EGUsphere, referee comment RC1
https://doi.org/10.5194/egusphere-2022-1319-RC1, 2023

Comment on egusphere-2022-1319

Thierry Fouquet (Referee)

Referee comment on "Generalized Kendrick analysis for improved visualization of atmospheric mass spectral data" by Mitchell W. Alton et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-1319-RC1, 2023

Apologies for being late with my comments, I had trouble submitting my post.

I read this article with great interest and did appreciate its form and the application of Kendrick analyses for a new type of mass spectral data and unusual instrument. I also acknowledge the efforts of the Authors to try finding differences between their proposed "scaled Kendrick mass defect" SKMD and the existing "REKMD" or "traditional" Kendrick analysis. If I strongly support new applications of this easy and powerful data processing/visualization tool and its implementation in more programs, I am nonetheless circumspect about the added value of the SKMD formula justifying the introduction of a new name and a seemingly new concept. Based on my comments below, I do strongly recommend not to try adding a new term to a long list of wrongly named methods but keep using either "REKMD" (also wrongly named, but published before) or better simply "Kendrick analysis", and shift the focus of this draft from the differentiation SKMD / REKMD which is non-existent to the application of the Kendrick analysis to their unusual data, providing examples of ion series assignments, separation of ion series of interest from congested mass spectra, deisotoping, binary comparisons, ….

1) The formula of "SKMD" look different from the formula(s) of "REKMD" proposed in the original article, but this notation KM=m/z*x/R has already been reported in J Mass Spectrom. 2019;54:933–947 (doi: /doi.org/10.1002/jms.4480), sadly not cited by the Authors.

2) The case x=1 in the formula KM=m/z*x/R has already been explained in the same article, and named "remainders" by several Authors (e.g. Anal Chem. 2019;91(10):6479â□□6486 or Anal Chem. 2018;90(14):8716â□□8718). The only change introduced by this preprint would then be the range of values 1<x<2/3R, and x>2R. Looking at the plot ΔSKMD and ΔREKMD in the Supporting Information (Fig S1), it seems that the expansions provided by SKMD are already achieved by REKMD in its linearity range. I have not been convinced by the examples of x=4 and x=40 (not truly REKMD in terms of range of divisors) in terms of gain of visualization / separation as

compared to the other plots (x=17, 20, 24) which are truly REKMD. Can the Authors find a case where values of x <2/3R or >2R provide a truly unique separation capability not achievable by REKMD with integer or non-integer x ? I haven't found any case myself yet.

3) The formula KM=m/z*x/R is still fundamentally a traditional Kendrick change of basis as proposed by Kendrick a while ago, simply choosing x instead of round(R) as the new reference mass. The formula KM=m/z*round(R)/R is a basic "rule of three", setting the mass of R at an integer value round(R) to define a new reference instead of the IUPAC convention m(12C)=12, re-calculating other mass accordingly. In the SKM formula, the Authors choose to set the mass of 16O (or other moieties) at 2, or 6, or 40 or any integer instead of 16 to define a new mass scale. The concept does not vary from what Kendrick proposed, so does this really deserve a new name ? I do agree that the same question should have been raised when the concept of REKMD has been introduced. That is the reason why I am strongly in favor of calling the whole method a "Kendrick analysis" with no other mention.

m(/z) 16O IUPAC = 15.9949 --> m(/z) 16O Kendrick base = 16

m(/z) IUPAC --> rule of three --> m(/z) Kendrick base = m/(z) IUPAC *16/15.9949

m(/z) 16O IUPAC = 15.9949 --> m(/z) 16O Kendrick base = x

m(/z) IUPAC --> rule of three --> m(/z) Kendrick base = m/(z) IUPAC *x/15.9949

4) REKMD and its latest variations as reported in detail in the same article J Mass Spectrom. 2019;54:933–947 takes the charge state of the ions into account (simply adding an integer to the formula to cancel z) so it is applicable for multiply charged ions, while SKMD in its current form would deal with singly charged ions only. More importantly, this additional integer in the formula of KM also made possible the generation of an infinite number of Kendrick plots with a pseudo continuous coverage of expansions using non-integer x (ie nearly 0 step between expansions allowing the finest tune to separate series) while SKMD does only provide fixed expansions varying linearly with a step dictated by the value of R itself, and no finer control available (the larger R, the larger the step).

5) REKMD or its variations has already been implemented - but not called as such - in numerous programs, free or commercial, such as MZMine2, MSRepeatFinder, SpectraScope, Kendo, or Mass Mountaineer not to mention in simple Excel spreadsheets. These programs do not mention any REKMD but simply incorporate the divisor "x" which can be changed by the user. It is a wonderful idea to keep implementing the Kendrick analysis with all its variations in other programs, but would that require to use a new name and a seemingly new concept which in fact produces the same results as those already reported ? Would it be clearer for users to have a tool called "Kendrick analysis" with no S or RE or no mass defect (cf last comment below), but simply this "x" textbox to play with the change of basis and expansions ?

6) As a last reason not to introduce another (S)KMD term, and as pointed out by one reader of the community in his comment, several Authors strongly recommend not to use the "KMD" term anymore. The values we are dealing with are not mass defects – and not mass excess either – but fractional mass calculated with no a priori knowledge about the ions. This point has been greatly explained by the Authors in this preprint, as opposite to other plots such as Van Krevelen diagrams which require the elemental composition of ions to be known prior to their generation. The calculation of true mass defect/excess ALSO requires the elemental composition of the ions to be known. The y-values of Kendrick plots are computed a) without knowing the compositions of the ions but only their m/z, and b) are comprised between -0.5 and 0.5 (or -1 and 0, or 0 and 1) regardless of m due to the aliasing of the formula m-round(m) (or floor or ceiling) while the mass defect/excess would keep increasing/decreasing with the number of atoms in a molecule above or below (-)0.5