panel of Fig. 6 shows an average spectrum of 6 single MS-diff spectra recorded around 09:55 UTC. The measured data bins and the cumulative sum are presented. The dotted black line marks the initial peak center (m/z 43.055) as utilized for the iterative peak identification. A visual separation of ions is not possible.

The left middle panel of Fig. 6 depicts a peak fit of a single model peak (calculated using m/z 32). After this fitting, a significant residual is left. By utilizing the maximum of the residual (m/z 43.019) and the center of the single peak fit (m/z 43.052) new initial conditions are calculated. Consequently, a double peak fit (right middle panel) perfectly characterizes the measured data. The residual is now negligible. The double peak fit results in two peak centers with m/z 43.015 and m/z 43.043, identified as $C_2H_3O^+$ and $C_2H_5N^+$ respectively. The dotted blue line visualizes the relative contribution of m/z 43.015 to the total signal (dashed black line, middle panel).

No contribution of $C_3H_7^+$ (m/z 43.054), a frequently measured SOA fragment, was detected during the experiment. It is possible,

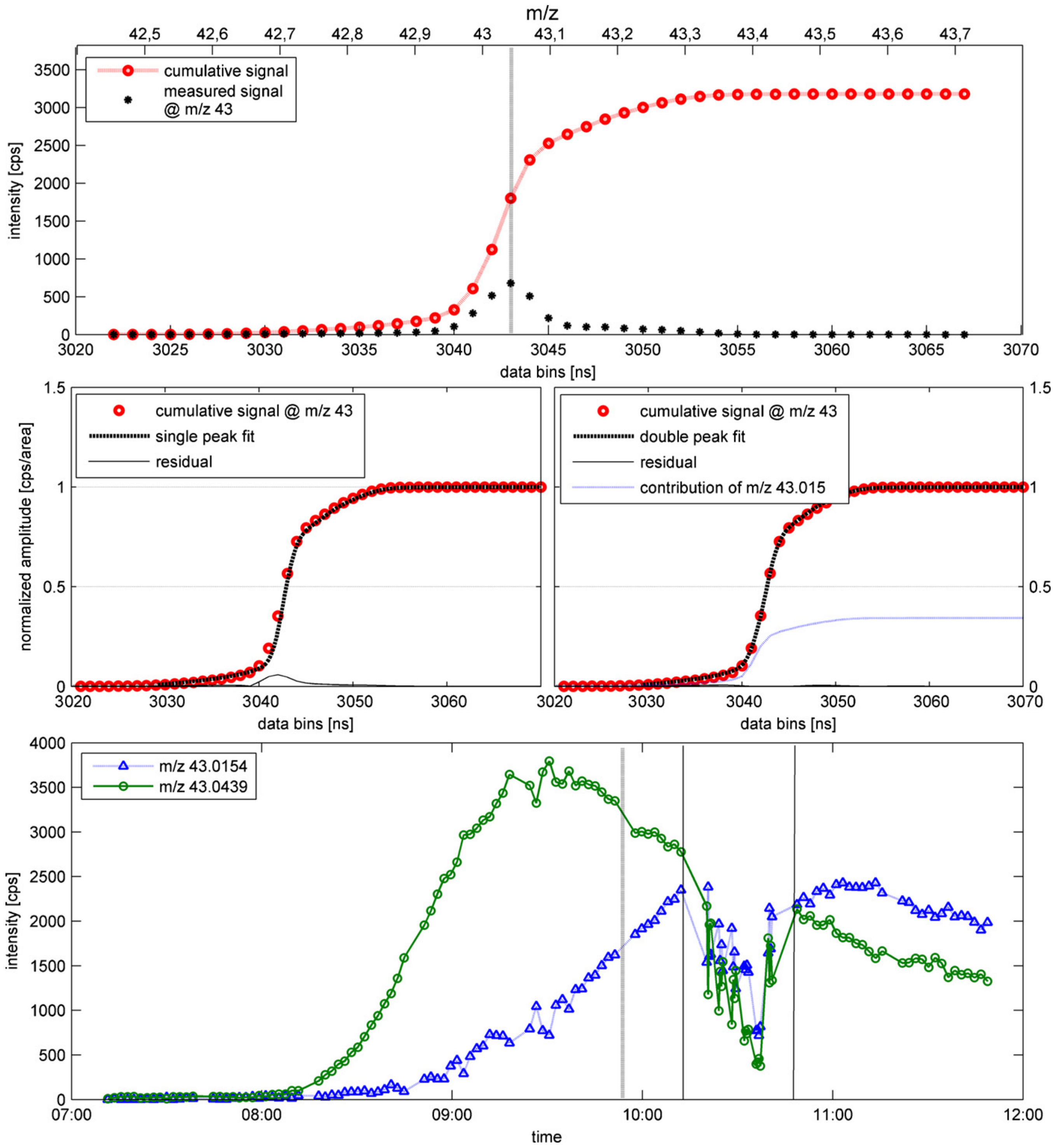

Fig. 6. Iterative peak detection and time series analysis. The top panel shows the measured data points and the cumulative sum of data points (top panel). The middle panels depict the corresponding peak fits of a single model peak (middle panel, left) and a double peak fit (middle panel, right) together with the corresponding residuals. Two ionic peaks could be identified, m/z 43.015 and m/z 43.043. The dotted blue line indicates the relative contribution of m/z 43.0154 to the total signal (dashed black line, middle panel, right). Time series of the two identified peaks during a smog chamber experiment (bottom panel). Between the solid vertical lines a thermogram has been performed.

that a small contribution of $C_3H_7^+$ is produced during the experiment. Nevertheless, $C_3H_7^+$ is separated from $C_2H_5N^+$ by less than 0.3 FWHM. According to the simulations, only peaks with similar intensities can be detected such close together.

Mass accuracies around 50 ppm are achieved, a factor 5–10 higher compared to the simulations. This is explained by an inseparable superposition of isobars that slightly shift the positions of the peak centers e.g., analysis of *m*/*z* 42 exposes an inseparable isobaric interference of the $C_2H_2O^+$ isotope $^{13}CCH_2O^+$ (*m*/*z* 43.013) to the $C_2H_3O^+$ (*m*/*z* 43.018) ion signal. According to the simulations, a mass resolving power of at least 4000 is necessary for a quantitative separation of these ions ($d = 0.4$ FWHM and similar peak intensities), a value that can be reached by the HR-TOF AMS operated in W-mode [13] but not on a C-TOF instrument with the standard data processing.

The bottom panel of Fig. 6 shows the complete time series of both ionic signals. The dotted vertical line indicates the average time of the average spectrum (top panel). The time evolution is very different for the two ions, indicating minimum interferences as a function of time between these two peaks. The intensity of the nitrogen containing ion $C_2H_5N^+$ increases shortly after the chamber roof is opened and photo-oxidation started. The oxygenated fragment $C_2H_3O^+$ starts to increase about 30 min later. Additional analysis of the thermogram shows a much higher volatility of $C_2H_5N^+$ compared to $C_2H_3O^+$ e.g., 70% of $C_2H_5N^+$ and only 40% of $C_2H_3O^+$ are evaporated at 383 K (data not shown). Therefore the peak fitting routines help to reveal chemico-physical properties of the aerosol by distinguishing two ionic signals, separated by only 0.65 FWHM.

6. Discussion and conclusion

A method for an improved analysis of C-TOF AMS data has been presented. By using cumulative peak fitting and iterative residual analysis, multiple isobaric peaks can be separated. The applied algorithms are explained and simulations have been performed to discover the constraints of this method. Finally an atmospheric simulation chamber experiment (photo-oxidation of trimethylamine under low NO_x conditions; max. aerosol loading = 35 $\mu g/m^3$) is used for demonstrating the method performance and discussing the constraints for real data analysis around *m*/*z* 43. In this region, two isobaric peaks could be separated (*m*/*z* 43.015 and *m*/*z* 43.043) and identified as $C_2H_3O^+$ and $C_2H_5N^+$ respectively. Nevertheless, indistinguishable isobaric interferences slightly shift the exact *m*/*z* ratios resulting in a worse mass accuracy as expected from the simulations. Still, 50 ppm accuracy is impressive for the identification of two isobaric peaks, separated by 0.65 FWHM, and an instrument with a mass resolving power of only 1000. For comparison, one data bin distance at *m*/*z* 43 corresponds to a mass difference of 350 ppm. According to the simulations, a mass resolving power of at least 4000 would be necessary to separate the isobaric interferences that are only separated by about 100 ppm (*m*/*z* difference of 0.0044).

This issue underlines the problem with static mass lists, as utilized by the standard HR-TOF AMS analysis tool PIKA [13]. Fixed fitting mass lists, containing unexpected ions, interfering ions of small intensity and inseparable ions (e.g., isotopes), can significantly change the fitting results. Currently, PIKA (v1.09) allows a fit of up to ~20 isobars around *m*/*z* 43. Four of them are used in the default mode, $^{13}CCH_2O^+$, $C_2H_3O^+$, $^{13}CC_2H_6{}^+$ and $C_3H_7{}^+$ respectively. Other ions have to be manually selected by the user. In addition, PIKA utilizes fixed peak centers without any tolerance range for fitting. Therefore, the expected mass scale calibration error of the HR-TOF AMS between 28 ppm (V-mode) and 7 ppm (W-mode) will proceed to an error in peak quantification.

To quantify the propagated error, the described algorithms have been applied with fixed mass centers t_i instead of an allowed variation of $\pm(\Delta_i/5)$. 28 ppm mass scale inaccuracy for the separation of 100 counts $^{13}CCH_2O^+$ and 10,000 counts $C_2H_3O^+$ in V-mode ($R = 2000$) were considered for simulation. The results are 460 counts of $^{13}CCH_2O^+$ and 9620 counts of $C_2H_3O^+$ and 1120 counts of $^{13}CCH_2O^+$ and 8960 counts of $C_2H_3O^+$, if the centers were shifted towards or away from the smaller peak respectively. These values demonstrate a fitting error up to 11% for the parent peak $C_2H_3O^+$ quantification propagated from a small inaccuracy of the mass scale calibration! Similar results have to be expected by using PIKA.

In addition the methodology developed here implies another advantage. Analysis of the time series of a fitted mass center allows identifying possible interferences. If the exact mass of a peak is stable during a certain time series it is likely to be a pure ionic signal, whereas a continuous change in exact mass indicates the growth of an undistinguished interfering ion. Such an exact mass analysis can be automated and utilized for elemental analysis.

In conclusion, the simulations give very detailed information about the co-contributions of the parent and child peaks. If time series with high dynamic peak variations are available, this information helps to identify optimal periods for precise exact mass identification with the lowest expected mass deviation possible.

Future work will focus on porting the algorithms to HR-TOF AMS and PTR-TOF datasets. In addition it will be evaluated, if the PTR-TOF routines developed for a fast evaluation of 10 Hz eddy covariance datasets by Müller et al. [15] can be used for a fast HR analysis of the P-TOF AMS data. By any interest in testing and/or further development of the applied algorithms the reader is welcome to contact the authors.

Acknowledgements

This work was carried out as part of the ADA-project (Atmospheric Degradation of Amines) and has received financial support from Statoil ASA, Vattenfall AB, Masdar and CLIMIT through contract 201604. The authors acknowledge the Centre National de la Recherche Scientifique (CNRS), the Institut National des Sciences de l'Univers (INSU), the Instituto Universitario Centro de Estudios Ambientales del Mediterráneo CEAM – UMH and Antoinette Boréave for the helpful support.

References

[1] Intergovernmental Panel on Climate Change (IPCC): Climate Change 2007 – The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC, Cambridge University Press, 2007.

[2] D.W. Dockery, C.A. Pope, X.P. Xu, J.D. Spengler, J.H. Ware, M.E. Fay, B.G. Ferris, F.E. Speizer, New England Journal of Medicine 329 (1993) 1753–1759.

[3] C.A. Pope, D.W. Dockery, Journal of the Air & Waste Management Association 56 (2006) 709–742.

[4] G.E. Likens, C.T. Driscoll, D.C. Buso, Science 272 (1996) 244–246.

[5] K.S. Carslaw, O. Boucher, D.V. Spracklen, G.W. Mann, J.G.L. Rae, S. Woodward, M. Kulmala, Atmospheric Chemistry and Physics 10 (2010) 1701–1737.

[6] A.H. Goldstein, I.E. Galbally, Environmental Science & Technology 41 (2007) 1514–1521.

[7] M. Hallquist, J.C. Wenger, U. Baltensperger, Y. Rudich, D. Simpson, M. Claeys, J. Dommen, N.M. Donahue, C. George, A.H. Goldstein, J.F. Hamilton, H. Herrmann, T. Hoffmann, Y. Iinuma, M. Jang, M.E. Jenkin, J.L. Jimenez, A. Kiendler-Scharr, W. Maenhaut, G. McFiggans, T.F. Mentel, A. Monod, A.S.H. Prevot, J.H. Seinfeld, J.D. Surratt, R. Szmigielski, J. Wildt, Atmospheric Chemistry and Physics 9 (2009) 5155–5236.

[8] J.T. Jayne, D.C. Leard, X.F. Zhang, P. Davidovits, K.A. Smith, C.E. Kolb, D.R. Worsnop, Aerosol Science and Technology 33 (2000) 49–70.

[9] J.L. Jimenez, J.T. Jayne, Q. Shi, C.E. Kolb, D.R. Worsnop, I. Yourshaw, J.H. Seinfeld, R.C. Flagan, X. Zhang, K.A. Smith, J.W. Morris, P. Davidovits, Journal of Geophysical Research 108 (2003) 8425.

[10] M.R. Canagaratna, J.T. Jayne, J.L. Jimenez, J.D. Allan, M.R. Alfarra, Q. Zhang, T.B. Onasch, F. Drewnick, H. Coe, A. Middlebrook, A. Delia, L.R. Williams, A.M. Trimborn, M.J. Northway, P.F. DeCarlo, C.E. Kolb, P. Davidovits, D.R. Worsnop, Mass Spectrometry Reviews 26 (2007) 185–222.

[11] J.L. Jimenez, M.R. Canagaratna, N.M. Donahue, A.S.H. Prevot, Q. Zhang, J.H. Kroll, P.F. DeCarlo, J.D. Allan, H. Coe, N.L. Ng, A.C. Aiken, K.S. Docherty, I.M. Ulbrich, A.P. Grieshop, A.L. Robinson, J. Duplissy, J.D. Smith, K.R. Wilson, V.A. Lanz, C. Hueglin, Y.L. Sun, J. Tian, A. Laaksonen, T. Raatikainen, J. Rautiainen, P. Vaattovaara, M. Ehn, M. Kulmala, J.M. Tomlinson, D.R. Collins, M.J. Cubison, E.J. Dunlea, J.A. Huffman, T.B. Onasch, M.R. Alfarra, P.I. Williams, K. Bower, Y. Kondo, J. Schneider, F. Drewnick, S. Borrmann, S. Weimer, K. Demerjian, D. Salcedo, L. Cottrell, R. Griffin, A. Takami, T. Miyoshi, S. Hatakeyama, A. Shimono, J.Y. Sun, Y.M. Zhang, K. Dzepina, J.R. Kimmel, D. Sueper, J.T. Jayne, S.C. Herndon, A.M. Trimborn, L.R. Williams, E.C. Wood, A.M. Middlebrook, C.E. Kolb, U. Baltensperger, D.R. Worsnop, Science 326 (2009) 1525–1529.

[12] F. Drewnick, S.S. Hings, P. DeCarlo, J.T. Jayne, M. Gonin, K. Fuhrer, S. Weimer, J.L. Jimenez, K.L. Demerjian, S. Borrmann, D.R. Worsnop, Aerosol Science and Technology 39 (2005) 637–658.

[13] P.F. DeCarlo, J.R. Kimmel, A. Trimborn, M.J. Northway, J.T. Jayne, A.C. Aiken, M. Gonin, K. Fuhrer, T. Horvath, K.S. Docherty, D.R. Worsnop, J.L. Jimenez, Analytical Chemistry 78 (2006) 8281–8289.

[14] J. Meija, J.A. Caruso, Journal of the American Society for Mass Spectrometry 15 (2004) 654–658.

[15] M. Müller, M. Graus, T.M. Ruuskanen, R. Schnitzhofer, I. Bamberger, L. Kaser, T. Titzmann, L. Hörtnagl, G. Wohlfahrt, T. Karl, A. Hansel, Atmospheric Measurement Techniques 3 (2010) 387–395.

[16] T. Titzmann, M. Graus, M. Müller, A. Hansel, A. Ostermann, International Journal of Mass Spectrometry 295 (2010) 72–77.

[17] J.D. Allan, J.L. Jimenez, P.I. Williams, M.R. Alfarra, K.N. Bower, J.T. Jayne, H. Coe, D.R. Worsnop, Journal of Geophysical Research-Atmospheres 108 (2003) 10.

[18] M. Graus, M. Müller, A. Hansel, Journal of the American Society for Mass Spectrometry 21 (2010) 1037–1044.

[19] T.F. Coleman, Y.Y. Li, Siam Journal on Optimization 6 (1996) 418–445.

[20] R.S. Brown, N.L. Gilfrich, Analytica Chimica Acta 248 (1991) 541–552.

[21] J. Coles, M. Guilhaus, Trac—Trends in Analytical Chemistry 12 (1993) 203–213.

[22] A.C. Aiken, P.F. DeCarlo, J.L. Jimenez, Analytical Chemistry 79 (2007) 8350–8358.

[23] A.C. Aiken, P.F. Decarlo, J.H. Kroll, D.R. Worsnop, J.A. Huffman, K.S. Docherty, I.M. Ulbrich, C. Mohr, J.R. Kimmel, D. Sueper, Y. Sun, Q. Zhang, A. Trimborn, M. Northway, P.J. Ziemann, M.R. Canagaratna, T.B. Onasch, M.R. Alfarra, A.S.H. Prevot, J. Dommen, J. Duplissy, A. Metzger, U. Baltensperger, J.L. Jimenez, Environmental Science & Technology 42 (2008) 4478–4485.

[24] J.H. Kroll, N.M. Donahue, J.L. Jimenez, S.H. Kessler, M.R. Canagaratna, K.R. Wilson, K.E. Altieri, L.R. Mazzoleni, A.S. Wozniak, H. Bluhm, E.R. Mysak, J.D. Smith, C.E. Kolb, D.R. Worsnop, Nature Chemistry 3 (2011) 133–139.

[25] P.J. Silva, M.E. Erupe, D. Price, J. Elias, G.J. Malloy Q., Q. Li, B. Warren, D.R. Cocker, Environmental Science & Technology 42 (2008) 4689–4696.

[26] X. Ge, A.S. Wexler, S.L. Clegg, Atmospheric Environment 45 (2011) 524–546.

[27] K.H. Becker, EUPHORE: Final report to the European commission, EV5V-CT92-0059, Bergische Universität, Wuppertal, Germany, 1996.