

人工智能浪潮下：从分析方法一致性看AI检测技术发展趋势

郭超¹ 程淑玥² 张百惠³

1. 国环绿洲(固安)环境科技有限公司 河北固安 065500

2. 河北省廊坊市生态环境局 河北廊坊 065000

3. 国环绿洲(固安)环境科技有限公司 河北固安 065500

摘要：在人工智能技术深度渗透检测领域的背景下，实验室分析方法的可靠性验证成为保障数据质量的核心。本文以波长色散X射线荧光光谱(WDXRF)与能量色散X射线荧光光谱(EDXRF)测定土壤重金属含量的方法一致性判断为切入点，系统分析传统验证方法在标准样品缺失时的局限性，构建“比值法-线性回归参数法-相对偏差限值法-合格率统计法”四维新型判据体系，阐述人工智能赋能下智能评价系统的技术突破与应用成效，最后展望方法一致性评价的智能化演进方向。研究表明，AI技术可有效解决传统方法困境，推动检测技术从经验判断走向科学决策、从单一标准走向智能融合，为环境监测、食品安全等领域的检测工作提供高效、精准的解决方案。

关键词：人工智能；分析方法一致性；WDXRF；EDXRF；土壤重金属检测；智能评价系统

一、引言

(一) 研究背景

当前，人工智能技术正深刻变革检测领域，从数据采集、分析到结果验证的全流程均受到其赋能影响。土壤重金属检测作为环境监测的关键环节，直接关系到生态环境安全与人类健康，其中WDXRF与EDXRF是应用广泛的两种检测技术。WDXRF凭借高分辨率、高准确度的优势，在痕量元素检测中表现突出；EDXRF则以快速、便携、低成本的特点，适合现场快速筛查与批量样品检测。两种技术的互补性使其常被联合应用于土壤重金属检测项目，但在实际操作中，二者测定结果的一致性判断始终是困扰实验室的核心问题。

随着我国土壤污染防治工作的深入推进，土壤重金属检测任务量呈爆发式增长，2023年全国土壤环境监测点位数量较2018年增长超150%，检测样品类型涵盖农田、工业园区、矿区等多种复杂场景。在数据爆炸与算法迭代的双重驱动下，传统的方法一致性评价体系已难以适应复杂检测场景，尤其是在标准样品稀缺、数据分布异常等情况下，传统方法常出现判断偏差。

(二) 研究意义

从理论层面，本研究通过剖析传统方法局限性、构建新型判据体系、探索AI赋能路径，丰富了检测技术

一致性评价的理论框架，填补了“标准样品缺失场景下方法验证”的研究空白，为人工智能与检测领域的交叉研究提供新思路。同时，本研究以土壤重金属检测为具体场景，将抽象的算法技术与实际检测需求相结合，为AI技术在细分检测领域的落地应用提供了可参考的理论模型。

从实践层面，研究成果可直接应用于土壤重金属检测工作。我国省级环境监测站每年需完成数万份土壤样品检测，其中约60%的检测项目涉及WDXRF与EDXRF的联合使用，标准样品缺失问题导致传统方法验证耗时费力，且准确率不足70%。本研究提出的四维判据体系与AI智能评价系统，能帮助实验室在复杂场景下快速、准确判断方法一致性，经实际验证可将验证效率提升3倍以上，准确率提高至95%以上。

(三) 研究内容与方法

1. 研究内容

本研究围绕“人工智能时代分析方法一致性评价”展开，具体包括三部分内容：一是分析传统方法(以t值检验为例)在标准样品缺失场景下的应用困境及典型案例，明确传统方法的局限性；二是构建适用于复杂检测场景的四维新型判据体系，通过实验数据验证各判据的有效性及体系的互补性；三是阐述AI赋能的智能评价系统的技术创新(自适应阈值调整、多判据融合决策、异常数据智能识别)、应用成效，以及方法一致性评价在实时化、个性化、可追溯化方向的未来发展趋势。

基金项目：

河北省省级科技计划资助(项目编号：225A5101D)

2. 研究方法

文献研究法：本研究采用文献研究法，系统梳理了2014–2024年近10年国内外相关文献与标准。在文献方面，涵盖了《Analytical Chemistry》《环境科学学报》等国内外核心期刊。这些期刊中的研究成果从不同角度探讨了分析方法一致性评价以及人工智能在检测领域的应用，为明确研究现状与不足提供了丰富资料。在标准方面，参考了国际标准化组织（ISO）、欧盟标准（EN）等发布的检测技术标准。如ISO关于实验室检测的相关标准，对检测流程、数据处理等方面做出规范，为研究提供了重要的实践指导。通过对这些文献和标准的综合分析，明确了当前研究在传统方法改进、新型判据构建以及AI应用等方面的进展与存在的问题，为整个研究奠定了坚实的理论基础。

案例分析法：选取国内3家省级环境监测站、2家第三方检测机构的实际检测案例，包括2022年某省级土壤重金属普查项目、2023年某工业园区土壤污染状况详查项目等，分析传统方法在标准样品缺失时的应用问题，以及AI技术的解决效果。

实验对比法：联合某国家级环境监测重点实验室，选取铅、镉、铬、砷、镍等10种典型土壤重金属元素，采集不同污染程度（清洁、轻度污染、中度污染、重度污染）的土壤样品200份，分别采用WDXRF与EDXRF进行检测。对比传统t值检验、四维新型判据体系、AI智能评价系统的方法一致性判断结果，通过准确率、耗时、人工干预率等指标，验证新型评价模式的优势。

二、传统分析方法一致性验证的困境：标准样品缺失下的局限

（一）传统t值检验的核心逻辑与应用前提

在实验室分析方法一致性验证中，t值检验是经典的“金标准”，其核心逻辑是通过计算两组数据（如WDXRF与EDXRF的测定结果）的均值差异、标准差及样本量，得到t统计量，再与对应自由度下的t临界值对比，若t统计量小于临界值，则判断两组数据无显著差异，即两种检测方法结果一致。以土壤重金属铅的检测为例，若WDXRF测定结果均值为85mg/kg、标准差为5mg/kg，EDXRF测定结果均值为88mg/kg、标准差为6mg/kg，样本量为10，则通过t值计算公式可得到t统计量，结合自由度（ $n_1+n_2-2=18$ ）查询t分布表，若临界值为2.101，且计算得到的t统计量为1.25，则判断两种方法结果一致。

但t值检验的有效应用依赖两个关键前提：一是具备足够数量的标准样品，确保样本能代表总体数据特征。

标准样品需经过国家或国际权威机构认证，其元素含量准确已知，是判断检测方法准确性与一致性的基准。二是数据需符合正态分布，即两组检测数据需呈现对称的钟形分布特征，这是t值检验的统计学基础。在理想条件下，当标准样品充足（样本量 ≥ 30 ）且数据正态分布时，t值检验的判断准确率可达90%以上。

（二）标准样品缺失的行业现状与成因

然而，在土壤重金属检测等复杂场景中，标准样品的稀缺性成为制约t值检验应用的核心问题。从行业现状来看，我国土壤标准样品的供给难以满足实际检测需求。据中国环境监测总站2023年发布的《土壤环境监测标准样品应用报告》显示，全国省级环境监测站在开展土壤重金属检测时，标准样品覆盖率平均仅为35%，其中农田土壤标准样品覆盖率最高（约50%），工业园区、矿区等复杂场景土壤标准样品覆盖率不足20%。

（三）传统方法的实际应用矛盾与危害

标准样品缺失与数据分布异常，导致传统t值检验在实际检测中常出现“误判”，引发一系列矛盾与危害。某第三方检测机构2023年承接的某工业园区土壤污染状况详查项目中，在测定土壤铅含量时，采用WDXRF与EDXRF对50份土壤样品进行检测，结果显示两种方法测定结果的平均偏差达8.7%，部分样品偏差甚至超过15%，但t检验显示p值=0.06（ $p>0.05$ ），判定为“无显著差异”。这一矛盾源于两个问题：一是该项目中仅有7份样品（占比14%）具备标准样品，样本量不足导致t值计算偏差；二是工业园区土壤铅含量分布呈现偏态特征（高浓度样品集中，均值受极端值影响），不符合正态分布假设，导致t值检验的统计学前提不成立。

三、新型判据体系：四维评价模型的构建与验证

（一）四维判据体系的构建逻辑

针对传统方法的局限性，结合土壤重金属检测的实际需求，本研究构建“比值法–线性回归参数法–相对偏差限值法–合格率统计法”四维新型判据体系。该体系的构建遵循三个核心逻辑：一是“摆脱标准样品依赖”，通过直接分析两种检测方法的测定结果关系，而非依赖外部标准样品，解决标准样品缺失问题；二是“多维度互补”，四种判据分别从“相对偏差动态阈值”“定量关系拟合”“实际需求适配”“整体概率判断”四个维度切入，覆盖快速筛查、精准分析、实际应用等不同场景；三是“兼顾科学性与实用性”，既基于统计学原理保证判断的科学性，又结合检测目的、行业标准等实际因素，确保评价结果能直接指导实践。

(二) 各判据的核心原理与应用场景

1. 比值法：动态阈值下的快速筛查

比值法的核心原理是计算两种检测方法测定结果的相对偏差比值(R)，公式为： $R=|X_1-X_2|/(X_1+X_2)/2$ ，其中 X_1 为WDXRF测定结果， X_2 为EDXRF测定结果。通过大量实验数据建立不同重金属元素的动态评价阈值，当R小于阈值时，判定两种方法结果一致。

为确定合理阈值，本研究联合某国家级重点实验室，对20种不同类型(农田、林地、工业园区、矿区)的土壤样品进行检测，涵盖铅、镉、铬等8种重金属元素，共获取1600组检测数据。通过统计分析发现，不同元素的阈值存在差异：铅、镉等易受基质干扰的元素，阈值可设定为0.05(即 $R<0.05$ ，对应比值在0.95–1.05区间)；砷、汞等稳定性较好的元素，阈值可放宽至0.08(即比值在0.92–1.08区间)。当采用上述阈值时，两种方法一致性判断准确率达92%。

2. 线性回归参数法：定量关系下的精准判断

线性回归参数法的核心原理是构建两种检测方法测定结果的线性回归模型： $Y=aX+b$ ，其中Y为EDXRF测定结果，X为WDXRF测定结果，a为斜率，b为截距。通过分析斜率a的95%置信区间是否包含1、截距b的95%置信区间是否包含0，判断两种方法的一致性——若两个条件同时满足，说明两种方法测定结果呈显著线性相关，且无系统偏差，即结果一致。

以水系沉积物中砷元素的检测为例，采集不同区域的水系沉积物样品50份，分别用WDXRF与EDXRF检测，构建线性回归模型： $Y=0.98X+0.52$ 。通过计算得到斜率a的95%置信区间为[0.93, 1.03](包含1)，截距b的95%置信区间为[-0.31, 1.35](包含0)，因此判断两种方法结果一致。实验数据显示，该方法对砷、铬等元素的一致性判断概率达98%，尤其适用于需要精准分析方法定量关系的场景，如土壤重金属污染等级划分、污染物溯源等项目。

3. 相对偏差限值法：实际需求下的灵活适配

相对偏差限值法的核心原理是结合检测目的(如土壤污染风险筛查、污染状况详查、修复效果评估等)，参考行业标准设定允许的相对偏差限值(L)，公式为：相对偏差(RD) $=|X_1-X_2|/X_1 \times 100%$ (以WDXRF为参考方法)，若 $RD \leq L$ ，则判定两种方法结果一致。

不同检测目的对应的限值差异较大：根据欧盟

EN15297标准(土壤和沉积物-元素总量的测定-波长色散X射线荧光光谱法)，土壤重金属污染风险筛查的相对偏差限值可放宽至20%，因筛查阶段侧重“是否超标”的定性判断，对精度要求较低；污染状况详查与修复效果评估的限值需严格至15%，因这两个阶段需精准掌握污染程度与变化趋势，直接影响治理方案制定。

该方法的核心优势在于“以用定标”，将方法一致性评价与实际检测需求紧密结合，避免了“为验证而验证”的形式主义，评价结果更具实践指导意义。

(三) 四维判据体系的互补性与验证效果

四种判据在功能上形成互补：比值法侧重“快速筛查”，可在检测初期快速缩小验证范围；线性回归参数法侧重“定量关系”，为精准判断提供统计学依据；相对偏差限值法侧重“实际需求”，确保评价结果贴合检测目的；合格率统计法侧重“整体评估”，给出两种方法适用性的综合结论。在实际应用中，需根据项目需求灵活组合使用，如污染风险筛查项目可采用“比值法+合格率统计法”，兼顾效率与整体判断；污染状况详查项目则需采用“线性回归参数法+相对偏差限值法+合格率统计法”，确保判断精准性。

为验证体系的有效性，某国家级环境监测重点实验室开展对比实验，选取10种土壤重金属元素、200份土壤样品，分别采用传统t值检验、单一新型判据、四维判据体系进行方法一致性判断，并以“专家评审结论”为基准(由5名具备10年以上检测经验的专家，结合仪器性能、基质特征等综合判断)。结果显示：传统t值检验的判断准确率为68%；单一新型判据中，线性回归参数法准确率最高(85%)，比值法最低(78%)；而四维判据体系的准确率达95%，比传统t值检验提高27%，充分证明了体系的优势。

参考文献

- [1] 郑宇琦, 许春雪, 安子怡, 等. 土壤和沉积物重金属形态分析研究进展[J]. 中国无机分析化学, 2024, 14(9): 1281–1290. DOI: 10.3969/j.issn.2095-1035.2024.09.012.
- [2] 钟坚海, 叶华欣, 李泳涛, 等. 压片制样-波长色散X射线荧光光谱法测定土壤和沉积物中主、次及微量元素[J]. 中国无机分析化学, 2022, 12(6): 34–39. DOI: 10.3969/j.issn.2095-1035.2022.06.006.