

《环境空气 挥发性有机物来源解析 正定矩阵因子分
解（PMF）模型应用技术指南》编制说明
（征求意见稿）

《环境空气 挥发性有机物来源解析 正定矩阵因子分
解(PMF)模型应用技术指南》编制组

二〇二五年八月

目 录

目 录	I
1 臭氧污染防治背景	1
2 文件制订的必要性	1
3 臭氧及关键前体物溯源技术发展现状	2
4 指南制订的原则与思路	4
5 主要技术内容及确定依据	5
6 实际应用检验	12
7 参考文献	17



《环境空气 挥发性有机物来源解析 正定矩阵因子分解 (PMF) 模型计算技术指南》编制说明

1 臭氧污染防治背景

《大气污染防治行动计划》(2013-2017)和《打赢蓝天保卫战三年行动计划(2018-2020)》实施以来,全国、重点区域和城市细颗粒物($\text{PM}_{2.5}$)浓度快速下降,二氧化氮(NO_2)浓度缓慢下降,臭氧(O_3)浓度波动上升,目前各种污染浓度仍远高于区域背景。

从污染程度来看,区域 O_3 污染问题日益凸显, O_3 浓度呈缓慢上升态势、且在空气质量超标天中占比逐步增加,已成为影响优良天数的重要因素。2021年京津冀区域 O_3 浓度为160.2微克/立方米,相比2015年上升8%,较全国平均 O_3 浓度偏高23.5微克/立方米,明显高于美国(以2020年日最大8小时第90百分位计,约为123微克/立方米)和德国、法国等欧盟国家;因 O_3 浓度超标(国家二级文件限值为160微克/立方米)导致的污染天数达487天,已基本与 $\text{PM}_{2.5}$ 相当,成为区域空气质量持续改善的重要制约因素。

从污染空间分布来看, $\text{PM}_{2.5}$ 和 O_3 浓度高值区具有较好的空间一致性,但发生的季节存在差异;如京津冀、苏皖鲁豫地区, $\text{PM}_{2.5}$ 浓度高值出现在冬季,夏季在同一区域出现 O_3 浓度高值。研究结果表明, $\text{PM}_{2.5}$ 和 O_3 同根同源,冬季表现为 $\text{PM}_{2.5}$ 污染,夏季表现为臭氧污染。同时,相比 $\text{PM}_{2.5}$ 污染主要集中在京津冀、汾渭平原、东北等区域, O_3 污染在长三角、珠三角、川渝、长江中游等区域同样显著。

因此,总体而言,随着大气污染治理行动的实施,我国大气污染特征发生深刻变化。重点区域以硝酸盐、臭氧为代表的二次污染已成为制约空气质量改善的关键因素。二次污染由关键前体物高强度排放决定,由大气强氧化性驱动,以关键前体物减排为抓手,将是持续改善 $\text{PM}_{2.5}$ 污染、协同防控 O_3 污染的关键途径。为实现空气质量持续改善,亟需先进科技引领,在提升污染成因科学认知的基础上,形成前体物排放、污染溯源、预测和精准调控能力。

2 文件制订的必要性

臭氧污染受到气象、气候、传输等多方面因素影响,但氮氧化物(NO_x)和挥发性有机物(VOCs)两个关键前体物居高不下的排放是区域臭氧污染的主因。

我国 NO_x 和人为源VOCs排放量总体处于高位。根据清华大学MEIC清单分析,与2010年相比,我国 NO_x 排放量呈现先增加(2012年达到峰值2800万吨)后逐年下降趋势,但远小于 SO_2 降幅;我国人为源VOCs排放呈波动增加趋势,增幅在4~12%之间,目前仍未进入稳定的下降通道。在全国尺度上,VOCs人为源排放约为天然源排放的1~2倍。总体而言,我国VOCs和 NO_x 排放量均处于高位,需进一步加大VOCs和 NO_x 减排力度,将排放量减至百万吨级别,才能实现环境空气质量的根本好转。

近地表大气 O_3 形成机理较为复杂,但是对其宏观机理的认识基本明确。 NO_2 在光照的

条件下反应生成 O_3 ， O_3 又氧化 NO 生成 NO_2 和 O_2 ，稳定的“ $NO-NO_2-O_3$ ”反应循环不会导致环境 O_3 浓度的上升，但是大气中共存的 VOCs 可在 OH 自由基氧化作用下生成 VOCs 中间体 (RO_2 , HO_2)，其与 O_3 竞争氧化 NO ，进而使二次产物 O_3 净增。近地表 O_3 浓度与前体物 VOCs 和 NO_x 呈显著的非线性响应关系。根据 O_3 生成与其前体物 VOCs 和 NO_x 敏感性，可将 O_3 生成分为 NO_x 控制区、VOCs 控制区和过渡区。已有研究显示，我国多数城市地区属于 VOCs 控制区，乡村、郊区等多属于 NO_x 控制区或过渡区。如 2020 年夏季开展夏季臭氧强化监督帮扶期间，石家庄、济南、徐州长时间处于 VOCs 控制区；临沂、连云港、阜阳长时间则处于 VOCs 和 NO_x 协同控制区。

现阶段我国在宏观层面的 VOCs 管控重点为石化、化工、溶剂使用、包装印刷、机动车等污染源，但由于其涉及的行业/领域（国民经济行业分类）远远多于 NO_x ，导致不同区域或城市在研判 VOCs 管控重点时难度较大，已开展的大量污染防治工作未带来预期的成效。同时，由于 VOCs 涉及的化学组分繁多，高活性组分不仅是臭氧的主要来源，也是对二次 $PM_{2.5}$ 有重要贡献。因此，亟需制定出符合我国 VOCs 监测体系的溯源技术方法指南，为“精准、科学”开展 VOCs 污染防治奠定科学基础，提升 $PM_{2.5}$ 和臭氧污染协同防控能力。

3 臭氧及关键前体物溯源技术发展现状

3.1 国内外溯源方法

臭氧源解析技术 (Ozone Source Apportionment Technology, OSAT) 和地区臭氧评估技术 (Geographic Ozone Assessment Technology, GOAT) 是常用的两种臭氧源解析技术，其主要方法是在传统敏感性分析方法或模型的基础上将网格模型中加入额外的反应性示踪剂，以全程追踪特定源排放的污染物。虽然相对传统模型，OSAT 等模型的精度已有显著提高，但也有研究指出，尽管这些示踪物质表面上可以代表真实的源解析，但在实际模拟计算中，当面临非线性物种浓度求解计算时，仍然需要进行额外的假设。由于已有的空气质量模型法通常需要结合气象模式、清单处理模型等多级模型，运算量大，因此，其解析过程往往用时较长，时效性较差，无法满足现场监管的需求。

受体模型法是常用的 VOCs 溯源方法，如化学质量平衡模型 (CMB) 和因子分析类模型 (PMF、PCA、UNMIX 等)。同时，也有研究直接对臭氧前体物排放的非线性规律进行总结，通过模糊规则挖掘 (fuzzy rules generation)、神经网络 (neural network) 或类神经模糊 (neuro-fuzzy) 等方法，建立了关于二次污染的源解析函数，通过少量 EMEP 模型结果确定函数参数，得到源解析的显式方程。CMB 模型主要是根据质量守恒定理，通过排放源化学组分的含量和源贡献浓度乘积的线性相加以表达污染物的化学浓度；该模型需在已知排放源源谱和受体成分的基础上，通过示踪元素法、最小二乘法等算法，解出各类源的浓度（排放强度）和贡献百分比。PMF 的基本原理为 $X=GF+E$ ， X 标识大气 VOCs 中测得的化学物种的浓度， G 标识源对环境受体的贡献， F 标识污染源中化学物种的质量分数（特征谱矩阵）， E 为残差；通过利用权重计算出 VOCs 中各化学组分的误差，结合源标识物及受体成分，在

源谱未知的情况下，利用最小二乘法获得受体 VOCs 中的主要污染源及贡献。目前，国内研究者已针对北京、天津、郑州、武汉等多个城市的 VOCs 来源开展研究，通常借助空气质量模型或 PMF 等方法；部分研究者会使用多种方法进行综合对比，以提高溯源结果的可靠性。

总体而言，为实现 VOCs 溯源的目的，受体 VOCs 组分浓度和污染源 VOCs 成分谱为两个必不可少的要素和前提。

3.2 我国 VOCs 监测现状

我国约有近 130 个城市开展了 VOCs 常规监测并实现了数据上传、联网，主要监测的 VOCs 组分为 PAMS 中的 57 种物质，包括烷烃、烯烃、芳香烃和乙炔。同时，也有百余个城市参照《环境空气 醛、酮类化合物的测定 高效液相色谱法》（HJ683-2014）开展了醛酮类物质的业务化手工监测。持续开展的业务监测为掌握我国 VOCs 变化规律和趋势奠定了基础。

相较于监测部门开展的常规监测，也有研究者在 PAMS 和醛酮类物质的基础上对 TO-15 或 TO-14 中 60 余种物质进行同时检测。该检测方法需要同时使用 PAMS 和 TO-15/14 混合标准气体，通常利用气相色谱-质谱联用仪进行组分浓度分析。对于质谱峰无对应文件物质的情况，通常对比 NIST 数据库进行半定量分析。通过定量和半定量相结合的方式，有研究测出的组分达到 200 种以上。随着 VOCs 研究的深入，国内外已有研究团队针对 IVOCs、SVOCs、LVOCs 进行研究，但目前相关的样品采集、检测分析方法尚不成熟，未纳入常规管理体系。

由于 VOCs 中包含的组分千差万别，不同研究者在开展检测之前会制定专门的检测方案，以明确目标检出组分、浓度范围等关键信息。因此，不同研究者在围绕不同区域开展溯源研究时，使用的 VOCs 检测方法往往不尽相同。

3.3 国内外 VOCs 源成分谱研究现状

VOCs 源成分谱的研究高度依赖检测分析技术的进步，是分物种清单、空气质量模式、环境/健康影响以及污染来源解析的重要基础。

国外研究机构针对各类污染源开展的 VOCs 源成分谱测试和分析相对较早。国际上使用最广泛的 VOCs 源成分谱为美国环保署（USEPA）发布的 SPECIATE 数据库，该数据库由 USEPA 下属国家暴露研究实验室、国家风险管理研究实验室、空气质量计划和文件办公室、传输和空气质量办公室共同建立和维护。2002 年发布的 SPECIATE 3.2 是首个网络版源成分谱数据库，之后历经 9 次补充和修订，目前已更新至 SPECIATE 5.1。数据库集成了 USEPA 下属机构的测试分析结果和学术机构公开发表的相关科研成果，包含的成分谱信息由 2002 年的 2071 条增加至目前的 6746 条，其中，有机气体成分谱信息 2561 条，涵盖约 30 个行业/领域、204 个组分。此外，也有学者汇总整理了欧洲各国的研究成果，形成欧洲 VOCs 源成分谱库，但系统性和持续性相比 SPECIATE 略有不足。

SPECIATE 数据库是 USEPA 建立国家分物种和温室气体排放清单、开展光化学污染模拟、应用和验证 VOCs 源解析模型、评估有毒有害污染物排放及其风险中最常用的工具，为掌握大气污染物和温室气体排放特征、PM_{2.5} 和 O₃ 污染演变规律、人群和生态系统的暴露风险等奠定了基础，有力支撑了国家和各州污染管控政策的评估、优化和实施。SPECIATE 数据库同时也是国内外研究者开展相关研究最常用的数据库。尽管经过了多次补充和修订，SPECIATE 数据库中数据仍然存在不确定性。USEPA 针对 SPECIATE 的评估结果显示，在州层面，不同州的 VOCs 源成分谱丰富程度不同，仍有相当一部分污染源缺乏本地化成果，如原油储存、天然气生产、燃气汽车等；同时，建筑涂料、工业涂装、农药使用等污染源面临数据陈旧的问题，需重新开展测试分析。

相较于美国已经建立完善的 VOCs 源成分谱样品采集、测试分析、数据收集和发布共享的体系，我国尚未构建国家层面的 VOCs 源成分谱库，但近几年相关研究成果的不断积累为其打下了良好的基础。近年来，我国学者针对原油炼制、焦化、涂料生产、汽车制造、食品制造、燃油汽车、柴油机械等污染源开展了大量现场测试，获得了一批宝贵的基础数据。2021 年年初，暨南大学团队搜集整理已公开发布的我国人为源 VOCs 源成分谱，形成一套较为全面的 VOCs 源成分谱数据集，包含 363 条源成分谱、涵盖 101 个细分行业/领域和 447 个 VOCs 组分。通过与 SPECIATE 的对比发现，我国与美国的 VOCs 源成分谱差异显著，例如燃油汽车源中甲醛占比相差 10% 左右，工业源之间的差距更大。

通过对已有研究成果的梳理和对比，发现我国 VOCs 源成分谱研究和应用主要存在以下三方面问题：同研究成果之间差异大，无法有效支撑污染防治、覆盖污染源有限，部分高活性污染源需加强管控、缺乏更新机制，面临数据时效性问题。

4 指南制订的原则与思路

4.1 基本原则

（1）科学与支撑原则

本文件用以规范 VOCs 溯源过程中的主要流程，文件中规定的各项要求符合现阶段溯源技术方法的数学原理和运行方法，相关结果可支撑 VOCs 的精准防控。

（2）客观公正性原则

文件的制订客观真实体现了 VOCs 溯源技术方法的优缺点，在充分调研、掌握 VOCs 溯源方法和结果适用性的基础上，结合 VOCs 监测和标识组分最新研究成果提出规范化的溯源流程和源贡献计算方法，使溯源过程和输出结果客观、公正。

（3）体系协调性原则

本文件的制订充分考虑到与其他技术指南、污染防治政策相衔接，避免交叉重叠，污染源分类、主要源的常用标识组分等与 VOCs 监测分析方法标准相适用、配套，满足环境监督管理对标准的要求。

4.2 总体思路

基于对国内外 PMF 模型相关文件规范、PMF 相关研究进展、PMF 应用情况的文献和资料调研，总结已有的研究成果和实际应用经验，总结凝练 PMF 模型应用中存在的难点疑点问题以及关键技术参数；以问题为导向，开展专项研究和验证分析，得到优化后的技术参数；借助专家咨询等方法，编制形成《环境空气 挥发性有机物来源解析 正定矩阵因子分解(PMF)模型计算技术指南》（公开征求意见稿）。

5 主要技术内容及确定依据

5.1 结构框架

本文件正文部分包括前言、适用范围、规范性引用文件、术语和定义、模型原理和计算流程、数据准备、基础计算、旋转计算、VOCs 源贡献计算、结果合理性判断共十部分。文件还包括资料性附录 A。

5.2 适用范围

广泛的资料调研结果显示，PMF 模型在 VOCs 溯源研究中有大量应用，相对比较成熟。VOCs 包含的组分繁多，不同监测方法的目标检出物种差异较大。已有研究使用国家业务化 VOCs 在线监测系统（57 种组分）和手工监测（通常为 116 种组分）开展 VOCs 溯源研究；前者的优点在于数据量大、且连续性强，可结合 VOCs 组分浓度昼夜变化规律等信息研判污染源，缺点在于缺乏部分典型 VOCs 组分，如含氧 VOCs（OVOCs）、卤代烃等；后者的优点在于组分种类覆盖了常见的 VOCs 组分，可有效提高溯源结果的精准性，缺点在于成本较高。

因此，本文件规定适用于环境空气 VOCs 来源解析工作中应用 PMF 模型对环境空气 VOCs 手工采样和在线监测分析数据进行 VOCs 来源计算。

5.3 术语和定义

本文件共有术语 13 个。其中：

- （1）环境空气，引用了《环境空气质量标准》（GB 3095-2012）。
- （2）挥发性有机物，引用了《挥发性有机物无组织排放控制标准》（GB 37822-2019）。
- （3）环境受体，参考了《大气颗粒物来源解析技术指南（试行）》（环发[2013]92 号）。
- （4）正定矩阵因子分解模型、受体化学组分、标识组分、因子、不确定度、Q 值、残差、基础计算、旋转计算等术语参考了生态环境部发布的《颗粒物来源解析 正定矩阵因子分解模型计算技术指南》（HJ 1353-2024）、美国 EPA《EPA positive matrix factorization (PMF) 5.0 fundamentals and user guide》以及《Source apportionment of volatile organic compounds: Implications to reactivity, ozone formation, and secondary organic aerosol potential》、《Measurement and minutely-resolved source apportionment of ambient VOCs in a corridor city during 2019 China International Import Expo episode》等相关研究论文。

5.4 模型原理和计算流程

(1) 模型原理

该部分介绍了 PMF 模型的基本数学原理。PMF 的原理是将受体 VOCs 组分浓度矩阵 $X_{(n \times m)}$ 因子化，分解为两个因子矩阵， $F_{(p \times m)}$ 和 $G_{(n \times p)}$ ，以及一个“残差矩阵” $E_{(n \times m)}$ ，见公式 1。

PMF 模型限定矩阵 G 和 F 中的组分都是正值，即非负限制。PMF 模型解析上述矩阵的方法是通过定义一个“目标函数”（Object function） Q ，并使这个目标函数的值最小，见公式 2。当目标函数 Q 值最小时，模型将受体浓度矩阵 X 分解成源贡献矩阵 G 和因子谱矩阵 F 。

(2) 计算流程

该部分介绍了 PMF 模型工作和运行的基本流程。应用 PMF 模型进行 VOCs 来源解析的计算流程包括数据准备及导入，基础计算及结果分析、旋转计算及结果分析和结合当地主要排放源信息进行合理性判断等环节，见图 1。

数据准备即对数据进行预处理，主要包括对数据进行清洗，剔除异常值、未检出数据、空白值等操作，PMF 模型对输入系统的数据有明确的约束（数据非负），数据清洗后选择样品数据，计算组分的不确定度等；依据 VOCs 组分的分布特征，选择进入 PMF 模型的 VOCs 组分。

经过数据准备后，进行以下计算。首先构建化学特征矩阵：将收集到的 VOCs 浓度数据组成矩阵，其中每一行表示一个样本（时间和空间点），每一列表示一个 VOCs 化合物，矩阵元素表示对应位置上 VOCs 的浓度值。其次依据信噪比、物种浓度、示踪性等原则，对模型中的组分进行权重设置，随后进行反复尝试确定模型解析的因子个数，对得到的结果进行判断分析，评估基础计算过程，识别因子为具体的源类并得到相应的源贡献值，初步运算结果可依据 Q 值是否收敛、观测值和模拟值散点图、时间序列图拟合情况，对基础计算结果进行误差评估等。如果模型计算结果不好，可回到基础计算设置中优化参数或重新设置进入模型的主要 VOCs 组分的权重，重新计算直至得到相对满意的结果。

经过基础计算后，再进行旋转计算，进一步优化计算结果。旋转计算中设置旋转计算的参数，反复尝试。和基础计算类似，需要对得到的结果进行判断分析，评估旋转计算过程，识别因子为具体的源类并得到相应的源贡献值，对旋转计算结果进行误差评估等。如果模型计算结果不好，调整旋转计算和基础计算的参数或进行数据处理，重新计算直至得到相对满意的结果。对于计算的结果需要进行合理性判断，最终得到源解析结。

进行 PMF 模型时，需要根据 VOCs 组分浓度和不确定数据、基础运算结果和旋转运算结果，不断对 VOCs 浓度数据进行剔除和参数设定，以得到相对较优的计算结果。在导入

VOCs 组分浓度数据时，特征组分应优先选择环境浓度高、源指示性强和监测数据相对完整的组分进行 PMF 运算。



图 1 PMF 模型的主要计算流程

5.5 数据准备

(1) 数据种类和数量要求

根据 PMF 模型开展环境空气 VOCs 来源解析的需求，需要输入环境空气 VOCs 质量浓度和化学组分浓度数据，VOCs 化学组分基本要求需要包括烷烃、烯烃、芳香烃等。参考已公开发表文献以及颗粒物来源解析技术指南等文件，为保证源解析结果的可靠性，本文件规定了应用 PMF 模型进行 VOCs 源解析计算至少需要的数量：每个受体监测点位不少于 100 个有效样品的数据量，VOCs 化学组分至少需要 57 种组分（见下表），应包括主要组分和标识组分。常用的标识组分包括苯、甲苯、乙苯、二甲苯、苯乙烯、乙烯、乙烷、乙炔等。为解析更多的源类，可增加醇类、酯类、酮类、醚类、卤代烃等有机标识组分。增加输入的 VOCs 组分和样品数量，将使 PMF 模型计算结果更稳定、更具指向性。

表 1 PAMS 目标检出物质（57 种）

序号	成分	EPA（ppb）	CAS 号码
1	乙炔	40	74-86-2
2	苯	30	71-43-2
3	正丁烷	40	106-97-8
4	正丁烯	30	106-98-9
5	顺 2 丁烯	35	107-01-7
6	反 2 丁烯	25	107-01-7
7	异丙基苯	40	98-82-8

序号	成分	EPA (ppb)	CAS 号码
8	环己烷	40	110-82-7
9	环戊烷	20	287-92-3
10	癸烷	30	124-18-5
11	1,3-二乙基苯	40	141-93-5
12	1,4-二乙基苯	25	105-05-5
13	2,2-二甲基丁烷	40	75-83-2
14	2,3-二甲基丁烷	50	79-29-8
15	2,3-二甲基戊烷	50	565-59-3
16	2,4-二甲基戊烷	40	108-08-7
17	十二烷	40	112-40-3
18	乙烷	25	74-84-0
19	乙基苯	25	100-41-4
20	乙烯	20	74-85-1
21	间-甲乙苯	25	620-14-4
22	邻-甲乙苯	30	611-14-3
23	对-甲乙苯	40	622-96-8
24	庚烷	25	142-82-5
25	正己烷	30	110-54-3
26	1-己烯	60	592-41-6
27	异丁烷	25	75-28-5
28	异戊烷	40	78-78-4
29	异戊二烯	40	78-79-5
30	2-甲基庚烷	25	592-27-8
31	3-甲基庚烷	25	589-81-1
32	2-甲基己烷	25	591-76-4
33	3-甲基己烷	25	589-34-4
34	2-甲基戊烷	20	107-83-5
35	3-甲基戊烷	40	96-14-0
36	甲基环己烷	30	108-87-2
37	甲基环戊烷	25	96-37-7
38	正壬烷	25	111-84-2
39	正辛烷	30	111-65-9
40	正戊烷	25	109-66-0
41	1-戊烯	25	109-67-1
42	顺-2-戊烯	35	627-20-3
43	反-2-戊烯	25	646-04-8
44	丙烷	40	74-98-6
45	丙基苯	30	103-65-1
46	丙烯	25	115-07-1
47	苯乙烯	40	100-42-5
48	联三甲苯	25	526-73-8
49	1,2,4-三甲苯	40	95-63-6
50	1,3,5-三甲苯	25	108-67-8
51	2,2,4-三甲基戊烷	30	540-84-1
52	2,3,4-三甲基戊烷	25	565-75-3
53	甲苯	40	108-88-3
54	十一烷	30	1120-21-4

序号	成分	EPA (ppb)	CAS 号码
55	间二甲苯	40	108-38-3
56	邻二甲苯	25	95-47-6
57	对二甲苯	40	106-42-3

(2) 数据质量要求

用于 VOCs 源解析工作的质量浓度数据，应采用国家、行业或地方方法文件获得；没有国家、行业或地方方法标准，可采用国际标准、国外标准或研究建立的方法，经确认后使用，同时监测需按照 HJ 630 的要求纳入质量管理体系，开展质量保证和质量控制工作并符合相关要求。VOCs 组分的分析需参照相关国家文件方法 HJ 683、HJ 750、HJ 1010 等。

通常需要对分析后的数据进行数据清洗与质控。主要包括以下几个方面：

1) 异常离群值的检查和处理：通过时间序列图等方式，对偶发性的异常高或低浓度数据进行检查，并结合经验判断数据详细情况，做好剔除或者保留。分析发现突高值与正常值差异较大，易导致结果偏差，故在本次数据筛选过程中做剔除处理；

2) 0 值的检查和处理：模型输入不允许出现任何 0 值，如果有 0 值存在，模型不可运行。对 0 值数据较少，以均值代替；如 0 值数据过多，则降低该因子进入模型以后的权重；

3) 低于检出限值的检查和处理：若该因子检出率太低，需要考虑剔除处理；另外，根据模型指导手册开展处理，浓度选用检出限的一半，不确定度选用 5/6 检出限；

4) 高于检出限的数据处理：浓度不需要处理，但根据指导手册不确定度计算需要确定 EF 值，结合测试平行样确定 EF 值；

5) 缺失数据的处理：PMF 模型数据输入时，不允许数据有空白值存在，在剔除异常值、未检出数据及空白值后，对空白单元格以“-999”形式补全，可在模型运转时以物种均值代替“-999”。

(3) 不确定度计算

不确定度是 PMF 模型中的关键输入数据，直接影响 PMF 模型计算结果。受体模型的不确定度主要来自于大气 VOCs 采集和化学成份测量的不确定性、源成份谱的共线性（即不同排放源可能有相似的源成分谱）以及对二次来源正确判定等问题。不确定度是在 PMF 模型计算过程中对组分的浓度矩阵进行加权计算。不同研究者在 PMF 模型计算过程中可能会采用不同不确定度的计算方法。本文件主要参考美国 EPA 《EPA positive matrix factorization (PMF) 5.0 fundamentals and user guide》，统一规定了不确定度的计算方法如下：

$$Unc = \frac{5}{6} \times MDL \quad (3)$$

$$Unc = \sqrt{(EF \times Conc)^2 + (0.5 \times MDL)^2} \quad (4)$$

式中：Unc——VOCs 组分的不确定度， $\mu\text{g}/\text{m}^3$ ；

MDL——VOCs 组分的检出限, $\mu\text{g}/\text{m}^3$;

Conc——VOCs 组分的质量浓度, $\mu\text{g}/\text{m}^3$;

EF (Error fraction) ——VOCs 组分的误差系数。

当 VOCs 组分的质量浓度小于等于 MDL 时, 采用公式 (3) 计算; 当 VOCs 组分的质量浓度大于 MDL 时, 采用公式 (4) 计算。误差系数的大小范围一般为 0.1~0.6。对于 VOCs 质量浓度或者组分质量浓度不稳定或者接近检出限, 采用较大的误差系数。当 VOCs 组分质量浓度数据缺失较多时, 可设定较大的误差系数。

5.6 基础计算

数据的模型检验。将数据导入到 PMF 模型, 通过 PMF 模型的信噪比、浓度散点图、时间序列图等功能, 分别查看不同组分之间关系、组分随时间变化情况。对于异常高或者低的组分数据, 经数据重复核实后属监测异常或其他原因的, 可予以剔除。同时对数据量进行核实。进入模型的 VOCs 组分数据, 物种不少于 30 种, 一般为 30-50 之间, 数据量要求达到 200 组以上, 由此保证模型结果的可靠性。

数据计算权重的选择。按照 PMF 模型计算要求, 需要对输入数据进行检验设定计算权重。PMF 模型中各 VOCs 组分均将会被赋予一个计算权重, 类型包括 Bad (差)、Weak (弱) 和 Strong (强) 三个水平。对于权重为 Strong 的组分, 按输入的浓度和不确定度参与计算; 对于权重为 Weak 的组分, 模型在计算时会自动增加该组分的不确定度; 对于权重为 bad 的组分, 则不参与计算。关于数据计算权重的设置要求, 在应用 PMF 模型进行 VOCs 源解析时, 需要使用示踪组分来识别污染源类, 因此建议将与 VOCs 来源相关的标识组分设定为高权重; 其他 VOCs 组分作为辅助分析, 可以设置成 weak。当某一组分的缺失数据较多时, 还可通过采用设定较大的不确定度来降低该组分的计算权重。由于不同研究工作中测试的 VOCs 组分差异较大, 后文中将对主要目标检出物种及主要示踪组分进行举例分析。

因子个数的确定。在经过初步检验之后, 可对数据进行模型计算。PMF 模型计算包括基础计算(Base Run)和旋转计算(Rotational Run), 其中基础计算为必须的计算操作, 根据计算结果, 选择开展旋转计算。在基础计算中, 需要设定因子(Factor)个数、运行次数、是否随机选择种子(Sed)等参数。基础计算中, 因子个数设定是一个重要步骤。较多的因子个数会导致无意义的因子分布, 而较少的因子个数会使混合源难以分离。因子个数需要基于对 VOCs 可能来源、样品数量、采集时间等基本认识以及模型反复计算来综合确定, 需要多次反复尝试。运行次数和种子等参数可选择模型默认值。建议从 4 个源类开始尝试, 有助于后续调整源类更快速。

在基础计算中, 首先需要根据城市基本情况 (比如根据当地排放源清单、污染源普查数据、工业布局、能源结构、机动车保有量等信息), 初步确定因子个数的范围。通过设置不同的因子个数进行尝试性计算, 验证和确定适合的因子个数。总结 PMF 模型计算的经验和

文献调研结果，提供了常用的 4 种数学验证方法供参考使用，具体如下：

(1) 根据 Q 值随因子个数变化关系判断。在设定合适的不确定度后，可逐渐增加因子数，观察 Q 值的变化。当因子数逐渐增加，而对应结果的 Q 值没有明显变化，则此时因子数增加为合理。当因子数增加到某一特定值（例如 P），Q 值开始有明显变化，则 p-1 则可能为适合的因子数。

(2) 根据组分的残差大小判断。组分的残差（尤其是标识组分）应在正负 3 以内，如果残差有较大的分布，则应重新检查因子个数。

(3) 根据多次计算结果的差异情况判断。多次运行不应该得到多个 Q 值或因子谱差异性比较大的结果(如 20 个结果中，有多个结果之间差异较大)，如果有，则说明因子数的设定可能不合理。

(4) 根据回归检验结果判断。将 G 矩阵与 VOCs 浓度数据进行回归检验。如果回归系数出现负值，说明因子个数设置可能过多；F 矩阵的每一个因子除以相应的回归系数后，如果加和大于 12，则说明选择的因子可能过少。

模型的数学验证是有一定的局限性。通过数学验证确定，仅代表所输入的数据计算得到的结果从数学意义上是合适的，不代表解析结果完全符合当地实际情况。因此还需要通过物理意义（如当地的实际情况）情况进一步辅助判断因子数的选择，比如根据当地排放源清单、污染源普查数据、工业布局、能源结构、机动车保有量等信息，最终确定合适的因子个数。

5.7 旋转计算

当模型基础计算得到的因子难以识别为实际源类的时候（例如解析出来不属于这个源的物种占比很高），就可使用旋转计算（Rotational run），使因子谱中的标识组分更加突出或相对较平均，促进因子识别为具体污染源类。旋转计算包括 F 矩阵峰值模型计算（Fpeak model run）和约束模型计算（Constrained model run）。其中 F 矩阵峰值模型计算是常用的旋转计算方法。在 F 矩阵峰值模型计算中，需要尝试不同的 Fpeak 参数。当 Fpeak 参数设为正值时，F 矩阵中的化学组分在某些因子中占比会更加突出；Fpeak 参数设为负值时，F 矩阵中的化学组分在各因子中的占比则相对较平均。依次通过选择“Rotational Tools”、“FpeakRotation & Notes”和“Model Runs”，在 Selected Fpeak Runs 处一般复选 0.5、-0.5、1.0、-1.0，按 Run 进行旋转运算。看到 Fpeak 运算结束后，同样出现了因子/贡献图（Profiles/Contribution）、因子指纹图（Factor Fingerprints）、G 空间图（G-Spaceplot）、因子贡献图（Factor contribution）、分析总结（Diagnostics），分别点击查看 Fpeak 之后的输出结果，并与前面的比较选择最佳结果。查看 Q 值在不同 Fpeak 参数下的变化情况，要求 Q 值的变化在 5% 以内，并且选择 Q 值拐点之前的 Fpeak 值对应的解析结果。

5.8 VOCs 源贡献计算

利用 PMF 模型计算得到的 G 矩阵（源贡献矩阵）与 VOCs 质量浓度做多元线性回归分析，计算出 G 矩阵各列（源）的回归系数。将各列的回归系数乘以各列，即为各污染源对

VOCs 的贡献，回归公式如下：

$$VOCs_j = \sum_{k=1}^p s_k g_{jk} \quad (5)$$

式中： $VOCs_j$ ——第 j 个样品的 VOCs 质量浓度；

s_k ——回归系数；

g_{jk} ——PMF 模型计算的 G 矩阵中第 k 源对第 j 个样品的归一化贡献；

p ——源个数；

$s_k g_{jk}$ ——计算得到的 k 源对第 j 个样品的贡献。

在城市环境空气 VOCs 来源解析工作中，为了表征城市不同区域的污染特征、代表城市 VOCs 污染来源的整体情况，往往会在城市典型区域设立多个监测点位，得到相应的 VOCs 组分浓度数据。结合 VOCs 反应性较强、不同点位差异较大的特性，提供针对多个点位的 PMF 源解析计算方法：将多个点位的源解析结果进行平均，则得到城市的解析结果。

5.9 结果合理性判断

源解析结果既要符合模型计算要求，还需要符合实际情况。

符合模型计算要求是模型模拟计算出的各 VOCs 组分浓度与实测测试结果接近。可通过直接比较 VOCs 组分模拟值与实测值、分析主要化学组分或标识组分的残差、对比 Q 计算值与 Q 理论值等方法，判断模型模拟结果的好坏。如果模拟值与实测值越接近、化学组分残差越小、Q 计算值与 Q 理论值越接近，则表明模拟计算结果更合理。

实际情况要求是经过识别的源谱特征及其源贡献合理。可通过 PMF 模型计算得到的因子谱与实测源成分谱相关性比较、源贡献与当地主要排放源信息比较分析等方法来评估模型结算结果的合理性。当地主要排放源信息主要来自大气污染物排放清单、环境统计等。

5.10 附录

附录 A 作为资料性附录提供了 VOCs 源类识别方法和主要源的标识组分。共提供了 21 个 VOCs 主要排放源的常用标识组分，该内容是总结近年来我国本土化 VOCs 源成分谱研究的综合成果而来，源成分谱的样品采集和测试分析区域涉及北京、河南、广东、福建、辽宁等省份，能较好的代表我国 VOCs 重点行业的实际情况。所提供的 21 个污染源与常规污染源分类有所不同，但与我国近年来 VOCs 相关的污染防治政策中分类方法基本保持一致，能较好的对应国家和地方重点管控对象。

6 实际应用检验

6.1 应用过程

为了验证本文件的源解析过程的科学性、结果的可信度，中国环境科学研究院的研究人

员开展方法验证工作。验证方法是参照本文件开展专项的环境受体 VOCs 采样，目标检出组分为 116 种（烷烃、烯烃、炔烃、芳香烃、卤代烃、OVOCs，见表 2），以此为基础进行 PMF 模型计算，获得源解析计算结果并与污染源信息进行比对分析。

编制组开展测试的地点为京津冀及周边地区“2+26”城市中某重点城市的工业园区，于 2023 年夏季，使用内表面钝化处理的苏码罐采集环境受体 VOCs 样品 126 个，采用 GC-MSD/FID 对 VOCs 样品进行定性和定量分析。该城市的夏季首要污染物为臭氧，秋冬季首要污染物为 PM_{2.5}，污染源包括工业源、机动车、柴油车/机、固定燃烧源、生物质燃烧、餐饮等。

表 2 环境受体 VOCs 检出组分

编号	名称	编号	名称
1	丙烷	59	反式-1,2-二氯乙烯
2	异丁烷	60	1,1-二氯乙烷
3	1-丁烯	61	甲基叔丁基醚
4	乙烯	62	乙酸乙烯酯
5	丁烷	63	甲基乙基酮
6	乙炔	64	顺式-1,2-二氯乙烯
7	反式-2-丁烯	65	乙酸乙酯
8	乙烷	66	己烷
9	顺式-2-丁烯	67	氯仿
10	异戊烷	68	四氢呋喃
11	正戊烯	69	1,2-二氯乙烷
12	戊烷	70	1,1,1-三氯乙烷
13	异戊二烯	71	苯
14	反式-2-戊烯	72	四氯化碳
15	顺式-2-戊烯	73	环己烷
16	2,2-二甲基丁烷	74	1,2-二氯丙烷
17	环戊烷	75	一溴二氯甲烷
18	2,3-二甲基丁烷	76	1,4-二氧杂环己烷，二氧六环
19	2-甲基戊烷	77	三氯乙烯
20	3-甲基戊烷	78	甲基丙烯酸甲酯
21	1-己烯	79	庚烷
22	甲基环戊烷	80	反式-1,3-二氯-1-丙烯
23	2,4-二甲基戊烷	81	甲基异丁基酮
24	2-甲基己烷	82	顺式-1,3-二氯-1-丙烯

25	2,3-二甲基戊烷	83	1,1,2-三氯乙烷
26	3-甲基己烷	84	甲苯
27	2,2,4-三甲基戊烷	85	甲基丁基酮
28	甲基环己烷	86	二溴一氯甲烷
29	2,3,4-三甲基戊烷	87	1,2-二溴乙烷
30	2-甲基庚烷	88	四氯乙烯
31	3-甲基庚烷	89	氯苯
32	辛烷	90	乙苯
33	壬烷	91	间对二甲苯
34	异丙基苯	92	溴仿
35	正丙苯	93	苯乙烯
36	间乙基甲苯	94	四氯乙烷
37	邻乙基甲苯	95	邻二甲苯
38	癸烷	96	4-乙基甲苯
39	1,2,3-三甲基苯	97	1,3,5-三甲苯
40	间二乙基苯	98	1,2,4-三甲苯
41	对二乙基苯	99	苄基氯
42	十一烷	100	1,3-二氯苯
43	十二烷	101	1,4-二氯苯
44	丙烯	102	1,2-氯苯
45	氟利昂-12	103	1,2,4-三氯苯
46	氯甲烷	104	萘
47	氟利昂-114	105	六氯-1,3-丁二烯
48	氯乙烯	106	甲醛
49	1,3-丁二烯	107	乙醛
50	溴甲烷	108	丙烯醛
51	氯乙烷	109	丙醛
52	丙酮	110	甲基丙烯醛
53	氟利昂-11	111	正丁醛
54	异丙醇	112	丁烯醛
55	1,1-二氯乙烯	113	戊醛
56	二氯甲烷	114	己醛
57	二硫化碳	115	苯甲醛
58	氟利昂 113	116	间甲基苯甲醛

6.2 应用结果

结合环境受体 VOCs 组分分布特征，共选择 30 种组分作为标识组分（表 3），共解析出主要源类 7 种，结果如下图所示。因子 1 中乙炔、苯、苯乙烯及氯甲烷等浓度及物种贡献较高，乙炔以苯主要为燃烧源的示踪物，因此识别该源类为燃烧源。因子 2 中异丙醇、二氯甲烷及氯仿浓度较高且贡献较大，同时伴有一定比例的烷烃贡献，其中异丙醇为有机化工的主要标识物，同时二氯甲烷及氯仿主要应用领域为工业，因此判定该源类为有机化工源。因子 3 由于异戊二烯贡献率占绝对优势，异戊二烯主要来源于植物源，还有部分来自机动车排放、工业源等来源，但该源类仅有异戊二烯占比较高，因此判定该源类为天然源。因子 4 中乙苯、间/对二甲苯及邻二甲苯贡献及浓度比例较高，主要作为涂料、油漆、粘合剂和清洗剂的典型污染物，因此识别该源类为工业涂装。因子 5 主要组分为异丁烷、丁烷、异戊烷、十一烷、十二烷、癸烷等浓度较高，异戊烷正丁烷是液化石油气挥发的典型示踪物，而短链烷烃在汽油车尾气中也有排放，识别因子 5 为汽油车。因子 6 中二硫化碳贡献及浓度均最高，且含有一定浓度的甲苯、氯仿及氯甲烷，二硫化碳为工业的示踪物，二硫化碳为典型的工业示踪物，且烷烃含量较低，因此识别该源类与因子 2 共同属于有机化工源。因子 7 中乙酸乙酯贡献较高，同时甲苯、2-甲基戊烷、3-甲基戊烷、环己烷含量较高，综合判定为低含量 VOCs 包装印刷。

表 3 标识组分

序号	名称	序号	名称	序号	名称
1	乙烷	11	癸烷	21	邻二甲苯
2	丙烷	12	乙烯	22	二氯甲烷
3	异丁烷	13	丙烯	23	1,2-二氯丙烷
4	丁烷	14	异戊二烯	24	乙醛
5	异戊烷	15	乙炔	25	丙酮
6	2-甲基戊烷	16	苯	26	异丙醇
7	3-甲基戊烷	17	甲苯	27	二硫化碳
8	十一烷	18	乙苯	28	丙烯醛
9	十二烷	19	间/对二甲苯	29	乙酸乙酯
10	环己烷	20	苯乙烯	30	甲醛

VOCs 源贡献计算结果表明：该工业园区的 VOCs 首要来源为汽油车，占比为 29%；其次有机化工，占比为 23%；燃烧源占比为 15%；包装印刷-低 VOCs 源类占比为 13%；天然源占比为 11%；工业涂装占比为 9%。

6.3 校验分析

使用该城市大气污染物排放清单对源解析结果进行校验，表 4 为 VOCs 排放贡献排序。综合对比可知，有机化工和机动车均排名前两位，燃烧源次之。包装印刷-低 VOCs 在源解析中排名较高的原因可能为：环境受体 VOCs 采样点位临近区县分布有较多纸版印刷等企业，此类企业主要使用水性油墨、植物油基油墨，属于低 VOCs 原辅材料，主要排放乙酸乙酯、异丙醇、异戊烷等；因而从排放量来看，其总排放量较小，但由于毗邻环境受体采样点位，因此在源解析结果中排名较为靠前。

表 4 不同污染源 VOCs 排放贡献

贡献排序	二级源分类	VOCs 排放贡献
1	有机化工	24.49%
2	道路移动源	21.64%
3	工业生物质锅炉	7.77%
4	非道路移动源	5.09%
5	固定燃烧源	4.74%
6	农药使用	4.49%
7	建筑涂料	2.44%
8	固废处理	1.55%
9	餐饮	1.12%
10	其他溶剂	0.85%
11	加油站	0.80%
12	工业涂装	0.68%
13	油气储存	0.15%
14	油气储运	0.12%
15	印刷印染	0.11%
16	其他工业	23.96%

校验结果表明，基于手工监测和推荐使用的标识物种开展的 PMF 模拟分析，所得源解析结果与污染排放分配基本一致。因此，参照本文件开展 VOCs 来源解析可得出稳定、有效的源贡献结果。

7 参考文献

- [1] 《环境空气质量标准》（GB 3095-2012）；
- [2] 《挥发性有机物无组织排放控制标准》（GB 37822-2019）；
- [3] 《大气颗粒物来源解析技术指南（试行）》（环发[2013]92 号）；
- [4] 《颗粒物来源解析 正定矩阵因子分解模型计算技术指南》（HJ 1353-2024）；
- [5] 《EPA positive matrix factorization (PMF) 5.0 fundamentals and user guide》；

- [6] 《Source apportionment of volatile organic compounds: Implications to reactivity, ozone formation, and secondary organic aerosol potential》 ;
- [7] 《Measurement and minutely-resolved source apportionment of ambient VOCs in a corridor city during 2019 China International Import Expo episode》 ;
- [8] Petros Koutrakis, Spyros N. Pandis, Eleni Samara *et al.*, Positive matrix factorization of PM_{2.5}: comparison and implications of using different speciation data sets. *Environmental Science & Technology*, 2012 7024-7032, DOI 为 10.1021/es302358g.

