|  |  |
| --- | --- |
| ICS | 73.040 |
| CCS | |  | | --- | |  |   D21 |

团体标准

P/CIQA -170—2024

铝土矿 铝硅铁含量的测定 近红外光谱法

Bauxite - Determination of aluminum, silicon, and iron content - Near infrared spectroscopy method

XXXX - XX - XX发布

XXXX - XX - XX实施

中国出入境检验检疫协会  发布

1. 前言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国出入境检验检疫协会综合质量服务标准化技术委员会（CIQA/TC12）提出并归口。

本文件起草单位：中国检验认证集团河北有限公司。

本文件参与单位：XXX。

本文件主要起草人：XXX、XXX、XXX、XXX。

本文件版权归中国出入境检验检疫协会所有。任何单位或个人未经许可，不得以营利为目的，印制、出版、翻译、转发或复制全文或部分文字。

本文件是首次发布。

铝土矿 铝硅铁含量的测定 近红外光谱法

* 1. 范围

本文件规定了采用近红外光谱法快速测定铝土矿中铝、硅和铁含量的方法。

本文件适用于铝土矿铝、硅和铁含量的快速测定。

本文件不适用于仲裁检验。

* 1. 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

YS/T 575.1-2007 铝土矿石化学分析方法 第1部分：氧化铝含量的测定 EDTA滴定法

YS/T 575.3-2007 铝土矿石化学分析方法 第3部分：二氧化硅含量的测定 钼蓝光度法

YS/T 575.4-2007 铝土矿石化学分析方法 第4部分：三氧化二铁含量的测定 重铬酸钾滴定法

YS/T 575.23-2021 铝土矿石化学分析方法 第23部分：元素含量的测定 X射线荧光光谱法

GB/T 2009-1987 散装矾土取样、制样方法

GB/T 8170 数值修约规则与极限数值的表示和判定

GB/T 29858 分子光谱多元校正定量分析通则

* 1. 术语和定义

GB/T 29858和本文件界定的术语和定义适用于本文件。

校正模型 calibration model

利用化学计量学方法建立的铝土矿样品近红外光谱与对应铝土矿测定值之间的数学模型。

仪器的标准化 instrument standardization

对一组红外仪器进行校正的过程。通过校正，这组设备用同一标定模型对同一样品进行测定时，获取相近的测定值。

模型鲁棒性 model robustness

模型在输入数据存在噪声、干扰或分布变化时，维持其稳定性的能力。

* 1. 原理

当近红外光照射到铝土矿样品表面时，矿物内部的化学键会在特定波长范围内吸收能量，产生特征吸收峰，例如三水铝石（γ-Al(OH)₃）和一水铝石（γ-AlO(OH)）的 O–H 振动模式在1400 nm和2200 nm附近具有显著吸收。采用漫反射近红外光谱采集样品的光谱数据，并结合光谱预处理（如去噪、平滑、降维）优化数据质量。随后，利用偏最小二乘回归（PLSR）、支持向量回归（SVR）及基于U-Net结构的深度学习等方法，构建光谱特征与铝土矿成分含量之间的线性或非线性映射模型，实现氧化铝（Al₂O₃）、二氧化硅（SiO₂）和氧化铁（Fe₂O₃）的快速、无损定量分析。

* 1. 仪器设备
     1. 近红外光谱仪

仪器应具有漫反射模式，其谱区范围为900–2450 nm（12900/cm-4000/cm），光谱分辨率（FWHM）≥6.2 nm，信噪比≥13000:1，波长重复性≤0.2nm，仪器应具备自检功能。或可采用满足上述光谱性能要求的其他同类设备。

* + 1. 样品杯

与近红外光谱仪配套的石英样品杯。

* + 1. 样品研磨、粉碎设备

选择适当的研磨或粉碎设备，制备符合粒度要求的铝土矿样品。

* 1. 样品
     1. 总则

本文件中的样品包括:用于近红外光谱仪校正模型的校正样品和验证样品、用于检验仪器稳定性检查的仪器质控样品和用于校正模型性能监控的监控样品。

* + 1. 校正样品和验证样品

校正样品和验证样品需具有代表性，涵盖各种影响因素，如考虑：

a）样品的硅铝铁范围应覆盖待测样品的预期检测范围；

b）需考虑地质成因的不同；

c）加工工艺、储存条件、物理状态（粒度、湿度、氧化程度）；

d）仪器差异。

* + 1. 仪器质控样品

质控样品应具备长期稳定性，物理化学性质尽可能与待测样品保持一致。

* + 1. 校正模型性能监控样品

校正模型性能监控样品宜从待测样品种随机选取，需采取一定的策略确保样品在校正范围内分布均匀。例如，将样品按含量高低进行分段后在每段中随机选择样品，或选择能涵盖常规范围的样品。

* 1. 校正模型的建立与验证
     1. 校正模型构建
        1. 样品选择

建模样品分为校正样品和验证样品，且应满足本文件规定的样品适用范围。

* + - 1. 校正样品集

在建立基于近红外光谱的铝土矿铝、铁、硅含量快速检测模型时，用于校正的样品数量建议不少于500个，以确保统计显著性和模型的泛化能力。校正集样品应全面覆盖待测样品的成分含量范围，若现有校正样品集未能完全覆盖待测样品的所有铝土矿成分含量变动范围，应补充新的、具有代表性的校正样品，以确保数据的完整性和代表性。为保证样本的均衡性，避免模型因某一区间样本过多而产生偏倚，建议将每个成分含量范围均分为5个区间，且每个区间内样品数量不少于50个。

* + - 1. 验证样品集

用于验证模型的验证样品数量建议不少于100个。验证集样品应该为校正样品集外的样品。验证样品的品质变化范围应涵盖校正样品的品质变化范围，各验证样品品质数据在整个变化范围内应均匀分布。

* + - 1. 测定参考值

按照表1规定的参考方法，测定校正样品和验证样品各项品质指标。

表1 参考方法

|  |  |  |
| --- | --- | --- |
| 项目 | 单位 | 参考方法 |
| 氧化铝（Al2O3） | % | YS/T 575.1-2007或YS/T 575.23-2021 |
| 二氧化硅（SiO2） | % | YS/T 575.3-2007或YS/T 575.23-2021 |
| 氧化铁（Fe2O3） | % | YS/T 575.4-2007或YS/T 575.23-2021 |

* + - 1. 建立模型

校正模型的构建流程主要包括数据清洗、数据恢复和模型建立三个关键阶段。

数据清洗阶段，采用如马氏距离等异常样本剔除方法对采集数据进行预处理，剔除数据中的噪声和异常值，提升数据的纯净度和可靠性，确保异常值对后续建模的干扰降至最低（参见附录A）。

数据恢复阶段，采用如标准正态转换等的光谱信息恢复技术，以进一步提升数据质量，为后续建模提供高质量的基础数据（参见附录B）。

模型建立阶段，搭建相应的建模环境，对校正样品的近红外光谱数据与参考值数据进行关联建模。通过选择合适的建模算法和参数，建立准确的预测模型，并计算模型的有效性指标，以评估模型的性能和适用性，模型评价指标见8.3。此外，还可以参考其他被广泛采用的判断模型性能的指标。

* + 1. 校正模型的验证

应用7.1.4所示的规定的参考方法测定验证样品的品质数据作为参考值。采用7.1.5建立的校正模型对验证样品的近红外光谱进行预测，得到各品质数据的测定结果。

* 1. 校正模型的性能评价
     1. 总则

近红外光谱检测样品的模型性能需通过预测精度、模型稳定性等多种指标综合评估，以确保其满足模型的准确性、鲁棒性与适用性要求。

* + 1. 结果可视化与散点图分析

可视化绘图是校正模型性能统计学评价的重要方法之一，如参考值-近红外测定值与残差-近红外测定值的散点图。数据散点图可以清晰地表现参考值与近红外测定值之间的偏差及异常值的情况，参考值-预测值关系图中，坐标越接近X=Y的参考线，预测效果越好；参考值-残差关系图中，坐标越接近Y=0,预测偏差越小。

残差按公式（1）计算：

（1）

式中：

——第i个残差：

——第i个样品参考值；

——第i个样品近红外测定值。

* + 1. 模型预测精度评价指标

模型预测的性能可用预测标准误差（SEP）和决定系数（）作为评价指标。

* + - 1. 预测标准误差

用于衡量模型在独立验证数据集上的预测性能的指标，评估模型对未知数据的预测能力，SEP按照公式（2）进行计算：

（2）

式中：

——第i个验证样品的近红外测定值；

——第i个验证样品对应的参考值；

——用于验证模型的样品数。

* + - 1. 决定系数

决定系数R2是指校正模型可解释的变异占总变异的比重，是评价模型拟合程度的一项指标，是相关系数R的平方。R2取值范围为0～1，R2越大，模型拟合越好，越接近1越好。决定系数按式（3）进行计算：

（3）

式中：

——第i个验证样品的近红外测定值；

——第i个验证样品对应的参考值；

——,…,的均值；

——用于验证模型的样品数。

* + 1. 模型稳定性指标

同一样品连续5次扫描结果的重复性限应满足12.1的要求，不同仪器、操作者或实验室间验证结果的再现性临界差应满足12.2的要求。

* 1. 测试步骤
     1. 样品制备

按照GB/T 2009方法制备铝土矿样品。

采用标准化的制备流程确保待测试样的制备与验证样品的制备保持一致。

样品制备后应尽快检测，防止性质变化影响光谱真实性。

* + 1. 光谱采集

依据仪器说明书启动近红外光谱仪，预热至稳定状态（建议≥30 min）,完成自检并验证基线平直度。

在无样品条件下，使用标准背景板采集背景光谱，频度与样品检测一致，确保环境光及仪器噪声校正。

将样品均匀铺展于样品盘中（厚度：1.5±0.2cm），避免颗粒分层或堆积。每个样品连续扫描5次，单样品5次扫描的吸光度绝对差值≤0.0004 AU,超限数据需剔除并补测至达标，取有效扫描结果的算术平均值作为最终光谱数据。

检测期间实验室温度应控制在23±2，湿度≤40%，避免日光直射及振动干扰。

* + 1. 结果评估

试样测定结果需在所使用的校正模型范围内方为有效。

若试样存在光谱异常，则近红外光谱测定结果不可采纳。

若对同一样品进行多次测量，且结果满足重复性限要求（见12.1）,则计算其算术平均值。

* 1. 仪器稳定性检查
     1. 用仪器质控样品检查仪器稳定性

为确保仪器硬件的稳定性及预防仪器故障，在待测试样测定时要对仪器质控样品进行检测，且每天至少进行一次仪器质控样品的检测。质控样品的制备需与待测样品的制备保持一致，并且以能够使其存储有效期最大化的方式进行存储。仪器质控样品需能够保持长时间的稳定，其稳定性要根据实际情况进行检测。仪器质控样品需与待测试样同时进行近红外分析，以保障不间断的质量控制。

对质控结果的波动进行记录并绘制质控图，同时，对有显著性变化趋势的质控图进行研究。

* + 1. 仪器诊断

对扫描型光谱仪，需至少每周检查一次其波长/波数的准确度和精密度，如果仪器商推荐，检查频度还需更高。并且，检查结果要与相应技术规格及要求进行对比。

需至少每周或按照厂商推荐的时间间隔对仪器的噪音进行检查。

* + 1. 联网仪器

如果几台仪器联网使用，需根据厂商的推荐对仪器之间的标准化给予特别关注。

* 1. 校正模型性能监控
     1. 总则

分析样品前需对校正模型的适用性进行检查。可采用在校正和验证过程中所使用的异常值剔除方法，如马氏距离和光谱残差等。这在大多数仪器中可被自动执行。

如果样品没有通过适用性检验，则不能用现有的校正模型进行检测，需对校正模型进行修正或升级。异常样品的检测可用于选择哪些样品需用参考方法赋值，然后用于升级校正模型。

如果校正模型能够适用于被分析样品，则该样品光谱可采用该校正模型进行测定。

校正模型性能需持续选用有参考方法赋值的监控样品进行验证，以确保校正处于稳定的最优状态并满足分析准确性要求。对近红外校正模型性能的验证频度，需足以确保该方法运行的稳定性受控于参考方法的系统偏差和随机误差。该频度尤其依赖于每天分析的样品数量和样品集的波动。

校正模型性能监控样品数量满足统计分析要求，需要至少20个样品（以满足变量的正态分布）。验证样品的分析结果可用于校正模型运行性能的首次监控，接下来每周5～10个样品即可满足对校正模型性能的监控。如果使用较少的监控样品，一旦有结果超出了控制限，则很难做出正确判断。

* + 1. 参考值和近红外测定值差异的控制图

结果宜采用控制图进行评估，控制图以分析样品数作为横坐标，参考值和近红外测定值间的差值作为纵坐标；(95%的置信概率)和（99.8%的置信概率）可作为警戒限和行动限，其中采用独立于校正样品的验证样品获得。

当校正和参考值的测定均按照规定标准执行时，20个数据点中只有1个在警戒限之外，1000个数据点中只有2个在行动限之外。

根据控制图检查系统偏差、偏差的正负和极端异常值。可使用休哈特控制图（Shewart）的一般规则对结果进行评估。注意过多规则同时使用时可能导致虚假报警。

可组合采用以下规则进行检查：

a）1个点超过了行动限；

b）连续3个点中有2个在警戒限之外；

c）连续9个点在0线的同一侧。

展现其他特征（如参考值和近红外测定值间差值的平均值）的控制图和其他评判规则也可以使用，以强化对结果的评估。

对结果评估时，要注意和所测得的近红外测定值与参考值的差异也包含参考值的不精密度。如果参考值的不精密度小于的1/3，则该影响可忽略不计。

为减少虚假报警的风险，宜以不同的顺序分别进行监控样品的近红外光谱分析和参考值分析，以避免日间系统偏差的影响。

如果样品经常超出警戒限，并且控制图显示都是随机误差（无趋势性或系统偏差），可能是由于在确定行动限和警戒限时使用了太过理想的。此时，通过频繁调整校正以勉强将结果限制于控制限范围内的做法不可取，需对进行重新评估。

稳定运行一段时间后的校正模型可能会出现失控的情况，此时需对校正模型进行升级。升级之前，首先要评估该失控发生的原因是否是参考值含量范围的变化、测定条件的变化（如更换新的操作者）、仪器出现漂移或发生故障等。有些情况下，只通过调整校正模型的常数项就能满足分析要求。其他情况下，可能需要重新运行校正程序。重新校正时，校正样品要包含对校正模型性能进行监控的监控样品，也可专门选择样品来进行重新校正。

在参考值的分析方法满足精密度要求，且近红外的检测条件与仪器性能保持稳定的条件下，偏差出现了显著性差异或值升高的情况，则可能是由于样品的化学、生物学或物理性质与校正样品相比发生了变化。

其他控制图（如z比分数）也可使用。

* 1. 精密度
     1. 重复性限

本文件规定的近红外光谱所测指标的重复性限大小与参考方法一致。

* + 1. 再现性临界差

本文件规定的近红外光谱法所测指标的再现性临界差的大小与参考方法一致。

* 1. 试验报告

试验报告应包括下列信息：

a) 实验室名称和地址；

b) 试验报告发布日期；

c) 本文件的编号；

d) 试样本身必要的详细说明；

e) 分析结果；

f) 测定过程中存在的任何异常特性和在本文件中没有规定的可能对试样或标准样品的分析结果产生影响的任何操作。

附 录 A

（资料性附录）

异常值处理方法示例

A.1 样品集数据异常

对采集到的NIRS样品数据，若某条光谱数据与待测成分性质之间的关系和其他光谱数据的对应关系差异较大，会对建立的校正模型的性能产生干扰。这类异常数据可通过马氏距离与光谱残差关系图进行筛选。

A.1.1马氏距离Mahalanobis Distance（MD）

马氏距离反应样品的离群性。某一样品距离样品群越远，该样品是异常点的可能性越大，按式（4）～（7）进行计算。

假设为光谱矩阵，为平均光谱，则第j个元素表示为：

（4）

原始光谱数据集矩阵的中心矩阵第j列表述为：

（5）

马氏矩阵为：

（6）

若M存在逆矩阵，则对样品X定义其对应的马氏距离为：

（7）

在实际处理实验光谱数据时，可以根据实验对数据的剔除要求，通过设置一定的阈值，将对应样本的MD大于该阈值的数据定为异常样本。

A.1.2欧氏距离 euclidean distance（ED）

欧几里得空间中两点间的直线距离。计算光谱数据的平均光谱作为中心矢量，按式（8）进行计算。计算单个未知样本与平均光谱之间的欧氏距离，按式（9）进行计算。

 （8）

 （9）

式中：

——校正样品总样本个数；

——校正样品中第*i*个样本的光谱。

A.2 预测结果数据异常

若某个样品的预测值与参考值有显著差异，对可通过参考值与预测值关系图进行筛选，明显离群的样品为异常。

附 录 B

（资料性附录）

异常值处理方法示例

B.1 Savitzky-Golay平滑

Savitzky-Golay（SG）通过局部多项式最小二乘拟合，用拟合值替代原始数据点，实现平滑效果。其核心是通过设定一个特定长度的窗口，在窗口内采用k阶多项式对数据点进行精确拟合，从而计算得到各数据点的平滑值。SG按式（10）进行计算：

（10）

式中：

——卷积平滑系数；

——为单个样品待拟合；

——拟合后的输出数据。

示例：窗口大小设为5，多项式阶数设为3。

B.2 标准正态变换（standard normal variate transformation,SNV）

利用SNV可以克服由于铝土矿颗粒不均匀导致的光谱信息干扰。其原理是将光谱数据转换为标准正态分布的变量，对每个数据点进行标准差归一化和均值中心化操作，消除光谱数据中的尺幅效应。在漫反射光谱测量中，易因颗粒度差异、表面粗糙度差异等引入噪音。SNV通过数据变换消除这些物理影响，增强光谱与样品化学性质之间的真实关联性。SNV按式（11）、（12）进行计算。

 （11）

 （12）

式中：

——待数据预处理的近红外光谱；

——光谱的平均吸光度；

——光谱第i个波数点对应的吸光度值；

——光谱的波数点总数。

附 录 C

（规范性附录）

近红外光谱测定铝土矿多元校正模型建立实施例

C.1校正模型建立的目的与适用性范围

应用傅里叶变换近红外漫反射光谱法检测被测铝土矿的硅、铝、铁，为原料的使用适时提供数据支持。

C.2样品收集

本文件适用于氧化铝（Al2O3）的测定范围在31.1%～57.66%，二氧化硅（SiO2）的测定范围在0.23%～21.29%，氧化铁（Fe2O3）的测定范围在4.19%～36.52%的铝土矿样本，超出本文件测定范围的铝土矿样品可以参照执行，共收集600条数据。根据为了保证检测结果的准确度，尽量通过模型内插预测铝土矿待测成分含量。

C.3数据划分

将600个样品数据划分为预测数据集和验证数据集。预测数据集用于校正模型建立，验证数据集用于校正模型的验证。校正模型数据和验证模型数据的成分含量尽量覆盖全范围，校正模型数据500个，验证模型数据100个。

C.4异常值剔除

为保证校正模型预测的精确度，本示例采用马氏距离法对异常值进行识别并剔除。

C.5数据处理

本示例利用标准正态（SNV）对光谱数据进行处理，以消除样品颗粒不均导致的光谱数据噪声。

C.6校正模型的建立与验证

使用基于U-Net架构的深度学习网络进行近红外光谱分析。如图C.1所示，该模型的整体架构由左侧的编码器、中间的解码器以及右侧的回归头组成。编码器部分包含四个下采样单元，每个单元由两个卷积块和一层最大池化层构成。每个卷积块首先通过卷积核大小为3的卷积层提取特征，随后依次进行批标准化和ReLU激活函数的非线性变换，生成卷积块的输出。最大池化层用于提取高级抽象特征，并在编码过程中逐渐增加通道维度，同时通过减半特征序列长度来减少计算负担，保证特征提取的质量。解码器部分由四个上采样单元组成，每个单元包含特征融合结构和两个卷积块。解码器通过跳跃连接将前一层和对应编码器输出的特征拼接，实现浅层细节信息和深层语义信息的融合。例如，解码器3的输入来自解码器4和对应尺度编码器3的输出。特征融合后，使用与编码器类似的卷积块进行处理，并引入cSENet通道注意力机制以增强特征的表征能力。回归头由三个独立的全连接层组成，学习不同参数的个性特征，并行输出氧化铝（Al₂O₃）、二氧化硅（SiO₂）和氧化铁（Fe₂O₃）的测定值。



图C.1 基于U-Net架构的深度学习模型示意图

C.7模型参数设置

初始学习率设置为0.001，并采用ReduceLROnPlateau学习率衰减策略，使用剩余20%样本的损失值动态调整学习率。每步训练输入至网络中的批量大小设置为32，最大迭代轮数设置为100。

C.8模型性能评估

本示例采用8.3模型预测精度评价指标作为测定值和参考值的准确性指标。该模型具体的模型评价指标见表C.1。

表C.1模型的评价指标

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 项目 | 单位 | 含量百分数 | R2 | SEP |
| 氧化铝（Al2O3） | % | 31.1～57.66 | 0.94 | 1.08 |
| 二氧化硅（SiO2） | % | 0.23～21.29 | 0.89 | 0.89 |
| 氧化铁（Fe2O3） | % | 4.19～36.52 | 0.98 | 0.99 |