

MassAnalyzer 1.0 软件使用说明

一、功能描述

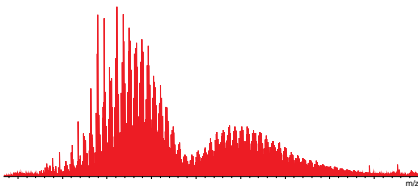
MassAnalyzer 软件是一种基于机器学习的质谱图分子式快速解析算法，MassAnalyzer 1.0 版本可用于计算含 Carbon(C), Hydrogen (H), Oxygen (O), Nitrogen (N), Sulfur (S), Chlorine (Cl), Bromine (Br), Fluorine (F) 和 Phosphorus (P)等九种元素的有机物负离子模式质谱图中质核比 (m/z) 对应的分子式, 目前支持负离子模式: M-H。程序可以通过 [MassAnalyzer: Machine learning-based mass spectrometry discovery \(sjtu.edu.cn\)](http://MassAnalyzer: Machine learning-based mass spectrometry discovery (sjtu.edu.cn)) 访问。在运行本程序前, 请仔细阅读本说明并设置好计算的参数, 确认输入质谱文件格式正确, 并留下邮箱, 计算结果会在 24h 内发送到您的邮箱。

二、软件输入格式

输入文件应为 excel 的 '.xlsx' 格式。文件应该包含四列, 标题分别为 'm/z', 'I', 'Res', 'S/N'。标题顺序可以改变但是大小写和符号必须正确, 其中 m/z 表示质核比, I 表示峰强度, Res.表示分辨率, S/N 表示信噪比。如果缺少分辨率和信噪比的信息, 该列置 0。具体 excel 形式可以参考网页中的 example download。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	m/z	I	Res.	S/N	注1、每一列标题为“m/z”, “I”, “Res. ”, “S/N”。										
2	100. 7077	4043925	1346738	17. 1	注2、标题顺序可以改变, 但是大小写必须正确。										
3	106. 2823	3481865	1285391	14. 4	注3、m/z是质核比, I是峰强度, 这两组数据必须有。										
4	106. 9479	3823333	904969	16	注4、Res. 是分辨率, S/N是信噪比, 这两组数据如果没有请置0。										
5	109. 0294	4677805	974085	20. 1	注5、上传时除第一行标题不能出现如何字母和汉字。										
6	110. 6183	8203261	899352	36. 7											
7	118. 5385	7229382	758144	32. 1											
8	118. 5419	23600582	9561066	109. 3											
9	126. 948	4281354	855590	18. 2											
10	126. 9512	4113162	835031	17. 4											
11	126. 9541	3482762	861577	14. 4											

图 1. MassAnalyzer 1.0 软件输入文件格式。



三、 参数设置

the maximum relative error 参数为最大相对误差，由质谱仪决定相对误差大小，本软件选用的单位为 ppm，即百万分之一，公式是 $Acc = (\text{实验值} - \text{理论值}) / \text{理论值} * 1000000$ 。请输入大于 0 的整数。

其中，the maximum number of nitrogen atoms, sulfur atoms and chlorine atoms 为解析分子式中最大允许 N, S, Cl 原子数目大小，需要输入不小于 0 的整数，建议略大于实验中已知有机物原子数目，如果样本中不含 S 或 Cl，输入 0 即可。Bromine (Br), Fluorine (F), Phosphorus (P) 为三个可选项，如果确认需要解析的分子式中含有这些元素，则勾选对应的元素。

四、 运行及输出文件说明

程序每次只能提交一个 excel 文件，但是可以在结果发送之前可以多次提交不同的文件。点击 submit 提交之后，结果会在 24h 内发送到您输入的电子邮箱中，邮件压缩包密码请联系 hbshen@sjtu.edu.cn 获取。每一个文件的结果都会保存为一个压缩包，其中包含 1 个 excel 表和 8 张对应的可视化图。

- Excel 表格中有 6 个 sheet，其中 result 为非同位素的分子式，也是作图的依据，resultC、resultS、resultO、resultH 为同位素峰的分子式，MolecularClass 为有机物的分类。
- 8 张图片的具体含义及图形示例如下：

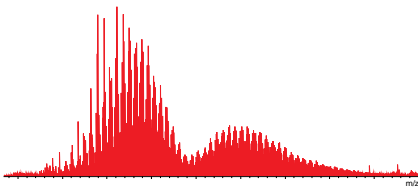
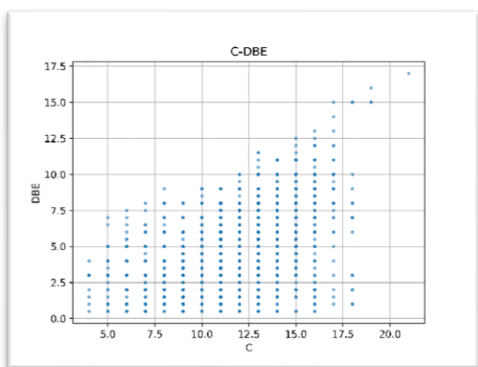
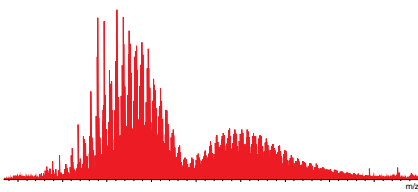
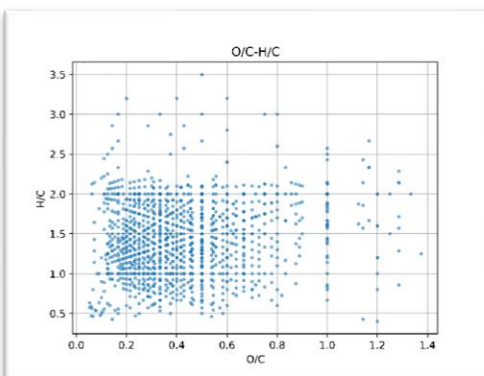


表 1. MassAnalyzer 软件输出的 8 张解析图片含义。

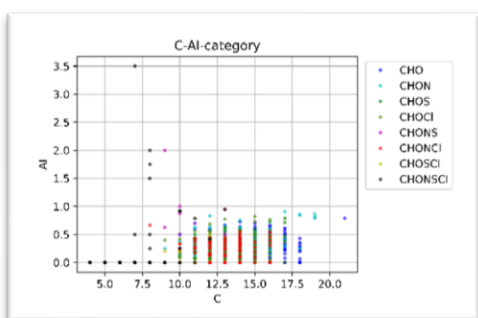
	结果图文件名	结果图名字	含义
1	plots_of_DBE_versus_carbon_numbers	C-DBE	碳原子数量和 DBE 关系示意图
2	plots_of_H_C_ratios_versus_O_C_ratios	O/C-H/C	碳氢比和碳氧比关系示意图
3	C-AI-category	C-AI-category	依据 category 分类的碳原子数量和 AI 关系示意图
4	C-Xc-category	C-Xc-category	依据 category 分类的碳原子数量和 Xc 关系示意图
5	mz-HC-category	mz-HC-category	依据 category 分类的质核比和碳氢比关系示意图
6	mz-KMD-category	mz-KMD-category	依据 category 分类的质核比和 KMD 关系示意图
7	Van_Krevelen_diagrams_relative_intensity	O/C-H/C-relative intensity	依据相对峰强度分类的碳氢比和碳氧比关系示意图
8	Van_Krevelen_diagrams_s_n	O/C-H/C-s/n	依据信噪比分类的碳氢比和碳氧比关系示意图



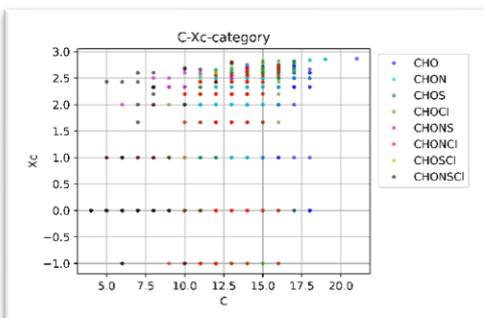
(1) C-DBE



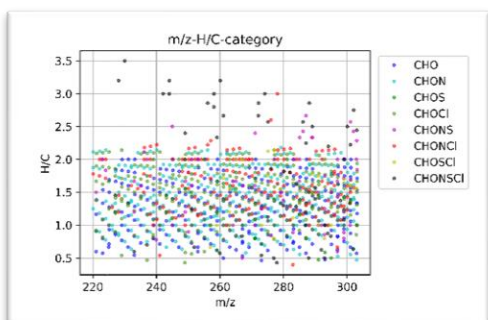
(2) O/C-H/C



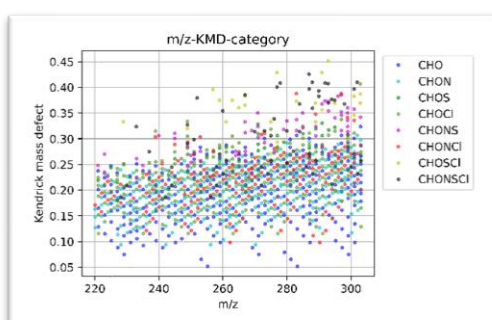
(3) C-AI-category



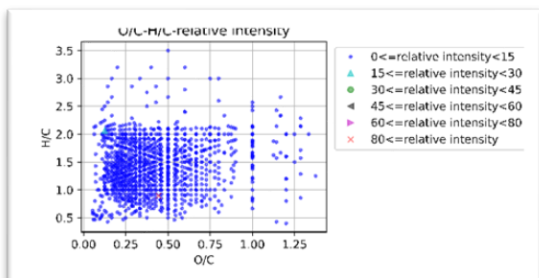
(4) C-Xc-category



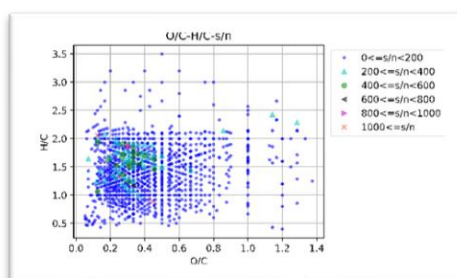
(5) m/z-HC-category



(6) m/z-KMD-category

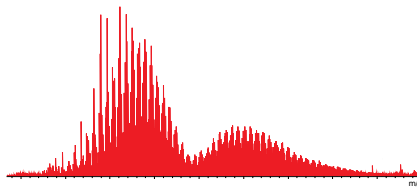


(7) O/C-H/C-relative intensity



(8) O/C-H/C-s/n

图 6. 计算 input_sample 的 8 张图片结果图示例。



五、维护和开发

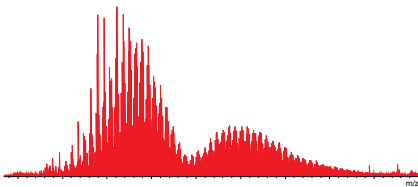
运行过程中如出现异常等问题请及时联系我们, 将输入文件及参数发送到邮箱。任何建议或者针对性需求, 如添加更多的元素或者绘制更多的可视化结果图, 请与我们联系。

六、联系方式

Email: hbshen@sjtu.edu.cn

电话: 021-34205320

地址: 上海市东川路 800 号上海交通大学电信群楼 2 号楼 529 室



附录: MassAnalyzer 软件输出的 excel 参数说明。

1、软件所使用原子质量

C: 12.0000000

H: 1.00782503207

O: 15.9949146

N: 14.003074

P: 30.973761998

S: 31.9720707

Cl: 34.968852682

Br: 78.9183376

F: 18.998403163

e: 0.00054858

C¹³: 13.0033548

H²: 2.0141018

O¹⁸: 17.9991603

S³⁴: 33.967867

2、输出参数所对应的公式

$$AI_{mod} = \frac{1 + C - 0.5 \cdot O - 0.5 \cdot H}{C - 0.5 \cdot O}$$

$$NOSC = 4 - \frac{(4 \cdot C + H - 2 \cdot O)}{C}$$

$$DBE = \frac{(1 + 2 \cdot C - H + N)}{2}$$

$$AI_{denominator} = C - O - S - N$$

$$AI = \frac{(1 + C - O - S - 0.5 \cdot H)}{AI_{denominator}}$$

$$Xc_{denominator} = DBE - (O + S)$$

$$Xc = \frac{(3 \cdot DBE - 2 \cdot (O + S))}{Xc_{denominator}}$$