# From mass to structure: modified van Krevelen diagram and adjusted indexes for high-resolution mass data of organic matter

Yang Xiaoqiu[1], Hechao Chen[1], Jing Wu[1], Yaofei Li[1], Yulin Zhang[1], and Qiwei Zhang[1]

[1]Jianghan University Key Laboratory of Optoelectronic Chemical Materials and Devices

July 10, 2023

**Abstract**

The traditional van Krevelen diagram, the nominal oxidation of organic carbon, and several aromaticity indexes generally considered C, H, O, N, P, and S factors, or just C, H, and O. However, these evaluated methods may provide incorrect interpretation for high-resolution mass data of organics containing halogen elements. Therefore, we adapted the van Krevelen diagram and the other indexes with halogen elements and verified with 2235 model organics. The results demonstrated that the modified van Krevelen diagram and the adjusted indexes were more accurate for the analysis of mass data, either in the compound classification or in the evaluation of oxidation state of organic matter.

**From mass to structure: modifiedvan Krevelen diagram and adjusted indexes for high-resolution mass data of organic matter**

Dear Editor:

The development of high-resolution mass spectrometry (HRMS) has led to deeper insight into the molecular characterization of organic matter. However, a large amount of data obtained by HRMS cannot be immediately converted into specific substances, especially for Fourier transform ion cyclotron resonance mass spectrometry. Some methods have been developed to further explore the information obtained by HRMS by interpreting the data and predicting structures based on exact molecular weight and element constituent patterns.[1] Of these methods, the van Krevelen diagram is widely used to classify of organic matter by plotting atomic ratios of H/C and O/C in two-dimensional ordinations.[2] In addition, some index are employed to evaluate molecular characterizations, such as the nominal oxidation state of carbon (NOSC)[3] and the aromaticity index (AI, AImod and Xc),[4,5]. While the traditional van Krevelen diagram and the indexes were mainly targeted at natural organic matter (NOM), the calculation equations normally considered C, H, O, N, P, and S factors, or just C, H, and O. However, for organic matter, particularly environmental contaminants and artificial chemicals, halogen elements are important elements and their chemical and toxicological properties cannot be overlooked. [6-8] Moreover, the application of HRMS in environmental pollutants and anthropogenic organics is increasing, [8,9] and the ability to more accurately interpret large amounts of data is essential. Therefore, we adapted the van Krevelen diagram and the other indexes with halogen elements for the analysis of organic matter with complex constituents.

To verify the modified van Krevelen diagram and the adjusted indexes, chemical formulas of 2235 organics were evaluated, which originated from a web database of disinfection byproducts (*http://dbps.com.cn/#/main)* . The results were compared and depicted as follows:

**The van Krevelen diagram**

The van Krevelen diagram is normally depicted as a two-dimensional plot using the molar ratio of hydrogen to carbon (H/C) as the ordinate and the molar oxygen-to-carbon ratio (O/C) as the abscissa.[2] The H/C

ratio separates compounds according to the degree of saturation, whereas O/C ratios separate compounds according to O classes.[10] Halogens are generally bonded to organic matter by replacing hydrogen, so their contribution to valence bond saturation should be equivalent to that of hydrogen. For halogen-containing organics, if only the H/C ratio is considered, the material structure will be misjudged. For example, the H/C ratio of trichloromethane is 1, and it is classified as an unsaturated compound by the traditional van Krevelen diagram; however, trichloromethane is a typical saturated organic matter. When halogen is introduced to the calculation of the H/C ratio, H(X)/C (H(X) is the sum of H and halogen elements) of trichloromethane is 4, which is an apparent characteristic of saturated compounds. Fig. 1(a) and (b) show the traditional van Krevelen diagram with H/C versus O/C and the modified van Krevelen diagram with H(X)/C versus O/C, respectively. The distribution of organics clearly shifts upward in the diagram. In particular, compounds with H/C ratios lower than 0.5 are classified as condensed aromatic hydrocarbons,[11,12]while almost no condensed aromatic compounds are found in the 141 compounds with H/C ratios lower than 0.5 in the database. These compounds are mostly multihalogen substituted organics with H(X)/C ranging from 0.67 to 4. The above results indicate that using the modified van Krevelen diagram plotted with H(X)/C versus O/C could provide a more accurate classification for the organic matter detected by HRMS.

**The oxidation of organic carbon**

Three indexes are used to evaluate the oxidation state of organic carbon, including Cox,[3] NOSC,[13] and CHO.[11] The equations of these three indexes are calculated as follows:

$$Cox = (2O+2S-H+3N-5P)/C \quad (1)$$

$$NOSC = -((-Z+4C+H-3N-2O+5P-2S)/C)+4 \quad (2)$$

$$CHO = (2O-H)/C \quad (3)$$

In these three equations, C, O, H, N, P, and S represent the stoichiometric numbers of the elements C, O, H, N, P and S, respectively. Z in Eqn. (2) is the net charge of organic matter. Cox and CHO are special cases of NOSC. When Z=0, Cox = NOSC, and when the stoichiometric numbers of the heteroatoms are zero, CHO = NOSC. However, even NOSC ignores the contribution of halogen to the oxidation state of organic carbon. According to the definition of NOSC and Cox,[3,13] we recommend the equation used to adjust NOSC, which is depicted as NOSC': NOSC' = -((-Z+4C+H-3N-2O+5P-2S-X)/C)+4. In the equation of NOSC', X is the sum of stoichiometric numbers of halogen elements (F, Cl, Br, I).

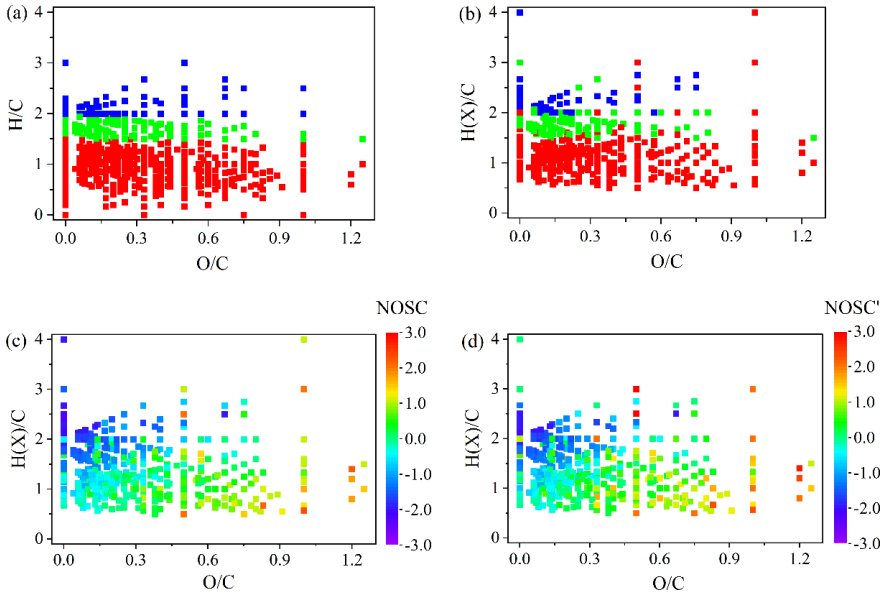


Fig. 1 (a) The traditional van Krevelen diagram; (b) The modified van Krevelen diagram; (c) The modified van Krevelen diagram associated with NOSC; (d) The modified van Krevelen diagram associated with NOSC'.

(In (a) and (b), the red dots symbol the compounds with H/C<1.5, the green dots are the compounds with 1.5[?]H/C<2.0, and the blue dots are the compounds with H/C[?]2.0; NOSC: the nominal oxidation state of carbon; NOSC': the adjusted nominal oxidation state of carbon)

Fig. 1(c) and (d) demonstrate the modified van Krevelen diagrams associated with NOSC and NOSC', respectively. It can be seen that the compounds with NOSC'>0 are slightly higher than those with NOSC>0. Among the 2235 organics in the database, approximately 21.3% with NOSC<0 in Fig. 1(c) change into that with NOSC'[?]0 in Fig. 1(d). Generally, compounds with NOSC>0 are more oxidized and less degradable, while those with NOSC<0 are more likely degraded. When the impact of halogen is not considered, the oxidation state of organic carbon is underestimated, and the degradation potential of organic matter is overestimated.

**Aromatic indexes**

Three methods were used for calculating the aromaticity index, including AI, AImod [4] and Xc.[5] Of them, AI was suggested for compounds in which oxygen was mainly present as carbonyl functional groups; AImod was adjusted on the basis of AI when oxygen could also be bound with π-bonds in carboxyl groups;[12]and Xc was aimed at aromatic compounds with long alkylation.[5] Considering the contribution of halogen elements to the saturation of organic matter, we adjusted these three indexes based on the original equations and put forward them as AI', AImod' and Xc', respectively. The equations used to calculate the tree adjusted indexes are as follows:

AI' = (1+C-O-S-0.5(N+P+H+X))/(C-O-N-S-P) AImod' = (1+C-0.5O-S-0.5(N+P+H+X)/(C-0.5O-N-S-P)
Xc' = [3((1+C-0.5(H+X)+0.5(N+P)-(O+S))-2]/(1+C-0.5(H+X)+0.5(N+P)-(O+S))

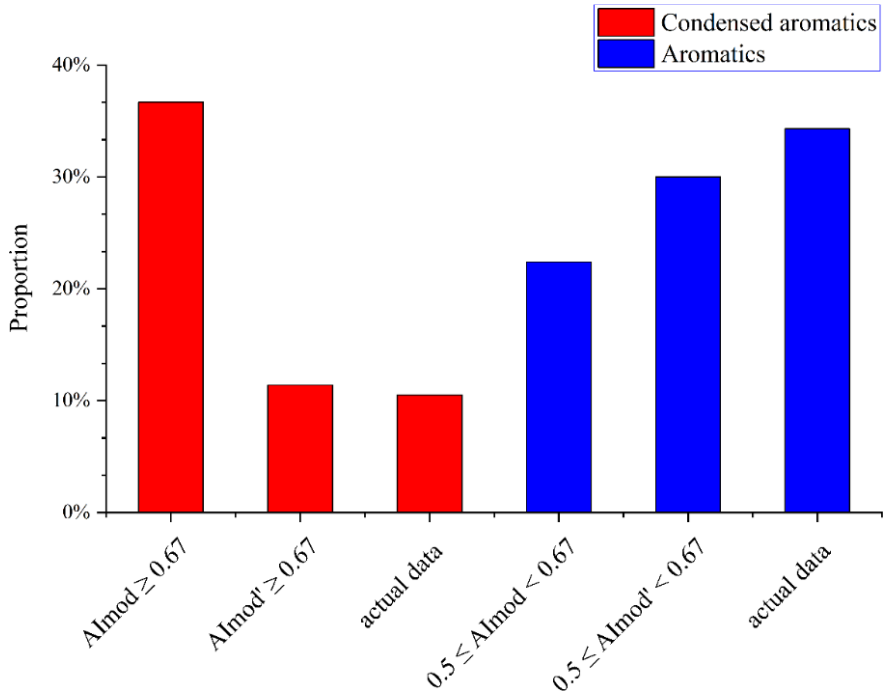


Fig. 2 Comparison of AImod and AImod' with actual data (organics with AImod or AImod'[?]0.67 are classified as condensed aromatics, and those with 0.5[?]AImod or AImod' <0.67 are categorized as aromatics)

According to the structures of the 2235 organics, a comparison between AImod and AImod' was chosen to determine the validity of the adjusted aromatic indexes (as shown in Fig. 2). The proportions of condensed aromatics and aromatics predicted by AImod' are closer to the actual data than those predicted by AImod.

In summary, introducing halogen to the interpretation models of HRMS data for organic matter improves the accuracy. The modified van Krevelen diagram, NOSC and aromatic indexes are equivalent for multiple organic matter, including NOM, emerging environmental pollutants, and artificial products.

**Acknowledgments**

The authors gratefully acknowledge the advice of Pro. Chenggang Zhou from China University of Geosciences. This work was supported by the Hubei Provincial Department of Education (General Guiding Project, B2022279).

*Hechao Chen[1], Jing Wu[1], Yaofei Li[1], Yulin Zhang[1], Qiwei Zhang[1], Xiaoqiu Yang[1]**

1. Key Laboratory of Optoelectronic Chemical Materials and Devices of Ministry of Education, School of Optoelectronic Materials & Technology, Jianghan University, Wuhan 430056, China

*Corresponding to X.Q. Yang, School of Optoelectronic Materials & Technology, Jianghan University, Wuhan 430056, China. Email:*xiaoqiuyang@jhun.edu.cn*

**References**

1. Zhang XX, Han JR, Zhang XR, et al. Application of Fourier transform ion cyclotron resonance mass spectrometry to characterize natural organic matter. *Chemosphere.* 2020;260.

2. Kim S, Kramer RW, Hatcher PG. Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van Krevelen diagram. *Anal Chem.* 2003;75(20):5336-5344.

3. Masiello CA, Gallagher ME, Randerson JT, Deco RM, Chadwick OA. Evaluating two experimental approaches for measuring ecosystem carbon oxidation state and oxidative ratio. *Journal of Geophysical Research: Biogeosciences.* 2008;113(G3).

4. Koch BP, Dittmar T. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter (vol 20, pg 926, 2006). *Rapid Commun Mass Spectrom.* 2016;30(1):250-250.

5. Yassine MM, Harir M, Dabek-Zlotorzynska E, Schmitt-Kopplin P. Structural characterization of organic aerosol using Fourier transform ion cyclotron resonance mass spectrometry: Aromaticity equivalent approach. *Rapid Commun Mass Spectrom.* 2014;28(22):2445-2454.

6. Richardson SD, Ternes TA. Water Analysis: Emerging Contaminants and Current Issues. *Anal Chem.* 2022;94(1):382-416.

7. Richardson SD, Plewa MJ, Wagner ED, Schoeny R, Demarini DM. Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: a review and roadmap for research. *Mutat Res.* 2007;636(1-3):178-242.

8. Peng H, Chen CL, Cantin J, et al. Untargeted Screening and Distribution of Organo-Iodine Compounds in Sediments from Lake Michigan and the Arctic Ocean. *Environmental Science & Technology.*2016;50(18):10097-10105.

9. Zhang HF, Yang M. Characterization of brominated disinfection byproducts formed during chloramination of fulvic acid in the presence of bromide. *Sci Total Environ.* 2018;627:118-124.

10. Wu ZG, Rodgers RP, Marshall AG. Two- and three-dimensional van Krevelen diagrams: A graphical analysis complementary to the Kendrick mass plot for sorting elemental compositions of complex organic mixtures based on ultrahigh-resolution broadband Fourier transform ion cyclotron resonance mass measurements. *Anal Chem.* 2004;76(9):2511-2516.

11. Mann BF, Chen HM, Herndon EM, et al. Indexing Permafrost Soil Organic Matter Degradation Using High-Resolution Mass Spectrometry.*Plos One.* 2015;10(6).

12. Koch BP, Dittmar T. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Commun Mass Spectrom.* 2006;20(5):926-932.

13. LaRowe DE, Van Cappellen P. Degradation of natural organic matter: A thermodynamic analysis. *Geochim Cosmochim Acta.*2011;75(8):2030-2042.