analytical chemistry

pubs.acs.org/ac Article

Improved Mass Accuracy and Isotope Confirmation through Alignment of Ultrahigh-Resolution Mass Spectra of Complex Natural Mixtures

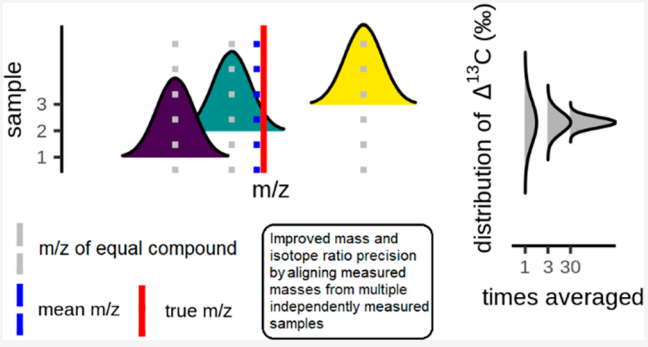
Julian Merder, Jan A. Freund, Ulrike Feudel, Jutta Niggemann, Gabriel Singer, and Thorsten Dittmar*

Cite This: *Anal. Chem.* 2020, 92, 2558−2565

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) is one of the state-of-the-art methods to analyze complex natural organic mixtures. The precision of detected masses is crucial for molecular formula attribution. Random errors can be reduced by averaging multiple measurements of the same mass, but because of limited availability of ultrahigh-resolution mass spectrometers, most studies cannot afford analyzing each sample multiple times. Here we show that random errors can be eliminated also by averaging mass spectral data from independent environmental samples. By averaging the spectra of 30 samples analyzed on our 15 T instrument we reach a mass precision comparable to a single spectrum of a 21 T instrument. We also show that it is possible to accurately and reproducibly determine isotope ratios with FT-ICR-MS. Intensity ratios of isotopologues were improved to a degree that measured deviations were within the range of natural isotope fractionation effects. In analogy to $\delta^{13}C$ in environmental studies, we propose $\Delta^{13}C$ as an analytical measure for isotope ratio deviances instead of widely employed C deviances. In conclusion, here we present a simple tool, extensible to Orbitrap-based mass spectrometers, for postdetection data processing that significantly improves mass accuracy and the precision of intensity ratios of isotopologues at no extra cost.



Natural organic matter in aquatic, sedimentary, and soil environments are highly complex mixtures, largely of low molecular weight substances (<1000 Da).[1−3] Marine dissolved organic matter (DOM), for example, consists of at least hundreds of thousands if not millions of different substances.[4] All these substances are interconnected within the global carbon cycle, due to past and recent metabolic or physicochemical processes, making their identification an important issue in the context of climate change.[5−7]

Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) has established itself as one of the state-of-the-art methods to analyze such complex mixtures.[8−10] The ultrahigh precision of detected masses is a crucial requirement for subsequent molecular formula attribution. Even a precision of <0.1 ppm cannot lead to unique molecular formula matches[8] as with increasing mass the number of possible matches is inflated via combinatorial explosion. On top of that, the resolution decreases with increasing mass to charge ratio.[11] This challenge has inspired ideas how to improve the accuracy of the molecular formula attribution e.g. using isotope patterns or ratios[8,9,12] or investigating homologous series.[8,13]

The measurement of a mass in FT-ICR-MS is always accompanied by two potential sources of errors that have a direct influence on the accuracy. The first is a systematic error, a shift from the true value to lower or higher values within measurements. A systematic error emerges due to the aggregation of multiple biases such as imprecise calibration of the instrument, directed electronic misadjustments, wrong information transfer or many other factors influencing the measurement. The systematic error is either positive or negative but always directed. The second is a random error, which is a fluctuation of the measured mass from the true mass by chance only. The total random error of a mass measurement can be a combined error of multiple undirected effects that are all of purely probabilistic nature, e.g. unpredictable fluctuations of the environment influencing the measurement. The random error can be reduced by arithmetic averaging multiple measurements of the same mass, since positive and negative deviations from the true values (for vanishing systematic error) within measurements cancel out. In fact, the random error is reduced by the factor of $1/\sqrt{n}$, where n represents the number of averaged measurements.[14] Usually,

Received: September 17, 2019
Accepted: December 30, 2019
Published: December 30, 2019

https://dx.doi.org/10.1021/acs.analchem.9b04234
Anal. Chem. 2020, 92, 2558−2565

multiple mass spectrometry scans are accumulated for the analysis of a single sample to take advantage of this averaging effect and to enhance sensitivity. But because of limited availability of ultrahigh-resolution mass spectrometers, most studies cannot afford analyzing each sample multiple times to further reduce the mass error through averaging after the actual measurement. Also, many environmental studies involve dozens or even hundreds of individual samples, increasing the pressure for measuring efficiency. In such sample sets, however, most molecular formulas are detected multiple times in even only partly related samples, because many compounds are globally ubiquitous even in contrasting environments.[15] Thus, in principle, one can consider the multiple detection of the same mass across different samples as replicate measurements and take advantage of the error reduction through averaging.

The same potentially holds true not only for masses but also for their signal intensities. Signal intensity, however, is sample specific, because identical compounds occur in varying concentrations in the environments.[15] Averaging signal intensities of detected masses across samples is therefore not useful, but the intensity ratio of isotopologues is determined by the natural abundance of isotopes and should as such be constant across samples in a first approximation. In analogy to the potential reduction of mass error through averaging also the natural abundance ratio of isotopologues should be approached more precisely through averaging of signal intensity ratios.

To make use of averaged masses and intensity ratios from multiple spectra, they have to be aligned first. This spectral alignment is a crucial step, because aligned mass values will slightly deviate from spectrum to spectrum due to the above-mentioned sources of errors. In this study we present a method of spectral-alignment and show that averaging masses over true or technical replicates gives rise to an additional reduction in mass error. For quantifying the mass error, we use the common metrics[16] mean absolute error (MAE) and root mean squared error (RMSE). Such a reduction of mass error has direct effects for molecular formula attribution because it allows reduction of the tolerance window that is used to find potential formula matches. Moreover, we show that for DOM spectra also isotope ratio deviances benefit from being averaged, making them an improved indicator of the correct molecular formula match as well.

EXPERIMENTAL SECTION

Data Acquisition. We used two different data sets in this study. One consisted of replicate measurements (n = 65) of the same sample collected from North Equatorial Pacific Intermediate Water near Hawaii. The DOM sample was obtained by solid-phase extracting >10,000 L of deep-seawater. Details on sample collection and processing were described in previously published work for this data set.[17,18] It represents the case example of measuring the same sample multiple times with values being considered as technical replicates reflecting only measurement errors. In the following we name this data "Deep Sea".

The second data set consisted of mass spectra of 137 DOM samples from the German Bight (North Sea) measured at different points in time. These samples span over a wide range of different proportions of fresh and marine waters, spring and summer plankton blooms, and winter conditions. DOM of the samples was obtained by solid-phase extracting the water samples according to the established protocol for such kind of samples.[19] In the following we name this data set "North Sea".

All DOM samples were measured on a Bruker Solarix 15 T FT-ICR-MS at the University of Oldenburg (Germany) in 50:50 methanol/water (v/v) at a DOM concentration of 20 mg carbon L^{-1}. Samples were directly infused into the electrospray ionization (ESI) unit and ionized in negative mode. Five hundred transients were summed per sample. The spectra were mass calibrated (linearly) using the Bruker Daltonics Data Analysis software package with an internal calibration list consisting of about 50 compounds present in the samples and covering the entire relevant mass range. The mass error for the calibration points was <0.1 ppm after calibration. For more details on instrument settings we refer to Riedel and Dittmar.[18]

Basic Data Processing. To eliminate noise from FT-ICR-MS spectra we applied the method detection limit (MDL).[18] Here, we used the 99.9% confidence limits of intensities of "noise" peaks as the upper limit. All ions with at least two isotopologues were singly charged, as evident from isotopologue mass spacing. Thus, we assume that all detected ions were singly charged and use the term mass synonymous for mass to charge ratio.

Peak Alignment. To average masses from different spectra, the masses that belong to the same molecular formula, but are found in different spectra, had to be identified and aligned correctly. In the following we explain in short the steps that lead to a cross tabulation of aligned masses.

Step 1: Across all spectra we identified the mass with the largest intensity (m_{ref}).

Step 2: We combined this peak with at most one peak per spectrum that is at minimal distance and within a suitable tolerance of 0.5 ppm. The distance between a mass, present in another spectrum than m_{ref} was calculated as

$$\frac{m - m_{ref}}{m_{ref}} \times 1e6 \text{ ppm}$$

Given a set of n spectra, the collection eventually comprised a mass cluster of at most n peaks (which was reached only when the mass was found in all spectra).

Step 3: All members of the cluster were removed from the raw data set and the corresponding cluster was transferred to a row of a cross tabulation with columns collecting cluster properties of interest, in particular, its mass average and masses and intensities of all cluster members.

Step 4: Returning to Step 1, we repeated all steps until all masses were processed.

The validation of a successful mass alignment was accomplished by comparing the mass differences and standard errors within mass clusters with those between mass clusters. The between mass cluster distances should be much larger and clearly separable from the distribution of mass differences within mass clusters which was always the case for our data sets.

Molecular Formula Attribution. For every mass cluster its average mass was used to attribute all possible molecular formulas within a tolerance range of 0.5 ppm. We used a master list, containing all chemically valid combinations of elements up to 1000 Da considering the elements (ranges) C (1−100),^{13}C (0−1), O (0−100), H (1−200), N (0−3), S (0−

https://dx.doi.org/10.1021/acs.analchem.9b04234
Anal. Chem. 2020, 92, 2558−2565

1), and P (0−1). We considered it chemically valid to encompass element combinations that summed up to even valence and lay in the range of certain elements (H ≤ 4C, H > 0.25C, O ≤ 2C, N ≤ C + 1, S ≤ C + 1, P ≤ C + 1). We excluded all formulas containing a ^{13}C that had no ^{12}C isotopologue.

Molecular Formula Confirmation. For each data set (Deep Sea and North Sea) we constructed three subsets of attributed molecular formulas that are inclusively ordered (All matches ⊇ Likeliest match ⊇ Isotope verified match).

These subsets were defined as follows:

Subset 1: "All matches"

This subset included all molecular formula attributions even multiple matches for a single mass.

Subset 2: "Likeliest match"

This subset was already more restricted than subset 1. It only included the likeliest molecular formula for a single mass obtained by applying the following sequence of filters. If we had found only one formula attribution for a mass this became automatically the likeliest match. For masses with more than one formula attribution we applied three criteria in a hierarchical order:

(1) Isotope verification: Here, we used a maximum $\Delta^{13}C$ (see eq 4) tolerance of ±1000 permille to flag a molecular formula as isotope verified. If only one molecular formula suggestion for the mass was isotope verified, we already had found the likeliest one. If multiple molecular formulas were isotope verified, we proceeded to the next filter with all isotope verified formula suggestions. For a mass without any isotope verified formula, we simply proceeded to the next filter retaining all molecular formulas.

(2) Maximum length of homologous series network: If we still had multiple formula matches for a mass, we counted the homologous series (CH_2) length that each formula suggestion of that mass exhibited in the complete data set and took the molecular formula with the maximum homologous series length as the likeliest match.[8]

(3) Minimal distance in ppm to the reference mass: If there was still more than one formula suggestion for a mass, we took the molecular formula with the closest distance to its reference mass as the likeliest match.

Subset 3: "Isotope verified match"

This subset was the most restrictive one. It included only the likeliest molecular formulas from subset 2 (likeliest match) that were isotope verified.

Quantification of Mass Error. For each data subset (index "s": all matches, likeliest matches, isotope verified), we calculated the well-known Mean Absolute Mass Error[16] (MAE) of the averaged masses (eq 1).

$$\mathrm{MAE}_s = \frac{\sum_{i=1}^{N_s}\left(\left|\frac{\bar{m}_i^s - M_i^s}{M_i^s}\right| \times 1e6\right)}{N_s}\ \mathrm{ppm} \tag{1}$$

with $\bar{m}_i^s$ = averaged mass i, N_s = number of masses in subset s, and M_i^s = reference mass of attributed molecular formula.

We preferred using MAE over the root mean squared error at this point, because it is more robust to outliers.[16]

To demonstrate how averaging spectra improves the mass precision, we computed the MAE for ubiquitous masses for different number of spectra, (n = 1, 3, 30). Three is a common number of technical replicates, while 30 may correspond to a typical number of samples taken along environmental gradients (e.g., a salinity gradient in an estuary). To estimate the distribution of the MAE, we used a bootstrapping approach (30-fold repeated sampling with replacement from the pools of spectra—Deep Sea: 65, North Sea: 137). Additionally, for both data sets and each of their respective subsets, we applied a Kruskal−Wallis test[20] to test for statistically significant differences across the selected values of "n".

The progressive error reduction by increasing the number of aligned spectra could also be followed through the root mean squared error (RMSE)[16] (defined in analogy to the MAE). This enabled us to calculate random and left-over systematic error (see below). To show this, we excluded all masses with an SNR below 20 and used only the remaining ubiquitous masses from the isotope verified subsets. We did this because we expected the mass error to have no SNR dependence anymore and the highest accuracy in formula attribution with the isotope verified subset, so that there were no extreme values corrupting the (nonrobust) RMSE and our model fit. For each remaining molecular formula, we averaged masses over "n" randomly drawn spectra. We swept "n" over the range from one to 65 for the Deep Sea data or 137 for the North Sea data, respectively, and computed the "n" specific RMSE_n:

The functional relation between RMSE and "n" could be described by the nonlinear parametric model for error partitioning[21] (eq 2).

$$\mathrm{RMSE}_n = \sqrt{\mathrm{err}_{\mathrm{sys}}{}^2 + \mathrm{err}_{\mathrm{rnd}}{}^2/n} \tag{2}$$

So that we could fit random "$\mathrm{err}_{\mathrm{rnd}}$" and systematic "$\mathrm{err}_{\mathrm{sys}}$" errors to the obtained RMSE_n values.

Savory et al.[21] have shown that there is a systematic mass error along the mass axis within spectra. This systematic error is not linearly increasing with mass but can be highly nonlinear with a clear autocorrelation structure along the mass to charge ratio axis. We employed a general additive model[22] (gam) to fit a nonparametric function of mass error (as response variable) versus mass (as explanatory variable). Using this result, we recalibrated the measured masses to reduce the mass error for each spectrum independently. With recalibrated spectra, we repeated all of the above analyses and compared them to related nonrecalibrated data.

Isotope Ratio Deviances. Isotope ratio deviances were calculated between molecular formulas differing only in one stable isotope (isotopologues). This was done to determine the number of carbon atoms in a compound, independently from the molecular mass, in order to have an independent confirmation of a molecular formula assignment. In the following such pairs of molecular formulas are denoted as "isotope pairs". In this study we only considered $^{13}C/^{12}C$ isotope pairs. Previous approaches in geochemistry defined a C deviance[8,23,24] as the isotope pair intensity ratio divided by the relative abundance of ^{13}C minus the number of carbons in the molecule to calculate an isotope ratio deviance (eq 3).

For this study, we calculated the isotope ratio deviances ($\Delta^{13}C$) in analogy to $\delta^{13}C$ where deviations from the average natural isotope abundance were related to the Vienna Pee Dee Belemnite[25,26] (VPDB) standard. This quantity is frequently used as a sample trait in biological, geochemical, or

https://dx.doi.org/10.1021/acs.analchem.9b04234
Anal. Chem. 2020, 92, 2558−2565

paleoclimatic studies.[27] Here, we used it on a molecule-specific basis for every isotope pair based on binomial probabilities. This approach is not new in mass spectrometry[28] but not yet well established in geochemistry where C deviance is more frequently applied.[8,23,24]

Underlying the binomial distribution is the idea that the molecule is assembled via sequentially drawing its "n" carbon atoms from an urn (a carbon pool). The urn is filled with a vast number of ^{13}C and ^{12}C isotopes with relative abundances given by p and $(1 - p)$, respectively. In this approach VPDB = $p/(1 - p)$ which is equivalent to p = VPDB/(1 + VPDB). When the relative abundance of ^{12}C and ^{13}C in the urn is known, we can calculate the probability to draw one ^{13}C and "$n - 1$" ^{12}C and the probability to draw "n" ^{12}C isotopes. We expect that the ratio of these two probabilities matches the intensity ratio for the isotope pair. This suggests measuring deviations from the expectation by the expression for $\Delta^{13}C$ shown in eq 4.

$$C\ \text{deviance} = \frac{\left(\frac{\text{Intensity}\ ^{13}C}{\text{Intensity}\ ^{12}C}\right)}{p} - n \tag{3}$$

$$\Delta^{13}C = \left(\frac{\left(\frac{\text{Intensity}\ ^{13}C}{\text{Intensity}\ ^{12}C}\right)}{\left(\frac{\text{expected}\ ^{13}C = \binom{n}{1} \times p^1(1-p)^{n-1}}{\text{expected}\ ^{12}C = \binom{n}{n} \times j^n(1-j)^0}\right)} - 1\right) \times 1e3‰ \tag{4}$$

with n = number of Carbon atoms; p = probability of ^{13}C; and j = probability of ^{12}C (here: $1 - p$).

We note that the matching of intensity and related probability ratios may also be expected for any number of ^{13}C isotopes found in the molecule. This means, we expect that $\Delta^{13}C$ calculated for an isotope pair containing zero and one ^{13}C is equal to $\Delta^{13}C$ calculated for isotope pairs involving zero and more than one ^{13}C per molecule, as both nominator and denominator change. We also assumed that isotopologues involving other elements (e.g., ^{18}O) share the same $\Delta^{13}C$ under the condition that they occur in both molecular formulas of the isotope pair. A generalization to a multinomial approach is not necessary for carbon. Due to low peak intensities for molecules containing more than one ^{13}C or other isotopologues such as ^{18}O, we focused on the isotope pairs with the largest intensities which are generally molecules containing only ^{12}C (monoisotopic ions) and one ^{13}C for molecules of low molecular weight (<1000 Da) (eq 4). Contrary to the commonly used C deviance, our approach provides a measure that can directly be compared to natural deviations ($\delta^{13}C$). We purposefully used the annotation Δ instead of δ to emphasize that we were not aiming at detecting natural isotope fractionations related to biological or geochemical processes. In analogy to $\delta^{13}C$ we calculated deviations from the VPDB standard, but in our study, we are not aiming at ruling out instrument error as a source for systematic deviations from that standard.

When comparing C deviance with the theoretically derived $\Delta^{13}C$, we see that C deviance scales linearly with the number of carbon atoms in the molecule (Figure S1, light green line); in fact, the following linear relationship (eq 5) between C deviance and $\Delta^{13}C$ holds true.

$$C\ \text{deviance} = \frac{\frac{\Delta^{13}C}{1000‰} + p}{(1 - p)} \times n \tag{5}$$

Assuming the average natural abundance ratio of $^{13}C/^{12}C$ in a compound containing 100 carbons, the C deviance approach overestimates the number of carbons of that compound by 1. Exemplarily, a $\Delta^{13}C$ of −36 permille would give an absolute C deviance below −1 for a molecular formula with more than 40 carbon atoms as indicated in Figure S1. Such a $\Delta^{13}C$ is within the range of natural isotope fractionation occurring in natural organic matter[29,30] by, for example, enzymatic processes. Consequently, using −1 as a threshold in C deviance would lead to false exclusions when used for molecular formula confirmation. C deviance is an erroneous estimator. Because of that, we strongly suggest using $\Delta^{13}C$ rather than the C-deviance as a measure for isotope ratio deviance because C deviance introduces a systematic error.

FT-ICR-MS signal intensities close to the analytical noise (low SNR) are affected by an intrinsically large error. We investigated the influence of SNR on observed $\Delta^{13}C$ values. For this, for each isotope pair we calculated the standard deviation of $\Delta^{13}C$ and related mean SNR values of the ^{12}C isotopologue over all aligned spectra of either the Deep Sea or North Sea data set. We fitted a model (eq 6) to decompose the error of $\Delta^{13}C$ into a SNR-dependent and an SNR-independent part.

$$\sigma_{\Delta 13C_{SNR}} = \sqrt{(\text{err}_{SNR}/\text{SNR}^b)^2 + \text{err}_{asymptotic}^{\ 2}} \tag{6}$$

Here, "err_{SNR}" and "$\text{err}_{asymptotic}$" represented both independent random errors, where the former represented the SNR-dependent and the latter the SNR-independent (residual) error. The exponent "b" was not supported by a theoretical reasoning but added, because a log−log plot suggested a power law relationship. From the Deep Sea data set we exemplarily took the four molecular formulas with the largest SNR values, so that the SNR-dependence was negligible, and after confirming that their $\Delta^{13}C$ values along spectra were normally distributed, we tested for statistically significant differences between the mean $\Delta^{13}C$ of these four formulas using an analysis of variance (ANOVA).

RESULTS AND DISCUSSION

Overview of Data Basis. The Deep Sea data subsets contained nearly twice the number of molecular formulas ("all matches" = 2590, "likeliest match" = 2489, "isotope verified" = 2072) compared to the North Sea data ("all matches" = 1730, "likeliest match" = 1712, "isotope verified" = 1195). In both cases we obtained only very few multiple formula assignments as indicated by the minor differences between the respective "all matches" and "likeliest match" subsets. The calculation time for the molecular formula attribution of ubiquitous masses was approximately 10−12 times faster compared to a formula attribution done for every spectrum separately (65 Deep Sea, 137 North Sea), because formula attribution had to be done only once. The factor was not directly proportional to the number of spectra, because of the time required for spectral alignment.

Reduction of Mass Error. Averaging masses following from mass alignment yielded a substantial reduction of mass error. As expected, we observed a pronounced MAE reduction for both data sets progressing with the number of aligned

https://dx.doi.org/10.1021/acs.analchem.9b04234
Anal. Chem. 2020, 92, 2558−2565

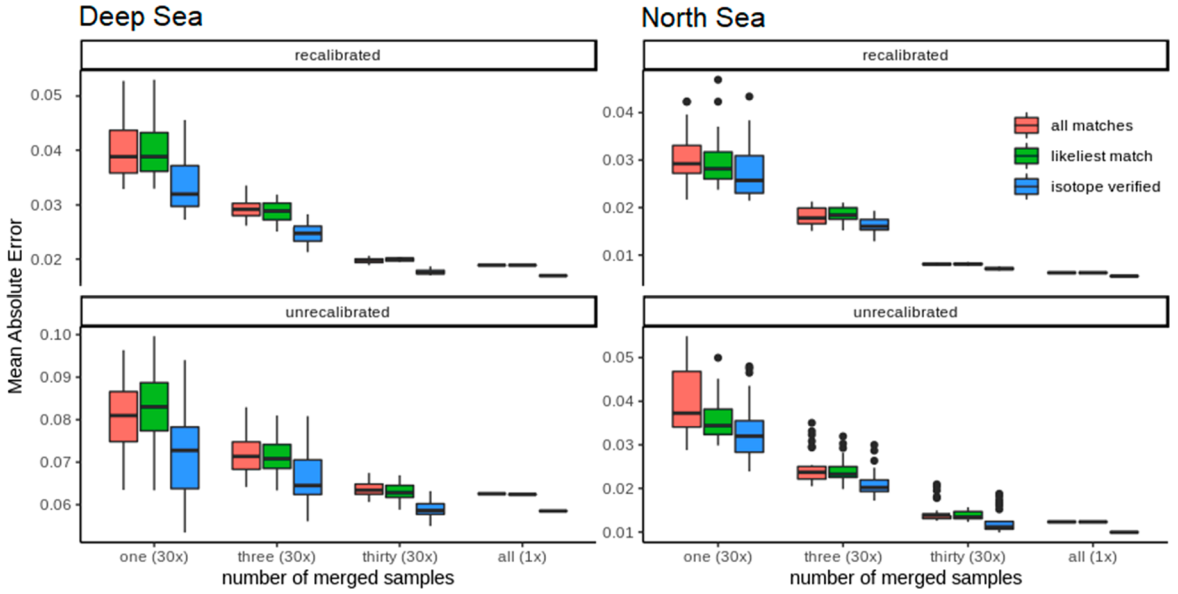


Figure 1. Mean absolute error (MAE) of attributed molecular formulas. The factors "one", "three", "thirty", and "all" correspond to the number of spectra over which equal masses were averaged and used for the MAE calculation.

spectra ("*n*"). For example, the mass error was more than halved when averaged 30 times, resulting in values ranging between 0.015 and 0.065 ppm (Figure 1). Kruskal–Wallis showed that spectral averaging leads to a statistically significant ($p < 0.01$) reduction of MAE (cf. Figure 1) for all subsets (all matches, likeliest matches, isotope verified).

The isotope-verified subsets showed by far the smallest MAE in each case (Figure 1). This underpins the utility of isotope verification for the analysis of FT-ICR-MS mass spectra. The window in the molecular formula attribution can lead to multiple matches for a mass. If the true formula contains elements not included in the formula attribution settings, e.g. Cu, all attributed molecular species must necessarily be wrong. This will most probably increase MAE, especially when the tolerance window is wide. Such formulas will generally not be isotope verified. Consequently, such formulas will be excluded from the isotope verified subset, thus reducing the MAE. A further reason follows from the fact that detected peaks of ^{13}C isotopologues will be adjoint to ^{12}C peaks with high SNR with the consequence of smaller MAE. In other words, 50% of peaks in an isotope-verified subset have a relatively high SNR, going hand in hand with higher mass precision. This ratio of peaks with high SNR is always smaller in the "likeliest match" subset, because it includes ^{12}C mass peaks with SNR values not high enough to exhibit isotopologues.

Finally, the recalibration of the mass axes of each of the individual spectra via general additive models (gam fit) reduced the MAE further for both North Sea and Deep Sea data sets to 0.01–0.02 ppm (Figure 1). The RMSE reduction of the isotope verified subset—with a minimum SNR of 20—of both Deep Sea and North Sea is clearly visible in Figure 2. As predicted,[21] the reduction follows a curve described by formula 2. As already shown for the MAE the mass recalibration clearly reduced the mass error (RMSE) further down for both data sets, but more pronounced for the Deep Sea data set. Still a small amount of systematic mass error remained even after our recalibration. This is clearly visible as the asymptotes of the curves (Figure 2), which represent the leftover systematic error which is always nonzero. This

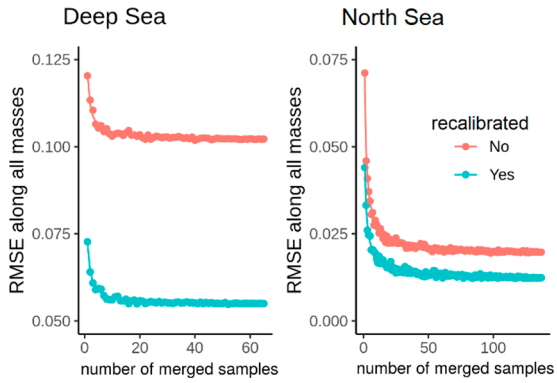


Figure 2. Data inherent mass error reduction by averaging equal considered masses along aligned spectra measurements. The saturation values represent a left-over systematic error. Note the different scale of the *y*-axes to highlight the shape of the curve.

systematic error must be considered as an average systematic error across many compounds; it is due to multiple factors including, but not limited to, imperfect recalibration or remaining wrong formula attributions.

Because the precision is improved by averaging of mass spectra of different samples, mass spectrometers using averaged measurements can become competitive to single measurements of instruments with much higher magnetic field strengths. Extrapolating the effect of magnetic field strength on mean absolute error reported for small molecules,[31] a mean absolute error of 0.01–0.02 ppm, that we reached by averaging 30 spectra using a 15 T instrument, would correspond to single spectrum data with a field strength greater than 21 T, if all other performance parameters are considered identical. Furthermore, mass error reduction would allow shrinking the tolerance window that is used to screen for formula matches. A smaller tolerance window goes hand in hand with a reduced rate of erroneous formula attributions.

Reduction of Error in Isotopologue Intensity Ratios. The distribution of isotope ratio deviances of the ubiquitous isotope pairs ranging below the deviance threshold within

https://dx.doi.org/10.1021/acs.analchem.9b04234
Anal. Chem. 2020, 92, 2558–2565

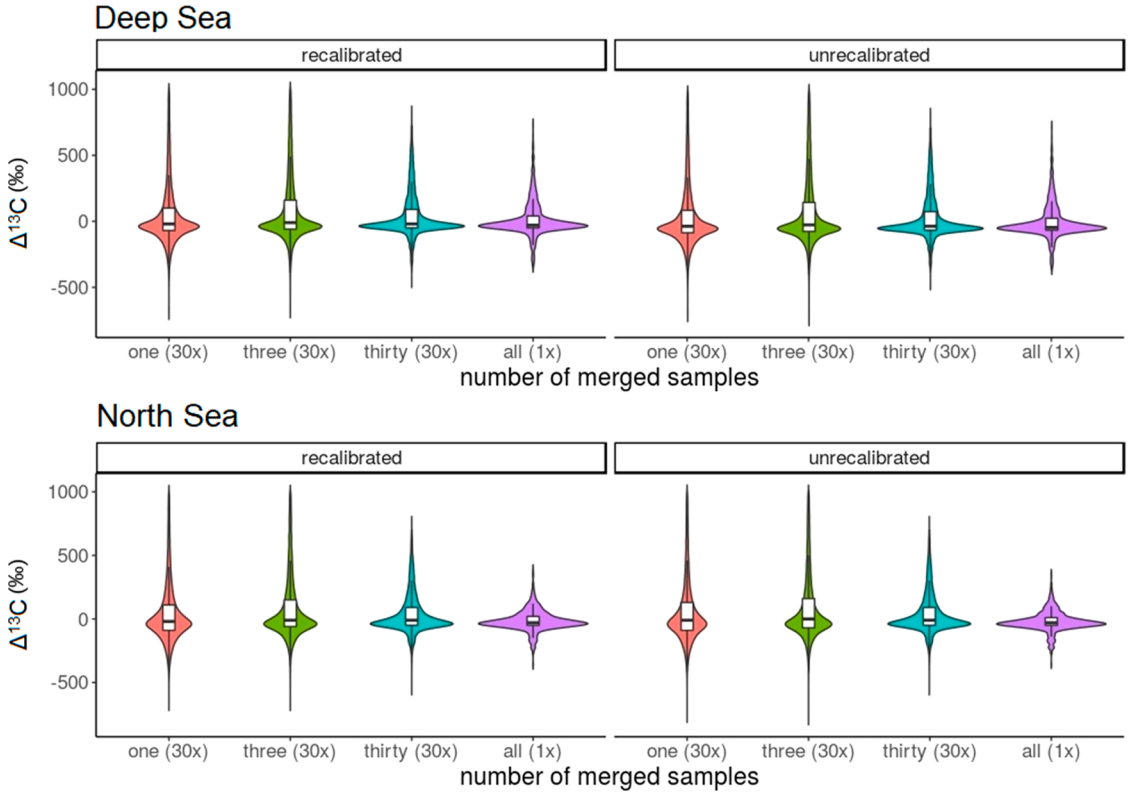


Figure 3. Distributions of $\Delta^{13}C$ of ubiquitous masses after mass merging and averaging of isotope ratios visualized in violin plots which are based on a rotated kernel density estimation of the data flanking a boxplot in the center. The factors "one", "three", "thirty" and "all" correspond to the number of spectra across which averaging took place. The (30×) corresponds to the number of bootstraps.

molecular formula attribution (±1000 permille) showed a clear contraction with growing "*n*" (Figure 3). As expected, the recalibration of masses did not affect the isotope ratio distributions because these were based solely on intensities. By contrast to mass error, the true isotope ratio deviance to be estimated was nonzero. Most natural substances or carbon pools have been subject to enzymatic isotope fractionation and exhibit negative $\delta^{13}C$ values usually in the range of −10‰ to −30‰, but isotope fractionation effects can also be much larger, e.g. in the case of methanogenesis.[32] By decreasing random error through averaging spectra, we considerably narrowed the distribution of $\Delta^{13}C$ close to $\delta^{13}C$ values observed for bulk DOM. Indeed, the $\Delta^{13}C$ values were comparable to observed bulk $\delta^{13}C$ values for natural organic matter in the literature[29,30] and were similar for both Deep Sea and North Sea (Figures 3 and 4). Instrumental fluctuations of intensity values were large (exceeding a factor of 2) and clearly related to instrument drifts, as exemplarily shown for molecular formula $C_{19}H_{24}O_9$ (Figure S2). By contrast, intensity ratios and, as such, $\Delta^{13}C$ values showed no clear autocorrelation or trend. This result is encouraging and indicates that it is possible to accurately and reproducibly determine isotope ratios of bulk DOM from averaging isotope ratios of individual compounds detected in FT-ICR-MS spectra. It also independently confirms that molecular formula assignment based on detected masses was accurate because the $\Delta^{13}C$ values converge to values we expect for DOM. Furthermore, the tolerance of ±1000 permille that we used within formula attribution could be reduced to ±400 permille for both data sets (Figure 4) when using all spectra. In an even more conservative setting, isobaric compounds, where only the

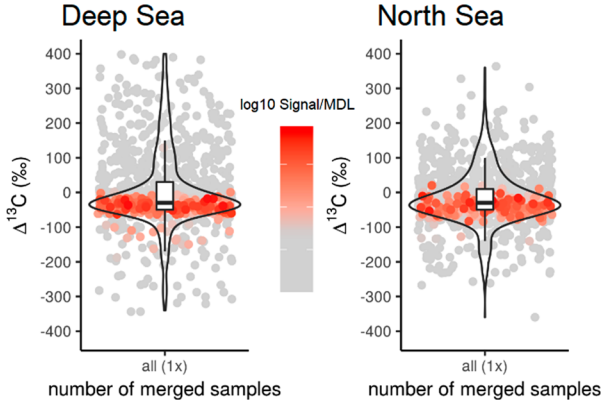


Figure 4. Violin plots of averaged $\Delta^{13}C$ values calculated along all spectra of the respective data sets for each isotope pair. Colors (red = high, gray = low) correspond to mean signal-to-noise ratios of the respective isotope pair across all spectra.

numbers of heteroatoms and hydrogens varies, not the numbers of carbons, can be accounted for by extending our method to other isotopologues, such as ^{15}N.

The SNR of the isotope pair had a systematic influence on the range of $\Delta^{13}C$ (Figures 4 and 5, here SNR based on the ^{12}C peak). This suggests that $\Delta^{13}C$ values are measured more precisely with higher SNR, which is consistent with the fact that instrumental noise in intensity is higher for low-SNR peaks. Our fitted model (eq 6) separating SNR dependent from SNR independent random error described both data sets (Figure 5) very well, although to a lesser extent the North Sea

https://dx.doi.org/10.1021/acs.analchem.9b04234
Anal. Chem. 2020, 92, 2558−2565

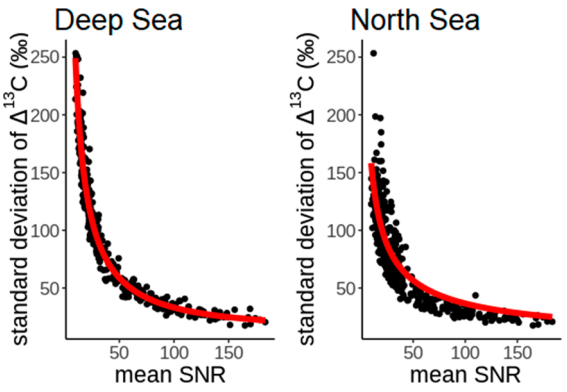


Figure 5. Variance decomposition of $\Delta^{13}C$ values observed in both data sets. Red line indicates model fit. The SNR of the ^{12}C component was used for each isotope pair.

data set. We observed an asymptotic SNR-independent error in the data of 10‰ for North Sea and 11‰ for Deep Sea.

Quantile plots (Figure S3) for the four masses with the highest mean SNR within the Deep Sea data set showed normally distributed $\Delta^{13}C$ values. This strengthens our assumptions that the left-over (asymptotic) error can be reduced via averaging spectra by a factor of $1/\sqrt{n}$. The mean $\Delta^{13}C$ of these four formulas (means with 95% confidence intervals: $C_{19}H_{24}O_9$ = −20.85 ‰ ± 3.8 ‰; $C_{19}H_{24}O_{10}$ = −22.53 ‰ ± 4.4 ‰; $C_{20}H_{24}O_{10}$ = −46.54 ‰ ± 2.8 ‰; $C_{20}H_{26}O_{10}$ = −43.17 ‰ ± 3.0 ‰) showed statistically significant differences (Figure S3) tested by ANOVA ($p < 0.01$). There is no reason to assume that elemental composition or structural features of a compound would lead to systematic instrument errors of $\Delta^{13}C$. Thus, it is possible that we approximated natural isotope ratios ($\delta^{13}C$) of individual molecules by determining $\Delta^{13}C$.

CONCLUSION

With this study we have shown that mass precision of FT-ICR-MS is strongly improved by alignment of spectra from different natural organic matter samples. Time-consuming and expensive replicate measurements are not necessarily required. As a rule of thumb, the best cost to benefit ratio was seemingly obtained with a set of 30 independent samples. More samples or replicate measurements of identical samples did not yield noteworthy improvements, even though studies involving steeper environmental gradients and compositionally more differentiated DOM arguably could benefit from larger data sets. The resulting improved accuracy of mass is a key to successful molecular formula attribution. Our method can not only be applied to improve future data sets but could also easily be used to reprocess existing data sets and be extended to other instruments; e.g. Orbitrap-based mass spectrometers. A further advantage is the reduction of computational effort because molecular formula attribution has to be done only once for the averaged masses and not for every spectrum separately. Furthermore, aligned mass spectra can be directly used to exclude noise in the data by accepting only masses present in at least a certain number of spectra for further processing steps[33]—a strategy that makes intuitive sense when a study involves replicate measurements or only weakly differentiated DOM samples. The smallest mean absolute mass error (MAE) was obtained for isotope-verified subsets of molecular formulas. Thus, only isotope-verified subsets should be used in studies where maximum accuracy of formula attribution is of highest importance. Averaging improves the mass precision and allows obtaining mean absolute errors that otherwise can only be reached by instruments with much larger magnetic field strengths. By averaging 30 spectra (samples) of our 15 T instrument, we reach a precision comparable to a single spectrum measurement of a 21 T instrument. Further improvements in mass accuracy are obtained by novel instrument technology, mainly novel ICR-cell design, absorption mode, and detecting second harmonic oscillations. Here we show that by simple postdetection data processing, significant improvement in mass accuracy can be obtained at little cost.

In addition to improving mass accuracy by aligning spectra and averaging, also the accuracy of intensity ratios of isotopologues was improved to a degree that measured deviations were within the range of natural isotope ratios. We recommend calculating isotope ratios as accurately as possible using binomial probabilities instead of the traditionally used, erroneous expression of C-deviance, which is increasingly biased with the number of carbon atoms in a molecule. In analogy to $\delta^{13}C$ in environmental studies, we introduce $\Delta^{13}C$ as an analytical measure for isotope ratio deviance. Contrary to raw intensities, $\Delta^{13}C$ values are only marginally influenced by instrumental fluctuations and they are reproducibly and accurately detectable. Spectra-averaged $\Delta^{13}C$ values of molecular formulas, especially when accounting for the influence of SNR, reach values in the range of natural organic matter.[29,30,34] However, there is considerable and consistent variation in $\Delta^{13}C$ of individual molecules. We encourage future studies to validate this observation with reference compounds with known $\delta^{13}C$. The determination of $\delta^{13}C$ on a molecular formula level would allow revolutionary new insights into the source and cycling of compounds through aquatic food webs and in biogeochemical cycles.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.9b04234.

Figure comparing $\Delta^{13}C$ and C deviance; figure showing instrumental drift of intensities exemplarily shown for $C_{19}H_{24}O_9$; figure with $\Delta^{13}C$ values of the four molecular formulas with the highest mean SNR in the Deep Sea data set (PDF)

AUTHOR INFORMATION

Corresponding Author

Thorsten Dittmar − *University of Oldenburg, Oldenburg, Germany, and Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg (HIFMB), Oldenburg, Germany*; Email: thorsten.dittmar@uni-oldenburg.de

Other Authors

Julian Merder − *University of Oldenburg, Oldenburg, Germany*; orcid.org/0000-0002-5958-7016

Jan A. Freund − *University of Oldenburg, Oldenburg, Germany*

Ulrike Feudel − *University of Oldenburg, Oldenburg, Germany*

https://dx.doi.org/10.1021/acs.analchem.9b04234
Anal. Chem. 2020, 92, 2558−2565

Jutta Niggemann − *University of Oldenburg, Oldenburg, Germany*
Gabriel Singer − *Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany, and University of Innsbruck, Innsbruck, Austria*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.9b04234

Author Contributions

JM, JF, GS, and TD designed the analysis. JM coded and conducted the statistical tests with input from JF, UF, GS, JN, and TD. JN analyzed and supplied the North Sea data. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

The alignment algorithm is implemented in a publicly available FT-ICR-MS data analysis pipeline called "ICBM-OCEAN"- available under www.icbm.de/icbm-ocean. The alignment code can be found at: https://github.com/JulianMerder/ICBM-OCEAN--Supplementary-Code-Modules.git.

■ ACKNOWLEDGMENTS

This study was carried out in the framework of the PhD research training group "The Ecology of Molecules" (EcoMol) supported by the Lower Saxony Ministry for Science and Culture. We thank Katrin Klaproth for highly qualified technical assistance with FT-ICR-MS analyses.

■ REFERENCES

(1) Wagner, C.; Sefkow, M.; Kopka, J. *Phytochemistry* **2003**, *62*, 887−900.
(2) Qi, Y.; Volmer, D. A. *Rapid Commun. Mass Spectrom.* **2019**, *33*, 2−10.
(3) Kujawinski, E. B. *Environ. Forensics* **2002**, *3*, 207−216.
(4) Zark, M.; Christoffers, J.; Dittmar, T. *Mar. Chem.* **2017**, *191*, 9−15.
(5) Hedges, J. I. *Mar. Chem.* **1992**, *39*, 67−93.
(6) Kellerman, A. M.; Dittmar, T.; Kothawala, D. N.; Tranvik, L. *Nat. Commun.* **2014**, *5*, 3804.
(7) Jiao, N.; Herndl, G. J.; Hansell, D. A.; Benner, R.; Kattner, G.; Wilhelm, S. W.; Kirchman, D. L.; Weinbauer, M. G.; Luo, T.; Chen, F.; Azam, F. *Nat. Rev. Microbiol.* **2010**, *8*, 593.
(8) Koch, B. P.; Dittmar, T.; Witt, M.; Kattner, G. *Anal. Chem.* **2007**, *79*, 1758−1763.
(9) Kind, T.; Fiehn, O. *BMC Bioinf.* **2006**, *7*, 234.
(10) Hertkorn, N.; Frommberger, M.; Witt, M.; Koch, B. P.; Schmitt-Kopplin, P.; Perdue, E. M. *Anal. Chem.* **2008**, *80*, 8908−8919.
(11) Stenson, A. C.; Marshall, A. G.; Cooper, W. T. *Anal. Chem.* **2003**, *75*, 1275−1284.
(12) Tolić, N.; Liu, Y.; Liyu, A.; Shen, Y.; Tfaily, M. M.; Kujawinski, E. B.; Longnecker, K.; Kuo, L.-J.; Robinson, E. W.; Paša-Tolić, L.; Hess, N. J. *Anal. Chem.* **2017**, *89*, 12659−12665.
(13) Hughey, C. A.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G.; Qian, K. *Anal. Chem.* **2001**, *73*, 4676−4681.
(14) Singer, G. A.; Fasching, C.; Wilhelm, L.; Niggemann, J.; Steier, P.; Dittmar, T.; Battin, T. J. *Nat. Geosci.* **2012**, *5*, 710−714.
(15) Zark, M.; Dittmar, T. *Nat. Commun.* **2018**, *9*, 3178.
(16) Chai, T.; Draxler, R. R. *Geosci. Model Dev.* **2014**, *7*, 1247−1250.
(17) Green, N. W.; Perdue, E. M.; Aiken, G. R.; Butler, K. D.; Chen, H.; Dittmar, T.; Niggemann, J.; Stubbins, A. *Mar. Chem.* **2014**, *161*, 14−19.
(18) Riedel, T.; Dittmar, T. *Anal. Chem.* **2014**, *86*, 8376−8382.
(19) Dittmar, T.; Koch, B.; Hertkorn, N.; Kattner, G. *Limnol. Oceanogr.: Methods* **2008**, *6*, 230−235.
(20) Kruskal, W. H.; Wallis, W. A. J. *J. Am. Stat. Assoc.* **1952**, *47*, 583−621.
(21) Savory, J. J.; Kaiser, N. K.; McKenna, A. M.; Xian, F.; Blakney, G. T.; Rodgers, R. P.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2011**, *83*, 1732−1736.
(22) Wood, S. N. *Generalized Additive Models: An Introduction with R*; Chapman and Hall/CRC, 2017.
(23) Leefmann, T.; Frickenhaus, S.; Koch, B. P. *Rapid Commun. Mass Spectrom.* **2019**, *33*, 193−202.
(24) Weber, R. J. M.; Southam, A. D.; Sommer, U.; Viant, M. R. *Anal. Chem.* **2011**, *83*, 3737−3743.
(25) Kendall, C.; Caldwell, E. A. In *Isotope Tracers in Catchment Hydrology*; Kendall, C., McDonnell, J. J., Eds.; Elsevier: Amsterdam, 1998; pp 51−86.
(26) Reference and Intercomparison Materials for Stable Isotopes of Light Elements; International Atomic Energy Agency: Vienna, 1995.
(27) Libes, S. M. *Introduction to marine biogeochemistry*, 2nd ed.; Academic Press: Amsterdam; Boston, 2009; p xiii, 909 p.
(28) Alonso, J. I. G.; Rodriguez-González, P. *Isotope Dilution Mass Spectrometry*; Royal Society of Chemistry, 2013.
(29) de Kluijver, A.; Schoon, P. L.; Downing, J. A.; Schouten, S.; Middelburg, J. *Biogeosciences* **2014**, *11*, 6265−6276.
(30) McCarthy, M. D.; Beaupre, S. R.; Walker, B. D.; Voparil, I.; Guilderson, T. P.; Druffel, E. R. M. *Nat. Geosci.* **2011**, *4*, 32−36.
(31) Karabacak, N. M.; Easterling, M. L.; Agar, N. Y. R.; Agar, J. N. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1218−1222.
(32) Summons, R. E.; Franzmann, P. D.; Nichols, P. D. *Org. Geochem.* **1998**, *28*, 465−475.
(33) Payne, T. G.; Southam, A. D.; Arvanitis, T. N.; Viant, M. R. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1087−1095.
(34) Lamb, A.; Wilson, G.; Leng, M. *Earth-Sci. Rev.* **2006**, *75*, 29−57.