Squirrel and Pika Error Estimates

* UMR error review
* HR ‘TotCountErr’
* HR ‘TotSigErr’

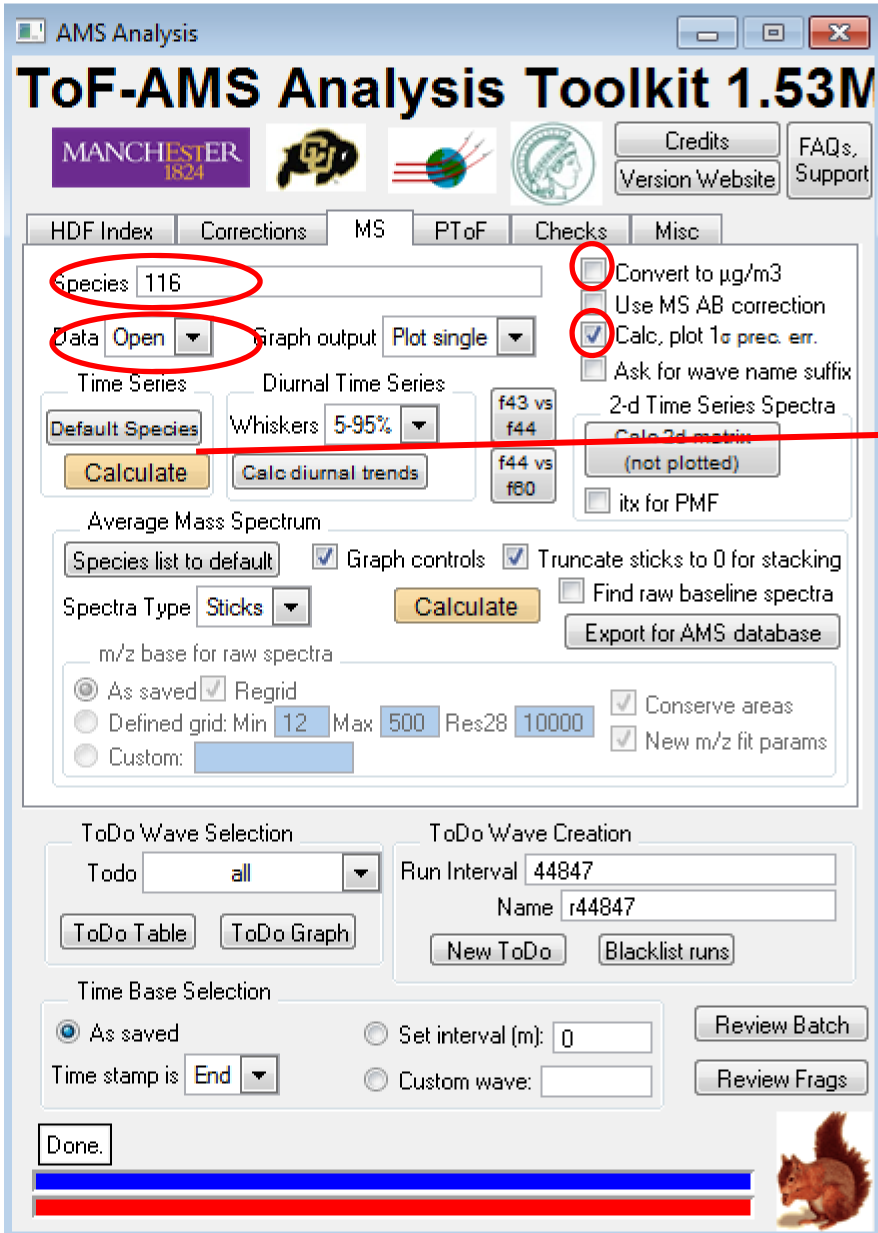
May 1, 2014,

Donna Sueper

Pika version 1.53N

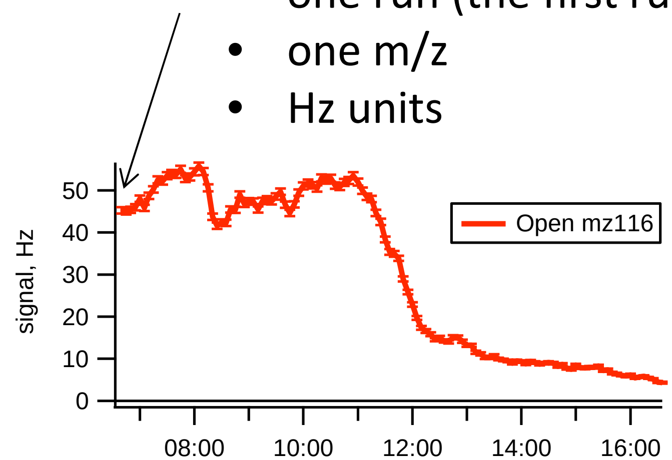
UMR error concepts and math derivations by James Allen,
http://cires.colorado.edu/jimenez-group/UsrMtgs/UsersMtg11/JDAerrorpresentation.pdf

How do I view the UMR errors that have been calculated in the corrections tab? …. One method



For simplicity, focus on

- one run (the first run)
- one m/z
- Hz units



m/z 116 signal for the first run for MSOpen is **45.245** Hz

m/z 116 error for the first run for MSOpen is **0.745** Hz

The next slide will explicitly show how the 0.745 Hz error was calculated for the first run.

Detailed example of UMR 1σ precision error calculation for MS at single run, single m/z

For this run we spent 63.5 s in MS Open; the AB correction factor for this run is 0.911 (unitless); single ion strength is 16.4 (bits); ToF pulsers were at 25000 (Hz); UMR integration width is 52.294 (ns); the signal above baseline of mz116 is 45.245 Hz; the signal of the baseline mz116 is 0.202 Hz. From the AMS panel input sigma is 1.2 (unitless) and electronic noise is 0.001 bits*sqrt(s). The ToF duty cycle correction is typically formulated as sqrt(28/mz of interest).

<u>For the open signal above baseline:</u>

OpenErr_a = abs(signal) = abs(45.245)	= 45.245 (Hz)
OpenErr_b = OpenErr_a / sampling time in s= 45.245/63.5	= 0.713 (Hz^2)
OpenErr_c = OpenErr_b * sigma^2 = 0.713*(1.2^2)	= 1.026 (Hz^2)
OpenErr_d = OpenErr_c / AB correction factor = 1.026/0.911	= 1.126 (Hz^2)
OpenErr_e = OpenErr_d / ToF duty cycle corr = 1.126*sqrt(28/116)	= 0.553 (Hz^2)

<u>For the electronice noise (not dependent on time-varying quantities nor open/closed):</u>

ElecNoise_err_a = ElecNoise^2 = 0.001^2	= 1e-6 (s *bits^2)
ElecNoise_err_b = ElecNoise_err_a / (single ion^2) = 1e-6 / (16.4^2)	= 3.787e-9 (s)
ElecNoise_err_c = ElecNoise_err_b * UMR integration width = 3.787e-9 * 52.294	= 1.944e-7 (unitless)
ElecNoise_err_d = ElecNoise_err_c /ToF pulsers = 1.980e-5 / 25000	= 7.777 e-12 (Hz)
ElecNoise_err_e = ElecNoise_err_d * ToF duty cycle corr^2 = 7.777 e-12 *(28 /116)	= 1.877*e-12 (Hz)

<u>For the open signal below baseline (do just as we did for signal above baseline):</u>

OpenBaseErr_a = abs(0.202)	= 0.202 (Hz)
OpenBaseErr_b = BaseErr_a/sampling time in s= 0.202/63.5	= 3.18e-3 (Hz^2)
OpenBaseErr_c = BaseErr_b * sigma^2 = 3.18e-3*(1.2^2)	= 4.58e-3(Hz^2)
OpenBaseErr_d = BaseErr_c / AB correction factor at this run = 4.58e-3/0.911	= 5.03e-3(Hz^2)
OpenBaseErr_e = BaseErr_d / ToF duty cycle corr = 5.03e-3/sqrt(28/116)	= 2.47e-3(Hz^2)

Total Open error = sqrt(OpenErr_e + ElecNoise_err_e^2+OpenBaseErr_e + ElecNoise_err_e^2)
= sqrt (0.553 + (1.877*e-12)^2 + 2.47e-3 + (1.877*e-12)^2) **= 0.745 (Hz) for UMR m/z 116 Open; signal =45.25**

Similar calculations for m/z 116 for the closed signal for this same run (using different sampling time in MS Closed mode) gives
gives 0.256 Hz Total cl closed error for the m/z signal of 2.332Hz

Total Diff Error = (Diff signal for this run is open – closed = 45.245 - 2.332 = 42.913 Hz Diff signal
sqrt((OpenErr_e+ElecNoise_err_e^2+OpenBaseErr_e+ElecNoise_err_e^2)+(ClosedErr_e+ElecNoise_err_e^2+OpenBaseErr_e+ ElecNoise_err_e^2))
= 0.788 (Hz) for UMR m/z 116 Diff; signal = 42.913

Ugh! That was painful… what have I learned?

* A good estimate is to divide the Hz signal by the time spent measuring and then find the square root.

For the previous calculation of MS open for this m/z and this run this is
~ sqrt(45/63) ~= sqrt(1) = 1 (est) ~= .7 (actual)

For the previous calculation of MS diff for this m/z and this run this is
~ sqrt ((45/63)^2 + (2/63)^2) ~= sqrt(1) = 1 (est) ~= .8 (actual)

* If you change your AB correction you really should recalculate your errors.

How are HR sticks different? … part 1

This is the 'pika' graph for the same run and same mz as in previous slides of UMR errors. Colored circles have been added to emphasis peak heights.

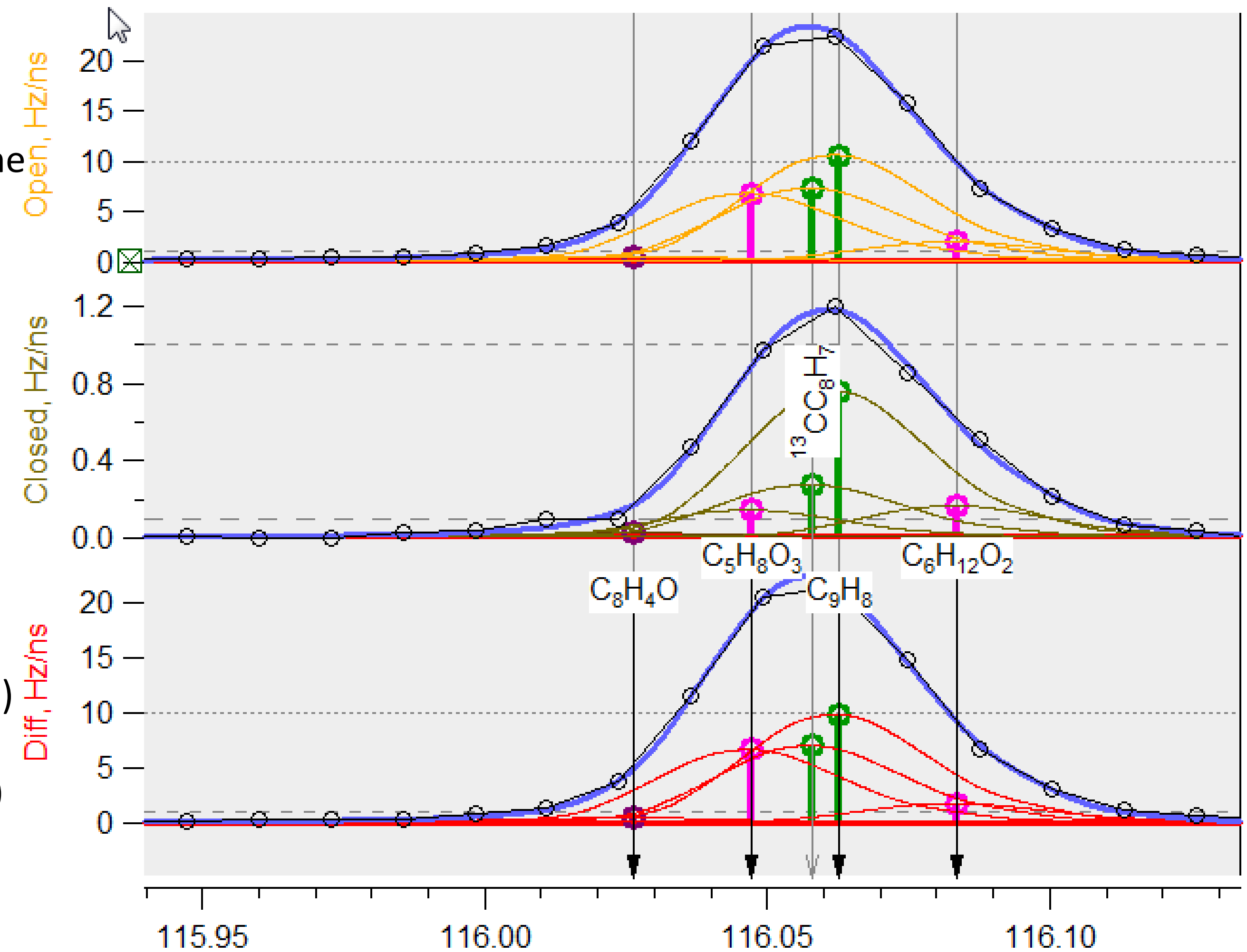
To convert from peak height to peak area we multiply by

HR ion signal = area under gauss-curve = PH*PSS*sqrt(pi)*PW*sqrt(28/116)

where the peak shape scalar (PSS) and the peak width (PW) have been predetermined by the user. The peak height (PH) is what is fit. Let's call this Equation 1.

For this m/z and this run
PW = 1.858, PSS = 1.034, PH=7.181

As an aside, this is V mode data.



OneRunHRSticksTable

590R X 1C | 12

ExactMassW	ExactMassText	Open Stick	Closed Stick	Open Peak Height	Closed Peak Height
116.026	C8H4O	0.744	0.036	0.449	0.022
116.047	C5H8O3	11.064	0.230	6.685	0.139
116.058	j13CC8H7	11.884	0.444	7.181	0.269
116.063	C9H8	17.294	1.242	10.451	0.750
116.084	C6H12O2	3.074	0.265	1.858	0.160

How are HR sticks different? …part 2

The HR stick calculation is heavily dependent on

(a) the m/z calibration, which directly affects

  (ai) HR ion identification and

  (aii) fitting

(b) relative proximity and heights of selected HR ions

and in general, to a somewhat lesser extent,

(c) the peak width,

(d) peak shape and

(e) the baseline estimate.

To address a –e a good check is to sum all the HR sticks for one integer m/z and compare to the UMR value. (See the Sq vs Pk button in the pika panel). The examination of isolated HR ions is also most useful.

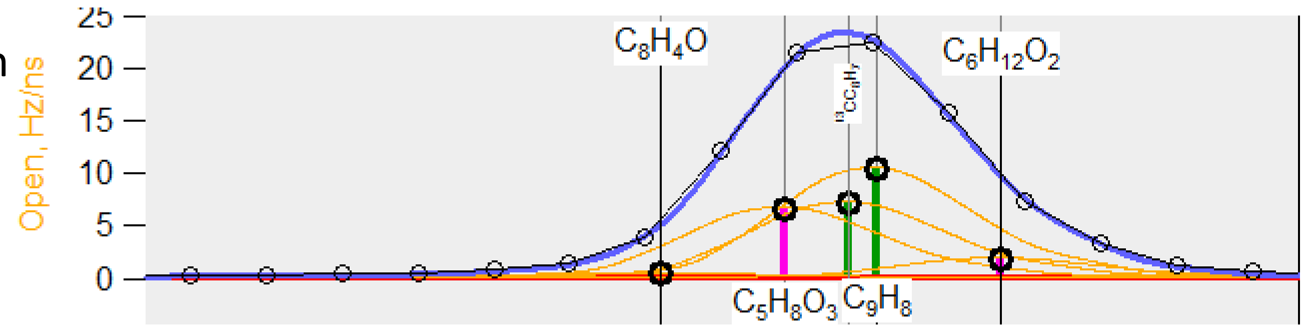
For low m/zs (<80 m/z) the list of possible HR ion formulas is small enough and the m/z distance between chosen adjacent HR ions is wide enough that typically (ai) is not an issue.

For high m/zs (120-150 m/z) use of the W mode signal can instruct the selection of ions, but this must be balanced with the amount of new information they may contain.

What is the HR error named HR*TotCountErr?

Briefly, these are the errors one gets by considering only the 'area under an HR ion fit'.

For simplicity we will focus on the open spectra. In this example there is one isotope, j13CC8H7, of comparable signal to the overlapping ions of C5H8O3 and C9H8.



The HROTotCountErr for C5H8O3 is calculated by the HR ion Hz stick; the HR ion stick is a scalar multiple of the peak height (black circles). With the HR ion stick value we just proceed with the mathematical calculations of the UMR methodology.
So C5H8O3 signal = 11.064 (Hz) and C5H8O3 error = 0.381 (Hz)

m/z	HR ion	Signal (Hz)	Error (Hz)
116.026	C8H4O	0.744	0.140
116.047	C5H8O3	11.064	0.381
116.058	j13CC8H7	11.884	0.394
116.063	C9H8	17.294	0.470
116.084	C6H12O2	3.074	0.219

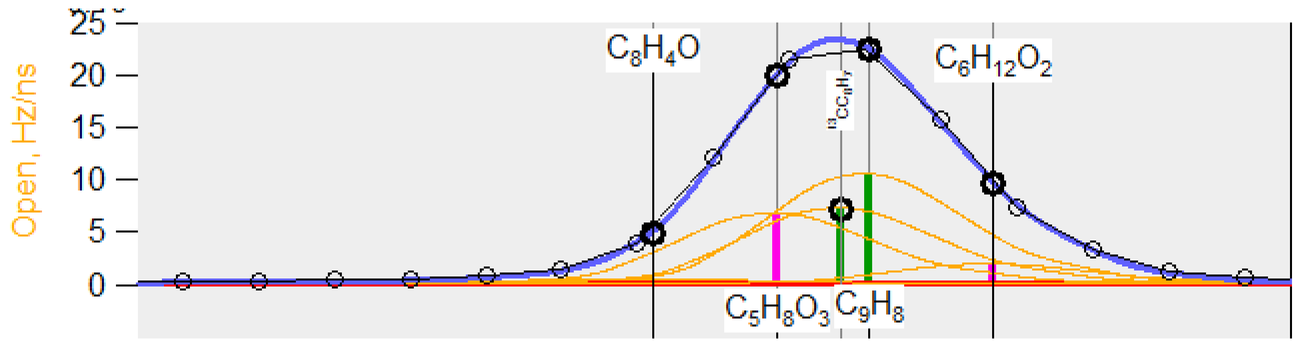
It is clear that this HROTotCountErrr C5H8O3 and C9H8 error is an underestimation due to the overlapping HR ion signals. In general, regions where HR ion peaks overlap are such that: greater overlap -> greater error

However signal for j13CC8H7 is constrained and not 'fit'. It's abundance is predetermined by the signal of the parent ion and the relative isotopic abundance. Colloquially, "it doesn't matter what the data, the black circles at m/z 116, indicate."

What is the HR error named HR*TotSigErr?

Briefly, these are the errors one gets by ‘using the blue trace as a surrogate for peak height’. Except if the HR ion is a constrained isotope, then it is simply it’s ‘real’ peak height.

The black circles for the non isotope HR ions represent a scalar of the number ions measured at an m/z position identified with an HR ion.



m/z	HR ion	Open Peak Height	Open signal (Hz)	HROTotCountErr (Hz)	Error estimate 'peak height'	HROTotSigErr (Hz)
116.026	C8H4O	0.449	0.743	0.140	4.919	0.240
116.047	C5H8O3	6.694	11.078	0.381	20.018	0.700
116.058	j13CC8H7	7.181	11.883	0.394	7.181	0.394
116.063	C9H8	10.446	17.285	0.470	22.476	0.720
116.084	C6H12O2	1.855	3.070	0.219	9.613	0.470

Conclusions

You should prefer the use of TotSigErr over TotCountErr Errors.

TotSigErr >= TotCountErr for any HR ion.
TotSigErr = TotCountErr when the HR ion is isolated.

The TotSigErr errors still do not account for many things such as uncertainty in m/z calibration, peak width, peak shape, baselines…

HR ions that are ‘too close together’ or ‘buried’ in the signal of an adjacent HR ion will have errors that are still currently underestimated.