Atmospheric
Measurement
Techniques

Peak-fitting and integration imprecision in the Aerodyne aerosol mass spectrometer: effects of mass accuracy on location-constrained fits

J. C. Corbin[1,2], A. Othman[3,a], J. D. Allan[4,5], D. R. Worsnop[6], J. D. Haskins[7], B. Sierau[1], U. Lohmann[1], and A. A. Mensah[1]

[1]ETH Zurich, Institute for Atmospheric and Climate Science, Zurich, Switzerland
[2]Laboratory for Atmospheric Chemistry, Paul Scherrer Institute, Villigen, Switzerland
[3]University of Zurich, Institute for Clinical Chemistry, Zurich, Switzerland
[4]National Centre for Atmospheric Science, University of Manchester, Manchester, UK
[5]School of Earth, Atmospheric and Environmental Sciences, University of Manchester, Manchester, UK
[6]Aerodyne Research Inc., Billerica, Massachusetts, USA
[7]Massachusetts Institute of Technology, Boston, Massachusetts, USA
[a]now at: Institute of Experimental and Clinical Pharmacology and Toxicology, University of Lübeck, Lübeck, Germany

Correspondence to: J. C. Corbin (joel.corbin@psi.ch)

Abstract. The errors inherent in the fitting and integration of the pseudo-Gaussian ion peaks in Aerodyne high-resolution aerosol mass spectrometers (HR-AMSs) have not been previously addressed as a source of imprecision for these or similar instruments. This manuscript evaluates the significance of this imprecision and proposes a method for their estimation in routine data analysis.

In the first part of this work, it is shown that peak-integration errors are expected to scale linearly with peak height for the constrained-peak-shape fits performed in the HR-AMS. An empirical analysis is undertaken to investigate the most complex source of peak-integration imprecision: the imprecision in fitted peak height, σ_h. It is shown that the major contributors to σ_h are the imprecision and bias inherent in the m/z calibration, both of which may arise due to statistical and physical non-idealities of the instrument. A quantitative estimation of these m/z-calibration imprecisions and biases show that they may vary from ion to ion, even for ions of similar m/z.

In the second part of this work, the empirical analysis is used to constrain a Monte Carlo approach for the estimation of σ_h and thus the peak-integration imprecision. The

estimated σ_h for selected well-separated peaks (for which m/z-calibration imprecision and bias could be quantitatively estimated) scaled linearly with peak height as expected (i.e. as n^1). In combination with the imprecision in peak-width quantification (which may be easily and directly estimated during quantification), peak-fitting imprecisions therefore dominate counting imprecisions (which scale as $n^{0.5}$) at high signals. The previous HR-AMS uncertainty model therefore underestimates the overall fitting imprecision even for well-resolved peaks. We illustrate the importance of this conclusion by performing positive matrix factorization on a synthetic HR-AMS data set both with and without its inclusion.

In the third part of this work, the Monte Carlo approach is extended to the case of an arbitrary number of overlapping peaks. Here, a modification to the empirically constrained approach was needed, because the ion-specific m/z-calibration bias and imprecision can generally only be estimated for well-resolved peaks. The modification is to simply overestimate the m/z-calibration imprecision in all cases. This overestimation results in only a slight overestimate of σ_h, while significantly reducing the sensitivity of σ_h to the unknown, ion-specific m/z-calibration biases. Thus, with only the mea-

sured data and an approximate estimate of the order of magnitude of m/z-calibration biases as input, conservative and unbiased estimates of peak-integration imprecisions may be obtained for each peak in any ensemble of overlapping peaks.

1 Introduction

The Aerodyne high-resolution aerosol mass spectrometer (HR-AMS; Jayne et al., 2000; DeCarlo et al., 2006) can provide continuous, time- and size-resolved measurements of particulate-matter (PM) composition (Canagaratna et al., 2007; Kimmel et al., 2011). In its different configurations, the HR-AMS has been used to investigate the composition and evolution of organic PM (Aiken et al., 2007; Jimenez et al., 2009; Zhang et al., 2011), the internal mixing and composition of black carbon (Cappa et al., 2012; Corbin et al., 2015a), and the external mixing of atmospheric aerosols (Freutel et al., 2013; Lee et al., 2015) among many other applications (Canagaratna et al., 2007).

HR-AMS signals are routinely quantified with the free, open-source "PIKA" software (Sueper et al., 2011), written in Igor Pro (WaveMetrics, OR, USA). PIKA allows the user to perform both basic data processing and complex analyses, including high-resolution peak analysis (DeCarlo et al., 2006), mass quantification (Allan et al., 2004), and elemental analysis (Aiken et al., 2007).

PIKA is widely used to prepare HR-AMS data for least-squares statistical models such as positive matrix factorization (PMF) (Paatero and Tapper, 1994; Ulbrich et al., 2009; Zhang et al., 2011) or simple linear regression (e.g. Lee et al., 2010). These statistical models are solved by minimizing the residuals between uncertainty-weighted data and the model, so both the data and their uncertainties must be quantified in order to obtain meaningful results. In the previous sentence, "uncertainty" refers only to the internal imprecision of the HR-AMS data set (discussed further below) and not to the overall accuracy of HR-AMS-based mass concentrations, elemental ratios, or other data products. These higher-order data products require additional calibration factors which dominate their overall inaccuracy of roughly ±25 % (Canagaratna et al., 2007; Bahreini et al., 2009; Aiken et al., 2008; Canagaratna et al., 2015). Such inaccuracies are not normally relevant for the statistical modelling of HR-AMS measurements and do not influence the peak-integration imprecisions discussed below.

The current version of PIKA (1.10C) estimates HR-AMS uncertainties from the square root of the number of estimated ion counts (DeCarlo et al., 2006), following techniques developed for unit-mass-resolution (UMR) versions of the AMS (Allan et al., 2003). These UMR AMSs estimate the signals at a given integer m/z by summing all ion counts near that signal. In contrast, PIKA estimates the signals of a given ion by fitting a Gaussian-like (pseudo-Gaussian)

function to background-subtracted data and integrating the fitted peak. This additional complexity introduces additional uncertainties but is necessary and useful because HR-AMS mass spectra are frequently comprised of incompletely resolved (overlapping) isobaric peaks, each represented by relatively few data points.

The significance and magnitude of peak-integration uncertainties in PIKA-analysed HR-AMS data have not been previously addressed. Müller et al. (2011) discussed peak fitting for the lower-resolution compact time-of-flight AMS (C-ToF-AMS), using a different algorithm than that used by PIKA (the PIKA algorithm is detailed in Sect. 2). Cubison and Jimenez (2015) used a PIKA-like approach to perform an extensive study of a two-peak system while assuming an ideal mass spectrometer. They provided a parameterization for imprecision estimation in such a system while considering the lowest-achievable m/z-calibration uncertainty and, relative to this study, a large number of data points per peak. The present manuscript approaches the problem from the opposite direction to Cubison and Jimenez (2015), beginning from an empirical analysis and ending with a method for estimation which is not limited to just two peaks. The two approaches are compared in further detail in Sect. 6.

This manuscript addresses peak-integration uncertainties in PIKA by using a test data set to explore and understand the origins of peak-integration uncertainties in PIKA, using methods that are intended to be applicable to any HR-AMS or other mass spectrometer. The results of this empirical analysis are then used to construct a Monte Carlo model of the PIKA peak-fitting procedure, which allows the magnitude of peak-integration uncertainties for well-resolved (isolated) peaks and for overlapping peaks to be estimated. This empirically based approach allowed several assumptions behind uncertainty estimation to be directly evaluated and may be applied to any new data set.

The manuscript is structured as follows. First, Sect. 2 outlines the sources of uncertainty in HR-AMS analysis and describes the details of a PIKA analysis. Then, Sect. 3 presents a theoretical discussion and empirical analysis of peak-integration errors, which arise from peak-width and peak-height uncertainties. The focus is on the simpler case of well-resolved peaks. Peak-width uncertainties may be directly estimated, but peak-height uncertainties are much more complex and thus are the focus of the remainder of this section. It is shown that peak-height uncertainties are mostly due to biases and imprecisions in the m/z calibration; these calibration errors are then quantitatively estimated for well-resolved peaks.

Based on the constraints established in the previous section, Sect. 4 applies a Monte Carlo approach to evaluate the imprecision in fitted peak heights due to m/z-calibration errors for well-resolved peaks. Even for these well-resolved peaks, it is shown that these imprecisions are significant, scale linearly with peak height, and are significantly influenced by m/z-calibration biases. The overall

Table 1. List of symbols and abbreviations. Symbols used only once in the text are omitted from this list.

AMS	Aerodyne aerosol mass spectrometer
HR-AMS	high-resolution AMS
PIKA	HR-AMS analysis software
PMF	positive matrix factorization
RMSE	root mean square error (Eq. 8)
UMR	unit (integer) mass resolution
A	the area of a peak
e	an error (imprecision, bias, or both)
f	a pseudo-Gaussian function (Eq. 3)
f_0	a unit-height peak function (Eq. 3 with $h = 1$)
f_T	a linear superposition of peaks f (Eq. 5 with $h = 1$)
G	a Gaussian function (Eq. 2)
h	the height of a peak
k_{DC}	mass-spectrometer duty cycle
m/z	mass-to-charge ratio
v	the peak-shape function (Eq. 3)
w	the width of a peak
x_i	the ith m/z bin of a mass spectrum
μ	the mode of a peak
σ_X	imprecision in X (for $X = h, w, \ldots$)

peak-integration imprecision is discussed, with an illustrative example for a specific peak. The impacts of this peak-integration imprecision on PMF are then demonstrated on a synthetic HR-AMS data set. Section 4 then extends the Monte Carlo approach to the case of overlapping peaks. Finally, Sect. 6 places our results in the context of previous work, discusses limitations, and addresses its practical application.

A number of mathematical symbols and abbreviations are used throughout; a list is provided in Table 1. Note that because of the details of the PIKA analysis (Sect. 2), m/z-calibration errors manifest as peak-location-prediction errors (μ-prediction errors); we therefore use the latter expression below.

Throughout this manuscript, the terms "imprecision" and "bias" are used when referring respectively to random and independent errors (averaging to 0) and to errors of constant value. The distinction between these two concepts varies naturally at different stages of the analysis: if peak fitting is biased by an effect that varies from peak to peak, then an imprecision in the overall set of fitted peaks will result. As we discuss below, this is the case when peak positions are constrained by an m/z calibration during fitting.

2 Background

2.1 Conceptual basis

Four conceptual categories contributing to HR-AMS uncertainties can be defined: interpretation, counting, instrumental, and analysis uncertainties. These uncertainties may be defined and addressed as follows.

1. Interpretation uncertainties arise when a given signal may arise from sources other than the analyte. For example, in the AMS, H_2O^+ ions may form when either gas-phase water, particle-bound water, or the thermal decomposition of oxygenated functional groups in organic PM (OM) (Aiken et al., 2008), inorganic PM (Chen et al., 2011), or BC (Corbin et al., 2014). The interpretation of an H_2O^+ ion as originating from any one of these sources therefore requires additional information or assumptions to be made. In the AMS, a fraction of the overall H_2O^+ signal is typically attributed to OM according to laboratory-measured fragmentation patterns (Allan et al., 2004) as detailed by Chen et al. (2011). This so-called fragmentation-table approach may introduce biases but not imprecision to the data set. The biases would be constant only for a constant aerosol composition.

   Interpretation uncertainties may also occur if an ion is misidentified or omitted in the PIKA software. Such errors are not considered in this work.

2. Counting uncertainties estimate the degree to which a count of n ions would vary if that count were repeated for the same system and the same time period (Taylor, 1997). They are therefore a measure of imprecision. Poisson uncertainties are included in the standard UMR-AMS and HR-AMS error models and are determined by translating mass-spectral peak areas to ion signal rates and applying the Poisson distribution to determine σ_n as $\sqrt{n}$ for a count of n ions over a specific time period (Allan et al., 2003; DeCarlo et al., 2006).

3. Instrumental uncertainties may arise due to electronic noise or to changes in the performance of various instrumental components. In the latter case, which for example may reflect changes in detector sensitivity or long-term performance fluctuations, the significance of such variations can be evaluated by Allan-variance analysis (Allan, 1966; Werle et al., 1993; Ng et al., 2011; Onasch et al., 2012). Whereas the Allan variance decreases with increasing averaging time for an ideal mass spectrometer (where only ion-counting uncertainties exist), it eventually increases with longer averaging times in real instruments. This increase reflects the introduction of additional error terms from slowly varying contributors such as electronics temperature or detector stability (Allan, 1966; Werle et al., 1993). Conversely, the minimum in a plot of Allan variance against averaging time indicates the maximum timescale over which instrumental uncertainties can be considered negligible.

   For an SP-AMS, Onasch et al. (2012) have shown that $\sigma_{counting}$ dominates instrumental uncertainties for averaging times below 100 s in a filtered-air sample. Since this is longer than typical HR-AMS and SP-AMS sampling times (internal backgrounds are taken automat-

ically), instrumental uncertainties should therefore be small relative to counting uncertainties. However, this may not be true for instrumental uncertainties that are dependent on signal loadings, such as those due to background measurements or detector saturation.

4. Finally, analysis uncertainties reflect the confidence with which mass-spectral peak areas can be determined. These uncertainties may comprise both biases and imprecisions. The remainder of this manuscript focusses on these uncertainties in PIKA, with an emphasis on the peak-integration imprecision which is most relevant to PMF (and other least-squares-minimization techniques). The next section describes the PIKA fitting procedure to provide a basis for this discussion.

2.2 Constrained peak fitting in PIKA

In PIKA (up to the current version, 1.11C), the signal intensity y due to a specific ion is considered to arise from a peak ϕ, together with some noise σ_y, for each measured ion-time-of-flight x (corresponding to different m/z):

$$y(x) = \phi(x) + \sigma_y. \tag{1}$$

To a first approximation, the peak ϕ has a Gaussian shape and may be modelled by

$$G = h \cdot \exp\left[\frac{-(x-\mu)^2}{w^2}\right], \tag{2}$$

where h is the peak amplitude (peak height), μ the mode and mean of the Gaussian (peak location), and w the standard deviation (SD; peak width). These parameters may be estimated by least-squares fitting to the data.

In practice, the Gaussian model is modified to account for peak broadening, skewness, tailing, or other instrumental non-idealities (DeCarlo et al., 2006) by defining a peak-shape function v. The v is determined empirically by averaging a large number of normalized, well-resolved peaks from a given data set and may be visualized as a one-dimensional vector of correction factors in normalized x space ($x_{\mathrm{norm.}} = [x-\mu]/w$). Including v in Eq. (2) defines the pseudo-Gaussian function (DeCarlo et al., 2006),

$$f = v \cdot G = h \cdot v \cdot \exp\left[\frac{-(x-\mu)^2}{w^2}\right],$$
$$f = h \cdot f_0. \tag{3}$$

To improve the robustness of the PIKA fitting routine against poorly resolved peaks and noisy data, some of the parameters in Eq. (3) are constrained during peak fitting. The v is determined as described above, μ is obtained from an m/z calibration, and w is from a peak-width parameterization, $w(m/z)$. These are obtained as follows.

The m/z calibration is achieved by fitting Eq. (2) to user-selected background peaks. The fitted peak locations μ in time-of-flight space are calibrated to their known m/z following the procedure described by DeCarlo et al. (2006). Separate m/z calibrations are obtained for every mass spectrum. In practice, these calibrations are applied to predict peak locations in ion time-of-flight space before fitting the raw data. We therefore refer below to "μ-prediction errors" rather than "m/z-calibration errors" except when specifically discussing m/z calibrations. Although all peaks were fitted in ion-time-of-flight space below, the discussion refers to peak m/z for clarity.

The $w(m/z)$ calibration is also obtained by fitting Eq. (2), but to a separate set of peaks. This set of peaks is carefully selected to ensure that only single, well-resolved peaks of unambiguous composition are considered. The fitted peak widths w are averaged before being parameterized by a (typically linear) calibration curve, as described by DeCarlo et al. (2006). In the present analysis, the $w(m/z)$-calibration procedure was slightly modified to improve its robustness against poor-quality peak fits, as described in detail in Corbin et al. (2015a) and briefly in the Appendix.

Thus, the three inputs to Eq. (3) are defined from three different calibrations on the data. These calibrations are fundamentally different: the $w(m/z)$ and v calibrations are determined once for the entire data set, while the m/z calibration is determined for each individual mass spectrum. Variability in the accuracy of the m/z calibration may therefore lead to variability in the analysis. Also, different peaks are used for all three calibrations, so that their prediction errors may also differ.

With the inputs described above, Eq. (3) is fitted to estimate h. An example is shown in Fig. 1, showing the CO_2^+ signal at m/z 44 in the lower panel and the fit residuals in the upper panel. The peak spans only 2–3 detector bins (corresponding to 4–6 ns), which was typical for m/z in the range 15–75 for this specific instrument. Other studies have reported slightly lower m/z resolutions and correspondingly a slighly higher number of detector bins representing each peak (e.g. DeCarlo et al., 2006; Sun et al., 2010; Ortega et al., 2013). The conclusions drawn at the end of this study are independent of this detail.

The fitted h is used by PIKA to estimate the area A of each peak via the Gaussian integral as $A(G) = hw\sqrt{\pi}$, which leads to (Sueper et al., 2011)

$$A = hw\sqrt{\pi} \cdot \frac{A_{f_0}}{A_{G_0}} \cdot k_{\mathrm{DC}}, \tag{4}$$

where A_{G_0} and A_{f_0} are the respective integrals of the standard Gaussian and pseudo-Gaussian distributions. These standard integrals are defined by Eqs. (2) and (3) with unit h and w, and $\mu = 0$. The factor k_{DC} is a correction for the mass-spectrometer duty cycle to account for the fact that lighter ions will be transported faster than slower ions into the mass spectrometer prior to time-of-flight measurement and, consequently, reside in the extraction region for shorter times (Hings, 2006). The ratio A_{f_0}/A_{G_0} is typically close to unity.
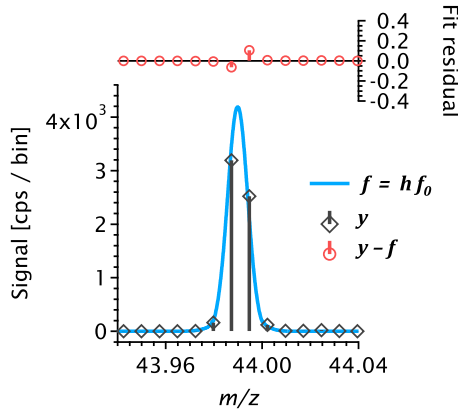
Figure 1. Example of a PIKA fit. The raw data (y; counts per second per bin), fit (f), and residuals ($y - f$) are shown.

After estimating A, PIKA currently estimates the uncertainty in A from the square root of the number of ions associated with the signal, according to an assumed Poisson distribution of errors (Allan et al., 2003).

In the HR-AMS, more than one peak is typically observed at each integer m/z. These peaks are modelled as a linear superposition f_T of pseudo-Gaussians with similar w (Sueper et al., 2011):

$$f_T = \sum_{i=1}^{j} h_i f_{0,i}(\mu_i, w).\qquad(5)$$

The uncertainty of fits to Eq. (5) with $j > 1$ (multiple overlapping peaks) is much more complex than the case where $j = 1$ ($\equiv$ well-resolved peaks).

The next section therefore focusses on the case of well-resolved peaks, before Sect. 5 discusses the overlapping-peak case.

3   Causes of peak-fitting uncertainty: well-resolved peaks in a real data set

3.1   Integration uncertainty for well-resolved peaks

The integration of fits to well-resolved pseudo-Gaussian peaks via Eq. (4) will lead to an uncertainty,

$$\left(\frac{\sigma_A}{A}\right)^2 = \left(\frac{\sigma_h}{h}\right)^2 + \left(\frac{\sigma_w}{w}\right)^2,\qquad(6)$$

which, if the fractions σ_h/h and/or σ_w/w are sufficiently large, may contribute significantly to the overall uncertainty of integrated mass-spectral signals.

In Eq. (6), the uncertainties in A_{f_0} and k_{DC} are considered negligible, since the integral A_{G_0} is known exactly; the integral A_{f_0} is also known exactly for a given v; and the uncertainty in k_{DC} depends only on the stability of the mass-spectrometer voltages, which fluctuate negligibly for the time

periods over which data are collected (Sect. 2.1). In addition, h and w have negligible covariance as they are determined via separate and independent calibrations (Sect. 2.2); the $w(m/z)$ calibration requires, and can be used to validate, the assumption that w is independent of h.

The value of σ_w may be directly estimated from the peak-width calibration procedure via the uncertainties in the calibration-fit coefficients. For the test data set discussed below, this approach indicated $\sigma_w/w \approx 2.5\%$. Such a percentage uncertainty is important as it scales with signal differently to the typically used Poisson uncertainty, increasing linearly with the number of ion counts rather than as the square root.

The estimation of σ_h is much more complex than σ_w. Whereas σ_w may be estimated directly from the calibration fit, σ_h reflects the imprecision of the constrained PIKA fitting procedure (Sect. 2.2), the constraints of which are predicted by two separate calibrations (peak width and location) as well as an empirically defined peak shape v.

3.2   Theoretical consideration of single-peak fitting errors

In this subsection we describe the theoretical basis on which the RMSE, defined below, has been interpreted in Sects. 3.3 and 3.4. The focus is on well-resolved peaks; overlapping peaks are addressed in Section 5.

Since the width, location, and shape of the PIKA pseudo-Gaussian fit function (Eq. 3) are predefined by different calibrations, the exponential term of Eq. (3) reduces to a transformation of the x variable:

$$f = h f_0(x, w, v, \mu),\qquad(7)$$

where f_0 is defined by Eq. (3) with $h = 1$ and with μ, w, and v given by the calibrations described in Sect. 2.2.

A fit to Eq. (7) is equivalent to a simple linear regression. (Indeed, the newest version of PIKA directly performs such a regression.) An uncertainty estimate for h could thus be obtained directly from the regression, but this estimate would not account for bias in the fit. Such bias must be accounted for, because when such bias varies from fit to fit (i.e. from mass spectrum to mass spectrum for the same peak), it produces an imprecision in h. Such a varying bias is to be expected, because a new f_0 is defined for each fitted ion. The accuracy of f_0 may vary in this context due, for example, to variations in the accuracy of the m/z-calibration equation which is obtained for each mass spectrum.

The influence of such variations on the imprecision of the fitted h values cannot be estimated directly from a single fit but must be considered in the context of the whole data set. It can be expected, however, that (in the absence of noise) errors in h due to f_0 will be proportional (scale linearly with h) since f_0 is scaled by h during fitting.

To quantify the error in a given fit of f, we use the SD of the fit residuals, the fit root-mean-square error (RMSE):

$$
\begin{aligned}
\text{RMSE} &= \sqrt{\frac{1}{j-1} \sum_{i=1}^{j} \left[y(x_i) - f(x_i) \right]^2} \\
&= \sqrt{\frac{1}{j-1} \sum_{i=1}^{j} \left[r_i \right]^2} = \sqrt{\frac{\text{SSR}}{j-1}}.
\end{aligned} \tag{8}
$$

As the RMSE of a single peak is the main quantity of interest below, the range $1 \le i \le j$ was limited to 2 SDs on either side of the peak centre ($\mu \pm 2w$), representing the trimmed variance of the residuals (e.g. Wilks, 2011).

The RMSE is not used to directly infer the error in the fitted h but rather as a diagnostic with which the causes of errors in f_0 may be elucidated and understood. Based on this understanding, the imprecision in h is then estimated.

For a sufficiently large sample, the expected value of the squared RMSE is the sum of the model variance and the squared model bias (Wilks, 2011). In the context of this discussion, this statement may be illustrated as follows. The RMSE represents the expected value of a fit residual, $\langle r_i \rangle$, which is related to the fitted and true peak functions $h_{\text{fitted},i} f_0$ and $h_{\text{true},i} \phi_0$ as

$$
\langle r_i \rangle = \langle h_{\text{fitted},i} f_0 - h_{\text{true},i} \phi_0 + \sigma_{y,i} \rangle, \tag{9}
$$

where $h_{\text{fitted},i}$ and $h_{\text{true},i}$ are scalars to the estimated and true unit-height peak functions f_0 and ϕ_0 respectively, and $\sigma_{y,i}$ represents noise in the data due to ion counting or background subtraction errors.

Letting $f_{0,i} = \phi_{0,i} + \varepsilon_i$, where ε_i is the error in f_0 at each data point resulting from imperfect knowledge of the true peak shape,

$$
\begin{aligned}
\langle r_i \rangle &= \langle h_{\text{fitted},i} \phi_0 + h_{\text{fitted},i} \varepsilon_i - h_{\text{true},i} \phi_0 + \sigma_{y,i} \rangle \\
&= h_{\text{fitted},i} \phi_0 + h_{\text{fitted},i} \langle \varepsilon_i \rangle - h_{\text{true},i} \phi_0 + \langle \sigma_{y,i} \rangle. \tag{10}
\end{aligned}
$$

For small peaks, counting uncertainties dominate the fit residual: $(h_{\text{true},i} \phi_0) / \sigma_{y,i} \to 0$, thus $\langle r_i \rangle \to \sigma_{y,i}$ in Eq. 10 (cf. Fig. 11).

For large peaks, the counting uncertainties $\sigma_{y,i}$ become negligible relative to $h_{\text{fitted},i} \langle \varepsilon \rangle$. In this situation, the slope of a plot of $\langle r_i \rangle$ against $h_{\text{fitted},i}$ will be

$$
\begin{aligned}
\frac{\mathrm{d} \langle r_i \rangle}{\mathrm{d} h_{\text{fitted},i}} &= \langle \varepsilon_i \rangle - \phi_{0,i} \left[\frac{\mathrm{d} h_{\text{true},i}}{\mathrm{d} h_{\text{fitted},i}} - 1 \right] \\
&= \langle \varepsilon_i \rangle, \tag{11}
\end{aligned}
$$

where $\mathrm{d} h_{\text{true},i} / \mathrm{d} h_{\text{fitted},i} = 0$ if the peak shape is independent of the peak intensity, as was already assumed (Sect. 2). This assumption is discussed further in Sect. 3.4.

Thus, for signals large enough for noise to be negligible, a plot of $\langle r_i \rangle$ (or an estimator of $\langle r_i \rangle$, such as the RMSE)
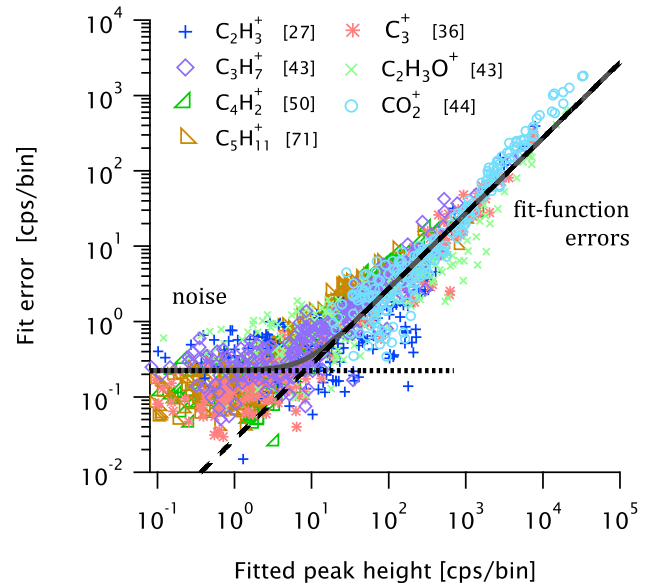

Figure 2. The RMSE of standard PIKA fits to well-resolved peaks in the test HR-AMS data set. Numbers in the legend show the integer m/z of these peaks. The dotted line indicates the lower limit of approximately constant RMSE for each ion, the dashed line the upper limit of constant relative RMSE, and the solid line the quadratic sum of the two.

against the fitted peak height h_{fitted} may show a linear slope where errors in the scaled peak-shape function f_0 dominate counting errors. The linearity arises because of the linear scaling of f_0 by h_f. If a unique f_0 is prescribed for each fit, as is the case in PIKA, then variability within the set of prescribed f_0 may lead to an overall imprecision σ_h in the set of fitted h_{fitted}. Below, we refer to h_{fitted} simply as h. For fits to well-resolved ions, we plot the RMSE against h to evaluate the factors that lead to the constant relative error defined in Eq. (11) in order to build an understanding on which the corresponding σ_h may be estimated in Sect. 4.

3.3 Identification of single-peak fitting errors using fit RMSE

The RMSE of a number of well-resolved peaks in the test data set are plotted in Fig. 2. Each point in the figure represents the RMSE of a fit of Eq. (7) to the data. The figure includes seven well-resolved peaks based on their being both well-resolved from any neighbours and present at a wide range of signal intensities. The mass spectra have been background subtracted before fitting as described in the Appendix. Background subtraction was performed to remove inconsequential differences between the ions, for example due to signals from background CO_2 gas. Other ions had very small background signals regardless.

The peaks in Fig. 2 span a range of different m/z, as indicated by the integer m/z values shown in the legend, and represent a range of different species. For example, C_3^+

ions were formed after BC vaporized at $\sim 4000\,\text{K}$, whereas $C_2H_3O^+$ and $C_3H_7^+$ ions were formed after organic PM vaporized below $873\,\text{K}$ ($600\,°\text{C}$) (Corbin et al., 2015a). In contrast to all other ions, these latter two organic ions are found at the same integer m/z and were therefore fitted simultaneously as a linear sum in PIKA (Eq. 5 with $j = 2$). This point is returned to in Sect. 5.

Two distinct trends are evident in Fig. 2. First, the RMSE approaches an asymptotic value for low peak heights ($\sim 1\,\text{counts s}^{-1}\,\text{bin}^{-1}$; cps bin^{-1}). This approximately constant RMSE ($\sigma_y \approx 0.2\,\text{cps bin}^{-1}$) is representative of noise in the data y. Second, the RMSE shows a constant slope for high-signal peaks, indicating an approximately constant relative RMSE for each peak (RMSE/$h \approx 3\,\%$). This illustrates the importance of fitting-function errors during the fit (Sect. 3.2). The constant RMSE and constant relative RMSE are respectively depicted by the dotted and dashed lines in Fig. 2, as well as their quadratic sum.

The constant RMSE/h for high signals suggests that σ_h/h may also be constant. Thus σ_h may also scale linearly with signal, as was the case for σ_w/w (Sect. 3.1), and may significantly influence the overall peak-integration uncertainty for high signals. The next sections therefore aim to obtain an understanding and quantification of σ_h/h for incorporation into the overall HR-AMS uncertainty.

3.4 Impacts of fit-parameter errors on the RMSE of single peaks

The source of the constant relative RMSE in Fig. 2 was argued via Eq. (7) to be most likely due to errors in the predefined pseudo-Gaussian function f_0. To elucidate which of the parameters defining f_0 have the greatest impact on the errors of the resulting fit, imprecisions or biases were manually added to each of several input parameters in the PIKA analysis procedure as follows.

The original magnitude of each error was estimated directly from the data. Based on these estimates, significantly larger uncertainties were added to the data, as specified in Table 2. In most cases, the errors specified in the table were used to define the SD of a Gaussian probability distribution from which a new error was sampled for each analysed peak.

Figure 3 plots the effects of these errors on the resulting RMSE for one exemplary ion, $C_3H_7^+$. This ion was chosen simply because it represented a large range of peak intensities (abscissas in Fig. 3); the trends seen in Fig. 3 were observed for all of the ions in Fig. 2 (not shown). The magnitudes of the errors used to generate the figure are highlighted in boldface in Table 2 and were chosen so as to give a visible change in the RMSE graphs (where possible), not to represent realistic errors.

With this approach, multiple potential sources of the constant relative RMSE term can be eliminated: noise in the predicted peak width w (Fig. 3b), errors in the slope of the subtracted linear baseline (Fig. 3d), and the use of a different

AMS vaporizer (Fig. 3e) did not alter the magnitude of the relative RMSE.

Conversely, two potential sources of the constant relative RMSE can be identified: noise in the predicted peak location μ (Fig. 3f) and errors in the peak-shape v (Fig. 3g and h). In these three cases (Fig. 3f–h), a relative RMSE of $\sim 15\,\%$ was observed.

The relative μ-prediction errors necessary to achieve this $\sim 15\,\%$ relative RMSE were 5–10-fold larger than typical μ-prediction errors; the errors in v necessary to achieve this $\sim 15\,\%$ relative RMSE were 2 orders of magnitude larger than the maximum distance between the best estimate of v and the data used to derive it, which was $0.4\,\%$. This suggests that m/z-calibration uncertainties play a larger role than peak-shape uncertainties in a real data set. This difference in sensitivity is not surprising given that μ is part of the exponential term in Eq. (3) while v is a linear term. In the next section, the fitting procedure was altered to prove that the majority of the fitting error was due to μ-prediction errors.

One additional implication of Fig. 3g and h is that RMSE/h is a function of v. Given that Fig. 2 showed that this fraction was nearly constant for high signals, we infer that the peak shape was nearly constant across this range of signals. If v was independent of signal intensity, then the individual ions comprising each peak interacted negligibly within the instrument, which provides justification for the assumed behaviour of overlapping peaks as a linear sum in Eq. (5).

3.5 Impacts of fit constraints on single-peak fitting errors

To explore the impact of errors in μ prediction on the fit RMSE, the fitting procedure was altered to allow μ and/or w to be varied during fitting. The peak shape v could not be meaningfully varied.

Allowing μ to vary by $\pm 20\,\text{ppm}$ during fitting reduced the fit RMSE by almost 1 order of magnitude at high signals (Fig. 4, red triangles), indicating that the majority of the fitting error was due to errors in the predicted μ (which result from both imprecisions and biases in m/z calibration). The RMSE at low signals was virtually unchanged, since it was dominated by noise in the data (Sect. 3.3). While Fig. 4 illustrates this result using $C_3H_7^+$, these conclusions were verified for all of the well-resolved ions discussed above.

Further relaxing the μ constraint to $\pm 50\,\text{ppm}$ did not further reduce the RMSE. Conversely, reducing the constraint to $\pm 10\,\text{ppm}$ resulted in a change intermediate between the $0\,\text{ppm}$ (fully constrained) and $\pm 20\,\text{ppm}$ constraints, suggesting that the true accuracy of the m/z calibration was on the order of $\pm 15\,\text{ppm}$. This value does not correspond to the imprecision of the m/z calibration, as discussed in the next section.

Allowing w to vary during the fit by $\pm 5\,\%$ ($2\,\sigma_w$) had no observable effect on the RMSE (Fig. 4). Although this re-

Table 2. Effects of manually introducing errors to the PIKA analysis procedure. Cases highlighted in bold are plotted in Fig. 3. Bracketed values are the estimated errors for the reference case.

Quantity	Error	Effect on constant RMSE	Effect on constant RMSE/h
Data[a] (x)	[~ 0.2], 1, **5**, 50 cps/bin	Increased	Negligible
Width[a] (w)	[< 2.5], 5, **25**, 50, 250 %	Negligible	Increased spread
Baseline height[b]	[< 0.1], 1, **5**, 50 cps/bin	Increased	Negligible
Baseline slope[a]	[10^{-5}], **10^{-3}**, 10^{-1} cps/bin $(m/z)^{-1}$	Negligible	Negligible
Location[a] (μ)	[~ 20], **50**, 100, 200 ppm	Negligible	Increased and increased spread
Peak shape[c] (v)	[Empirical,] **positive** or **negative** skewness	Negligible	Increased

[a] Added as Gaussian noise. [b] Added as the absolute value of Gaussian noise. [c] Skewed by multiplying the empirical peak-shape function with a normal cumulative distribution function, resulting in a 25 % smaller peak area (see Fig. 3g and h).
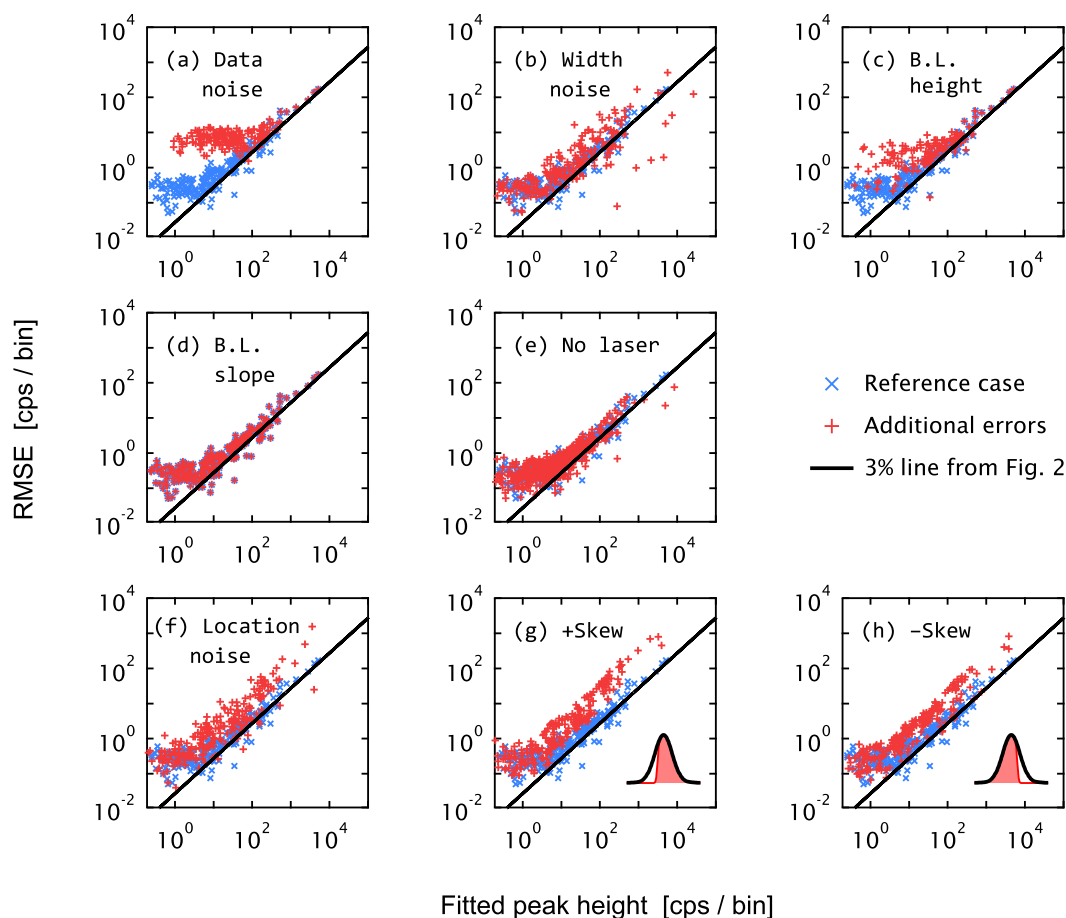

Figure 3. Response of the RMSE of a representative ion ($C_3H_7^+$) to the addition of very large errors at various stages of analysis (**a–d, f–h**). The added errors are defined by the highlighted values in Table 2. The laser vaporizer of a dual-vaporizer SP-AMS (see Corbin et al., 2015a) was on for all data except those of panel (**e**). B.L. is baseline.

sult may appear to contradict the fact that a significant uncertainty in the w calibration was estimated during calibration (Sect. 4.3), it is fully consistent with the only minor changes in RMSE observed when large errors were added to w in Sect. 3.4 (Fig. 3b). Thus, although uncertainties in w do not strongly influence the fitted h, they lead to uncertainties during the peak integration by Eq. (4) (cf. Eq. 6).

When both w and μ were allowed to vary during fitting, the RMSE behaved similarly to the case where only μ was allowed to vary (i.e. the red triangles in Fig. 4). In this case, some linear dependence of the RMSE on the fitted peak height remained, suggesting that f_0 still had some influence on the RMSE. This remaining dependence is attributed to errors in the peak-shape factor v. Since the empirically defined

Table 3. Summary of h-fitting biases and imprecisions from data and simulation. The second column, μ-prediction error, reflects errors in m/z calibration as estimated from the test data set for each ion (Fig. 5). The third to fifth columns show different Monte Carlo estimates of the error in fitted h as bias, and imprecision pairs (b_h, σ_h). Respectively these different estimates are the best estimate, an estimate where b_h is not taken into account, and the best estimate when broader peaks were simulated (resulting in twice the number of measurements per peak and representing potential differences between instruments).

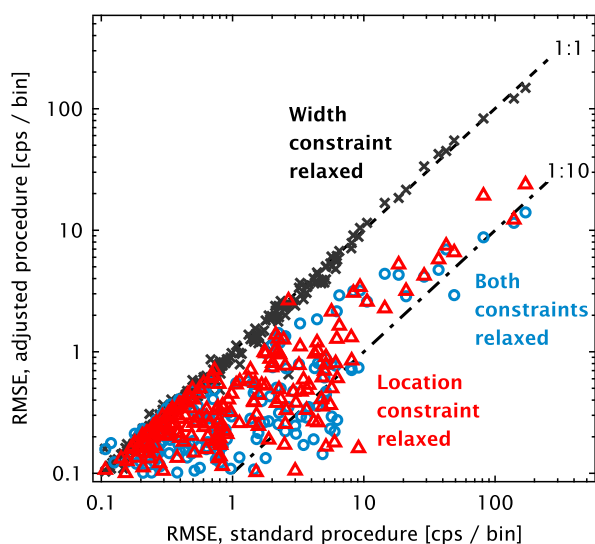| Ion	μ-prediction error	Percentage error in fitted h: bias, imprecision		
		from best-estimate μ errors	with μ imprecision only	with broader peaks
$C_2H_3^+$	4.6 ± 9.5	$-0.35, 1.06$	$-0.20, 0.93$	$-0.08, 0.13$
$C_3H_7^+$	-10 ± 7.5	$-0.65, 1.64$	$-0.17, 0.98$	$-0.19, 0.21$
$C_4H_2^+$	-14 ± 5.7	$-1.06, 2.10$	$-0.11, 0.78$	$-0.31, 0.23$
$C_5H_{11}^+$	-15 ± 5.0	$-1.03, 2.46$	$-0.12, 0.79$	$-0.49, 0.31$
C_3^+	4.8 ± 12	$-0.51, 1.52$	$-0.41, 1.43$	$-0.16, 0.24$
$C_2H_3O^+$	-5.9 ± 8.3	$-0.39, 1.36$	$-0.22, 1.06$	$-0.13, 0.17$
CO_2^+	6.9 ± 9.0	$-0.12, 1.41$	$-0.20, 1.13$	$-0.15, 0.21$

Figure 4. Changes in the RMSE of $C_3H_7^+$ in response to changes in the fitting procedure. The normal PIKA fitting procedure (abscissa) is compared to procedures where μ was allowed to vary by ± 20 ppm (red $\triangle$) and where w was allowed to vary by ± 5 % (grey $+$), or both were allowed to vary (blue $\circ$). Lines highlight the trends expected for no decrease in the RMSE (1 : 1) and a factor-of-10 decrease (1 : 10).

v (a vector of more than 100 values) could not be allowed to vary during fitting as μ and w were in Fig. 4, its influence on the RMSE trends could not be explored in this context. If the remaining RMSE trend was due to errors in v, these errors had a much-smaller impact on the RMSE (and therefore the goodness of fit of Eq. 3) than the errors in μ prediction, so they are not discussed further.

3.6 Estimation of μ-prediction (m/z) errors from single-peak fits: consequences for h-imprecision estimation

When the predicted peak location $\mu_{\text{predicted}}$ was allowed to vary during fitting (Sect. 3.5) the relative RMSE was significantly reduced in the high-signal regime. The final fit therefore represented a significantly better model of the data, which is interpreted as an improvement due to a reduced error in the μ. The fitted peak location, μ_{fitted}, is therefore interpreted as a good approximation to the true peak location, μ_{true}, and used to estimate errors in μ prediction as $e_\mu \approx (\mu_{\text{fitted}} - \mu_{\text{predicted}})$ in the current section. Limitations of this estimation approach are discussed in Sect. 6.

The dark-shaded data in Fig. 5 show the distribution of the estimated μ-prediction errors for peaks outside of the noise regime, defined as $h_{\text{fitted}} > 20$ cps/bin from Fig. 2. For this analysis, μ was allowed to vary by ± 40 ppm. If the aggregated e_μ data are fitted by a Gaussian function, the mean appears to be virtually 0 (fitted value: -0.5 ± 0.8 ppm), suggesting zero bias in $\mu_{\text{predicted}}$. However, the bias for each individual ion was not 0, as shown by the light-shaded data in the same figure. The μ_{fitted} for each ion showed significant biases (mean different from 0) as well as imprecisions (significant spread relative to the mean). The magnitude of these biases and imprecisions were estimated by a Gaussian fit in each case, with the fitted values shown in Table 3. The imprecisions in $\mu_{\text{predicted}}$ are generally of similar magnitude to the biases, and the biases show both positive and negative values. These biases must therefore be taken into account when estimating h-fitting imprecision for individual ions.

Figure 6 plots the data for m/z 43, $C_2H_3O^+$ (m/z 43.0184), and $C_3H_7^+$ (m/z 43.0548) that were shown in Fig. 5 as a scatter plot. As these two ions are of similar m/z, but are nonetheless well resolved with a separation of ~ 7 standard deviations (cf. Cubison and Jimenez, 2015), it was hypothesized that their μ-prediction biases may have
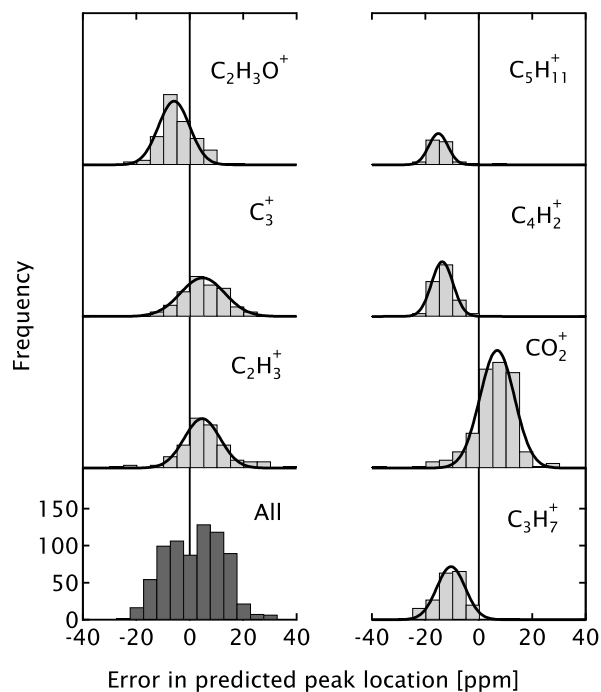
Figure 5. Distribution of prediction errors in μ estimated from the test data set for each of the ions shown in Fig. 2, including only peaks outside of the noise regime ($h_{\text{fitted}} > 20\,\text{cps bin}^{-1}$). Light-shaded data represent specific ions and dark-shaded data the aggregate of all ions. The light-shaded data are scaled similarly, so that their relative areas represent the relative number of data used to construct them. The mode and widths of the Gaussian fits to each histogram are given in Table 3.

been approximately equal. Such approximate equality would have allowed μ-prediction biases to be fitted as a single free parameter during fitting of Eq. (5). Alternatively, a single value for the μ-prediction bias could have been assumed during Monte Carlo analysis (Sect. 4).

The lack of a 1 : 1 trend in Fig. 6 (indeed, a slight negative trend is evident) suggests that the μ-prediction biases for nearby peaks cannot be assumed to be equal. This unexpected result may be related to the sensitivity of the fits to the location of the mode of a peak relative to the nearest detector bin, which is discussed in Sect. 4.1. The method by which e_μ was estimated may also play a role, but to some degree this method must also represent the sensitivity of the fitting algorithm to the data. Since this sensitivity also influences the ultimate quantity of interest, σ_h, this conclusion would remain true regardless of the accuracy and precision of the e_μ estimate. Moreover, we emphasize that ion-dependent biases in m/z calibration are not unique to this study, as discussed in Sect. 6.

Although it is readily apparent that a μ-prediction bias may lead to a bias in the fitted h for a given ion, it is less apparent that such a bias also significantly increases σ_h for any given μ-prediction imprecision. This occurs because the
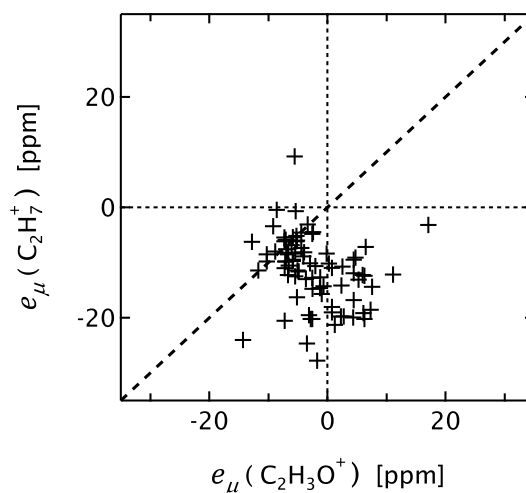

Figure 6. Scatter plot of the estimated errors in peak-location prediction (same data as in Fig. 5), e_μ, for the ions $C_2H_3O^+$ and $C_3H_7^+$, both found at $m/z\,43$.

slope of a peak increases with increasing distance from its mode (for small distances), such that the variation in fitted height is larger for peaks fitted further away from their mode. The magnitude of this effect is illustrated in Fig. 10, which is discussed below. It is therefore essential consider μ-prediction biases when estimating the magnitude of σ_h.

4   Monte Carlo estimation of σ_h for well-resolved peaks

A Monte Carlo approach for the estimation of peak-fitting uncertainties was developed based on the above discussion. With this approach, σ_h is directly evaluated by sampling μ-prediction errors (both m/z-calibration biases and imprecisions), repeatedly fitting a selected peak, and calculating the standard deviation of the resulting set of fitted h. Below, this approach is described in more detail and used to investigate the μ-prediction errors observed in the test data set. Then, the conditions under which the approach may be generalized to the case of overlapping peaks are explored.

The Monte Carlo approach used to evaluate σ_h in the test data set was evaluated by simulating peaks via Eq. (3) using the m/z axis of a mass spectrum from the test data set. For each simulated peak, a different m/z axis was randomly selected. The simulated peaks were generated using the peak shape v and width w expected for this data set, and the Gaussian noise described in Sect. 3.3 was added to the data. The simulated peaks were then fitted to Eq. (3) with μ intentionally constrained to erroneous values. These erroneous μ constraints were obtained by sampling from a Gaussian distribution with mean and SD given by the fits in Fig. 5. For the Monte Carlo estimation of σ_h for standard analyses, this approach must be modified slightly because m/z-calibration biases are unknown, as discussed further below.
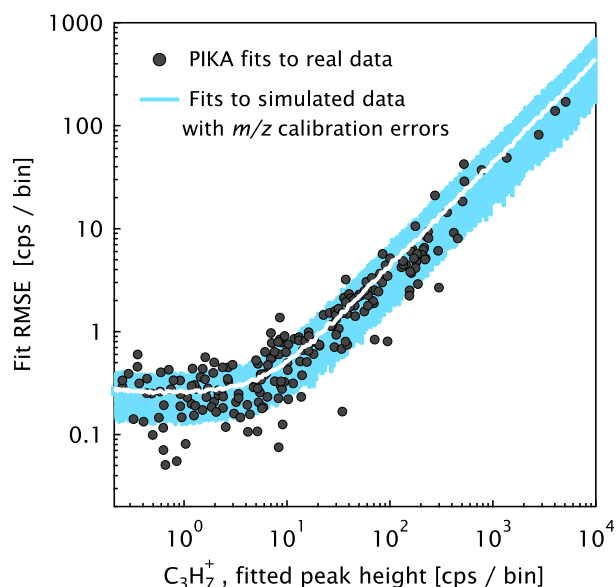
Figure 7. The RMSE of $C_3H_7^+$ fits in the test data set (symbols) and to simulated data (line, mean; shading, SD). Peaks of the empirically determined peak shape were simulated on the same m/z-axes as the measurements and were fitted using Eq. (3) after adding noise (0.35 cps bin^{-1}) and emulating bias and imprecision errors in the location constraints (Table 3). For each point, multiple simulations on randomly selected m/z axes were performed to estimate the standard deviation of the RMSE (blue shading).

The RMSE of these fits to the simulated data are plotted in Fig. 7 for $C_3H_7^+$. Each simulated peak height was fitted 100 times, each time with a different sample of μ-prediction and data-noise errors and on a different m/z axis. Increasing this number of samples did not affect the results. The figure shows the resulting mean (white line) and SD (light-blue shading) of the RMSE, which are in excellent agreement with the data.

It was noted in Sect. 3.5 that some small errors in v were likely present. Although no physical basis for the functional form of such errors was available, the impact of errors in v was investigated by adding arbitrary functions to the true peak shape. This allowed the trends in Fig. 4 for the relaxed-μ constraint case to be reproduced but made no visible difference to Fig. 7. Errors in v were therefore omitted from the model.

4.1 Imprecision for well-resolved peaks with known μ-prediction errors

The blue points in Fig. 8 (case A) show the σ_h corresponding to the RMSEs shown in Fig. 7. The imprecision increases as h becomes smaller due to the noise in the data but reaches an asymptotic value as h becomes larger. As noted above, these points represent random samples from the entire test data set.

In practice, it is desirable to estimate σ_h on a case-by-case basis for each peak in each mass spectrum. Such an estimate
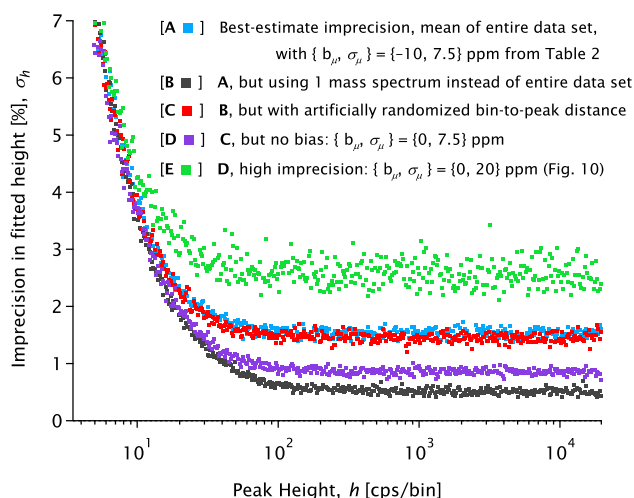

Figure 8. The imprecision in $C_3H_7^+$ fitted peak heights, σ_h, estimated via the Monte Carlo approach described in the text. Although counting imprecisions have been omitted so as to illustrate the influence of the different approaches (**a–e**), Gaussian noise with standard deviation 0.35 cps bin^{-1} was added to each mass spectrum to allow comparison with Fig. 7. The different approaches (**a–c**), detailed in the figure legend, illustrate the need to consider the position of the peak relative to the sampling grid ("bin-to-mode distance") when simulating peaks comprised of relatively few data points (this peak is illustrated in the upper-left panel of Fig. 10). The approaches (**d**) and (**e**) illustrate the σ_h that would be estimated in the absence of exact knowledge of the m/z-calibration bias.

is shown by case B in Fig. 8, which was obtained by performing the Monte Carlo approach described above except the random selection of different m/z axes was not performed. Different m/z axes showed different asymptotic values; the selected axis was chosen for its visual distinction. This difference does not represent real variability in the data: again, only the m/z axis was taken from the data set and the peak itself was simulated. Rather, the difference was due to a shift in the relative distance between the data points representing the peak and the m/z axis (the "bin-to-mode distance"), which is arbitrary and can vary between mass spectra. Case C demonstrates this fact by repeating case B while shifting the mode location of the peak by a randomly selected fraction of the bin spacing between 0 and 1 for each Monte Carlo fit. The result is an estimate of σ_h that is comparable to the entire-data-set case and represents fit-function errors rather than the bin-to-mode distance.

For overlapping peaks, the μ-prediction imprecision and bias is generally not known. Case D repeats case C but without the μ-prediction bias. The result is an imprecision estimate that is much lower. For case E, a much higher imprecision of 20 ppm has been used, based on the recommendation from Sect. 4.2 that the estimated μ-prediction imprecision be increased to allow unbiased estimation of σ_h when μ-prediction biases are unknown.
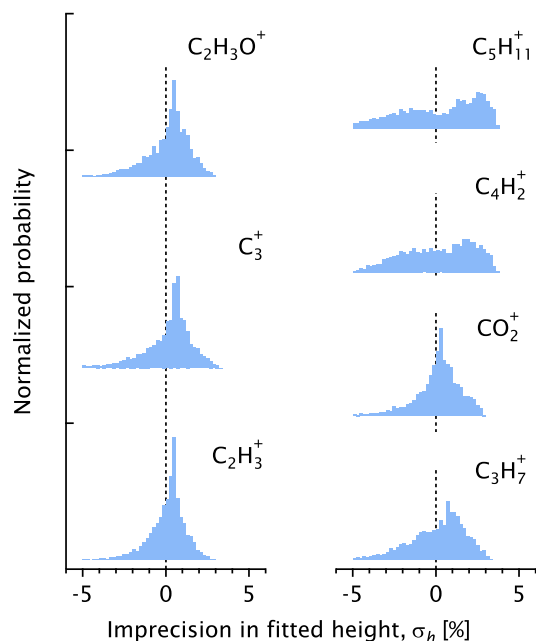
Figure 9. Normalized and centred probability distributions of the simulated imprecision in fits to well-resolved peaks from the test data set, using the μ-prediction biases and imprecisions given in Table 3. Distributions were centred by subtracting the mean simulated error, so that only fitting imprecisions (relevant to PMF) are shown. Simulations were performed on randomly selected m/z axes from the data set.

In Fig. 8, the standard deviation of multiple fits was used to represent σ_h. However, the distribution of errors comprising σ_h was not normally distributed, as shown in Fig. 9. Figure 9 includes all seven of the well-resolved peaks discussed above and only represents signals much larger than the noise level, since each followed the same asymptotic behaviour shown in Fig. 8 and since counting errors must ultimately be combined with peak-integration errors at low signal-to-noise ratios (Fig. 11). A large number of iterations (5000) were performed for each peak so that histograms could be plotted in detail. The mean fitting error (bias) has been subtracted from each distribution so that the plot represents σ_h. The errors are normalized to the absolute peak height, as they were a constant fraction of the height (as expected from Sect. 3). For all ions, the limiting case of very high ion signals was simulated. The reader is referred to Cubison and Jimenez (2015) for a discussion of the effects of ion-counting uncertainties on fitting imprecision.

Each probability distribution in Fig. 9 shows a very different skewness. These differences are largely controlled by the μ-prediction bias and imprecision, which may cause the predicted peak location to be consistently away from the peak mode. Away from the mode, the peak itself slopes more steeply and a small μ-prediction imprecision leads to a large variability in the fitted height. This fact means that a greater

imprecision results from a greater bias in the m/z calibration. The μ-prediction bias is therefore an important component of the imprecision of the fitted h. Without including this bias in estimating σ_h, the resulting imprecision estimates would be considerably different (Table 3).

The impact of the μ-prediction biases discussed above on σ_h will be reduced for broader (lower m/z-resolution) peaks, since such peaks are less steeply sloped. Broader peaks are also represented by more points in the mass spectrum, which has a major impact on the imprecision: for the case of $C_3H_7^+$, increasing the peak width by a factor of 1.5 (which doubled the number of data points representing $> 1\%$ of the peak height from 4 to 8) decreased the imprecision from 1.6 to 0.2 % (Table 3). This effect would not necessarily be reduced by increasing the number of points per peak without changing the mass-spectral resolution (Hilmer and Bothner, 2011) but instead only by reducing the μ-prediction errors.

4.2 Imprecision for well-resolved peaks with unknown μ-prediction errors

A non-negligible m/z-calibration bias is anticipated for any given mass-spectral peak, since ion trajectories are typically imperfectly described by the calibration function as discussed in Sects. 3.6 and 6. We explored the sensitivity of σ_h to this bias by calculating σ_h for the three exemplary peaks shown in the top row of Fig. 10, for a wide range of μ-prediction bias and imprecisions. Figure 10 thus illustrates the range of imprecision for which a given m/z-calibration bias becomes negligible: for the range of m/z-calibration biases where σ_h remains similar along a vertical cross-section of the middle panels of the figure, the actual value of the bias term need not be known.

To place Fig. 10 in context, the μ-prediction biases and imprecisions for ions in the test data set (Table 3) are shown in the central column. In the left- and right-hand side columns, only the μ-prediction errors representing the specified ions are shown. The central column was generated using a much-higher data-point density than was observed in the test data set and represents a higher data-sampling rate, as may be observed in other instruments. It also equivalently represents a higher m/z ion or a lower mass-spectral resolution instrument. The upper row of plots in Fig. 10 includes a horizontal dashed line to illustrate the peak width (w in Eq. 3). Also shown in this row are red vertical bars spanning ± 17 ppm, which is the standard deviation of the aggregated error estimates for well-resolved peaks in the test data set (shown in the lower-right plot of Fig. 5). The grey vertical bars span ± 10 ppm to illustrate the range of biases observed in Table 3. (The position of a peak relative to the grey bars will, by definition, change only between ions and between data sets, not between samples.)

The second row of plots in Fig. 10 plots σ_h for each of the three cases. Because the peak is nearly Gaussian, the plot is virtually symmetrical about the zero-bias line. Neverthe-
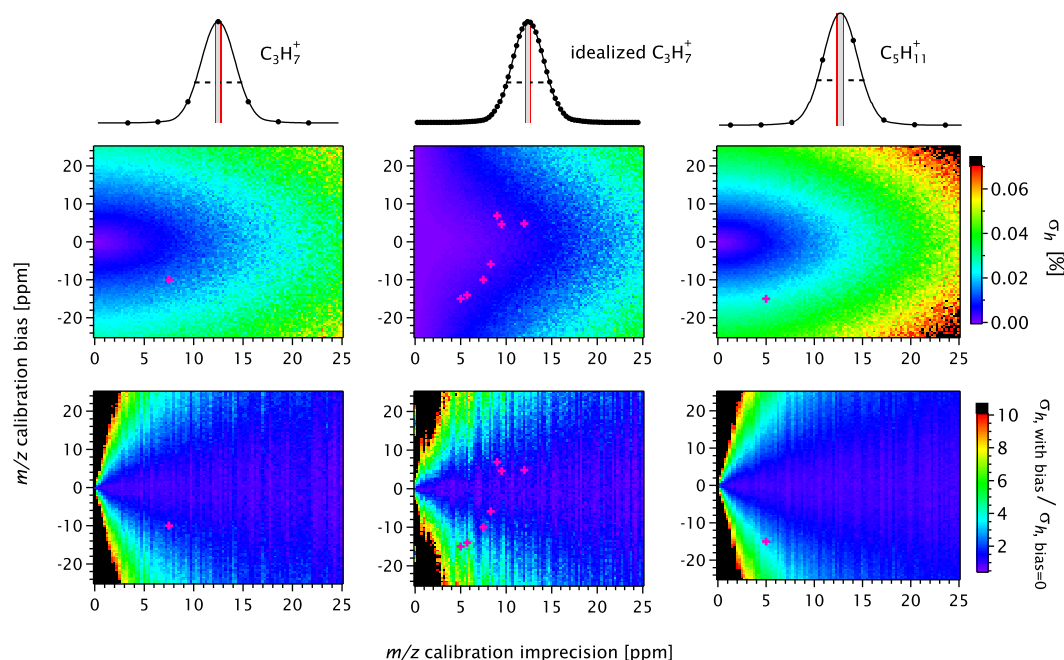
Figure 10. Response of the peak-fitting imprecision σ_h to variations in m/z calibration imprecision and mass accuracy, i.e. to variations in μ-prediction imprecision and bias. Each column corresponds to the simulated ion named and depicted in the first row respectively: $C_3H_7^+$ simulated on a normal m/z axis, $C_3H_7^+$ simulated with much a greater data point density (reflecting a broader peak and/or faster sampling rate), and $C_5H_{11}^+$ simulated on a normal m/z axis. The second row of plots depicts the Monte-Carlo-estimated σ_h (standard deviation of 300 fits with varying bin-to-mode distance and negligible counting noise) for each imprecision and bias (represented via the standard deviation and mean of a Gaussian distribution) specified on the horizontal and vertical axes respectively. The third row of plots depicts the effect of bias on σ_h: each point on the third-row images are normalized to $\sigma_{h,b=0}$, i.e. the horizontal line of zero bias from the respective second-row plot. The pink + represent the μ-prediction bias and imprecision estimated from the data, shown in the left and right columns for the respective ions, and in the central column for all seven ions in Fig. 9. The apparent noise is due to the variation of the bin-to-mode distance (see text). Vertical cross-sections on the second and third rows of plots correspond to the range of σ_h which may exist when μ-prediction bias is unconstrained.

less, we have plotted both negative and positive biases because this is the space in which the true, unknown biases of all peaks will lie.

The third row of plots in Fig. 10 shows the effect of μ-prediction biases on σ_h, relative to the no-bias case. That is, this row shows $\sigma_{h,\text{with bias}}$ divided by $\sigma_{h,\text{bias}=0}$, where the former quantity is the σ_h obtained when considering μ-prediction biases and the latter quantity is the σ_h obtained when neglecting those biases. It is clear that the larger the m/z-calibration imprecision, the smaller the importance of the m/z-calibration bias.

Thus, the bias need not be known or estimated for the accurate Monte Carlo estimation of σ_h; it is only necessary that the m/z-calibration imprecision employed for the Monte Carlo estimate is large enough that σ_h lies within the regime where the influences of m/z-calibration biases are negligible. This will result in σ_h values that are overestimated by a few percent but that are consistent between peaks with different m/z-calibration bias. For the biases observed in Fig. 10, a reasonable estimate for an m/z-calibration imprecision in

this regime may be 20 ppm; an example of the resulting σ_h is shown in shown in Fig. 8, case E.

Consistent estimates of σ_h between peaks with different calibration biases is important during least-squares fitting. This statement may be justified by the conclusions of Paatero and Hopke (2003), in particular the conclusion that underestimating the imprecision of certain variables during PMF can have significantly harmful effects. Those authors also concluded that overestimating uncertainties may have potentially positive effects. Both conclusions provide clear support for the present method.

An approach to estimating σ_h which does not require ion-specific estimates of m/z-calibration biases or imprecisions is also important during fits to multiple overlapping peaks, since the estimation of m/z-calibration biases in such a context (Sect. 3.6) may be difficult or impossible. This is further discussed in Sect. 5.
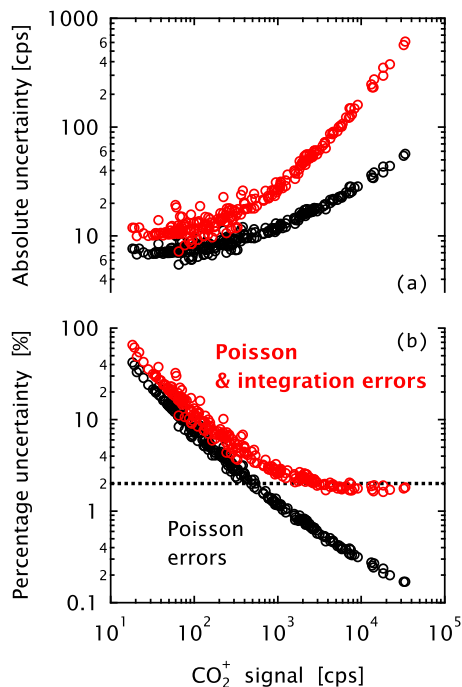
Figure 11. Example of the relationship between the uncertainty in the integrated area of for a single, well-resolved peak and its absolute area. The example uses real data for CO_2^+ and shows two cases; first considering only counting uncertainties (red) and second considering both counting and peak-integration uncertainties (black), following Eq. (13). The upper panel shows the absolute imprecision; the lower panel the relative imprecision. The horizontal dotted line illustrates the more-realistic asymptotic behaviour of the new approach.

4.3 Overall peak-integration imprecision: example for a well-resolved peak

With σ_w/w determined as $\sim 2.5\%$ (Sect. 3.1) and σ_h/h estimated by the Monte Carlo method described above (Sect. 4), for both well-resolved and overlapping peaks, the overall imprecision in the peak integration discussed in Sect. 3.1 can be estimated from Eq. (6) as

$$\left(\frac{\sigma_A}{A}\right)^2 = \left(\frac{\sigma_h}{h}\right)^2 + \left(\frac{\sigma_w}{w}\right)^2, \tag{12}$$

where σ_h/h is approximately constant (on the order of a few percent) for a well-resolved peak within a given data set and may become much larger as peak overlap becomes important (Cubison and Jimenez, 2015).

σ_A is considered independent of the Poisson counting uncertainties σ_p discussed in Sects. 2.1 and 5, so that the overall HR-AMS uncertainty may be expressed in units of ion counts per detector bin as

$$\sigma_{AMS} = \sqrt{\sigma_A^2 + \sigma_p^2}. \tag{13}$$

For well-resolved peaks, the first term in Eq. (13) scales linearly with signal, whereas the second term scales with the square root of the signal. Therefore, even for well-resolved peaks, the first term dominates at high signals. That is, in this case peak-integration imprecision dominates ion-counting imprecision. For overlapping peaks, peak-integration imprecision may become very large for much smaller signals.

While σ_h/h is most accurately estimated on a case-by-case basis using the Monte Carlo approach outlined above, we illustrate the importance of peak-integration errors by taking $\sigma_h/h \approx 4\%$ as a representative value for well-resolved peaks in the current data set (estimated based on Sect. 4.1), which together with $\sigma_w/w \approx 2.5\%$ gives

$$\sigma_A \approx (4.7\%)\,A \tag{14}$$

for well-resolved peaks in this data set. This is shown in Fig. 11, which has been obtained by programming Eq. (13) into PIKA using the value for σ_A given above. (As is standard in PIKA, the calculations have accounted for additional details related to sampling time, baseline noise, and mass-spectral duty cycle (Allan et al., 2003; Sueper et al., 2011). Due largely to sampling-time correction (Allan et al., 2003), the asymptote of Fig. 11b is less than 4.7 %.)

Fig. 11b illustrates the essential difference between fitting and counting errors. Whereas the relative uncertainty of Poisson errors ($\sim \sqrt{n}$) falls to 0 as the signal increases, the relative uncertainty of fitting errors ($\sim n$) tends to an asymptotic value as signal increases. This feature strongly affects the relative importance of high-signal data during uncertainty-weighted fitting. This importance will be even greater for peaks affected by overlap, for which fitting errors may increase rapidly and nonlinearly.

4.4 Influence on positive matrix factorization results

The arguments presented above clearly show that m/z-calibration limitations lead to a linear (fractional) imprecision term during the constrained peak-integration procedures of PIKA. This imprecision significantly increases the overall imprecision in a resulting data matrix; in particular, fractional imprecisions become more significant than ion-counting uncertainties for high signals, and overlap errors may be larger still. Thus, omitting this imprecision term from the uncertainty provided to PMF leads to an overweighting of higher-signal data, which may bias the PMF solution (Paatero and Tapper, 1994; Paatero and Hopke, 2003).

The importance of high signals means that fitting errors may be especially important for the high aerosol concentrations that may be measured at the roadside (e.g. Dallmann et al., 2014), within combustion plumes (e.g. Cubison et al., 2011), in highly polluted cities like Beijing or Mexico City (e.g. Zhang et al., 2007), or during source studies (e.g. Elsasser et al., 2013; Timko et al., 2014; Corbin et al., 2015a).

Given the variable importance of high signals within different data sets, and the importance of instrument-specific

parameters to the peak-integration imprecision as discussed above, an absolute statement of the relevance of these uncertainties in PMF is impossible. Nevertheless, as an example we performed PMF on a synthetic data matrix to demonstrate the significance of addressing peak-fitting errors.

The synthetic data matrix was constructed using the PMF solution presented by Corbin et al. (2015a), which corresponds to the test data set discussed herein. In detail: the factors (mass spectra) and loadings (time series) of the Corbin et al. (2015a) PMF solution were recombined into an error-free synthetic data matrix. Simulated noise was added to this matrix to represent Poisson and peak-fitting imprecisions (given by Eq. 13) by sampling from Poisson- and Gaussian-noise generators (Igor Pro, version 6.3) with an SD of 5 % of the peak height. For simplicity, peak-overlap errors were not included. PMF was performed as described in Corbin et al. (2015a). In particular, low signal-to-noise variables were downweighted (Paatero and Hopke, 2003) and the robust mode of PMF was used (Paatero, 1997), consistent with standard practice in the AMS community (Zhang et al., 2011).

When this synthetic matrix was factorized using only a Poisson imprecision term to weight the data (i.e. with the largest signals overweighted), the residual matrix showed significant outliers for the highest signals. For example, the largest residual outliers for CO^+, which was the highest-signal ion in the synthetic matrix, corresponded to spikes in the signal of that ion. These residual outliers were not present when the synthetic matrix was factorized using the correct imprecision model, Eq. (13).

The appearance of the highest CO^+ signals as spikes in the residual matrix when using the incorrect imprecision model is highly significant. Under normal circumstances, the analyst may have regarded such spikes as "outliers" reflecting transient signals or data-analysis problems, even though (in this case) they were purely the result of underestimated uncertainties. Properly weighting these spikes caused the r^2 between the output and input time series of the lowest-signal factor to increase from 0.54 to 0.74, with similar but smaller increases for the higher-signal factors, which were retrieved with $r^2 > 0.91$ in the wrongly weighted case and with $r^2 > 0.999$ in the correctly weighted case. Thus, the influence of high signals on the retrieval of low-signal PMF factors may thus directly influence the mass concentrations estimated by HR-AMS–PMF. This may significantly bias the mass concentrations of specific HR-AMS factors, particularly when they are retrieved in the presence of higher-signal factors. No exact number can be provided for this effect, as it will depend on the relative magnitude of the signals as well as the mass-spectral profiles of each of the retrieved components.
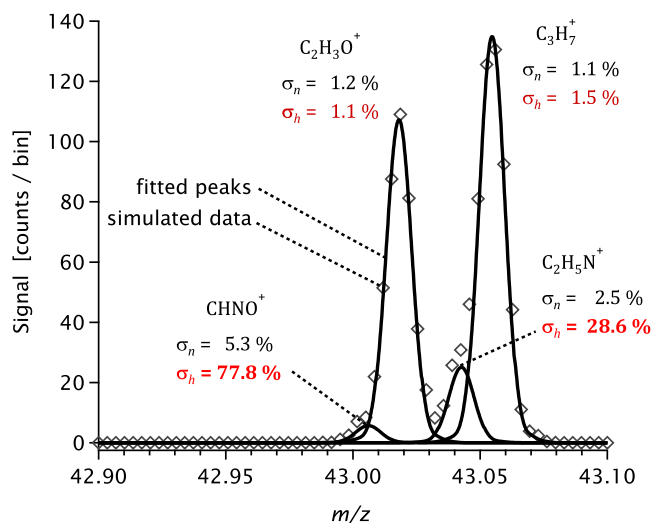

Figure 12. Overlapping-peak uncertainties estimated by the method outlined in this manuscript. Four model peaks are shown, defined with the same shape as observed in the test data set. Prior to fitting, counting noise was added to each data point and representative mass calibration errors (μ-prediction errors, see text) were added to each constrained peak location. The resulting imprecision in fitted peak height σ_h was estimated by Monte Carlo with effective imprecision 20 ppm, as labelled for each peak. The respective peak-integration imprecisions due to ion counting are labelled as σ_n. The ordinate units are counts bin^{-1}.

5 Monte Carlo estimation of σ_h for overlapping peaks with unknown μ-prediction errors

It was concluded in Sect. 4.2 that intentionally overestimating m/z-calibration imprecision by a modest amount allows the importance of m/z-calibration biases to be neglected in estimating σ_h for well-resolved peaks and that σ_h is then overestimated by only a few %.

This approach is also applicable, without modification, to the case of multiple overlapping peaks. Fig. 12 shows an example of this method, illustrating the peak-fitting and counting imprecisions (σ_h and σ_n) for four overlapping peaks. The four overlapping peaks were synthetically generated in a similar manner to the peaks described above (Sect. 4), except that counting imprecision was added to the counts in each bin by sampling from a Poisson-noise generator. Errors in μ-prediction (in this context, the sum of μ-prediction bias and a specific value of μ-prediction imprecision) were generated as $+5$, -7, $+10$, and -2 ppm for each of the four peaks (in order of increasing m/z); these are considered reasonable given the biases and imprecisions in Table 3. Errors in peak-width prediction (Sect. 3.1) were not included for the sake of generality.

The figure shows that, for the two larger signals, peak-fitting imprecisions were on the order of counting imprecision. For the two smaller peaks, peak-fitting imprecisions dominate, with the leftmost peak having a large enough

uncertainty to be considered unquantifiable. The influence of counting imprecision on σ_h has not been rigorously addressed here. Counting imprecision will only influence σ_h when the actual number of counted ions is small enough to deform the peak shape, in which case σ_n will dominate σ_h regardless of such deformation.

6   Discussion

6.1   Context and previous work

The largest uncertainty in this analysis was due to the estimation of the biases and imprecisions in μ prediction (which reflect errors in m/z calibration) directly from the data (Sect. 3.6). We therefore emphasize that the ion-to-ion variability in μ-prediction biases which strongly influence σ_h are not unique to our study.

Biases in m/z calibration, also referred to as mass accuracy, are a general feature of time-of-flight mass spectrometers in atmospheric science (DeCarlo et al., 2006; Cappellin et al., 2010; Junninen et al., 2010; Müller et al., 2011; Yatavelli et al., 2012) and otherwise (e.g. Vorm and Mann, 1994). This means that AMS-specific details such as particle vaporization prior to electron-impact ionization are not their sole cause. These biases may reflect a departure of the ion trajectories from their predicted paths (Guilhaus et al., 2000) but may also result when m/z calibration is performed on low-signal peaks (Cubison and Jimenez, 2015). When the former mechanism is the cause, numerical studies (Cubison and Jimenez, 2015) may underestimate m/z-calibration uncertainties and thus underestimate σ_h.

Cappellin et al. (2010) introduced additional fit coefficients to the standard m/z-calibration equation and successfully reduced the magnitude of calibration biases. However, this approach is not feasible for all data sets, as it requires multiple peaks to be reliably available for calibration. Moreover, the biases remaining after applying such a modified calibration function would still need to be known for accurate Monte Carlo estimation of σ_h; reducing m/z-calibration bias does not remove the need to account for it. Nevertheless, an improved calibration would reduce the overall σ_h.

6.2   Reducing σ_h

The above conclusions that the major causes of fitting errors in PIKA are errors in peak-location prediction raises the question of whether the fit procedure itself might be improved.

It would be preferable to allow a priori knowledge of calibration uncertainties to be incorporated into the fitting procedure, following Bayesian theory (Gelman et al., 2013). In such an approach, realistic uncertainty distributions could be directly applied to the constraint of μ and translated directly into uncertainty distributions of the resulting h. However, this does not solve the problem of m/z-calibration biases being

unknown for the case of overlapping peaks. The development of such an approach is beyond the scope of the present work, in part because the Igor Pro software in which PIKA has been developed over the past decade does not provide the necessary framework. We note that although the relaxation of the μ constraint in Sects. 3.5 and 3.6 allowed for uncertainties in μ to be accounted for, the implicit probability distribution assigned to these uncertainties was an unphysical, uniform distribution with unrealistic, discontinuous edges. For the typical case where multiple overlapping peaks are fitted, this approach is not considered robust.

In addition to an improved fitting procedure, an improved calibration procedure would be an obvious recommendation for reducing μ-prediction errors. An accurate calibration requires both a consistent calibrant signal and a consistent peak shape. These requirements were not met by any of the mass-spectral peaks: the only consistent signals were those of background ions from the gas phase, which were relatively few and displayed significantly different peak shapes than particulate signals, likely owing to their originating from generally different regions of the AMS ionization chamber. An improved calibration might make use of an internal standard, for example a polyfluorinated organic (DeCarlo et al., 2006), to be fitted by the same pseudo-Gaussian function used for data analysis (Eq. 3). However, at some point the limited temporal resolution of the ion detector (Hilmer and Bothner, 2011) and the variability of ion flight paths also limit the accuracy of the m/z calibration.

6.3   Computation time required for Monte Carlo estimation of σ_h

By definition, the Monte Carlo approach invariably requires more time than a direct approach. In large HR-AMS data sets, thousands of mass spectra may be recorded, each containing hundreds of peaks to be fitted. On modern personal computers, fitting may consequently take hours, so performing 100 Monte Carlo simulations during exploratory data analysis can be impractical. A compromise between rapid and reliable results is therefore desirable.

In their work, Cubison and Jimenez (2015) developed a parameterization designed to estimate σ_h based on the separation of two peaks, for the reasons noted above. This parameterization does not account for the constant relative imprecision expected for well-resolved peaks (Sect. 4.1). Additionally, the parameterization has been designed for the two-peak system.

To minimize the computation time required for the Monte Carlo approach, the following algorithmic approach could be used. First, the approximation $\sigma_h / h \approx$ constant could be applied for all well-resolved peaks. A percentage value of σ_h / h could be estimated using a subset of the data set; weighted least-squares fits will typically be insensitive to the exact value. For cases where two peaks are close enough to possibly overlap, the parameterization of Cubison and Jimenez

(2015) may be used to quantitatively define when their overlap is likely to influence σ_h (for example, when the parameterization indicates $\sigma_h/h > 5\%$ for any two peaks). When overlap is significant, a Monte Carlo estimation of σ_h can be initiated. Further, once a Monte Carlo estimate is obtained for a given set of peak heights, the result may be stored and applied to subsequent peaks of similar height ratios; where more restrictive definitions of similarity will be more accurate but slower. The validity of these approximations should be directly evaluated after initial data exploration by performing full Monte Carlo calculations for each group of peaks in the data set.

6.4 Other sources of uncertainty

While only peak-integration and ion-counting uncertainties were addressed in the discussion above, a number of other AMS-specific uncertainties can be identified.

Craven et al. (2012) discussed and addressed a number of HR-AMS uncertainties with regard to background signals and overlapping ions. Craven et al. (2012) also developed a data-smoothing approach to evaluate data quality for PMF, rejecting noisy variables. Such an approach may remain useful for the very-slowly changing systems studied by Craven et al. (2012), as it may account for imprecisions other than those caused by peak integration; however, it did not result in improved PMF uncertainty estimates but rather the selective rejection of low signal-to-noise variables. More generally, smoothing approaches cannot be applied to the rapidly changing signals observed during roadside studies, aircraft studies, or laboratory studies on rapidly changing sources, all of which may entail the measurement of the high PM loadings for which peak-integration uncertainties are most important.

Other AMS-specific uncertainties include the fundamental uncertainty involved in converting electronic signals at the detector to ion counts. This conversion is performed after estimating the signal intensity of a single ion, a process complicated by the signal-thresholding applied by the data acquisition software (version 4.0.9). We performed this single-ion measurement procedure on each of the measurement days described herein and obtained results varying by $\sim 20\%$. We emphasize that this 20% would correspond to a potential bias in the integrated areas of all ions, and is fundamentally different to the peak-integration imprecisions discussed extensively above. Moreover, as it is not clear whether this number represents instrumental or procedural variability (the standard procedure resulted in different ions being used for single-ion measurement on each day), it has not been included in Eq. (13). In addition, it is worth noting that the newest version of the HR-AMS includes significant hardware improvements which address this issue.

A second HR-AMS-specific scenario arises when the intensity of less-abundant isotope peaks are predicted based on the intensity of more-abundant isotope peaks. This procedure propagates fitting errors across integer m/z, and the corresponding uncertainty should also be propagated to these and neighbouring peaks, as discussed in Corbin et al. (2015a).

Finally, not all ions follow the peak shape established by the PIKA calibration procedure. In particular, thermally generated ions such as K^+ (Drewnick et al., 2006; Corbin et al., 2015a) or other slowly evaporating species (Salcedo et al., 2010; Craven et al., 2012) may lead to additional uncertainties for this reason.

7 Conclusions

Peak-integration uncertainties in the analysis of HR-AMS data by PIKA originate from uncertainties in peak-width prediction and in peak-height fitting. The former uncertainty may be easily estimated from the peak-width calibration procedure; the latter by an empirically constrained Monte Carlo approach.

Peak-fitting uncertainties depend most strongly on errors in the m/z calibration used to predict the peak location μ in ion-time-of-flight space. For well-resolved peaks, fitting uncertainties are sensitive to imprecisions in μ, to biases in μ, to the number of data points representing a given peak, and to the position of the peak centre relative to the nearest data point in m/z space. The first two of these factors appear to be most important, but all may be accounted for by empirically based Monte Carlo estimation.

Since peaks are fitted in PIKA by linearly scaling a predefined function, peak-fitting errors for well-resolved peaks also scale linearly with peak height. This leads to a constant-relative-imprecision term in the overall peak-integration uncertainty. Since a constant relative imprecision scales linearly with the ion count n as n^1, but counting uncertainties scale as $n^{0.5}$, this constant-relative imprecision term dominates counting uncertainties at high signal intensities. For example, in an exemplary data set with an estimated $\sim 4\%$ imprecision in fitted peak height, 2.5% imprecision in peak width, and 5% imprecision in integrated peak area, the relative imprecision term dominated counting uncertainties for well-resolved ions with areas of ~ 1000 cps. (Sample averaging times may influence this value.) Peak-integration uncertainties for well-resolved peaks will therefore be especially important for the high aerosol concentrations that may be measured near pollution sources or within highly polluted cities.

In a synthetic data set, including the constant-relative-imprecision term during PMF led to a significant improvement in the accuracy of the solution. Thus, although peak-integration uncertainties are much smaller than the uncertainties inherent in AMS-measured mass concentrations or elemental ratios, neglecting peak-integration uncertainties during PMF of data sets containing high-signal ions may significantly bias mass concentrations or elemental ratios of the retrieved factors. The dependence of the relative imprecision on the m/z calibration and mass-spectral resolution indicate

that it should be estimated for each data set or at least each new voltage configuration of a given mass spectrometer. The software used to perform this estimation in the present study, written in Igor Pro, is available upon request and will be incorporated into PIKA.

Finally, peak-fitting errors may also increase rapidly when peaks overlap significantly, potentially becoming much larger than the uncertainties of well-resolved peaks. It was shown that the fitting imprecision for overlapping peaks may be estimated directly by a minor modification to the Monte Carlo approach described above, that is, by intentionally overestimating the input imprecision such that biases may be neglected. This overestimate results in a moderate overestimate of the peak-fitting imprecision but avoids the ion-to-ion biases that would otherwise result from unquantifiable m/z-calibration biases.

Appendix A: Description of test data set

The data set used for evaluating and testing fitting uncertainties represents the mass spectra of fresh, aged, and filtered-and-aged aerosols emitted from a beech-wood combustion stove. Up to six batches of wood were burnt consecutively on 3 consecutive days in these experiments. A complete description of the experimental setup and instrument configuration is given in Corbin et al. (2015a).

The wood-combustion aerosols were vaporized and ionized in an Aerodyne HR-AMS equipped with a soot-particle (SP) vaporization module (Onasch et al., 2012). The data-acquisition card sampled once per ns. The majority of signals presented in the paper reflect SP-AMS-mode measurements, which were used because signals were consistently higher with the SP laser on and the analysis in this work applies mainly to higher signals. The peak-width calibration reported above (2.5 %) applies to both laser on and laser off data. SP signals were higher because (i) the majority of PM emitted by the stove was refractory BC (Corbin et al., 2015b) and (ii) the AMS is more sensitive to organic coatings when they are vaporized by the SP laser due to its physical position (Willis et al., 2014). As shown in Fig. 3e, the laser state had no effect on the peak-fitting results discussed herein.

Appendix B: Analysis of test data set

All data were analysed in Igor Pro (Version 6.2, Wavemetrics, OR, USA) using a modified version of PIKA, derived from PIKA 1.10H, and custom code. The modifications to PIKA consisted of improvements to the peak-width calibration procedure, the selective introduction of errors to the analysis, algorithmic improvements to the peak-fitting procedure, and the implementation of peak-integration uncertainties as part of the overall PIKA error calculation. The first modification improved the robustness of the peak-width calibration procedure by replacing the mean peak width with a trimmed mean, followed by a weighted fit to the data, as detailed in Corbin et al. (2015a).

In this paper, "diff" HR-AMS data have been presented. Diff data represent the difference of "open" measurements (comprising signals from particulate, gaseous, and background species) and "closed" measurements (comprising background species from gases and slowly evaporating material). The same trends seen in the diff data were seen in analogous plots for the open and closed data; however, the noise regime of the RMSE was much less noticeable in these data. Diff data were used to allow the two regimes of the RMSE to be clearly highlighted and to remove inconsequential differences due to different background levels, for example of CO_2^+ due to gaseous carbon dioxide.

The following peaks were used for m/z calibration: CH^+, Ar^+, CO_2^+, $^{182}W^+$, $^{184}W^+$, and $^{186}W^+$. While this list is not ideal, no other peaks were consistently present with sufficient signal for use as calibrant ions. Some peaks in the test data were well resolved but excluded for having anomalous peak shapes. These peaks included gas-phase ions present at high signals (e.g. N_2^+ or O_2^+), ions which were present at very low m/z and were therefore represented by only two detector bins (C^+ and CH^+) such that reliable fits could not be performed, and ions which were known to follow anomalous vaporization–ionization physics (e.g. K^+, discussed further in Corbin et al., 2015a).

Edited by: J. Schneider

References

Aiken, A. C., DeCarlo, P. F., and Jimenez, J. L.: Elemental analysis of organic species with electron ionization high-resolution mass spectrometry, Anal. Chem., 79, 8350–8358, doi:10.1021/ac071150w, 2007.

Aiken, A. C., DeCarlo, P. F., Kroll, J. H., Worsnop, D. R., Huffman, J. A., Docherty, K. S., Ulbrich, I. M., Mohr, C., Kimmel, J. R., and Sueper, D.: O/C and OM/OC ratios of primary, secondary, and ambient organic aerosols with high-resolution time-of-flight aerosol mass spectrometry, Environ. Sci. Technol., 42, 4478–4485, doi:10.1021/es703009q, 2008.

Allan, D. W.: Statistics of atomic frequency standards, Proc. IEEE, 54, 221–230, doi:10.1109/PROC.1966.4634, 1966.

Allan, J. D., Jimenez, J. L., Williams, P. I., Alfarra, M. R., Bower, K. N., Jayne, J. T., Coe, H., and Worsnop, D. R.: Quantitative sampling using an aerodyne aerosol mass spectrometer 1. Techniques of data interpretation and error analysis, J. Geophys. Res., 108, 4090, doi:10.1029/2002JD002358, 2003.

Allan, J. D., Delia, A. E., Coe, H., Bower, K. N., Alfarra, M. R., Jimenez, J. L., Middlebrook, A. M., Drewnick, F., Onasch, T. B., and Canagaratna, M. R.: A generalised method for the extraction of chemically resolved mass spectra from aerodyne aerosol mass spectrometer data, J. Aerosol Sci., 35, 909–922, doi:10.1016/j.jaerosci.2004.02.007, 2004.

Bahreini, R., Ervens, B., Middlebrook, A. M., Warneke, C., de Gouw, J. A., DeCarlo, P. F., Jimenez, J. L., Brock, C. A., Neuman, J. A., Ryerson, T. B., Stark, H., Atlas, E., Brioude, J., Fried, A., Holloway, J. S., Peischl, J., Richter, D., Walega, J., Weibring, P., Wollny, A. G., and Fehsenfeld, F. C.: Organic aerosol formation in urban and industrial plumes near Houston and Dallas, Texas, J. Geophys. Res., 114, D00F16, doi:10.1029/2008jd011493, 2009.

Canagaratna, M. R., Jayne, J., Jimenez, J., Allan, J., Alfarra, M., Zhang, Q., Onasch, T. B., Drewnick, F., Coe, H., Middlebrook, A., Delia, A., Williams, L., Trimborn, A., Northway, M., DeCarlo, P., Kolb, C., Davidovits, P., and Worsnop, D.: Chemical and microphysical characterization of ambient aerosols with the aerodyne aerosol mass spectrometer, Mass Spectrom. Rev., 26, 185–222, doi:10.1002/mas.20115, 2007.

Canagaratna, M. R., Jimenez, J. L., Kroll, J. H., Chen, Q., Kessler, S. H., Massoli, P., Hildebrandt Ruiz, L., Fortner, E., Williams, L. R., Wilson, K. R., Surratt, J. D., Donahue, N. M., Jayne, J. T., and Worsnop, D. R.: Elemental ratio measurements of organic compounds using aerosol mass spectrometry: characterization, improved calibration, and implications, Atmos. Chem. Phys., 15, 253–272, doi:10.5194/acp-15-253-2015, 2015.

Cappa, C. D., Onasch, T. B., Massoli, P., Worsnop, D. R., Bates, T. S., Cross, E. S., Davidovits, P., Hakala, J., Hayden, K. L., and Jobson, B. T.: Radiative absorption enhancements due to the mixing state of atmospheric black carbon, Science, 337, 1078–1081, doi:10.1126/science.1223447, 2012.

Cappellin, L., Biasioli, F., Fabris, A., Schuhfried, E., Soukoulis, C., Märk, T. D., and Gasperi, F.: Improved mass accuracy in PTR-TOF-MS: Another step towards better compound identification in PTR-MS, International J. Mass Spectrom., 290, 60–63, doi:10.1016/j.ijms.2009.11.007, 2010.

Chen, Q., Liu, Y., Donahue, N. M., Shilling, J. E., and Martin, S. T.: Particle-phase chemistry of secondary organic material: modeled compared to measured O:C and H:C elemental ratios provide constraints, Environ. Sci. Technol., 45, 4763–4770, doi:10.1021/es104398s, 2011.

Corbin, J. C., Sierau, B., Gysel, M., Laborde, M., Keller, A., Kim, J., Petzold, A., Onasch, T. B., Lohmann, U., and Mensah, A. A.: Mass spectrometry of refractory black carbon particles from six sources: carbon-cluster and oxygenated ions, Atmos. Chem. Phys., 14, 2591–2603, doi:10.5194/acp-14-2591-2014, 2014.

Corbin, J. C., Lohmann, U., Sierau, B., Keller, A., Burtscher, H., and Mensah, A. A.: Black carbon surface oxidation and organic composition of beech-wood soot aerosols, Atmos. Chem. Phys., 15, 11885–11907, doi:10.5194/acp-15-11885-2015, 2015a.

Corbin, J. C., Keller, A., Sierau, B., Lohmann, U., and Mensah, A. A.: Wood-stove and pellet-burner organic aerosol emissions aged in a continuous-flow photoreactor, Aerosol Sci. Tech., 49, 1037–1050, doi:10.1080/02786826.2015.1079586, 2015b.

Craven, J. S., Yee, L. D., Ng, N. L., Canagaratna, M. R., Loza, C. L., Schilling, K. A., Yatavelli, R. L. N., Thornton, J. A., Ziemann, P. J., Flagan, R. C., and Seinfeld, J. H.: Analysis of secondary organic aerosol formation and aging using positive matrix factorization of high-resolution aerosol mass spectra: application to the dodecane low-NO$_x$ system, Atmos. Chem. Phys., 12, 11795–11817, doi:10.5194/acp-12-11795-2012, 2012.

Cubison, M. J. and Jimenez, J. L.: Statistical precision of the intensities retrieved from constrained fitting of overlapping peaks in high-resolution mass spectra, Atmos. Meas. Tech., 8, 2333–2345, doi:10.5194/amt-8-2333-2015, 2015.

Cubison, M. J., Ortega, A. M., Hayes, P. L., Farmer, D. K., Day, D., Lechner, M. J., Brune, W. H., Apel, E., Diskin, G. S., Fisher, J. A., Fuelberg, H. E., Hecobian, A., Knapp, D. J., Mikoviny, T., Riemer, D., Sachse, G. W., Sessions, W., Weber, R. J., Weinheimer, A. J., Wisthaler, A., and Jimenez, J. L.: Effects of aging on organic aerosol from open biomass burning smoke in aircraft and laboratory studies, Atmos. Chem. Phys., 11, 12049–12064, doi:10.5194/acp-11-12049-2011, 2011.

Dallmann, T. R., Onasch, T. B., Kirchstetter, T. W., Worton, D. R., Fortner, E. C., Herndon, S. C., Wood, E. C., Franklin, J. P., Worsnop, D. R., Goldstein, A. H., and Harley, R. A.: Characterization of particulate matter emissions from on-road gasoline and diesel vehicles using a soot particle aerosol mass spectrometer, Atmos. Chem. Phys., 14, 7585–7599, doi:10.5194/acp-14-7585-2014, 2014.

DeCarlo, P. F., Kimmel, J. R., Trimborn, A., Northway, M. J., Jayne, J. T., Aiken, A. C., Gonin, M., Fuhrer, K., Horvath, T., and Docherty, K. S.: Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer, Anal. Chem., 78, 8281–8289, doi:10.1021/ac061249n, 2006.

Drewnick, F., Hings, S. S., Curtius, J., Eerdekens, G., and Williams, J.: Measurement of fine particulate and gas-phase species during the New Year's fireworks 2005

in Mainz, Germany, Atmos. Environ., 40, 4316–4327, doi:10.1016/j.atmosenv.2006.03.040, 2006.

Elsasser, M., Busch, C., Orasche, J., Schön, C., Hartmann, H., Schnelle-Kreis, J., and Zimmermann, R.: Dynamic changes of the aerosol composition and concentration during different burning phases of wood combustion, Energy Fuels, 27, 4959–4968, doi:10.1021/ef400684f, 2013.

Freutel, F., Drewnick, F., Schneider, J., Klimach, T., and Borrmann, S.: Quantitative single-particle analysis with the Aerodyne aerosol mass spectrometer: development of a new classification algorithm and its application to field data, Atmos. Meas. Tech., 6, 3131–3145, doi:10.5194/amt-6-3131-2013, 2013.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B.: Bayesian Data Analysis, CRC Press, Florida, USA, 2013.

Guilhaus, M., Selby, D., and Mlynski, V.: Orthogonal acceleration time-of-flight mass spectrometry, Mass Spectrom. Rev., 19, 65–107, 2000.

Hilmer, J. K. and Bothner, B.: Physical signal modulation of time-of-flight mass analyzers increases precision and decreases noise, Rapid Commun. Mass Sp., 25, 795–805, doi:10.1002/rcm.4923, 2011.

Hings, S. S.: Characterization and Field Deployment of a Novel Quantitative Time-of-Flight Aerosol Mass Spectrometer (ToF-AMS), PhD thesis, Johannes Gutenberg-Universität Mainz, Mainz, 2006.

Jayne, J. T., Leard, D. C., Zhang, X., Davidovits, P., Smith, K. A., Kolb, C. E., and Worsnop, D. R.: Development of an aerosol mass spectrometer for size and composition analysis of submicron particles, Aerosol Sci. Tech., 33, 49–70, doi:10.1080/027868200410840, 2000.

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prévôt, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., E, Dunlea, J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U., and Worsnop, D. R.: Evolution of organic aerosols in the atmosphere, Science, 326, 1525–1529, doi:10.1126/science.1180353, 2009.

Junninen, H., Ehn, M., Petäjä, T., Luosujärvi, L., Kotiaho, T., Kostiainen, R., Rohner, U., Gonin, M., Fuhrer, K., Kulmala, M., and Worsnop, D. R.: A high-resolution mass spectrometer to measure atmospheric ion composition, Atmos. Meas. Tech., 3, 1039–1053, doi:10.5194/amt-3-1039-2010, 2010.

Kimmel, J. R., Farmer, D. K., Cubison, M. J., Sueper, D., Tanner, C., Nemitz, E., Worsnop, D. R., Gonin, M., and Jimenez, J. L.: Real-time aerosol mass spectrometry with millisecond resolution, Int. J. Mass Spectrom., 303, 15–26, doi:10.1016/j.ijms.2010.12.004, 2011.

Lee, A. K. Y., Willis, M. D., Healy, R. M., Onasch, T. B., and Abbatt, J. P. D.: Mixing state of carbonaceous aerosol in an urban environment: single particle characterization using the soot particle aerosol mass spectrometer (SP-AMS), Atmos. Chem. Phys., 15, 1823–1841, doi:10.5194/acp-15-1823-2015, 2015.

Lee, T., Sullivan, A. P., Mack, L., Jimenez, J. L., Kreidenweis, S. M., Onasch, T. B., Worsnop, D. R., Malm, W., Wold, C. E., and Hao, W. M.: Chemical smoke marker emissions during flaming and smoldering phases of laboratory open burning of wildland fuels, Aerosol Sci. Tech., 44, i–v, doi:10.1080/02786826.2010.499884, 2010.

Müller, M., George, C., and D'Anna, B.: Enhanced spectral analysis of C-TOF aerosol mass spectrometer data: iterative residual analysis and cumulative peak fitting, Int. J. Mass Spectrom., 306, 1–8, doi:10.1016/j.ijms.2011.04.007, 2011.

Ng, N. L., Herndon, S. C., Trimborn, A., Canagaratna, M. R., Croteau, P., Onasch, T. B., Sueper, D., Worsnop, D. R., Zhang, Q., and Sun, Y.: An Aerosol Chemical Speciation Monitor (ACSM) for routine monitoring of the composition and mass concentrations of ambient aerosol, Aerosol Sci. Tech., 45, 780–794, doi:10.1080/02786826.2011.560211, 2011.

Onasch, T. B., Trimborn, A., Fortner, E. C., Jayne, J. T., Kok, G. L., Williams, L. R., Davidovits, P., and Worsnop, D. R.: Soot particle aerosol mass spectrometer: development, validation, and initial application, Aerosol Sci. Tech., 46, 804–817, doi:10.1080/02786826.2012.663948, 2012.

Ortega, A. M., Day, D. A., Cubison, M. J., Brune, W. H., Bon, D., de Gouw, J. A., and Jimenez, J. L.: Secondary organic aerosol formation and primary organic aerosol oxidation from biomass-burning smoke in a flow reactor during FLAME-3, Atmos. Chem. Phys., 13, 11551–11571, doi:10.5194/acp-13-11551-2013, 2013.

Paatero, P.: Least squares formulation of robust non-negative factor analysis, Chemometr. Intell. Lab., 37, 23–35, doi:10.1016/S0169-7439(96)00044-5, 1997.

Paatero, P. and Hopke, P. K.: Discarding or downweighting high-noise variables in factor analytic models, Anal. Chim. Acta, 490, 277–289, doi:10.1016/S0003-2670(02)01643-4, 2003.

Paatero, P. and Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, Environmetrics, 5, 111–126, doi:10.1002/env.3170050203, 1994.

Salcedo, D., Onasch, T. B., Aiken, A. C., Williams, L. R., de Foy, B., Cubison, M. J., Worsnop, D. R., Molina, L. T., and Jimenez, J. L.: Determination of particulate lead using aerosol mass spectrometry: MILAGRO/MCMA-2006 observations, Atmos. Chem. Phys., 10, 5371–5389, doi:10.5194/acp-10-5371-2010, 2010.

Sueper, D., Jimenez, J. L., Aiken, A., and DeCarlo, P.: PIKA ToF-AMS High Resolution Analysis Software, available at: http://cires1.colorado.edu/jimenez-group/wiki/index.php/ToF-AMS_Analysis_Software (last access: 2015-10-29), 2011.

Sun, Y. L., Zhang, Q., Anastasio, C., and Sun, J.: Insights into secondary organic aerosol formed via aqueous-phase reactions of phenolic compounds based on high resolution mass spectrometry, Atmos. Chem. Phys., 10, 4809–4822, doi:10.5194/acp-10-4809-2010, 2010.

Taylor, J. R.: An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements, University Science Books, Sausalito, California, USA, 1997.

Timko, M. T., Albo, S. E., Onasch, T. B., Fortner, E. C., Yu, Z., Miake-Lye, R. C., Canagaratna, M. R., Ng, N. L., and Worsnop, D. R.: Composition and sources of the organic particle emissions from aircraft engines, Aerosol Sci. Tech., 48, 61–73, doi:10.1080/02786826.2013.857758, 2014.

Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., and Jimenez, J. L.: Interpretation of organic components from Positive Matrix Factorization of aerosol mass spectrometric data, Atmos. Chem. Phys., 9, 2891–2918, doi:10.5194/acp-9-2891-2009, 2009.

Vorm, O. and Mann, M.: Improved mass accuracy in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of peptides, J. Am. Soc. Mass Spectrom., 5, 955–958, doi:10.1016/1044-0305(94)80013-8, 1994.

Werle, P., Mücke, R., and Slemr, F.: The limits of signal averaging in atmospheric trace-gas monitoring by tunable diode-laser absorption spectroscopy (TDLAS), Appl. Phys. B-Lasers O., 57, 131–139, 1993.

Wilks, D.: Statistical Methods in the Atmospheric Sciences, International Geophysics, Elsevier Science, Oxford, UK, 2011.

Willis, M. D., Lee, A. K. Y., Onasch, T. B., Fortner, E. C., Williams, L. R., Lambe, A. T., Worsnop, D. R., and Abbatt, J. P. D.: Collection efficiency of the soot-particle aerosol mass spectrometer (SP-AMS) for internally mixed particulate black carbon, Atmos. Meas. Tech., 7, 4507–4516, doi:10.5194/amt-7-4507-2014, 2014.

Yatavelli, R. L., Lopez-Hilfiker, F., Wargo, J. D., Kimmel, J. R., Cubison, M. J., Bertram, T. H., Jimenez, J. L., Gonin, M., Worsnop, D. R., and Thornton, J. A.: A chemical ionization high-resolution time-of-flight mass spectrometer coupled to a micro orifice volatilization impactor (MOVI-HRToF-CIMS) for analysis of gas and particle-phase organic species, Aerosol Sci. Technol., 46, 1313–1327, doi:10.1080/02786826.2012.712236, 2012.

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Allan, J. D., Coe, H., Ulbrich, I., Alfarra, M. R., Takami, A., Middlebrook, A. M., Sun, Y. L., Dzepina, K., Dunlea, E., Docherty, K., DeCarlo, P. F., Salcedo, D., Onasch, T. B., Jayne, J. T., Miyoshi, T., Shimono, A., Hatakeyama, S., Takegawa, N., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Williams, P., Bower, K., Bahreini, R., Cottrell, L., Griffin, R. J., Rautiainen, J., Sun, J. Y., Zhang, Y. M., and Worsnop, D. R.: Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes, Geophys. Res. Lett., 34, L13801, doi:10.1029/2007GL029979, 2007.

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Ulbrich, I. M., Ng, N. L., Worsnop, D. R., and Sun, Y.: Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: a review, Anal. Bioanal. Chem., 401, 3045–3067, doi:10.1007/s00216-011-5355-y, 2011.