

附件3

《环境空气 颗粒物来源解析 正定矩阵因子分解（PMF）
模型计算技术指南（征求意见稿）》
编制说明

《环境空气 颗粒物来源解析 正定矩阵因子分解（PMF）
模型计算技术指南》标准编制组

二〇二二年十月

目 录

1 项目背景.....	1
1.1 任务来源.....	1
1.2 工作过程.....	1
2 标准制订的必要性分析.....	2
3 国内外 PMF 模型计算相关进展.....	3
3.1 国内外相关研究进展.....	3
3.2 国内外相关标准方法.....	8
4 标准制订的基本原则和技术路线.....	8
4.1 标准制订的基本原则.....	8
4.2 标准制订的技术路线.....	9
5 方法研究报告.....	10
5.1 适用范围.....	10
5.2 规范性引用文件.....	10
5.3 术语和定义.....	10
5.4 模型原理和计算流程.....	11
5.5 数据准备.....	12
5.6 基础计算.....	15
5.7 旋转计算.....	18
5.8 颗粒物源贡献结果计算.....	19
5.9 结果合理性判断.....	20
5.10 附录.....	20
6 典型 PMF 模型软件操作步骤.....	21
6.1 基本情况.....	21
6.2 WALSPMF 1.0 计算步骤.....	21
6.3 EPA PMF 5.0 计算步骤.....	27
7 实际应用检验.....	38
7.1 应用过程.....	38
7.2 应用结果.....	38
8 参考文献.....	39

《环境空气 颗粒物来源解析 正定矩阵因子分解（PMF） 模型计算技术指南（征求意见稿）》编制说明

1 项目背景

1.1 任务来源

2021年，生态环境部办公厅印发《关于开展2021年度国家生态环境标准项目实施工作的通知》（环办法规函〔2021〕312号），下达了编制《环境空气 颗粒物来源解析 正定矩阵因子分解（PMF）模型计算技术指南》的任务。项目统一编号：2021-78。

1.2 工作过程

1.2.1 前期工作

2019年起，中国环境监测总站承担了生态环境部大气环境管理司的专项工作“大气环境质量管理”项目，开展正定矩阵因子分解（PMF）模型计算研究，编写《环境空气颗粒物来源解析正定矩阵因子分解（PMF）模型计算技术指南（试行）》。

2019年1月~4月，成立《环境空气颗粒物来源解析正定矩阵因子分解（PMF）模型计算技术指南（试行）》编制组，由中国环境监测总站、南开大学相关技术人员组成；编制组调研正定矩阵因子分解（PMF）模型的研究进展、各地开展PMF模型源解析工作情况，总结已有的研究成果，编写《环境空气颗粒物来源解析正定矩阵因子分解（PMF）模型计算技术指南（试行）》初稿。

2019年4月25日，编制组在北京召开“环境空气颗粒物源解析（模型计算和移动源采样）技术研讨会”，生态环境部大气环境管理司相关负责人参会，邀请来自南开大学、暨南大学、中国环境科学研究院、西安市环境监测站、宝钢环境监测站、济南市环境监测站等专家进行技术研讨。与会者对项目研究内容、技术路线、技术指南初稿等进行研讨，提出修改意见以及下一步工作要求。

2019年5月~9月，编制组对初稿进行补充完善，形成指南的文本及编制说明。2019年10月~11月，编制组开展实际应用检验，根据应用检验结果对文本及编制说明进一步修改。

2020年6月8日，编制组在北京召开“环境空气颗粒物源解析技术研讨会”，生态环境部大气环境管理司相关负责人参会，邀请来自北京大学、中国环境科学研究院、北京市环境保护监测中心、武汉市环境监测中心、西安市环境监测站等专家对文本及编制说明进行论证。编制组根据专家意见进行修改，形成文本及编制说明。

1.2.2 成立标准编制小组

2021年，接到标准制修订任务后，成立《环境空气颗粒物来源解析正定矩阵因子分解（PMF）模型计算技术指南》编制组，由中国环境监测总站、南开大学、中国环境科学研究院相关技术人员组成。

1.2.3 编写标准征求意见稿和编制说明

根据标准制修订的内容和格式要求，编制组对前期的文本及编制说明进行补充和修改，形成标准征求意见稿和编制说明。

1.2.4 标准征求意见稿审查

2021年8月9日，中国环境监测总站召开标准征求意见稿内部审查会，邀请北京大学、中国环境科学研究院、上海市环境监测中心、北京市环境监测中心等专家进行审查。专家组通过标准征求意见稿的内部审查，形成如下审查意见：（1）该标准对环境空气颗粒物来源解析模型计算具有指导意义，编制组提交的材料完整，标准编制内容科学、全面、可操作性强；（2）建议进一步完善适用范围、术语、定义、软件版本引用、附录等相应内容；（3）建议加强 PMF 源解析共线性问题的处理。编制组根据专家意见，进一步调研文献资料，对标准征求意见稿和编制说明进行修改完善。

2021年12月3日，生态环境部大气环境管理司相关负责人组织召开标准征求意见稿及开题报告材料审查会，邀请来自北京大学、中国环境科学研究院、中国科学院大气物理研究所等单位的9位专家进行审查。专家委员会一致通过该标准的开题论证和征求意见稿技术审查，对该标准的征求意见稿，形成如下技术审查意见：（1）标准主编单位提供的材料齐全、内容完整、格式规范，制订的标准具有科学性、适用性和可操作性；（2）建议与《环境空气 颗粒物来源解析 受体模型法监测数据处理与检验技术规范》同步发布，互相引用和借鉴，与《环境空气 颗粒物来源解析 化学质量平衡（CMB）模型计算技术指南》统一相关术语定义；（3）建议明确“5 数据准备”中数据种类和数量要求；修改附录 B 颗粒物源的标识组分等；（4）建议对文本进行规范性修改和完善。编制组根据专家意见进行修改完善，形成标准征求意见稿及编制说明。

2 标准制订的必要性分析

环境空气颗粒物来源解析工作（简称“源解析工作”）是环境空气颗粒物污染防治的基础和前提。2014年，原环保部办公厅《关于开展第一阶段大气颗粒物来源解析研究工作的通知》，全国直辖市、省会城市（除拉萨外）、计划单列市需要开展环境空气颗粒物来源解析研究工作。自2014年起，各城市相继开展颗粒物源解析工作，至2016年全部完成。由于环境空气颗粒物污染防治的长期性，以及颗粒物来源的变化，各地有必要继续开展源解析工作，为颗粒物污染防治提供支撑。如，根据《京津冀及周边地区2017年大气污染防治工作方案》，京津冀及周边地区“2+26”重点城市需开展环境空气颗粒物来源解析研究工作。2018年6月27日，《打赢蓝天保卫战三年行动计划》（国发〔2018〕22号）要求常态化开展重点区域和城市源排放清单编制、源解析等工作，形成污染动态溯源的基础能力。常态化开展源解析工作已成为颗粒物污染防治工作中的重点工作之一。然而，根据各城市颗粒物源解析工作的开展情况，源解析工作尚处于业务化研究阶段，离业务化运行还有一定的差距。

目前亟需建立一套源解析监测技术方法体系，为科学规范地开展源解析研究工作提供重要支持，推动源解析工作的业务化。该技术方法体系包括固定源、移动源和开放源等主

要源类采样技术、环境空气颗粒物采样技术、颗粒物及组分（无机元素、水溶性离子和 EC/OC 等）分析方法、源解析模型计算方法等。关于污染源采样方法、颗粒物化学组分分析方法，陆续出台标准方法，相对较为完善。源解析模型计算是源解析工作的重要环节，其科学性和规范性直接影响源解析结果的精准性和可比性。目前关于源解析模型计算方法，虽然自 2014 年大范围开展源解析工作以来，各地环保部门持续投入人力物力至源解析工作中，但是由于缺乏相关的标准方法，各地环保部门还是不能掌握模型操作方法。而各高校研究所对源解析模型研究和认识不同，存在采用的计算方法不统一、参数选择存在差异等问题，降低了不同城市源解析结果的科学性和可比性。

目前基于受体模型的源解析技术是各地开展源解析工作的主要技术手段，得到了广泛应用。受体模型源解析中，主要采用的模型有正定矩阵因子分解模型（PMF）、化学质量平衡模型（CMB）等。PMF 模型由于不需要输入详细的源成分谱，且能简单快速地获得高时间分辨率的源解析结果，被广泛地应用于各地解析研究工作。目前国内还没有 PMF 模型计算相关的标准方法。因而，有必要编制 PMF 模型计算技术指南，规范源解析计算过程，提高源解析工作的科学性和可比性。

3 国内外 PMF 模型计算相关进展

3.1 国内外相关研究进展

3.1.1 PMF 模型发展历史和原理

应用于环境空气颗粒物来源解析工作的受体模型包括 PMF、CMB、UNMIX、PCA 等，可分成两大类：（1）源未知类受体模型，包括 PMF、UNMIX、PCA 等，在模型计算不需要输入源类信息，（2）源已知类受体模型，包括 CMB 等，在模型计算需要输入详细的源类信息，比如颗粒物化学组分信息。PMF、CMB、UNMIX 是美国 EPA 推荐用于源解析的受体模型，应用比较广泛。在我国源解析工作中，应用最广泛的受体模型为 PMF 和 CMB。

PMF(Positive Matrix Factorization)模型即正定矩阵因子分解法，它是在因子分析的基础上发展起来的。1985 年，Thurston 和 Spengler^[1]首次提出绝对因子（APCS）的概念，将 PCA 模型与多元线性回归相结合而形成 PCA/MLR 模型，能够定量计算出源类对受体的贡献值。该模型通过 PCA 得到因子载荷和得分矩阵定性判断和分析污染源的数量和类型，然后根据 APCS 使用多元线性回归计算源贡献值。PCA/MLR 模型自提出以后得到了广泛的应用^[2-4]。然而，在应用过程中经常会遇到估算的源成分谱和贡献值出现负值，计算过程中无法估算不确定度问题，会产生有歧义的结果的问题。针对这些问题，在因子分析的基础上，Paatero 和 Tapper^[5]于 1994 年发表论文首先提出 PMF 模型的原理，接着于 1997 年提出 PMF 模型基于最小二乘法的运算方法^[6]。此后，Paatero 和 Hopke 不断对 PMF 模型进行改进，使得 PMF 模型在国内外大量研究中得以应用^[7,8]。PMF 模型可实现对因子谱和源贡献矩阵加以非负处理，并且在输入受体数据过程中加入受体浓度的标准偏差对受体数据进行不确定度估计，使得计算的结果更加可靠。

PMF 模型不需要输入详细的源成分谱，基于大量的受体观测值，根据因子的标识组分提取出不同源类的因子，识别出源类并解析各源类的贡献值。PMF 模型和其他源未知受体模型（UNMIX、PCA）的特点见下表所示。

表1 PMF、UNMIX、PCA 比较

名称	定量/定性	优势	局限性
正定矩阵因子分析法 (PMF)	定量	添加不确定性数据对数据进行逐点分析；不需要详细的源成分谱信息，可较好处理缺失及不精确的数据；对因子分解矩阵进行非负约束，使得到的源成分谱和源贡献率不会出现负值，结果更加合理	没有直观的因子数显示；需要设置多个因子数进行多次运算以获得目标矩阵 Q 的最小值
UNMIX 模型	定量	源贡献结果均为正值；直接对样本浓度数据进行计算，无需对数据进行复杂的转换；可以直观显示解析结果的好坏	需要数据质量较高（高于检出限）；缺失值较少（≤ 75%）的数据；需要的样本数据量较大
主成分分析法 (PCA)	定性	源定性识别结果相对可靠；操作简单	识别源数量有限，并非直接针对数据本身，而是对数据产生的偏差进行分析；无法定量给出源贡献

PMF 模型的基本原理^[5]是：先利用权重对颗粒物中各化学组分的误差进行计算，再利用最小二乘法来计算出颗粒物的主要污染源及其贡献率。PMF 模型将原始矩阵 $X_{(m \times n)}$ 因子化，分解为两个因子矩阵， $F_{(p \times n)}$ 和 $G_{(m \times p)}$ ，以及一个“残差矩阵” $E_{(m \times n)}$ ，如下式表示：

$$X_{(m \times n)} = G_{(m \times p)} \times F_{(p \times n)} + E_{(m \times n)} \quad (1)$$

式中， $X_{(m \times n)}$ 表示受体矩阵；每个样品中每种组分的浓度； $F_{(p \times n)}$ 表示因子谱矩阵； $G_{(m \times p)}$ 表示源贡献矩阵，每类源对每个样品颗粒物的贡献， $E_{(m \times n)}$ 是残差矩阵。 m 、 n 、 p 分别代表样品、组分、源类数目。

PMF 定义了一个“目标函数” Q ，通过迭代使这个目标函数的值最小：

$$Q = \sum_{i=1}^n \sum_{j=1}^m (e_{ij} / s_{ij})^2 \quad (2)$$

式中， s_{ij} 是第 i 个样品中第 j 个化学成分的标准偏差，或者不确定性。

与其他模型相比，PMF 具有不需要输入源成分谱数据，计算出的因子成分谱和源贡献为非负值，用数据标准偏差来对模型计算进行优化，可处理缺失数据和误差大的数据等特点。PMF 模型软件主要有两种类型，包括 DOS 版本 PMF 模型软件和图形界面 PMF 模型软件。Paatero 开发的个人 DOS 版本 PMF 软件主要包括 PMF2（二维因子分析模型）和 PMF3（三维因子分析模型），其没有可视化的图形界面和结果的分析图表，操作方法不易掌握，但其运行速度较快，所以在科学研究中应用较为广泛。图形界面 PMF 模型软件操作相对简单而且直观，更适合于业务化应用。

国外, 美国 EPA 开发的图形界面 EPA PMF 模型软件, 并且不断进行修改完善, 从一开始的 EPA PMF 1.1、EPA PMF 3.0, 直到 2014 年的 EPA PMF 5.0 版。EPA PMF 5.0 是目前 PMF 开发的最新版本, 可在美国 EPA 官网免费下载使用该模型。和 PMF 3.0 相比, EPA PMF 5.0 进行了以下改进, 包括增加了两种误差估算的方法, 并且可以根据使用者对污染源物理特征的理解, 加入对因子谱和源贡献矩阵的拉伸限制, 使得模型计算的解析结果更加具有物理意义, 增强了模型的适用性。

国内, 南开大学开发了图形界面的 PMF 模型软件, 即 WALSPMF 模型。该模型是在传统 PCA 模型的基础上, 结合非负限制迭代算法和加权交互最小二乘算法等运算机理研发的^[9]。WALSPMF 模型软件可在南开大学相关网站上免费下载, 需要安装 MATLAB 软件才能使用。与 EPA PMF 模型相比, 该模型使用基于 PCA 法解析初始因子, 并利用 WAL 算法, 对结果收敛加以限制, 得到一个最优的解析结果。WALSPMF 是以 PCA 法为基础, 各因子间正交关系较为明显, 避免了多个因子具有相同的特征组分问题, 增强了因子的识别特性^[11]。

3.1.2 PMF 模型研究应用

PMF 模型被广泛地应用于国内各地源解析研究工作中对于离线监测得到的颗粒物及挥发性有机化合物 (VOCs) 相关数据进行源解析计算。宋宇等人^[10]于 2002 年运用 PMF 模型对细粒子的来源进行了分析, 发现主要来源有 6 类: 城市扬尘、建筑水泥尘、生物质燃烧源、二次源、机动车尾气尘和煤烟尘。Xie 等人^[11]于 2004 年运用 PMF 模型对北京市环境空气中的 PM₁₀ 进行了来源解析, 结果表明, 北京空气中的可吸入颗粒物主要受城市扬尘、土壤风沙尘、煤烟尘、二次硫酸盐源、二次硝酸盐源、生物质燃烧源和机动车尾气尘的影响。Song 等人^[12]于 2004 年运用 PMF 模型解析了北京市 PM_{2.5} 的来源, 得出北京市冬季 PM_{2.5} 的主要来源为煤烟尘, 夏季主要贡献源类为二次源类, 此外机动车尾气尘、道路尘和生物质燃烧源类的贡献不容忽视。Xie 等人^[13]于 2010 年应用 EPA PMF1.1 研究了宝鸡市冬春两季 PM₁₀ 的主要贡献源类, 主要识别出 4 类源: 土壤风沙和道路尘、生物质燃烧源、微生物排放、化石燃料燃烧和真菌孢子的混合物源。Cai 等人^[14]于 2010 年利用 PMF 模型基于上海市 VOCs 的浓度, 结合当地主要工业源的分布及气象条件, 解析了 VOC 的主要 7 种源类。郇涛等人^[15]于 2010 年利用 PMF 模型对北京市 PM₁₀ 和 SO₂ 监测网络进行分析与评价, 给监测点的设置提供的信息, 优化监测网络。Cheng 等人^[16]于 2011 年运用 PMF 模型识别了 PM₁ 的主要来源, 研究表明, 机动车尾气尘是 PM₁ 的主要贡献源类, 源贡献分担率为 38%, 其次是二次气溶胶 (22%), 废物焚化炉/生物质燃烧源 (16%), 残油燃烧 (12%), 工业废气 (7%) 和城市扬尘 (5%)。成海容等人^[17]于 2012 年利用 PMF 模型解析了武汉市城区大气 PM_{2.5} 的来源, 结果表明, 其主要来源及源贡献分担率分别为机动车尾气尘 (27%)、二次硫酸盐和二次硝酸盐源 (27%)、工厂排放 (26%) 和生物质燃烧源 (20%)。2013 年, 程水源等人^[18]利用 PMF 模型对北京市 PM_{2.5} 中的碳组分进行源解析, 结果表明在采暖季, 北京 PM_{2.5} 中碳组分主要来源于机动车尾气尘、煤烟尘以及生物质燃烧源。南开大学 Tian 等人^[19]于 2013 年利用 PMF 模型解析了成都市 PM₁₀ 和 PM_{2.5} 的主要来源, 结果表明, PM₁₀ 和 PM_{2.5} 受相似源类的影响, 主要包括土壤风沙尘、机动车尾气尘、二次硫酸盐源、二次硝酸盐源和建筑水泥尘。

以上文献调研显示 PMF 模型可被用于开展环境空气中 PM_{1.0}、PM_{2.5}、PM₁₀、VOC 等来源解析，尤其是在环境空气颗粒物源解析中已比较成熟。应用的城市包括北京、上海、武汉、成都、宝鸡市等，解析出颗粒物来源包括城市扬尘、建筑水泥尘、生物质燃烧源、二次源、机动车尾气尘和煤烟尘、二次硫酸盐和二次硝酸盐源等。

除了对离线监测数据开展模型计算外，PMF 模型还可用于颗粒物在线监测数据的模型计算研究，实现颗粒物在线源解析^[20]。王晓浩等利用在线监测数据与 PMF 联用技术研究上海郊区冬季 PM_{2.5} 的来源解析^[21]。王古月等对镇江谏壁桥点位 2015 年 9 月底至 11 月初的在线 PM_{2.5} 样品进行研究，捕捉到多次重度污染天气，属于较为典型的秋季污染特征，其中，主要污染源是机动车尾气（28.2%），第二是燃煤（23.9%），第三为扬尘（11.5%）^[22]。高健等同时使用 PCA，PMF，ME2 和 Unmix 等多种受体模型分析了北京市 1 小时分辨率在线数据，比较了不同模型的解析结果^[23]。结果表明，PMF 和 ME2 模型更容易得到收敛解；但是由于 Si 和 Al 等标识组分的缺失，往往使扬尘等源类被低估。周敏等^[24]对大气中细粒子及其化学组分进行了在线连续观测，基于在线数据运用正定矩阵因子分析法（PMF）、化学质量平衡法（CMB）和多元线性模型（ME2）这 3 种受体模型开展颗粒物源解析并进行相互验证。结果表明，3 种模型获得的二次硫酸盐、二次硝酸盐、二次有机碳、移动源的源解析结果较接近，说明模型对这 4 类源的模拟较好。ME2 和 PMF 模型对燃煤源、扬尘源的拟合结果要好于 CMB；工业源则是 CMB 的结果更好。

以上文献调研显示，目前可以用于颗粒物的在线源解析研究工作的输入数据主要为在线颗粒物成分数据，如基于在线离子色谱测量的水溶性离子、基于在线热光法测量的有机碳和元素碳、基于在线 XRF 测量的重金属元素等。相比于经典源解析方法中成分的测量方式，在线测量方法由于其采样、样品处理及分析过程的高时间分辨率，在一定程度上牺牲了精准度，或者由于仪器设计原理的限制而只能测量部分粒径范围的颗粒物，在反映 PM_{2.5} 族群的代表性上有所欠缺。此外，由于在线分析方法的限制，在经典源解析样品分析中可以分析且对某些源类有特征指示性的某些关键物种（如土壤尘或扬尘指示物硅、铝等）较难实现在线分析。因此，PMF 模型在颗粒物在线源解析应用中存在着不少主要技术问题，还需要不断进行研究解决。

3.1.3 PMF 模型业务化应用情况

除了在学术研究领域得到了广泛的应用外，近几年 PMF 模型还在我国环保部门的颗粒物来源解析业务中也得到了比较多的实际应用，为环境空气颗粒物污染防治提供重要支撑。比如，2015 年南宁市开展了环境空气颗粒物来源解析，利用 PMF 模型解析了 PM₁₀ 和 PM_{2.5} 的来源，两个粒径段的颗粒物均识别了 6 个污染源，分别是扬尘源、生物质锅炉排放源、机动车源、燃煤源、硫酸盐、硝酸盐；2016 年扬州市开展了 PM_{2.5} 来源解析研究，利用 PMF 模型解析得到 6 个因子：二次源、燃煤源、扬尘、移动源、工业源、建筑尘。

2018 年，中国环境监测总站在全国环境监测系统开展了大气颗粒物源解析工作现状调查^[25]，调查的具体对象为地级及以上级别环境监测站，共发放调查 366 份，收到反馈 326 份。总体上，结果显示（见表 2）仅 13 个监测站具备一种或一种以上颗粒物源解析模型计算能力，占 4%；直辖市站具备相关能力的比例相对较高，为 50%；一般地级市站不具备相关模型计算能力。对于全国 3 大重点区域（见表 3），京津冀及周边地区监测站具备一种或

一种以上模型计算能力的比例相对较高，为 11%，长三角地区监测站具备模型计算能力的比例为 5%，珠三角地区和汾渭平原监测站不具备模型计算能力。上述结果表明全国各监测站模型计算能力具备比例很低。各级监测站的传统工作集中于参照标准规范对不同环境介质中污染物进行采样和分析测试，较少开展应用 CMB 或 PMF 模型进行源解析计算，缺少相应的模型计算技术指南或是其中一个重要的限制因素。

表2 各级监测站模型计算能力

监测站	反馈材料数量	具备模型计算能力	
		数量	占比
直辖市站	4	2	50%
省（区、兵团）站	28	7	25%
省会城市站	26	2	8%
计划单列市站	5	1	20%
一般地级市站	263	0	0%
总计	326	13	4%

表3 重点区域监测站模型计算能力情况

区域	反馈材料数量	具备模型计算能力	
		数量	占比
京津冀及周边地区	28	3	11%
长三角地区	40	2	5%
汾渭平原	11	0	0%
珠三角地区*	9	0	0%
总计	88	5	6%

注*：珠三角地区大气颗粒物源解析工作由广东省站统一组织。

3.2 国内外相关标准方法

PMF 是美国 EPA 在源解析工作中推荐使用的受体模型之一。美国 EPA 发布了 PMF 指导手册（《A Guide to Positive Matrix Factorization》）^[26]和 PMF5.0 使用手册（《EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide》）^[27]，其中 PMF5.0 使用手册详细介绍了 PMF 模型的原理、PMF 5.0 的功能、使用及实际应用案例。

欧盟发布的《European Guide on Air Pollution Source Apportionment with Receptor Models》^[28]，包含 PMF 模型的使用方法。

国内尚没有相关的标准方法。

4 标准制订的基本原则和技术路线

4.1 标准制订的基本原则

科学性原则 所制定的标准能够满足源解析工作的特定技术需求，所选择的技术方法要科学合理。

适用性原则 所制定的标准要能够在不同地区的源解析工作中使用，在技术上可行，具有可操作性。

规范性原则 所制定的标准对关键技术指标进行统一规定，提高源解析结果的可比性。

4.2 标准制订的技术路线

通过调研国内外 PMF 模型相关标准规范、PMF 相关研究进展、PMF 在源解析工作的应用情况，总结已有的研究成果和实际应用经验，编写《环境空气颗粒物来源解析正定矩阵因子分解（PMF）模型计算技术指南》（初稿）。

针对 PMF 模型应用中存在的难点疑点问题以及关键技术参数，继续调研已有的研究成果，必要时开展研究，得到优化后的计算方法及技术参数，形成《环境空气颗粒物来源解析正定矩阵因子分解（PMF）模型计算技术指南》（验证稿）。

选择典型城市环保部门以及高校研究所，进行实际应用，查看方法的可操作性和重复性，并提出修改意见。根据应用中反馈的问题，对指南进行修改完善，形成并提交技术指南及编制说明的征求意见稿。

具体技术路线图如下：

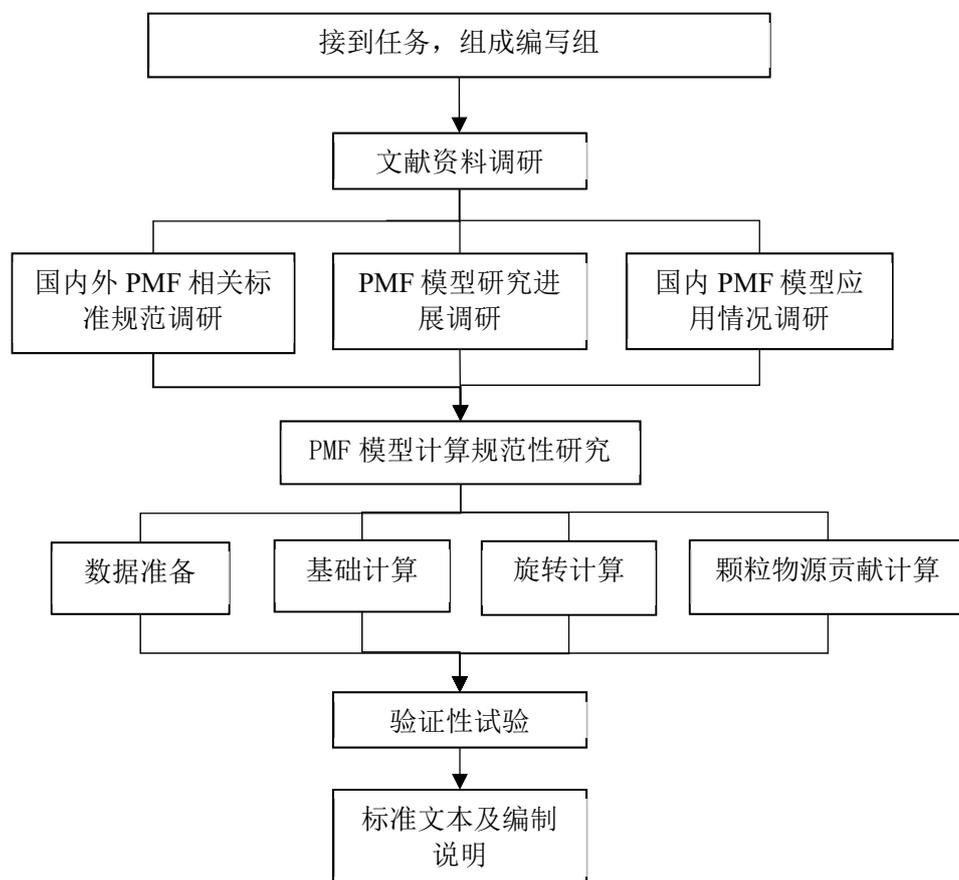


图 1 技术路线

5 方法研究报告

5.1 适用范围

本标准对正定矩阵因子分解模型（Positive Matrix Factorization, PMF）模型计算过程的主要技术环节进行规定，包括数据准备、数据输入和初步检验、基础计算、旋转计算、源解析结果计算、结果合理性判断等。

大量资料调研显示应用 PMF 模型对手工采样分析数据的计算在各地业务化工作中有大量应用，相对比较成熟。而对环境空气颗粒物在线数据的 PMF 模型计算受制于输入数据的准确性，业务化应用还不是很成熟。虽然通过在线颗粒物监测方法，实现用最少的人力得到大量高时间分辨率数据，但是在线方法数据精准度与手工数据有较大差距，而且对某些源类有特征指示性的关键物种（如硅、铝等）难以实现在线分析而无法取得相应数据，影响源解析结果准确性，仍需要进一步研究^[20,29,30]。

因此，本标准规定适用于环境空气颗粒物来源解析工作中应用 PMF 模型对环境空气颗粒物手工采样分析数据进行颗粒物来源计算。

5.2 规范性引用文件

本标准涉及的规范性引用文件如下：

HJ 618 环境空气 PM₁₀ 和 PM_{2.5} 的测定重量法

HJ 646 环境空气和废气气相和颗粒物中多环芳烃的测定气相色谱-质谱法

HJ 647 环境空气和废气气相和颗粒物中多环芳烃的测定高效液相色谱法

HJ 656 环境空气颗粒物（PM_{2.5}）手工监测方法（重量法）技术规范

HJ 657 空气和废气颗粒物中铅等金属元素的测定电感耦合等离子体质谱法

HJ 777 空气和废气颗粒物中金属元素的测定电感耦合等离子体发射光谱法

HJ 799 环境空气颗粒物中水溶性阴离子（F⁻、Cl⁻、Br⁻、NO₂⁻、NO₃⁻、PO₄³⁻、SO₃²⁻、SO₄²⁻）的测定离子色谱法

HJ 800 环境空气颗粒物中水溶性阳离子（Li⁺、Na⁺、NH₄⁺、K⁺、Ca²⁺、Mg²⁺）的测定离子色谱法

HJ 829 环境空气颗粒物中无机元素的测定能量色散 X 射线荧光光谱法

HJ 830 环境空气颗粒物中无机元素的测定波长色散 X 射线荧光光谱法

HJ××× 环境空气 颗粒物来源解析 受体模型法监测数据处理与检验技术规范

《环境空气颗粒物来源解析监测技术方法指南》（环办函[2020]8号）

《大气颗粒物来源解析技术指南（试行）》（环发[2013]92号）

5.3 术语和定义

环境受体的定义参考了《大气颗粒物来源解析技术指南（试行）》（环发[2013]92号），正定矩阵因子分解模型、受体化学组成、污染源化学成分谱、源贡献、源标识组分、因子、不确定度、Q 值、残差、基础计算、旋转计算等术语主要参考美国 EPA《EPA positive matrix factorization (PMF) 5.0 fundamentals and user guide》、《大气颗粒物来源解析

原理技术及应用》、《大气颗粒物来源解析受体模型的发展和应用》以及相关研究论文[20,29,31,32]。

5.4 模型原理和计算流程

5.4.1 模型原理

该部分内容介绍了 PMF 模型计算的数学原理。PMF 模型基本原理是将受体组分浓度矩阵 $X_{(n \times m)}$ 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，模型会将受体浓度矩阵 X 分解成 G 矩阵和 F 矩阵，分别是 PMF 模型计算的源贡献度矩阵和因子谱（源谱）矩阵。

按照上述原理设计并开放使用的软件，或通过采购获得相关软件，经测试后可以用于 PMF 模型计算。

5.4.2 计算流程

总结 PMF 模型的一般计算流程，主要包括数据准备、数据初步检验、基础计算运行及结果判断分析和误差评估、旋转计算及结果判断分析和误差评估、结果合理性判断等环节。PMF 模型计算，需要通过反复优化计算参数，得到相对较优结果的过程。模型计算流程图见图 2 所示。

根据模型的要求，开展监测分析，得到相应的数据。数据准备即对数据进行预处理，选择样品和颗粒物组分，计算组分的不确定度等；数据检验即查看统计分析结果、确定组分权重等。当分析结果中有异常数据时，需要重新进行数据处理。

经过数据准备的环节后，进行基础计算。基础计算中，设置基础计算的参数、反复尝试确定因子个数，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，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调整参数或数据处理，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，最终得到源解析结果。

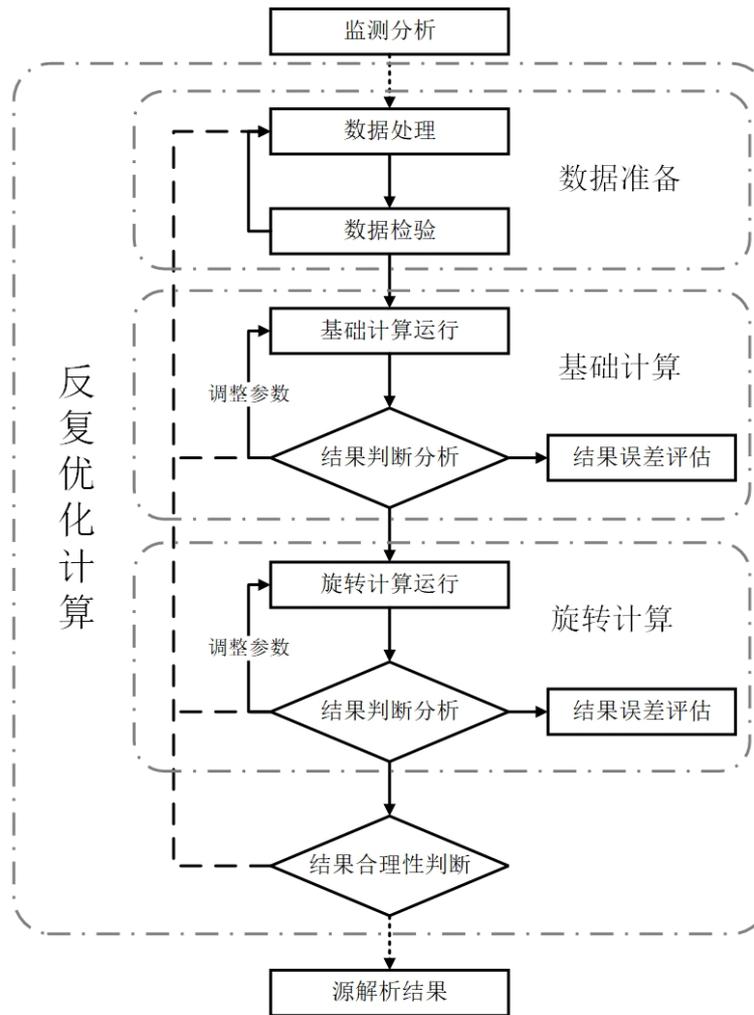


图2 PMF 模型的主要计算流程

5.5 数据准备

5.5.1 数据种类和数量要求

根据PMF模型开展环境空气颗粒物来源解析的需求，需要输入环境空气颗粒物质量浓度和化学组分浓度数据，颗粒物化学组分基本要求需要包括无机元素、碳组分、水溶性无机离子等。参考“关于印发《环境空气颗粒物来源解析监测技术方法指南（试行）》的通知（环办函[2014]1132号）”、“关于发布《大气颗粒物来源解析技术指南（试行）》的通知（环发[2013]92号）”，为保证源解析结果的稳定可靠，本标准规定了应用PMF模型进行颗粒物源解析计算至少需要的数据数量：每个监测点位不少于100个有效样品的数据量，颗粒物化学组分至少需要10种组分，应包括主要组分和标识组分。常用的组分包括OC、EC、SO₄²⁻、NO₃⁻、Cl⁻、NH₄⁺、Na、Mg、Al、Si、K、Ca、Mn、Fe、Cu、Zn等。为解析更多的源类，有条件的可以选择多环芳烃、左旋葡聚糖、正构烷烃、有机酸、脂肪酸、甾醇类等有机组分进行计算。增加输入的颗粒物组分和样品数量，将使PMF模型计算结果更稳定。

5.5.2 数据质量要求

输入的数据质量是取得可靠的PMF模型解析结果的基础。因此，本标准规定用于颗粒物源解析工作的监测数据，应采用国家、行业或地方方法标准获得；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。采用标准方法除了保证良好的数据质量外，还使不同城市或区域的数据可比。总结已有的相关标准规范，颗粒物及相关化学组分的分析需参照相关国家标准方法HJ 618、HJ 646、HJ 647、HJ 656、HJ 657、HJ 777、HJ 799、HJ 800、HJ 829、HJ 830和《环境空气颗粒物来源解析监测技术方法指南》等。

5.5.3 数据格式要求

参考EPA PMF 5.0等PMF模型的要求，PMF模型需要输入两个数据文件：（1）颗粒物及组分质量浓度的数据文件，（2）颗粒物及组分质量浓度不确定度的数据文件。为了使模型计算顺利进行，两个数据文件内容需要对应，格式符合模型软件的要求。比如EPA PMF 5.0模型软件要求第一行为组分名称，组分名称必须是唯一，中间不能含有逗号；第一列为样品编号，如样品ID、日期或时间、样品ID加上日期时间等。

5.5.4 不同滤膜来源数据的合并

源解析工作中，为满足颗粒物化学组分分析的需求，颗粒物需分别通过无机滤膜和有机滤膜平行采样。石英滤膜一般用于分析碳组分、水溶性离子以及有机组分；有机滤膜一般用于分析无机元素。采用不同滤膜监测分析得到的颗粒物质量浓度数据存在一定的差异。在PMF模型计算中，不同滤膜得到的颗粒物监测数据需经过处理后合并。关于不同滤膜得到的数据，本标准规定了两种滤膜得到颗粒物相关数据的处理和合并方法：（a）先将滤膜上颗粒物化学组分质量浓度除以对应滤膜的颗粒物质量浓度，得到相应的质量百分占比含量；（b）再将每个样品的各组分质量百分占比含量乘以相同的颗粒物质量浓度（如石英滤膜和有机滤膜的颗粒物质量浓度均值），合并成一套数据（单位 $\mu\text{g}/\text{m}^3$ ）。

5.5.5 数据初步检验

对于每个颗粒物样品的质量浓度和化学组分浓度数据，需要经过检验才能用于模型计算。目前，常用的检验方法有颗粒物化学组分质量平衡、阴阳离子电荷平衡等方法进行检验。

常用的颗粒物化学组分质量平衡检验是将化学组分加和浓度和颗粒物质量浓度进行比较。化学组分加和的计算结果，通常用百分比来表示，即每种化学组分的质量浓度除以颗粒物质量浓度的所得百分比的加和，以质量百分数为单位，可反映实际监测与理论上的偏差。由于监测的化学组分主要为标识组分和特征组分，且监测数量有限，因此，所有化学组分的含量之和应小于 100%。化学组分的加和不能太高（超过 90%左右），也不能太低（低于 40%左右）；若不在此范围内，可能是某种或某些组分的测定结果偏高或偏低，故而影响了总体加和结果，应根据分析记录再次分析化学组分浓度，若化学组分的加和还是

过高或过低，建议该条数据剔除。总结各地源解析工作经验，本标准要求的每个颗粒物样品的化学组分质量加和与颗粒物质量浓度的比值一般在 0.5~0.7 之间。

颗粒物阴阳离子电荷平衡检验是比较每个颗粒物样品的阴离子总电荷与阳离子总电荷。按照 HJ 799、HJ 800 等要求测定颗粒物中的阴阳离子，包括 Na^+ 、 NH_4^+ 、 K^+ 、 Mg^{2+} 、 Ca^{2+} 、 SO_4^{2-} 、 NO_3^- 、 Cl^- 、 F^- 等。常用的阴离子和阳离子平衡计算公式分别见公式（3）和（4）。

$$AE = (2 * [\text{SO}_4^{2-}] / 96) + ([\text{NO}_3^-] / 62) + ([\text{Cl}^-] / 35.5) + ([\text{F}^-] / 19) \quad (3)$$

式中：AE——颗粒物中负电荷的物质的量浓度（mmol/L）

$[\text{SO}_4^{2-}]$ ——硫酸根离子的质量浓度（mg/L）

$[\text{NO}_3^-]$ ——硝酸根离子的质量浓度（mg/L）

$[\text{Cl}^-]$ ——氯离子的质量浓度（mg/L）

$[\text{F}^-]$ ——氟离子的质量浓度（mg/L）

$$CE = ([\text{Na}^+] / 23) + ([\text{NH}_4^+] / 18) + ([\text{K}^+] / 39) + (2 * [\text{Mg}^{2+}] / 24) + (2 * [\text{Ca}^{2+}] / 40) \quad (4)$$

式中：CE——颗粒物中正电荷的物质的量浓度（mmol/L）

$[\text{Na}^+]$ ——钠离子的质量浓度（mg/L）

$[\text{NH}_4^+]$ ——铵根离子的质量浓度（mg/L）

$[\text{K}^+]$ ——钾离子的质量浓度（mg/L）

$[\text{Mg}^{2+}]$ ——镁离子的质量浓度（mg/L）

$[\text{Ca}^{2+}]$ ——钙离子的质量浓度（mg/L）

总体上，对 AE 和 CE 进行一元线性回归，当相关系数 $r \geq 0.8$ 时，认为阴阳离子接近平衡，斜率应该在 0.7~1.2 范围内，不在该范围的数据说明可能部分阴阳离子未检出，需对阴阳离子进行复测。本标准要求的对于每个颗粒物样品，其阴离子总电荷与阳离子总电荷的比值一般在 0.7~1.2 之间。

5.5.6 不确定度计算方法

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 $\text{PM}_{2.5}$ 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份谱）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。本标准主要参考美国 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：

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

$$Unc = \sqrt{(EF \times Conc)^2 + (0.5 \times MDL)^2} \quad (6)$$

式中：Unc——颗粒物分析项目的不确定度， $\mu\text{g}/\text{m}^3$ ；

MDL——颗粒物分析项目的检出限， $\mu\text{g}/\text{m}^3$ ；

Conc——颗粒物分析项目的质量浓度， $\mu\text{g}/\text{m}^3$ ；

EF（Error fraction）——颗粒物分析项目的误差系数。

当 Conc 小于等于 MDL 时，采用公式（5）计算；当 Conc 大于 MDL 时，采用公式（6）计算。

关于误差系数（EF），不同的研究学者在 PMF 模型计算中设定有所差异^[33-38]。对特定的研究问题，误差系数可设置为 0.05~0.2、0.1、0.2~0.5。例如，Liu 等考虑了采样期间扩散和重力损失等误差，对于粒径小于 25 nm 的颗粒物，误差系数取 0.15；粒径在 25-600 nm 的颗粒物，误差系数为 0.1，粒径大于 600 nm 的颗粒物，误差系数为 0.15^[33]。蔡凡涛等对南京市 PM_{2.5} 污染事件进行解析时，将所有组分的误差系数均设置为 0.1^[34]。总结 PMF 模型实际使用经验，推荐误差系数设置范围为 0.1~0.6。对于组分不稳定或者组分浓度接近检出限，给予较大的误差系数。当某一组分的缺失数据较多时，可设定较大的误差系数。

5.6 基础计算

5.6.1 基础计算运行

数据的模型检验。将数据导入到 PMF 模型，通过 PMF 模型的浓度散点图、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，经核实没有特定污染（如沙尘天气等），可予以剔除。

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 模型中各化学组分的计算权重包括 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 模型进行颗粒物源解析时，需要使用标识组分来识别污染源类，因此建议将与颗粒物来源相关的标识组分设定为高权重；其他颗粒物组分作为辅助分析，可以设置成 weak。因此，规定将颗粒物的关键标识组分（如 OC、EC、SO₄²⁻、NO₃⁻、NH₄⁺、Al、Si、Ca、K、Fe 等组分）权重设为 Strong。其他非关键标识组分（如：Cl⁻、Na、Mg、Cu、Cd、Cr 等组分）通常建议选择 Strong 权重计算，但在测量不确定性较高时，可设置成 Weak 参与计算。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。

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（Seed）等参数。基础计算中，因子个数设定是一个重要步骤^[39,40]。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颗粒物可能来源、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。

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 矩阵与颗粒物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 1.2，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。

5.6.2 基础计算结果判断

通过基础计算，得到多个计算结果，需进一步判断分析。通过 Q 理论值（ Q_{Theo} ）、Q 计算值（ Q_{True} ）、是否收敛（Converged）等参数，初步选择查看特定的计算结果。参考美国 EPA 《EPA positive matrix factorization (PMF) 5.0 fundamentals and user guide》，主要从 Q 计算值（ Q_{True} ）与 Q 理论值（ Q_{Theo} ）差异、观测值/预测值差异分析、残差分析（Residual analysis）等对计算结果进行判断。

总结源解析研究进展^[41-43]及实际经验，确定以下 3 种判断方法：

Q 计算值（ Q_{True} ）与 Q 理论值（ Q_{Theo} ）越接近（85%~115%之间），表明模拟计算结果更合理。

在观测值/预测值差异分析中，可通过分析输入值和预测值的相关性，来判断模型计算好坏。如果某组分的输入值和预测值之间有强相关性，则表明该组分计算较好；反之，则考虑是否应该降低该组分计算权重或从计算中排除。此外，还可以在时间序列上分析组分的输入值和预测值的差异。对于预测值明显高于输入值，需要进一步分析确定是否排除该数据。

在“Residual Analysis（残差分析）”中查看每一个组分的加权残差（通过不确定度加权），比如残差直方图中不同组分加权残差的百分比，判断各组分模型计算结果的好坏。判断方法如下：如果残差直方图显示残差范围在（-3，+3）之间，且呈正态分布，则该组分的模型计算结果在数学意义上较好；如果组分具有许多大尺度残差或显示非正常曲线，则表示该组分在数学意义上计算较差。

5.6.3 基础计算的结果分析

基础计算的结果主要包括因子谱（F矩阵）和因子贡献（G矩阵）。因子谱展示了不同化学组分在各因子中的占比及浓度。因子贡献显示了不同因子对各组分的平均贡献及时间序列贡献（点线图）。

因子的识别是 PMF 模型计算的关键。可以通过分析不同因子谱中占比较高的标识组分来将因子识别为具体的污染源类。总结源解析研究成果及工作经验^[29,30,32,44-47]，确定颗粒物源类识别方法及主要源类的标识组分，作为资料性附录供参考使用。颗粒物源类识别方法包括：①根据通过资料调研收集的各种源类别的主要标识组分来识别；②根据源类特有的时间变化特征来识别；③根据当地实际监测得到的主要源类源谱来识别；④根据通过模型进一步优化计算得到的因子来识别。主要源类的标识组分见表 4。

表4 颗粒物主要源的标识组分

编号	源类别	标识组分
1	燃煤源	OC、EC、Al、SO ₄ ²⁻ 、Cl ⁻ 、Ca（或Ca ²⁺ ）、As、藿烷、甾烷、烷基芘、多环芳烃等
2	机动车源	OC、EC、Zn、藿烷、甾烷、晕苯、荧蒽、芘等
3	扬尘源	Si、Al、Ca（或Ca ²⁺ ）等
4	二次硫酸盐	SO ₄ ²⁻ 、NH ₄ ⁺
5	二次硝酸盐	NO ₃ ⁻ 、NH ₄ ⁺
6	生物质燃烧源	K（或K ⁺ ）、OC、EC、左旋葡聚糖、植物甾醇、萜类物质等
7	工业源	Fe、Cu、Cr、Mn、Pb、Zn、Al、Sb、V等（视各地具体工业情况而定）
8	餐饮源	OC、胆固醇、十六烷酸、十八烷酸、豆甾醇、β-谷甾醇、壬醛、9-十六烯酸等
9	海盐源	Na（或Na ⁺ ）、K（或K ⁺ ）、Cl ⁻ 等
10	烟花燃放源	Sr、K（或K ⁺ ）、OC等

因子贡献（G矩阵）的时间序列贡献显示了污染源在时间序列上的贡献变化情况，可用于通过因子谱很难区分的源类。特定污染源会在时间序列上有特征的贡献变化情况，从而得到识别。比如机动车源和燃煤源（其主要标识组分均为 OC、EC 等），在某些地区，燃煤源存在显著的季节变化（比如冬季采暖中燃煤的大量使用），而机动车源季节变化不明显。

5.6.4 基础计算的误差评估

在得到模型基础计算结果之后，需要结果的误差评估，确定计算结果的不确定度，从而判断源解析结果的准确性。参考美国 EPA 《EPA positive matrix factorization (PMF) 5.0 fundamentals and user guide》，结果误差评估方法主要有 Bootstrapping（BS）误差评估、

Displacement (DISP) 误差评估、Bootstrap- Displacement (BS-DISP) 误差评估。3 种误差评估方法的适用性有所差别^[34,36,49]。BS 误差评估和 BS-DISP 误差评估适合解释源谱的不确定性，并且能够说明是否出现由于因子数目过多引起的数据过拟合。BS 误差评估结果通常不能捕捉旋转模糊度，但可以帮助识别重复性不高的因子。DISP 则要求具有较强的稳健性 (robust)，否则可能导致没有 DISP 误差评估结果。此外，DISP 分析给定的不确定性区间比较短。BS-DISP 的一个缺点是它的计算时间较长，可能需要几十个小时。因此，推荐采用 BS 误差评估方法评估模型计算结果 (不确定度) 的误差。

通过分析多次 BS 计算得到的因子和基础计算得到的因子匹配程度、Q 值的分布情况、各化学组分计算结果分布情况等，评估结果的可靠性。主要的评估方法如下：

在 BS 误差评估中，通过分析得到的因子和基础计算得到的因子匹配程度，评估因子中主要化学组分分配的合理性。如果两者匹配不到 80%，则这个因子的主要组分分配可能不当，需要通过调整因子数量或提高数据质量等提高匹配程度。

在 BS 误差评估中，通过 Q 修正值的分布情况，评估基础计算中 Q 值的合理性。BS 计算可得到 Q 修正值的最小、最大、中位数，以及第 25 和第 75 百分位数值。基础计算得到的 Q 值一般要求在第 25 和第 75 百分位数值之间。

在 BS 误差评估中，通过分析每个化学组分计算结果的分布情况，评估化学组分计算结果的不确定度。比如一个化学组分的箱型图上下距 (25 分位数与 75 分位数) 较宽 (超过 20%)，则表明该化学组分计算结果的不确定度可能较高；如果一个化学组分的箱型图上下距较窄 (小于 20%)，则说明该化学组分模型计算结果的不确定度可能较低，模型计算结果较好。

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，可以尝试使用旋转计算 (Rotational run)，对计算结果进一步优化，使因子谱中的标识组分更加突出，有利于因子识别为具体污染源类。PMF 模型计算中，旋转计算包括 F 矩阵峰值模型计算 (Fpeak model run)、约束模型计算 (Constrained model run)。F 矩阵峰值模型计算是常用的旋转计算方法^[50,51]。对于标识组分的占比不够突出 (某一标识组分在各因子中的占比比较平均)，不利于源类的识别的情形，可以通过 F 矩阵峰值旋转计算，调整提取出的 F 矩阵中的各因子中组分占比，从而使因子中的标识组分的占比更加突出，从而识别为具体源类。

F 矩阵峰值模型计算的基本原理如下：

一对因子矩阵 (G 和 F) 可以被转换为另一对具有相同 Q 值的矩阵 (G*和 F*)，称为“旋转”即 $G^*=GT$ and $F^*=T^{-1}F$ 。T 矩阵是一个 $p \times p$ 的非奇异矩阵，其中 p 是因子的个数。在 PMF 中，这不是严格意义上的旋转，而是 G 和 F 矩阵的线性变换。

F 矩阵峰值模型计算中，需要尝试不同的 Fpeak 参数。当 Fpeak 设为正值时，F 矩阵中的组分在某些因子中占比会更加突出；Fpeak 设为负值时，F 矩阵中的组分在各因子中的占比则相对较平均。设置不同的 Fpeak 参数时，Q 值不能变化太大，Q 值的变化在 5%以内认为时可以接受的，选择 Q 值拐点之前的 Fpeak 值对应的解析结果。参照“5.6.2、5.6.3、5.6.4”进行相应的结果判断、分析和误差评估。

PMF 共线性问题及解决方法。其他受体模型一样，PMF 存在共线性问题。PMF 计算过程没有考虑实际源类的物理意义，导致被提取的因子与实际的污染源成分谱之间存在一定差异，提取的单个因子可能包含不同污染源类的信息，出现了共线性问题。现有比较有效的方法是采用一些复合受体模型克服共线性问题，而开展进一步解析，比如正定矩阵分解/多元线性引擎 2-物种比值模型（PMF/ME2-SR）^[57]、偏目标转换-正定矩阵分解模型（PTT-PMF）^[58]。PMF/ME2-SR 模型是基于 PMF/ME2 科学计算平台，将污染源特征组分的比值作为提取过程的约束条件纳入模型，使得提取的因子更具物理意义。该模型可以求解 PMF 模型中各种多线性和拟多线性问题，通过实施外部约束，使解析结果朝着目标值或比值拉伸因子元素，在一定程度上使结果朝着更客观的实测源谱靠近，降低了 PMF 结果易受主观因素影响的不确定性。PTT-PMF 模型是在 PMF/ME2 的基础上做出的进一步改进。它是一种将源成分谱数据纳入模型，通过部分目标转换方法实现的新型源解析模型。具体来说，该模型将测量得到的实测一次源谱及虚拟的二次源谱作为目标因子，选择并固定所有源的源标识组分，通过一个方程式的拉动，实现因子谱的目标转换，使该模型提取的因子谱更接近实际的源成分谱，从而增加了解析结果的真实性和可靠性。

5.8 颗粒物源贡献结果计算

PMF 模型计算中可通过两种方法确定污染源对颗粒物的贡献：

（1）在“数据准备”中将颗粒物质量浓度数据和化学组分数据一起输入到 PMF 模型进行计算，将颗粒物质量浓度设置为总变量（Total variable）。这样就可以直接通过 PMF 模型得到污染源对颗粒物的源贡献。

（2）利用 PMF 模型计算得到的 G 矩阵（源贡献矩阵）与颗粒物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。回归计算可利用 Excel 或 SPSS 软件进行回归分析。将各列的回归系数乘以各列，即为各污染源对颗粒物的贡献，回归公式如下：

$$PM_{2.5(j)} = \sum_{k=1}^p s_k g_{jk} \quad (6)$$

式中： $PM_{2.5(j)}$ ——第 j 个样品的颗粒物质量浓度；

s_k ——回归系数；

g_{jk} ——PMF 模型计算的 k 源对第 j 个样品的源贡献，

p——源个数。

在城市环境空气颗粒物来源解析工作中，为了提高源解析结果的代表性，往往会在城市建成区设立多个监测点位，得到相应的颗粒物监测数据。总结源解析工作经验以及已有的研究进展^[52-55]，提供针对多个点位的 PMF 源解析计算方法：将多个点位受体数据进行合并，将合并后数据纳入 PMF 模型计算，从而所有样品的源解析结果；将每个点位所对应样品的源解析结果进行平均，得到该点位的源解析结果。

完成 PMF 模型计算后，填写“附录 B PMF 源解析计算记录表”，并附上参与计算的全部数据。

5.9 结果合理性判断

对于源解析结果的合理性，既要符合模型计算要求，还需要符合实际情况。

总结 PMF 模型使用经验及研究进展^[29,32,43-45,56]，提供两类关于结果合理性判断的方法，包括数学意义和物理意义上的合理性判断。

符合模型计算要求是模型模拟计算出的各组分浓度与实测测试结果接近。可具体通过直接比较颗粒物和化学组分模拟值与实测值、分析主要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、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。

符合实际情况要求是经过识别的源谱特征及其源贡献合理。通过比较因子谱与实测源谱（相关性高于 0.7）、PMF 平均源贡献排序与 CMB 模型计算结果一致性，分析源贡献变化趋势及其日变化趋势等来评估模型计算结果的合理性。此外，还需要通过其他监测结果或污染源信息等来辅助判断结果的合理性，可参考的方法如下：

（1）PM_{2.5}和 PM₁₀的浓度比值。如果 PM_{2.5}/PM₁₀ 比值较高，表明监测区域受细颗粒物排放源的影响可能较高；反之，则表明监测区域受粗颗粒物排放源的影响较高。因此，通过分析不同季节（或区域）的 PM_{2.5}/PM₁₀ 比值，结合当地排放源、气象特征，可帮助分析不同时空条件下的粗颗粒物和细颗粒物的主要污染来源。

（2）在线监测的 SO₂/NO_x 浓度比值。如果 SO₂/NO_x 的比值较高，表明监测区域受燃煤来源的影响可能比机动车来源高；反之，表明监测区域受机动车来源的影响较高。因此，通过分析不同季节（或区域）的 SO₂/NO_x 的比值，结合当地排放源、气象特征，可帮助分析不同时空条件下的燃煤来源和机动车来源等贡献的高低。

（3）在线监测 O₃、CO 等污染物浓度。O₃ 能反映光化学反应活性，可分析不同季节（或区域）的 O₃ 浓度，如果 O₃ 浓度较高，表明该季节（或区域）二次硫酸盐的贡献可能较其他季节（或区域）高；CO 主要由机动车或煤炭不完全燃烧排放，可分析不同季节（或区域）的 CO 浓度，如果 CO 浓度较高，表明散煤燃烧或机动车的贡献可能较其他季节（或区域）高。

（4）根据当地主要污染源信息，分析源解析结果的合理性。根据当地机动车保有量分析机动车源贡献的合理性，根据当地燃煤量分析燃煤源贡献的合理性，根据当地在建工地、在建道路等施工信息分析施工扬尘、道路扬尘的贡献。

（5）对于特殊污染源，需结合当地实际情况来分析。对于特定工业源，需要依据当地的工业布局及特点，分析工业源贡献的合理性；对于生物质燃烧源（如秸秆燃烧），需要考虑季节性变化等特点；对于海盐粒子源，需要考虑监测区域的地理位置等。

上述判断方法是总结源解析工作经验得到的，可用于辅助选择合适的源解析结果，还需要结合当地的实际情况，进行具体问题具体分析。

5.10 附录

附录A作为资料性附录提供了颗粒物源类识别方法和颗粒物主要源的标识组分。其中颗粒物源类识别方法包括资料调研总结的各种源类主要标识组分及主要源类的标识组分、特定源类的时间变化特征分析、与实际监测得到的源谱比较、约束模型计算。该内容是总

结近年来我国源解析的实践经验而来。颗粒物主要源的标识组分分别给出了二次硫酸盐、二次硝酸盐、二次源、扬尘源、机动车源、燃煤源、生物质燃烧源、工业源、海盐源、烟花燃放源的主要标识组分，作为源类识别的参考资料。

附录 B 作为资料性附录提供了 PMF 模型计算记录表，包括开展 PMF 模型计算时需要记录的输入数据、基础模型运行、旋转计算运行、输出记录、因子谱、因子贡献等。

6 典型 PMF 模型软件操作步骤

6.1 基本情况

现有的 PMF 模型软件包括 DOS 版本和图形界面版本。DOS 版本包括赫尔辛基大学 Paatero 教授开发的个人 DOS 版本，该软件为收费软件，需要向 Paatero 教授申请购买。图形界面版本包括南开大学开发 WALSPMF 1.0、美国 EPA 开发的 EPA PMF5.0 等。软件为公开免费软件，可在相关网站下载使用。相比 DOS 版本，图形界面版本操作相比简单，适合业务化应用。因此，编制组总结了 WALSPMF 1.0 和 EPA PMF5.0 的主要操作步骤，供参考。

6.2 WALSPMF 1.0 计算步骤

6.2.1 软件下载

可在从南开大学国家环境保护城市空气颗粒物污染防治重点实验室下载，下载地址：<https://env.nankai.edu.cn/air/main.htm>。

6.2.2 运行环境

Win XP、Win7、Win8、Win10（32 bit or 64 bit system）。运行程序之前，需安装 Matlab（2015b 及以上版本）。

6.2.3 数据准备

模型输入文件使用 excel 文件，对数据导入文件有严格要求。用户可以修改输入文件的名称。WALSPMF 模型需要输入一个数据文件，该输入文件中包含三个工作表，如图 3~5 所示：（1）颗粒物及组分质量浓度的数据文件（单位： $\mu\text{g}/\text{m}^3$ ），（2）颗粒物及组分质量浓度不确定度的数据文件（单位： $\mu\text{g}/\text{m}^3$ ），（3）PM 总质量的浓度（合计）（单位： $\mu\text{g}/\text{m}^3$ ）。注意不要更改三个工作表的名称。

三个工作表内容需要对应。在“concentration”工作表中：第一行为组分名称，组分名称必须是唯一，中间不能含有逗号。其余内容为环境数据集，如图 3 所示。在“uncertainty”工作表中：第一行为组分名称，保持与前面工作表中名称一致，其余内容为环境数据集的不确定性，如图 4 所示。在“TOT”工作表中：第一行为组分名称，保持与前面工作表中名称一致。其余内容（第一列）为 PM 总质量的浓度。具体格式请参照模型提供的“input.xls”导入文件格式。如图 5 所示。

1	S04	N03	Cl	NH4	EC	OC	Al	As	Ba	Br	Ca	Cu	Fe	K	Mn	Pb	Db
2	26.33127	10.36647	1.102581	10.80773	13.0402	32.89267	5.437436	0.001619	0.007444	0.0179	4.207578	0.040589	3.715535	2.238386	0.107788	0.046608	0.001619
3	27.58169	14.24344	1.120944	12.3469	13.73474	34.04381	3.893708	0.001557	0.007545	0.018293	4.025165	0.037154	2.893699	2.136618	0.08242	0.046747	0.001557
4	28.17244	17.30126	1.31177	13.16665	15.32911	38.18578	4.804991	0.001772	0.008769	0.021329	4.716187	0.042944	3.489439	2.521733	0.100351	0.052979	0.001772
5	30.22822	12.31878	0.984992	12.97433	15.37807	37.7092	3.725095	0.001615	0.006864	0.016286	3.749026	0.039187	2.864882	1.902134	0.078538	0.049816	0.001615
6	26.97743	9.939551	1.070559	10.93035	15.26339	37.9134	5.02654	0.00167	0.007337	0.017507	4.169257	0.043012	3.588335	2.151017	0.10119	0.051156	0.00167
7	26.95557	13.87623	0.782154	12.6042	12.84461	32.22389	5.55628	0.001667	0.00573	0.013022	3.445838	0.040169	3.744677	1.730908	0.106133	0.043439	0.001667
8	25.11706	12.40496	0.960457	11.15128	14.53584	36.29273	5.536721	0.001695	0.006737	0.015787	3.954564	0.043178	3.827821	2.016742	0.108276	0.048953	0.001695
9	24.23841	15.38181	0.819328	11.91528	14.47387	35.64481	4.269608	0.001585	0.005904	0.013657	3.397715	0.039024	3.110376	1.881118	0.085469	0.046746	0.001585
10	28.31873	10.9218	0.876288	12.04932	14.81842	36.47257	4.209992	0.001594	0.006227	0.014545	3.541688	0.039508	3.09826	1.767951	0.085216	0.047978	0.001594
11	22.71821	11.15734	0.896993	9.986264	14.91354	37.09921	5.462861	0.001652	0.006349	0.01478	3.807623	0.043436	3.79894	1.907862	0.10638	0.049346	0.001652
12	22.29276	10.0496	0.883909	9.474767	16.57307	40.17725	2.971248	0.001445	0.006171	0.014655	3.436654	0.038597	2.493779	1.674506	0.064981	0.051357	0.001445
13	30.27722	16.19236	1.187544	13.81775	13.83996	34.65075	4.892918	0.001728	0.008041	0.019371	4.35919	0.040535	3.458244	2.329015	0.099994	0.048392	0.001728
14	25.6279	12.88567	0.751524	11.82118	14.48578	35.40184	3.56826	0.001499	0.005474	0.01261	3.116639	0.036792	2.716298	1.512009	0.072949	0.045689	0.001499
15	26.1424	19.87304	1.518447	12.78101	13.54828	34.08734	4.431233	0.00167	0.009816	0.024437	5.064583	0.039158	3.217328	2.824417	0.09593	0.049627	0.00167
16	25.35816	13.86131	1.032272	11.59338	11.75655	29.44802	4.140477	0.001463	0.006955	0.016812	3.756926	0.034419	2.930879	2.015023	0.084939	0.041259	0.001463
17	30.60966	12.57977	1.391339	12.47573	16.90393	41.44101	3.285804	0.001653	0.009149	0.022603	4.728387	0.041057	2.727375	2.520696	0.076248	0.056245	0.001653
18	28.98828	14.77666	1.175164	12.95106	13.63876	34.325	5.382655	0.001734	0.007959	0.019132	4.40222	0.041632	3.719285	2.350718	0.107958	0.048167	0.001734
19	23.47092	13.86589	0.569319	11.67795	10.85386	26.77975	3.476924	0.001318	0.004298	0.009647	2.485503	0.030132	2.481125	1.21113	0.068287	0.035064	0.001318
20	26.99603	10.71682	0.681816	11.81375	14.84344	36.13412	3.421585	0.001487	0.005087	0.011547	2.941331	0.036825	2.645596	1.386833	0.069796	0.045965	0.001487
21	24.58778	12.09446	1.023579	10.79391	12.83484	32.45399	5.762554	0.00164	0.007019	0.016673	4.062038	0.041107	3.87939	2.138461	0.112272	0.045263	0.00164
22	25.92405	13.395	1.11307	11.58595	9.630184	24.84265	5.040825	0.001473	0.007371	0.017958	4.000036	0.033728	3.338106	2.222731	0.100317	0.037194	0.001473
	concentration	uncertainty	TOT														

图3 WALSPMF 1.0 输入文件浓度工作表

1	S04	N03	Cl	NH4	EC	OC	Al	As	Ba	Br	Ca	Cu	Fe	K	Mn	Pb	Db
2	2.633127	1.036647	0.110258	1.080773	1.30402	3.289267	0.543744	0.000162	0.000744	0.00179	4.207578	0.040589	3.715535	2.238386	0.107779	0.046608	0.000162
3	2.758169	1.424344	0.112094	1.23469	1.373474	3.404381	0.389371	0.000156	0.000755	0.001829	4.02517	0.037154	2.8937	2.13662	0.08242	0.046747	0.000156
4	2.817244	1.750126	0.131177	1.316665	1.532911	3.818578	0.480499	0.000177	0.000877	0.002133	4.71619	0.042944	3.48944	2.52173	0.10035	0.052979	0.000177
5	3.022822	1.231878	0.098499	1.297433	1.537807	3.77092	0.37251	0.000161	0.000686	0.001629	3.74903	0.03919	2.86488	1.90213	0.07854	0.049816	0.000161
6	2.697743	0.993955	0.107056	1.093035	1.526339	3.79134	0.502654	0.000167	0.000734	0.001751	4.16926	0.04301	3.58834	2.15102	0.10119	0.05116	0.000167
7	2.695557	1.387623	0.078215	1.26042	1.284461	3.222389	0.555628	0.000167	0.000573	0.001302	3.44584	0.04017	3.74468	1.73091	0.10613	0.04344	0.000167
8	2.511706	1.240496	0.096046	1.115128	1.453584	3.629273	0.553672	0.00017	0.000674	0.001579	3.95456	0.04318	3.82782	2.01674	0.10828	0.04895	0.00017
9	2.423841	1.538181	0.081933	1.191528	1.447387	3.564481	0.426961	0.000159	0.00059	0.001366	3.39771	0.03902	3.11038	1.68112	0.08547	0.04675	0.000159
10	2.831873	1.09218	0.087629	1.204932	1.481842	3.647257	0.420999	0.000159	0.000623	0.001454	3.54169	0.03951	3.09826	1.76795	0.08522	0.04798	0.000159
11	2.271821	1.115734	0.089699	0.998626	1.491354	3.709921	0.546286	0.000165	0.000635	0.001478	3.80762	0.04344	3.79894	1.90786	0.10638	0.04935	0.000165
12	2.229276	1.00496	0.086391	0.947477	1.657307	4.017725	0.297125	0.000144	0.000617	0.001465	3.43665	0.0386	2.49378	1.67451	0.06498	0.05136	0.000144
13	3.027722	1.619236	0.118754	1.381775	1.383996	3.465075	0.489292	0.000173	0.000804	0.001937	4.35919	0.04054	3.45824	2.32902	0.09999	0.04839	0.000173
14	2.56279	1.288567	0.075152	1.182118	1.448578	3.540184	0.356826	0.00015	0.000547	0.001261	3.11664	0.03679	2.7163	1.51201	0.07295	0.04569	0.00015
15	2.61424	1.987304	0.151845	1.278101	1.354828	3.408734	0.443123	0.000167	0.000982	0.002444	5.06458	0.03916	3.21733	2.28244	0.09593	0.04963	0.000167
16	2.535816	1.386131	0.103227	1.159338	1.175655	2.944802	0.414048	0.000146	0.000696	0.001681	3.75693	0.03442	2.93088	2.01502	0.08494	0.04126	0.000146
17	3.060966	1.257977	0.139134	1.247573	1.690393	4.144101	0.32858	0.000165	0.000915	0.00226	4.72839	0.04106	2.72738	0.25207	0.07625	0.05625	0.000165
18	2.898828	1.477666	0.117516	1.295106	1.363876	3.4325	0.538265	0.000173	0.000796	0.001913	4.40222	0.04163	3.71928	2.35072	0.10796	0.04817	0.000173
19	2.347092	1.386589	0.056932	1.167795	1.085386	2.677975	0.347692	0.000132	0.00043	0.000965	2.4855	0.03013	2.48112	1.21113	0.06829	0.03506	0.000132
20	2.699603	1.071682	0.068182	1.181375	1.484344	3.613412	0.342159	0.000149	0.000509	0.001155	2.94133	0.03682	2.6456	1.38683	0.06998	0.04597	0.000149
21	2.458778	1.209446	0.102358	1.079391	1.283484	3.245399	0.576255	0.000164	0.000702	0.001667	4.06204	0.04111	3.87939	2.13846	0.11227	0.04526	0.000164
22	2.592405	1.3395	0.11307	1.158595	0.963018	2.484265	0.504083	0.000147	0.000737	0.001796	4.00004	0.03373	3.33811	2.22273	0.10032	0.03719	0.000147
	concentration	uncertainty	TOT														

图4 WALSPMF 1.0 输入文件不确定性工作表

1	TOT				
2	161.895				
3	155.7397				
4	177.1812				
5	161.4602				
6	167.0179				
7	166.6873				
8	169.5329				
9	158.5211				
10	159.357				
11	165.2131				
12	144.4818				
13	172.8096				
14	149.8812				
15	167.005				
16	146.3479				
17	165.3477				
18	173.3532				
19	131.8324				
20	148.7081				
21	163.9603				
22	147.2896				
	concentration	uncertainty	TOT		

图5 WALSPMF 1.0 输入文件 PM 总质量浓度工作表

6.2.4 模型运行

打开 Matlab，路径导入到 WALSPMF 的文件夹，右键选择“WASLPMF.p”文件，选择 Run 运行模型。如图 6 所示。其中，a 为受体数据导入版块，b 为基础运算版块，c 为 Fpeak 旋转运算版块。



图6 WALSPMF 1.0 的用户界面

6.2.5 基础运算

(1) 模型基础计算操作步骤

单击“输入”按钮导入受体数据，组分数量与样品数量自动获取，以验证输入文件是否正确。随后选择参与计算的物种及提取的因子数量。导入受体数据后，可以在基础运算版块点击运行，得到基础模型运算结果，如图 7 所示。

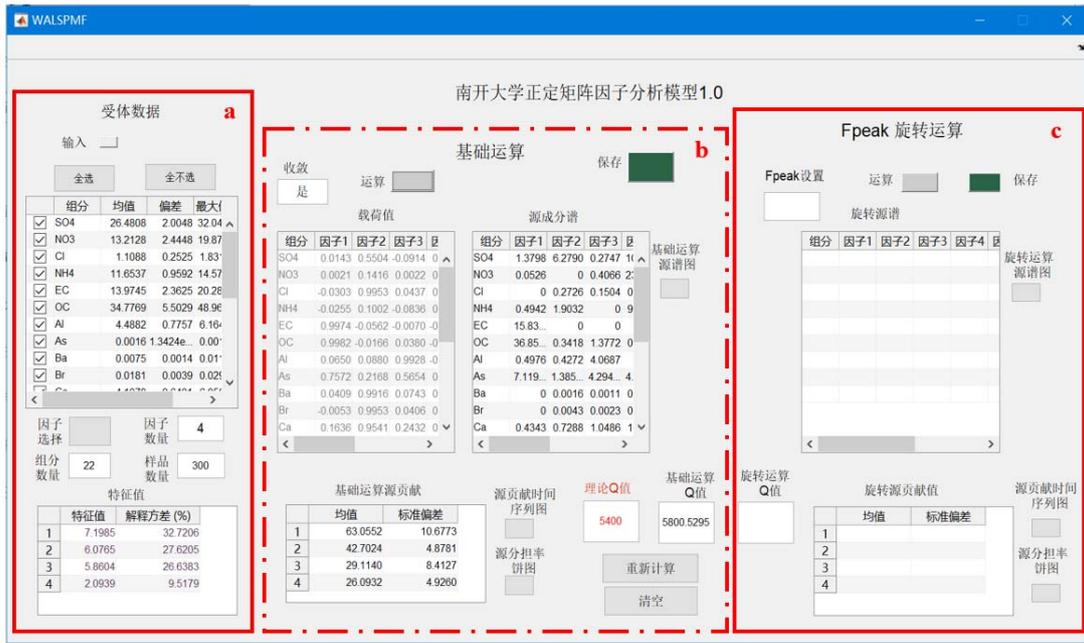


图7 WALSPMF 1.0 的基础计算运行结果展示

(2) 模型基础运行结果展示

基础运行结果包括：基础运算源谱数据和图（图 8）源贡献时间序列图（图 9）；源分担率数据和饼图（图 10）以及相对应的旋转运算结果展示。此外，模型提供理论 Q 值和基础计算 Q 值，用户可以根据两者的差值大致判断结果好坏。

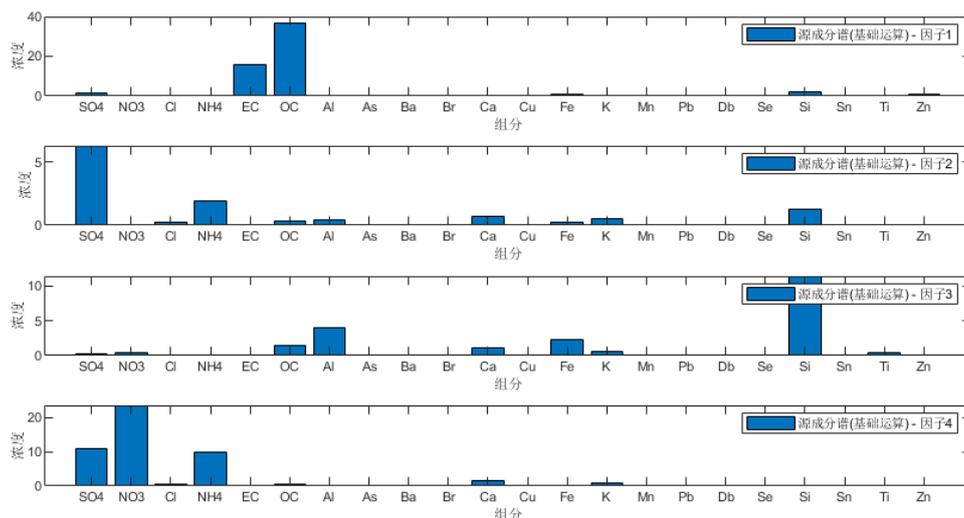


图8 基础运算源谱图

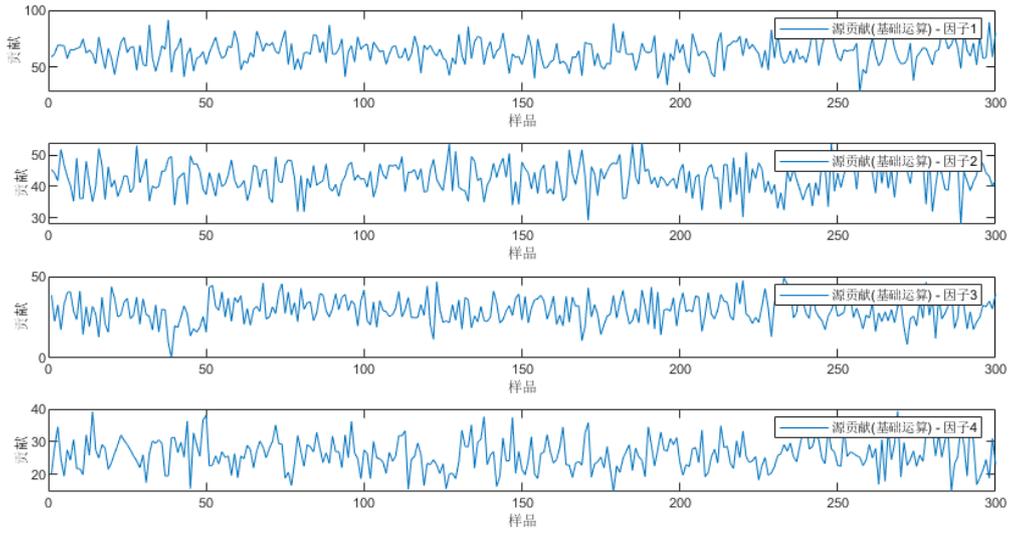


图9 源贡献时间序列图

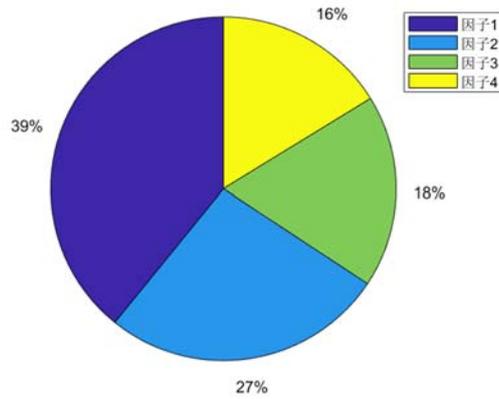


图10 源分担率饼图

6.2.6 旋转计算

(1) 模型旋转计算操作步骤

在基础运算结果的基础上，可以在旋转运算版块设定 F_{peak} 值。一般先设置 F_{peak} 值为 -1 到 1 之间，点击运算，得到旋转运算的结果，如图 11 所示。

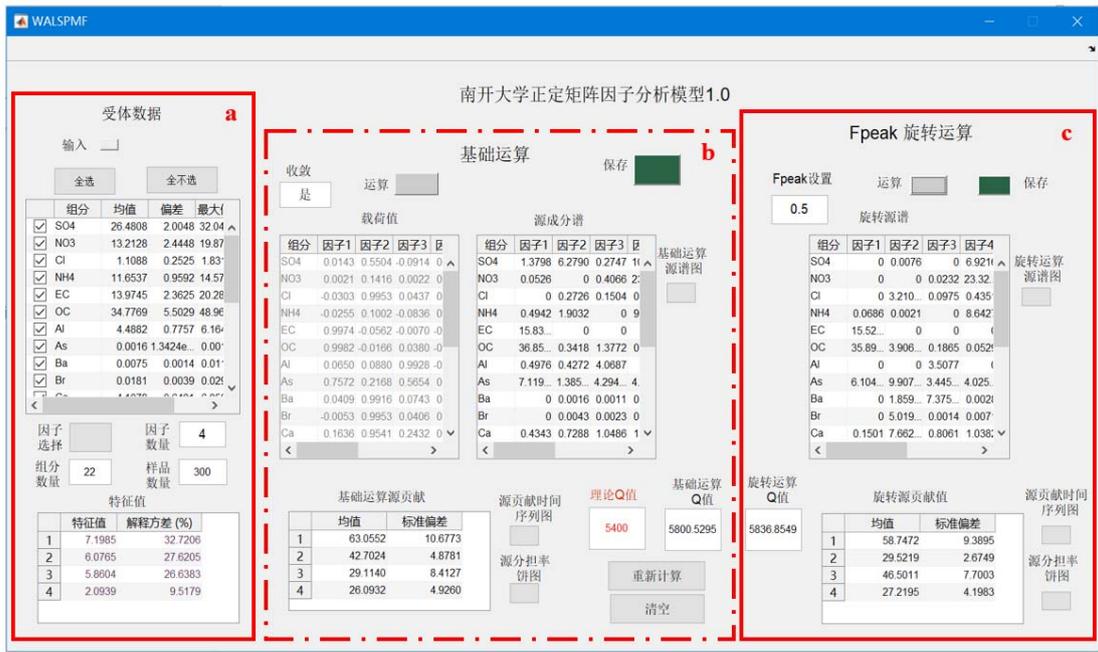


图11 WALSPMF 1.0 的旋转计算运行结果展示

(2) 模型旋转运行结果展示

旋转运行结果包括：旋转运算源谱数据和图（图 12）源贡献时间序列图（图 13）；源分担率数据和饼图（图 14）以及相对应的旋转运算结果展示。此外，模型提供了旋转后的计算 Q 值，用户可以根据该值与基础运算结果中理论 Q 值的差值大致判断旋转程度的好坏。

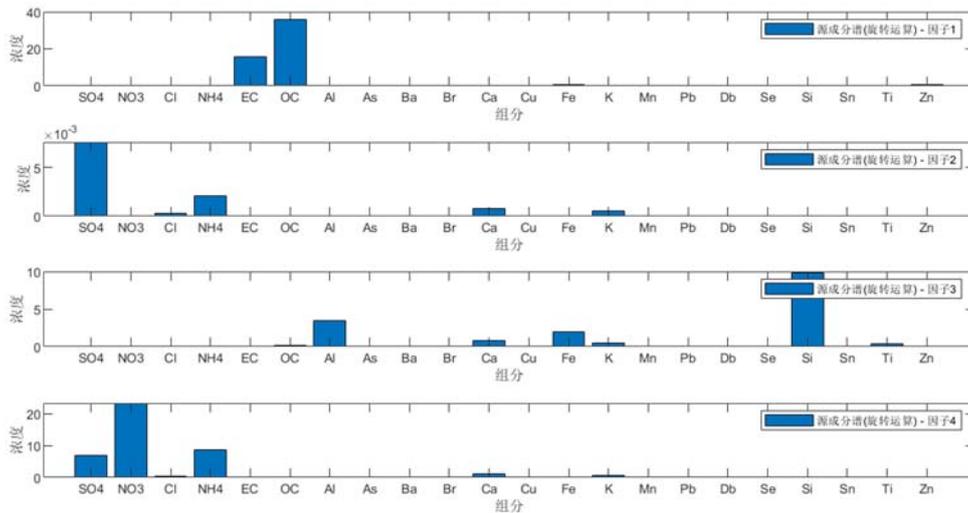


图12 旋转运算源谱图

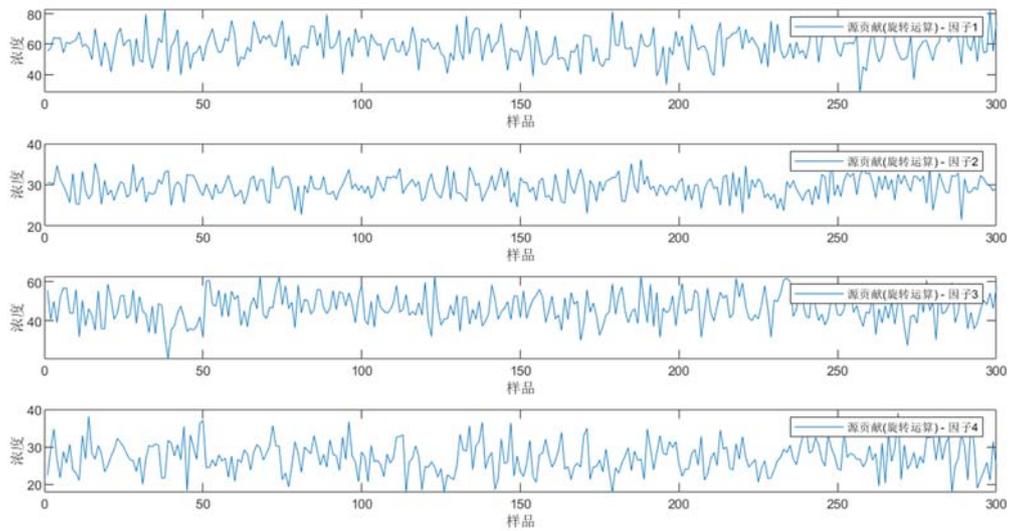


图13 源贡献时间序列图

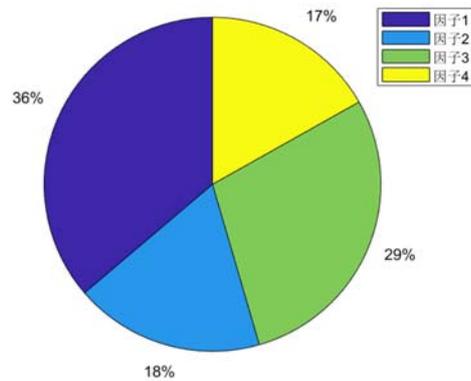


图14 源分担率饼图

6.2.7 输出结果导出

点击分别点击基础运算部分和 Fpeak 旋转运算部分中的保存结果按钮，输出文档会保存为“baserunoutput（基础运算结果）”和“Fpeakoutput（旋转运算结果）”的两个“.xls”的文件。文件中分别有：源成分谱、G矩阵、源贡献、源贡献均值、Q值和 Fpeak 值的工作簿，可供用户查看结果。

6.3 EPA PMF 5.0 计算步骤

6.3.1 软件下载

可在美国 EPA 网站免费下载，下载地址：<https://www.epa.gov/air-research/positive-matrix-factorization-model-environmental-data-analyses>。

6.3.2 运行环境

计算机（Win7 以上系统，3.1 GHz 以上处理器，4 GB 以上 RAM，需安装 framework 4.0 以上版本）。

6.3.3 数据准备

3 种格式的数据文件可输入至 EPA PMF5.0 模型，包括 tab 间隔的 txt 文件、逗号间隔的 csv 文件和 excel 文件，推荐使用 excel 文件。

PMF 模型需要输入两个数据文件：（1）颗粒物及组分质量浓度的数据文件，（2）颗粒物及组分质量浓度不确定度的数据文件。两个数据文件内容需要对应。第一行为组分名称，组分名称必须是唯一，中间不能含有逗号。第一列为样品编号，如样品 ID、日期或时间、样品 ID 加上日期时间等。两个数据文件均不允许有数据缺失、0 值和负值的情况。对于样品部分组分浓度未检出的情况，可用 1/2 检出限来代替；对于样品数据中部分组分数据缺失的情况，建议输入-999 来标识，对应的不确定度可暂用 5/6 检出限代替（模型计算时会自动替代）。

6.3.4 数据导入和初步检验

6.3.4.1 数据导入

启动 PMF 模型软件，进入“Model Data”界面（见图 15）。在“Data Files”的“Input Files”下选择样品组分浓度文件和不确定度文件的输入路径，点击“Load”（导入）选择相应的数据文件。在“Output Files”中，设置确定输出文件的路径、文件类型和输出文件前缀（Output File Prefix），输出文件前缀（Output File Prefix）只能是数字或字母。如对数据文件进行了修改，则需要重新加载输入文件。

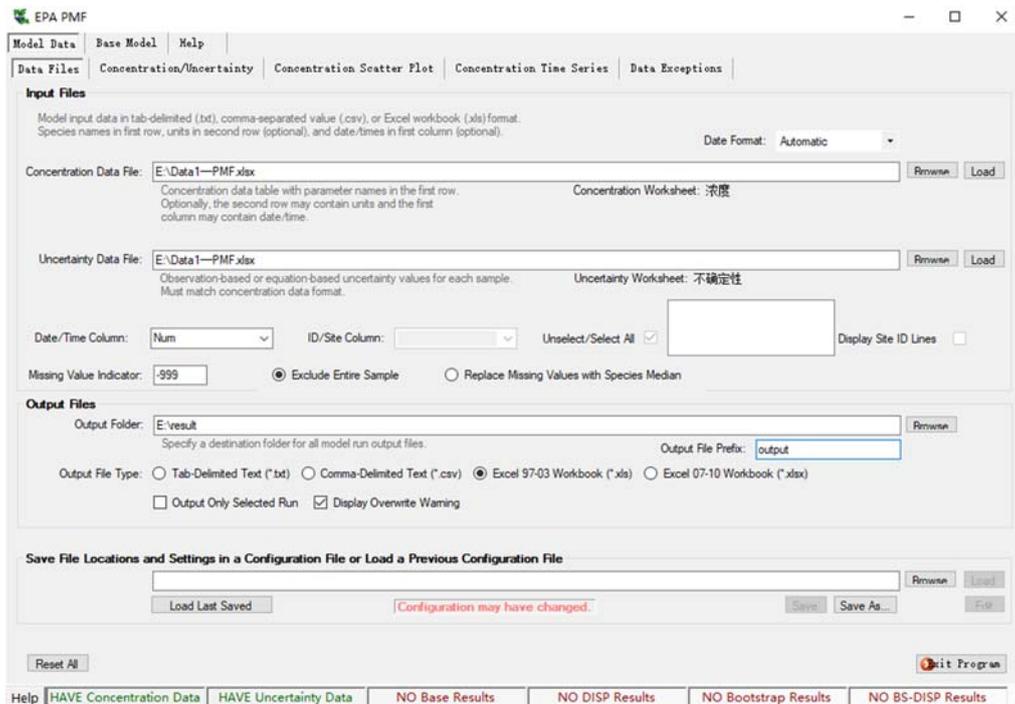


图15 数据导入的设置界面

PMF 模型中可以保存新的配置文件或者载入之前的配置文件。配置文件中的信息包括“Data Files”中的各类设置（例如，输入文件，输出文件位置和输出文件类型），“Concentration/Uncertainty”（浓度/不确定度）的设置以及“Base Model Run”（模型基础计算）、“Fpeak Rotation”（F 矩阵峰值旋转）、“Constrained Model Runs”（模型约束计算）中的所有运行参数。

6.3.4.2 初步检验

（1）在“Concentration/Uncertainty”中对受体化学组分数据的权重、异常值等进行设置或处理（界面见图 16）。根据 S/N（信噪比）值确定 PMF 中组分的计算权重，权重包括 Bad（差）、Weak（弱）和 Strong（强）三个水平。对于被选择 Strong 的组分，按输入的浓度和不确定度参与计算；对于被选择 Weak 的组分，模型在计算时会自动增加该组分的不确定度，进而降低该组分浓度的权重；对于被选择 bad 的组分，则不参与计算。建议将颗粒物的关键标识组分（如 OC、EC、SO₄²⁻、NO₃⁻、NH₄⁺、Al、Si、Ca、K、Fe 等组分）设为 Strong。其他非关键标识组分（如：Cl⁻、Na、Mg、Cu、Cd、Cr 等组分）建议设置 Strong，但是在分析不确定性较高时，可设置成 Weak 参与计算。当某一组分的缺失数据较多时，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。此外，计算各因子对颗粒物的贡献时，需要将颗粒物质量浓度数据输入模型，并将颗粒物质量浓度设置为 Total Variable（总变量），其计算权重默认为 Weak。

（2）在“Concentration Scatter Plot”（浓度散点图）中可选择任意两个组分作为 X 和 Y 轴，查看两个组分之间的散点图以及相关性。组分之间如果存在较高的线性关系表明组分可能来自同一污染源。

（3）在“Concentration TimeSeries”（浓度时间序列图）中显示化学组分浓度的时间变化情况，可用于查看组分浓度的时间变化情况，确定是否存在异常值（过高或过低）。可最多同时选择查看 10 个组分。对于所选定的异常数据，可以在图上单击鼠标右键或在某一时间段数据上拖动鼠标选择异常数据点，将其排除在模型计算之外（注：该操作不会在原始输入文件中删除数据）。

（4）在“Data Exceptions”（数据异常处理）中详细记录“Model Data”中对数据所做的更改，包括将组分列为“Weak”或“Bad”，在“Concentration Time Series”排除的异常数据等。单击鼠标右键可选择保存数据异常处理信息。

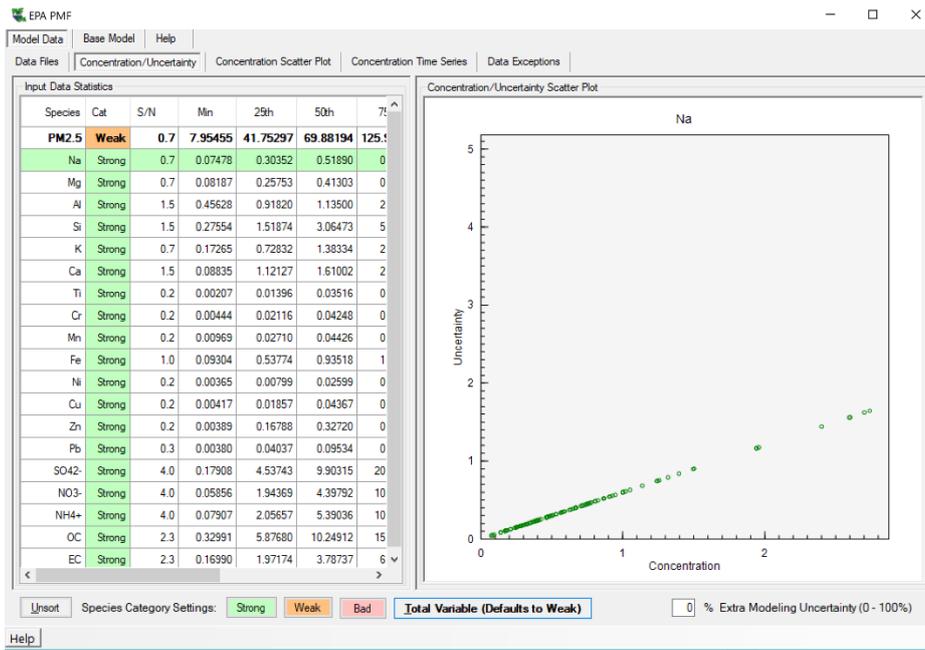


图16 数据初步检验界面

6.3.5 基础计算

6.3.5.1 基础计算的操作步骤

在经过初步检验之后，可对数据进行模型计算。PMF 模型计算分为模型基础计算（Base Model Run）和约束模型计算（Constrained Model Run），其中模型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 EPA PMF 5.0 模型基础计算中，需要设定因子（Factor）个数、运行次数、是否随机选择 seed（种子）等运行参数（图 17，a）。

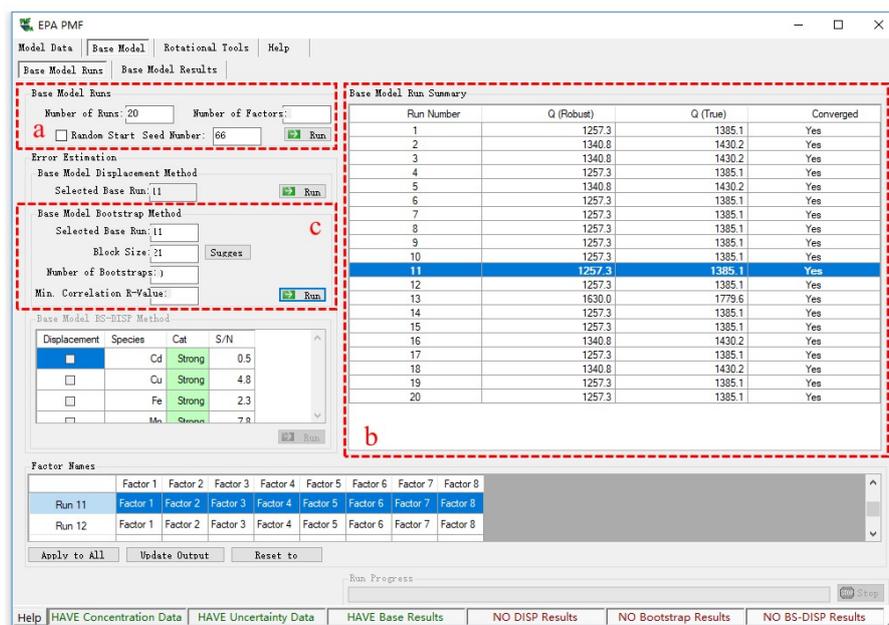


图17 模型基础计算界面

(a: 模型基础计算, b: 基础计算结果概况, c: 基础计算结果评估)

在模型基础计算中，需要设置运行次数（Number of Runs），该数值必须介于 1~999 之间的整数，模型默认值是 20。由于 PMF 模型的特点是一套数据可以得到多套计算结果。通过设置运行次数，模型就输出多少个计算结果，供用户选择。

设置种子（Seed），即模型基础计算中每次迭代的起点，建议选择默认。PMF 模型原理决定了模型会有多个符合公式（1）要求的结果。

设置因子数量，即设定“源类”的个数，该数量必须介于 1~999 之间的整数。因子的数量需要基于对颗粒物可能来源、样品数量、采集时间等认识来综合确定，需要多次反复尝试。此外也可以通过数学验证和实际情况等方面进行验证和确定。

6.3.5.2 模型基础计算结果及分析

（1）基础计算结果展示

在“Base Model Run Summary”（基础计算结果概况）（图 17，b）中，可每次计算结果的 Q（Robust）、Q（True）、是否收敛（Converged）等。点击选择一次计算结果（图 17，b），在“Base Model Results”可以查看详细的计算结果，包括“残差分析”、“预测值与观测值对比分析”、“因子谱”、“因子贡献”等（见图 18）。

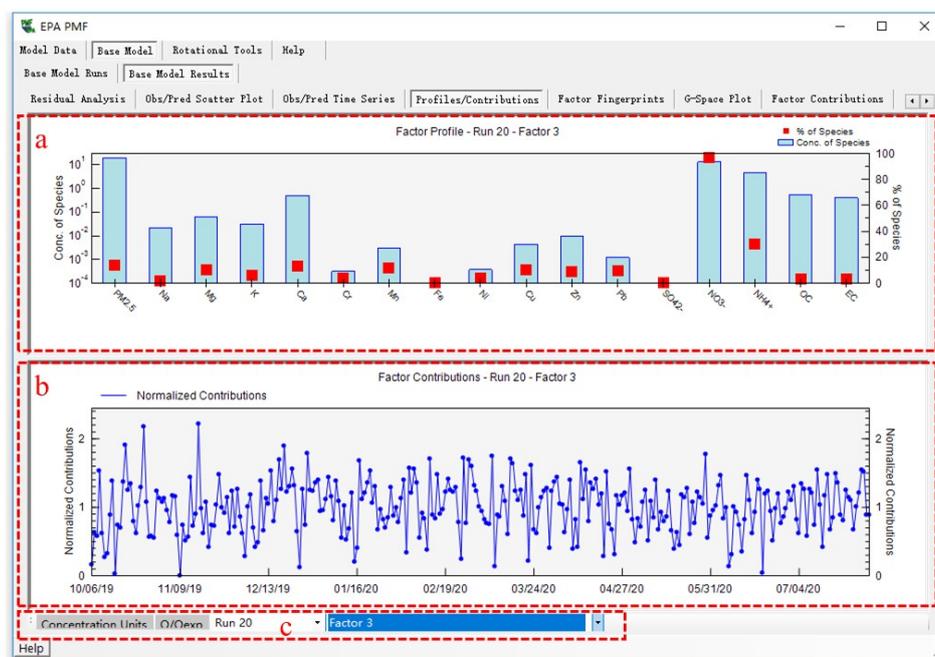


图18 基础计算结果展示中“Profiles/Contributions (因子谱/因子贡献)”界面

PMF 模型基础计算的结果主要包括因子谱和因子贡献，分别在“Profiles/Contributions（因子谱/贡献）”和“Factor Contribution（因子贡献）”中查看。在“Profiles/Contributions（因子谱/因子贡献）”界面中，因子谱图（即 F 矩阵，见图 18，a）展示了不同化学组分在各因子中的占比及浓度。用户可以通过分析不同因子谱中何种标识组分占比较高来将因子识别为具体的污染源类。比如，某因子中 Si、Al 等组分占比较高，则该因子可以初步识别为扬尘源类；如果某因子中 NH₄⁺、SO₄²⁻等组分占比较高，则该因子可以初步识别为二次硫酸盐。颗粒物源类识别方法及主要源类的标识组分见附录 D。

源贡献时间序列图（G矩阵，见图 18，b）则展示了污染源在时间序列上的贡献变化情况。在 EPA PMF 5.0 中，给出的源贡献（G 矩阵）是归一化后的源贡献变化序列（源贡献均值为 1），而不是浓度值。此外，可以根据源贡献序列识别源类。对于有些源类通过因子谱很难区分的情况，如机动车源和燃煤源（其主要标识组分均为 OC、EC 等），则可通过源贡献图进行区分，在某些地区，燃煤源存在显著的季节变化（冬季采暖），而机动车源季节变化不明显。

用户可以在“Profiles/Contributions（因子谱/因子贡献）”界面右击鼠标，选择“因子名称”，输入因子的名称，来对因子进行标注。点击“Q/Q_{exp}”（见图 18，c），查看每一个组分所对应的 Q/Q_{exp}。Q/Q_{exp} 是指所有组分的比例残差的平方之和除以组分数。通过检验 Q/Q_{exp}，可以了解 PMF 结果的综合残差。当某个组分或者某个样品的 Q/Q_{exp} 较大时（如大于 2），说明该组分或者样品的残差比较大，模拟结果不好，可能需要设置更多的因子（源类）。

“Factor Contribution”（因子贡献）中，给出了不同源类对各组分的平均贡献（饼图）及时间序列贡献（点线图），用于评估源贡献结果。

（2）基础计算结果分析

模型计算结果分析过程如下：

1) “Residual Analysis（残差分析）”中查看每一个组分的加权残差（通过不确定度加权）。通过残差直方图（见图 19），可以了解该组分加权残差的百分比，进而判断各组分模型计算结果的好坏。如果残差直方图显示残差范围在（-3，+3）之间，且呈正态分布，则该组分的模型计算结果在数学意义上较好；如果组分具有许多大尺度残差或显示非正常曲线，则表示该组分在数学意义上计算较差。

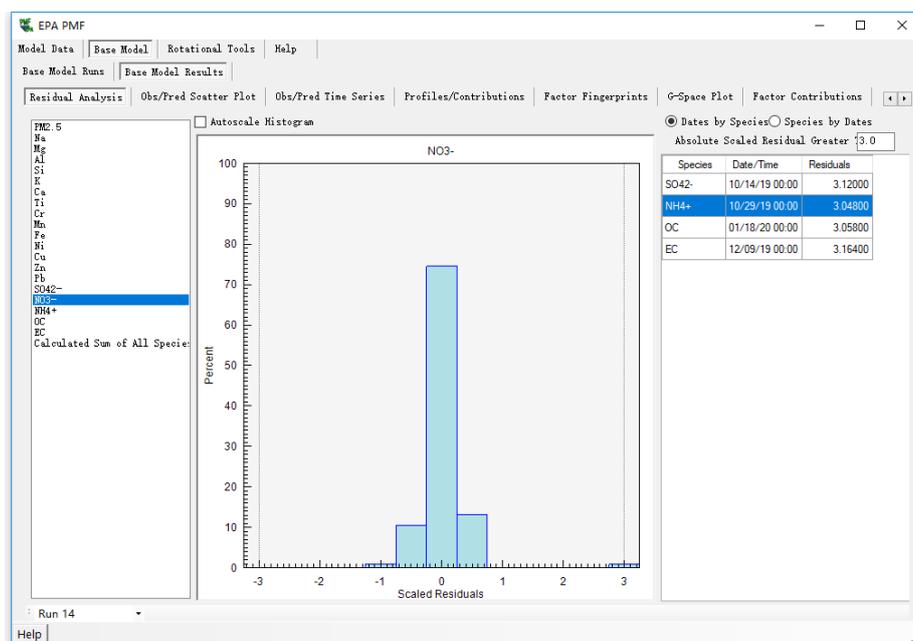


图19 基础计算结果分析中“Residual Analysis（残差分析）”界面

2) 在“Obs/Pred Scatter Plot (观测值/预测值散点图)”中, 可分析输入值和预测值的相关性。如果某组分的观察值和预测值之间有强相关性的组分, 则表明该组分计算较好; 反之, 则考虑是否应该降低该组分计算权重或从计算中排除。

3) 在“Obs/Pred Time Series (观测值/预测值时间变化图)”中, 可在时间序列上分析输入值和预测值的差异, 对于预测值明显高于输入值, 需要通过进一步分析确定是否排除该数据。

4) 在“Factor Fingerprints (因子谱)”界面中, 模型给出了各组分在每个因子的占比情况(柱状图), 用户可直观评估因子谱。

5) 在“G-Space Plot”中, 可显示两个因子的散点图, 用于了解各源贡献之间的关系。

(3) 基础计算的结果输出

模型基础计算完成后, 将自动创建输出文件。输出文件将使用“输出文件前缀”框中指定的前缀, 保存到“数据文件”屏幕的“输出文件夹”框中指定的目录中。创建的输出文件数取决于所选输出文件的类型。如果选择输出制表符分隔 (*.txt) 或逗号分隔 (*.csv) 的文件类型, 可生成五个输出文件: 诊断文件(* _diag), 因子贡献(* _contrib), 因子谱(* _profile), 残差(* _resid)和结果对比(* _run_comparison); 如果选择输出 Excel 工作簿 (*.xls) 的文件类型, 可生成两个输出文件: * _diag 和* _base。

各输出文件的内容如下:

* _diag 文件包含用户输入和模型诊断信息的记录(与“诊断”显示相同)。

* _contrib 文件包含了源贡献矩阵信息, 默认显示归一化贡献(源贡献均值为 1, 无量纲); 如果指定了总变量(颗粒物浓度), 则以质量单位显示贡献。

* _profile 文件包含了因子成分谱矩阵, 主要包括三种输出形式: 1) 质量浓度矩阵, 结果中各数据为各因子中不同化学组分的质量浓度 ($\mu\text{g}/\text{m}^3$); 2) 横向百分比归一化矩阵, 即各组分在所有因子中占比加和为 100; 3) 纵向百分比归一化矩阵, 即各因子中, 所有组分占比加和为 100。

* _resid 文件包含每个基本运行的残差(常规和加权残差), 用于在残差分析屏幕上生成图形和表格。

* _run_comparison 文件包含所有 PMF 运行中每个因子的组分分布摘要, 并与最低 Q (稳健) 运行进行比较。

* _base 文件在同一 Excel 工作簿内的工作表中包含* _contrib, * _profile, * _resid 和* _run_comparison。仅当用户选择“Excel Workbook”作为输出文件类型时, 才会显示此输出文件。

6.3.5.3 基础计算的结果误差评估

在得到模型基础计算结果之后, 需要评估结果误差, 以此判断结果的优劣程度。EPA PMF 5.0 的结果误差评估方法主要有 Bootstrapping (BS) 误差评估、DISP 误差评估、BS-DISP 误差评估等。推荐采用 BS 误差评估方法评估模型计算结果(不确定度)的误差。

BS 误差评估设置包括 3 个步骤(设置界面见图 17, c):

(1) 选择需要评估的模型基础计算结果。

(2) 设置每次“重采样” (Block Size) 的数量, 模型会根据输入数据自动给出该参数, 用户也可以自行设置。

(3) 设置 BS 运行的次数和最小相关性系数 (R)。对于初步分析, BS 运行次数可设置 50 次, 以快速衡量解的稳定性, 最终分析建议执行 100 次, 以确保统计数据稳定性。最小相关性系数 (R) 用于将一个 BS 运行因子分配给一个基本运行因子的最小 Pearson 相关系数, 默认值为 0.6。

通过 BS 误差评估, 可在“Base Model Bootstrap Results”下得到两种形式的结果: Bootstrap Box Plots 结果 (图 20) 和 Bootstrap Summary 结果 (图 21)。其中 Bootstrap Box Plot (图 20) 中的箱型图可以直观的描述组分模型计算结果的不确定度。每个因子中, 某些组分的箱型图上下距较宽, 则表明该组分模型计算的结果的不确定度可能较高; 如果某些组分的箱型图上下距较窄, 则说明该组分模型计算结果的不确定度可能较低。

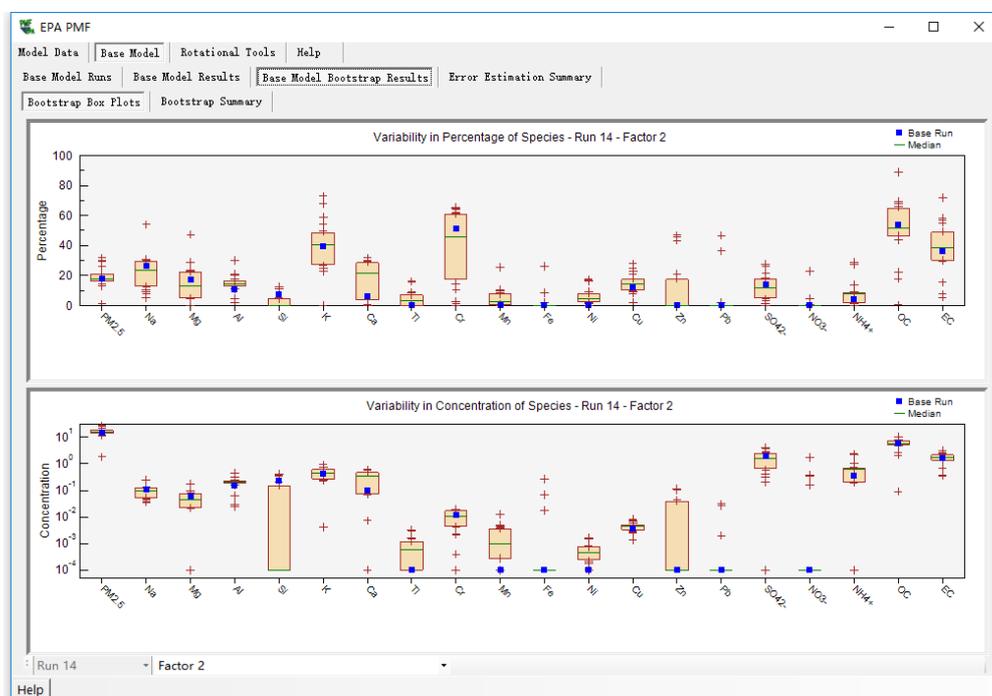


图20 BS 误差评估结果 “Base Bootstrap Box Plots”

在 Bootstrap Summary 中, 展示了每个 BS 因子和基础因子 (Base Factor) 匹配程度 (图 B.7, a)。如果匹配不到 80%, 则这个因子的主要组分分配可能不当。通过调整因子数量或提高数据质量等提高匹配程度。另外还可以查看 Q(稳定)值的分布情况 (图 B.7, b), 包括最小、最大、中位数, 以及第 25 和第 75 百分位数的 Q(稳定)值。模型基础计算得到的 Q 值一般要落在第 25 和第 75 百分位数之内。

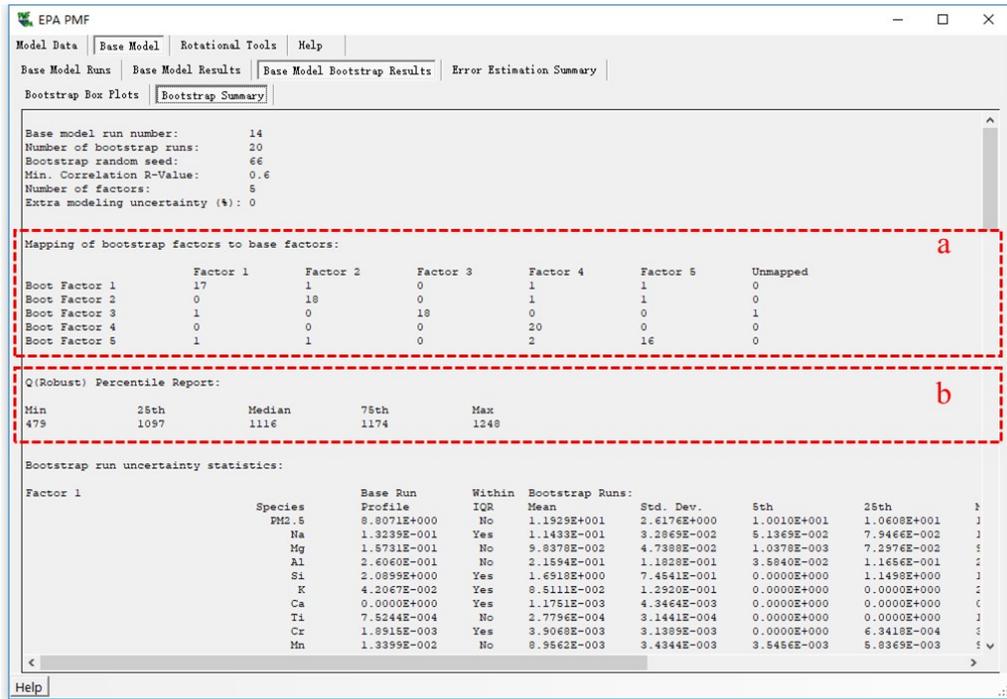


图21 BS 误差评估结果 “Base Bootstrap Summary”

6.3.6 旋转计算

EPA PMF 5.0 软件中，在 F 矩阵峰值模型计算参数设置界面（图 22，a），选择需要进行旋转计算的模型基础计算结果，再设置 Fpeak 参数。当 Fpeak 设为正值时，F 矩阵中的组分在某些因子中占比会更加突出；Fpeak 设为负值时，F 矩阵中的组分在各因子中的占比则相对较平均。设置 Fpeak 时，Q 值不能变化太大，Q 值的变化在 5% 以内认为时可以接受的；如果过大，说明 Fpeak 值（绝对值）设置过高，则 Fpeak 应往 0 值方向调整。Fpeak 参数建议设置在 -5 到 5，不能设置为 0。

通过尝试不同 Fpeak 值，得到相应的 Q 值，最后得到 Fpeak 值与 Q 值的关系图（图 23）。通过该图可选择合适的 Fpeak 值，建议选择 Q 值拐点之前的 Fpeak 值对应的解析结果，此时的 F 矩阵中的丰量组分会更突出，因此各因子之间差异性更大，便于区分源类。

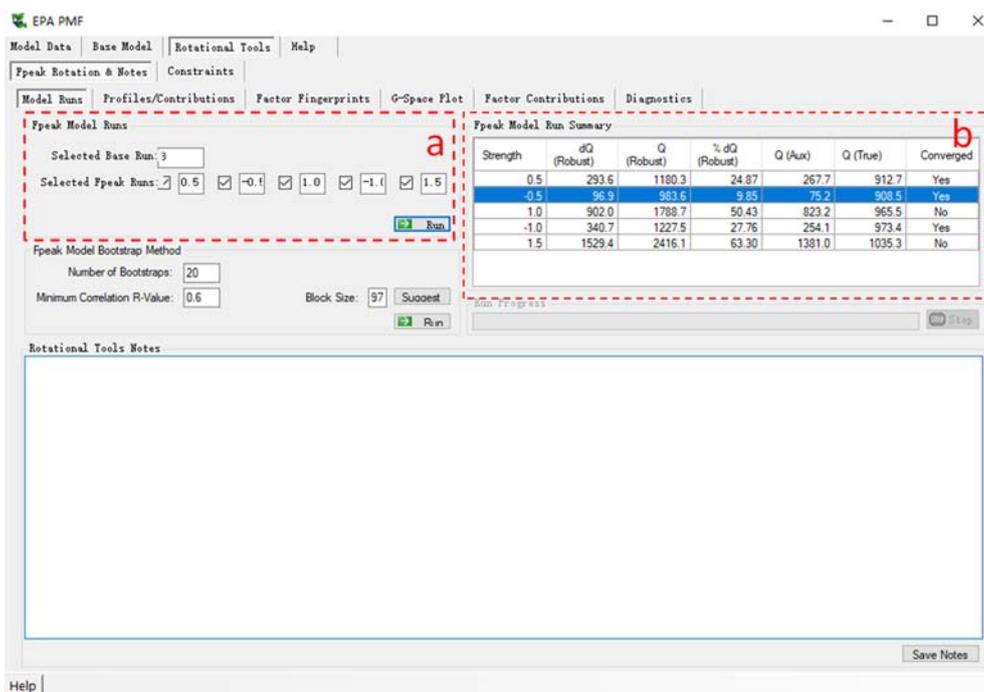


图22 Fpeak 旋转设置及计算结果界面

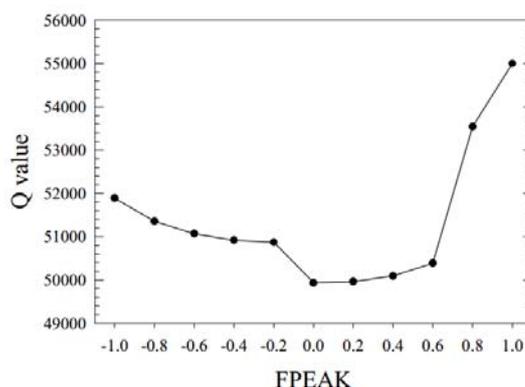


图23 Fpeak 值与 Q 值关系示例图

在 Fpeak 旋转计算得到的总体结果（图 21，b）中，在收敛（Yes）的结果中选择较小的 %dQ 值，若 %dQ 值不在 5% 以内时，需要进一步设置 Fpeak 参数。Fpeak 旋转计算后的详细结果展示与“模型基础计算结果展示”相似，包括“Profiles/Contributions(因子谱/因子贡献)”（见图 24）”、“Factor Fingerprints（因子指纹）”、“G-Space Plot（G 矩阵散点图）”、“Factor Contributions（因子贡献）”、“Diagnostics（诊断）”。



图24 经 Fpeak 旋转后的结果展示界面

6.3.7 颗粒物源贡献结果计算

在 EPA PMF 5.0 中，在“1 数据准备”中将颗粒物质量浓度数据和化学组分数据一起输入到 PMF 模型进行计算，将颗粒物质量浓度设置为总变量(Total Variable)。最后可通过 EPA PMF 5.0 模型直接得到污染源对颗粒物的源贡献（见图 25）。

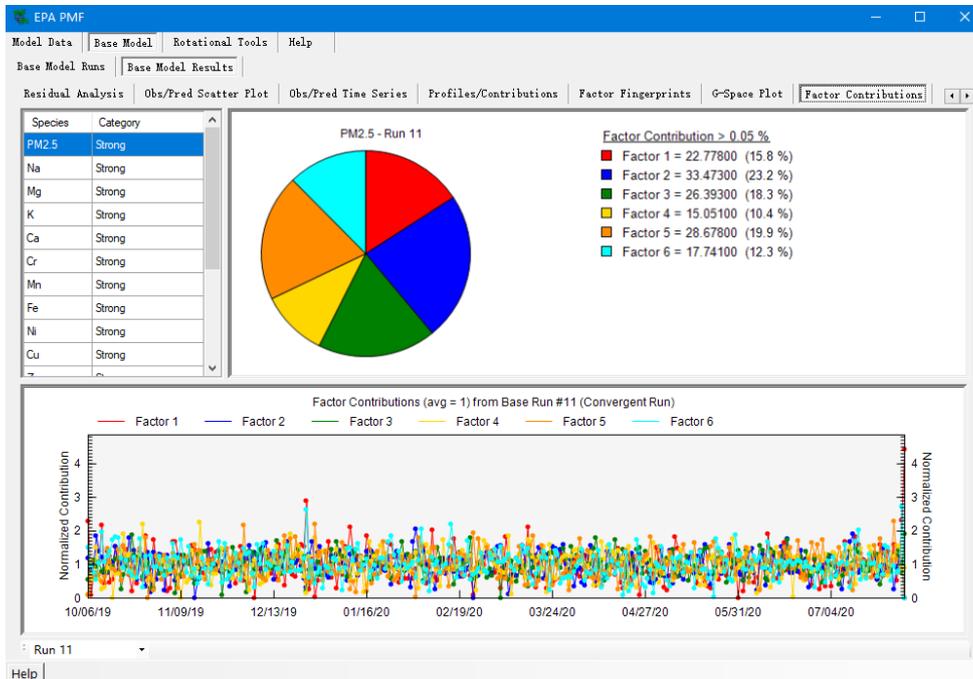


图25 污染源对颗粒物的源贡献

6.3.8 参考资料

《EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide》

7 实际应用检验

7.1 应用过程

为了进一步考察本标准的科学性和适用性，中国环境监测总站、福建省福州市环境监测中心站、西安市环境监测站共 5 名技术人员开展方法验证工作。验证方法是参照本标准对预先准备的一套实际样品的颗粒物源解析组分数据进行 PMF 模型计算，比较计算结果。

编制组准备的数据为某城市环境空气 PM_{2.5} 滤膜样品的监测结果，其中 OC/EC 的分析方法为热光反射法，水溶性离子的分析方法为离子色谱法，无机元素的分析方法为 ICP-MS 法。最终参与模型计算的组分包括 Na、Mg、Al、Si、K、Ca、Ti、Cr、Mn、Fe、Ni、Cu、Zn、Pb、SO₄²⁻、NO₃⁻、NH₄⁺、OC 和 EC 等 19 个组分。该城市的污染源类可分为燃煤、扬尘、机动车、二次硫酸盐和硝酸盐源类。

此外，编制组还向验证人员提供了该城市的一些信息，便于更好地对计算结果进行分析判断：测试数据来自中国某城市，采样点所在区县基准年三产占比为：第一产业（5.5%）、第二产业（45.7%）、第三产业（48.8%）。工业构成以轻工业为主，自然风光优，建成区面积 401 平方公里，建筑业企业房屋施工面积 921.68 万平方米，房屋建筑竣工面积 93.41 万平方米。能源消耗量 48 万吨标准煤；运输企业运力总量 7286 车。该地区首要污染物主要为细颗粒物（PM_{2.5}）和臭氧（O₃_{8h}）。

7.2 应用结果

5 名技术人员的验证结果见表 5 所示。验证中均采用了约束模型计算，设定的旋转系数从-0.4~0.1 之间，所得到的 Q_(Ture)和 Q_(Robust)差异较小。参照本标准，对同一数据进行 PMF 模型计算，得到 4 类主要污染来源：二次源、燃煤、机动车、扬尘。5 个验证结果的污染源贡献占比存在一定差异，但是排序一致，均为二次源>燃煤>机动车>扬尘，源贡献占比分别为 21.7%~31.8%、20.0%~23.5%、18.5%~20.7%、7.2%~13.2%，RSD 分别为 16.1%、6.3%、4.3%、26.5%。验证结果显示，本标准可供监测人员开展 PMF 模型源解析计算，由于经验不同，不同人员的计算结果会存在一定差异，但是源贡献排序可达到一致。

根据验证工作中的反馈意见，标准编制组进一步修改完善标准文本和编制说明，最终形成了《环境空气颗粒物来源解析正定矩阵因子分解（PMF）模型计算技术指南》文本和编制说明的征求意见稿。

表5 PMF 模型计算验证结果

PMF 模型计算		验证结果 1	验证结果 2	验证结果 3	验证结果 4	验证结果 5
计算方法	约束模型计算	是	是	是	是	是
	旋转系数	-0.4	0.1	-0.05	-0.1	-0.1
	Q _(Ture)	1114.2	1109.2	1101.5	1102	1102
	Q _(Robust)	1155.2	1133	1102.3	1105.2	1105.2
源贡献占比 (%)	二次源	21.7	24.9	31.8	31.6	30.3
	燃煤	20.0	23.5	21.8	21.4	20.5
	机动车	18.9	20.7	19.0	19.3	18.5

	扬尘	7.2	8.4	13.2	13.2	12.6
	其他	19.9	22.5	14.3	14.6	14.0

8 参考文献

- [1] Thurston GD, Spengler JD. A Quantitative Assessment of Source Contribution to Inhalable Particulate Matter Pollution in Metropolitan Boston[J]. Atmospheric Environment,1985, 19: 9–25.
- [2] Karar K, Gupta AK. Source apportionment of PM₁₀ at residential and industrial sites of an urban region of Kolkata, India[J]. Atmospheric Research, 2007, 84:30-41.
- [3] Wang ZS, Wu T, Shi GL, et al. Potential Source Analysis for PM₁₀ and PM_{2.5} in Autumn in a Northern City in China[J]. Aerosol and Air Quality Research, 2012, 12:39–48.
- [4] Xu J, Tian YZ, Zhang Y, et al. Source apportionment of perfluorinated compounds (PFCs) in sediments: Using three multivariate factor analysis receptor models[J]. Journal of Hazardous Materials 2013, 260:483-8.
- [5] Paatero P, Tapper U. Positive matrix factorization: a non-negative factor model with optimal utilisation of error estimates of data values [J]. Environmetrics, 1994, 5: 111-126.
- [6] Paatero, P. Least squares formulation of robust non-negative factor analysis. Chemometrics and Intelligent Laboratory Systems, 1997, 37: 23-35.
- [7] Hwang I, Hopke PK. Estimation of source apportionment and potential source locations of PM_{2.5} at a west coastal IMPROVE site[J]. Atmospheric Environment 2007, 41:506-18.
- [8] Khairy MA, Lohmann R. Source apportionment and risk assessment of polycyclic aromatic hydrocarbons in the atmospheric environment of Alexandria, Egypt[J]. Chemosphere 2013, 91:895–903.
- [9] Shi, GL, Xu, J , Peng, X , et al. Using a new WALSPMF model to quantify the source contributions to PM_{2.5} at a harbour site in China[J]. Atmospheric Environment, 2016, 126:66-75.
- [10] 宋宇,唐孝炎,方晨,等.北京市大气细粒子的来源分析[J]. 环境科学,2002,23(6):11 - 16.
- [11] Xie, S.D., Liu, Z., Chen, T., et al. Spatiotemporal variations of ambient PM₁₀ source contributions in Beijing in 2004 using positive matrix factorization. Atmospheric Chemistry Physics,2008, 8:2701–2716.
- [12] Song, Y., Tang, X.Y., Xie, S.D., et al. Source apportionment of PM_{2.5} in Beijing in 2004. Journal of Hazardous Materials,2007,146, 124–130.
- [13] Xie, M.J., Wang, G.H., Hu, S.Y., et al. Polar organic and inorganic markers in PM₁₀ aerosols from an inland city of China—Seasonal trends and sources. Science of the Total Environment, 2010, 408(22):5452–5460.
- [14] Cai, C.J., Geng, F.H., Tie, X.X., et al. Characteristics and source apportionment of VOCs measured in Shanghai, China. Atmospheric Environment,2010,44,5005-5014.

- [15] 郇涛,谢绍东,薄宇,等. 正矩阵因子分解法评价北京市 PM₁₀ 和 SO₂ 监测网络[J]. 环境科学, 2010,31(3):560-565.
- [16] Cheng, Y., Zou, S.C., Lee, S.C., et al. Characteristics and source apportionment of PM₁ emissions at a roadside Station. *Journal of Hazardous Materials*, 2011,195:82–91.
- [17] 成海容,王祖武,冯家良,等. 武汉市城区大气 PM_{2.5} 的碳组分与源解析[J]. 生态环境学报,2012,21(9):1574-1579.
- [18] 程水源,刘超,韩力慧. 北京及周边城市典型区域 PM_{2.5} 化学组分特征与来源分析[J]. 中国环境科学学会学术年会,2013.
- [19] Tian, Y.Z., Wu, J.H., Shi, G.L., et al. Long-term variation of the levels, compositions and sources of size-resolved particulate matter in a megacity in China. *Science of the total environment*, 2013, 463-464(1):462-468.
- [20] 高健,李慧,史国良,等. 颗粒物动态源解析方法综述与应用展望[J]. 科学通报, 2016, 61(27):3002.
- [21] 王晓浩, 赵倩彪, 崔虎雄. 基于在线监测的上海郊区冬季 PM_{2.5} 来源解析[J]. 南京大学学报(自然科学), 2015, 51: 517-523
- [22] 王古月,薛风云. 镇江市秋季大气颗粒物 PM_{2.5} 实时在线来源解析[J]. 环保科技,2017,23(03):28-31.
- [23] Gao J, Zhang Y, Zhang M, et al. Photochemical properties and source of pollutants during continuous pollution episodes in Beijing[J]. *J Environ Sci*, 2014, 26: 44–53
- [24] 周敏. 上海大气 PM_{2.5} 来源解析对比:基于在线数据运用 3 种受体模型[J]. 环境科学,2020,41(05):1997-2005.
- [25] 陈焯, 刘方, 薛荔栋,等. 重点区域大气颗粒物源解析监测发展现状. 2019 中国环境科学学会科学技术年会论文集 (第三卷) .
- [26] Hopke, P.K. A Guide to Positive Matrix Factorization. 2000. Clarkson University-Department of Chemistry. <https://www3.epa.gov/ttn/amtic/files/ambient/pm25/workshop/laymen.pdf>
- [27] Gary Norris, Rachele Duvall. EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide. 2014. U.S. Environmental Protection Agency. https://www.epa.gov/sites/default/files/2015-02/documents/pmf_5.0_user_guide.pdf
- [28] Claudio A. Belis, Bo R. Larsen, Fulvio Amato, et al. European Guide on Air Pollution Source Apportionment with Receptor Models. 2014. European Commission. http://source-apportionment.jrc.ec.europa.eu/Docu/EU_guide_on_SA.pdf
- [29] 张延君,郑玫,蔡靖,等.PM_{2.5} 源解析方法的比较与评述[J].科学通报,2015,60(02):109-121.
- [30] Wang F, Yu H, Wang Z, et al. Review of online source apportionment research based on observation for ambient particulate matter[J]. *Science of The Total Environment*, 2020, 762:144095.
- [31] 郑玫,张延君,闫才青,等.中国 PM_{2.5} 来源解析方法综述[J].北京大学学报(自然科学版),2014,50(06):1141-1154.
- [32] Zheng M, Yan C, Wang S, et al. Understanding PM_{2.5} sources in China: Challenges and

- perspectives[J]. *National Science Review*, 2017, 4(6):801-803.
- [33] Liu Z, Wang Y, Hu B, et al. Source apportionment of fine particle number and volume concentration during severe haze pollution in Beijing in January 2013[J]. *Environmental Science & Pollution Research*, 2016, 23(7):6845-6860.
- [34] 蔡凡涛,尚玥,戴维,等. PM_{2.5} 化学组分连续观测在污染事件源解析中的应用[J].*环境科学*, 2021,42(10): 4575-4581.
- [35] Hui L, Liu X, Tan Q, et al. Characteristics, source apportionment and contribution of VOCs to ozone formation in Wuhan, Central China[J]. *Atmospheric Environment*, 2018, 192, 55-71
- [36] Zhang X, Yin Y, Wen J, et al. Characteristics, reactivity and source apportionment of ambient volatile organic compounds (VOCs) in a typical tourist city[J]. *Atmospheric environment*, 2019, 215:116898.1-116898.13.
- [37] Liu Y, Wang H, Jing S, et al. Characteristics and sources of volatile organic compounds (VOCs) in Shanghai during summer: Implications of regional transport[J]. *Atmospheric environment*, 2019, 215(Oct.):116902.1-116902.14.
- [38] Gu Y, Liu B, Li Y, et al. Multi-scale volatile organic compound (VOC) source apportionment in Tianjin, China, using a receptor model coupled with 1-hr resolution data[J]. *Environmental Pollution*, 2020, 265(Pt A):115023.
- [39] Yi A, Ms A, Xi A, et al. Characterization and sources of volatile organic compounds (VOCs) and their related changes during ozone pollution days in 2016 in Beijing, China[J]. *Environmental Pollution*, 2019, 257:113599
- [40] Zheng H, Kong S, Yan Y, et al. Compositions, sources and health risks of ambient volatile organic compounds (VOCs) at a petrochemical industrial park along the Yangtze River[J]. *Science of The Total Environment*, 2019, 703:135505.
- [41] Yong Jie Li, Yele Sun, Qi Zhang, et al. Real-time chemical characterization of atmospheric particulate matter in China: A review[J]. *Atmospheric environment*, 2017, 158, 270-304.
- [42] Philip K. Hopke. Review of receptor modeling methods for source apportionment, *Journal of the Air & Waste Management Association*, 2016, 66:3, 237-259.
- [43] A Y Z, A J C, B S W, et al. Review of receptor-based source apportionment research of fine particulate matter and its challenges in China - ScienceDirect[J]. *Science of The Total Environment*, 2017, 586:917-929.
- [44] Hopke P K, Dai Q, Li L, et al. Global review of recent source apportionments for airborne particulate matter[J]. *Science of The Total Environment*, 2020, 740:140091.
- [45] 缙亚峰,余欢,王成,等. PM_{2.5} 化学组成观测设计对 PMF 源解析结果影响综述[J].*环境化学*,2020,39(07):1744-1753.
- [46] Gao J., Dong S.H., Yu H.F., et al. Source apportionment for online dataset at a megacity in China using a new PTT-PMF model. *Atmos. Environ.*, 2020, 229, 11.
- [47] Huang, R.J., Zhang, Y.L., Bozzetti, C., et al. High secondary aerosol contribution to particulate pollution during haze events in China. *Nature*, 2014, 514, 218-222.

- [48] Fang, T., Guo, H., Verma, V., et al. PM_{2.5} water-soluble elements in the southeastern United States: automated analytical method development, spatiotemporal distributions, source apportionment, and implications for health studies. *Atmos. Chem. Phys.* 2015, 15, 11667-11682.
- [49] Pentti Paatero, Shelly Eberly, Steven G. Brown, et al. Methods for estimating uncertainty in factor analytic solutions[EB/OL]. [2021-11-18].
http://cfpub.epa.gov/si/si_public_file_download.cfm?p_download_id=519156
- [50] Hopke P K. Recent Advances in Air Pollution Mixture Resolutions[J]. *Microchemical Journal*, 2020, 163(3):105907.
- [51] Hopke P K. Approaches to reducing rotational ambiguity in receptor modeling of ambient particulate matter [J]. *Chemometrics and Intelligent Laboratory Systems*, 2021. 210, 104252
- [52] 兰紫娟,江家豪,林理量,等. 深圳市典型混合功能区 PM_{2.5} 源解析研究[J].*中国环境科学*. 2021,41(09):4001-4008.
- [53] 冯婷,蔡一鸣,李振亮,等.重庆市典型城区 PM_{2.5} 含碳气溶胶季节变化和来源解析[J].*环境科学学报*,2021,41(05):1703-1717.
- [54] 赵清,李杏茹,王国选,等.运城秋冬季大气细粒子化学组成特征及来源解析[J].*环境科学*,2021,42(04):1626-1635.
- [55] 邓林俐,张凯山,殷子渊,等.基于 PMF 模型的 PM_{2.5} 中金属元素污染及来源的区域特征分析[J].*环境科学*,2020,41(12):5276-5287.
- [56] Wang F, Yu H, Wang Z, et al. Review of online source apportionment research based on observation for ambient particulate matter[J]. *Science of The Total Environment*, 2020, 762:144095.
- [57] Liu G R, Shi G L, Tian Y Z, et al. Physically constrained source apportionment (PCSA) for polycyclic aromatic hydrocarbon using the Multilinear Engine 2-species ratios (ME2-SR) method[J]. *Science of the Total Environment*, 2015, 502:16-21.
- [58] Jie G A, Sd A, Hy B, et al. Source apportionment for online dataset at a megacity in China using a new PTT-PMF model[J]. *Atmospheric Environment*, 2020, 229,117457.