

江 苏 省 地 方 标 准

DB 32/T XXXX—XXXX

高光谱近感水质监测技术规范

Technical specifications for hyperspectral proximal sensing inversion
of water quality monitoring

征求意见稿

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX – XX – XX 发布

XXXX – XX – XX 实施

目 次

前 言 III

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 总体要求和原则 2

 4.1 总体要求 2

 4.2 总体原则 2

5 技术流程 3

6 设备选型与布设 3

 6.1 选型要求 3

 6.2 布设环境要求 4

7 数据获取与预处理 4

 7.1 数据组成 4

 7.2 数据基本要求 5

 7.3 实测水环境参数获取与预处理 5

 7.4 高光谱反射率获取与预处理 6

 7.5 水环境参数与高光谱反射率同步匹配 6

8 水环境参数反演模型构建与应用 7

 8.1 数据集划分 7

 8.2 模型构建 7

 8.3 模型评估与优选 8

 8.4 模型应用 9

 8.5 模型更新与维护 9

9 质量控制与评估 10

 9.1 质量控制 10

 9.2 质量评估 11

10 安全要求 11

 10.1 设备安全 11

 10.2 数据安全 11

 10.3 野外作业安全 11

11 参数拓展与持续改进 11

 11.1 参数拓展 11

 11.2 持续改进 12

附 录 A （资料性） 光谱预处理 13

附 录 B （资料性） 光谱角计算 14

附 录 C （资料性） 模型精度评价指标 15

 C.1 决定系数（ R^2 ） 15

 C.2 均方根误差（RMSE） 15

 C.3 均方误差（MSE） 15

 C.4 平均绝对误差（MAE） 15

参 考 文 献 17

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。

本文件由江苏省生态环境厅提出、组织实施并归口。

本文件起草单位：中国科学院南京地理与湖泊研究所、南京中科深瞳科技研究院有限公司。

本文件主要起草人：李娜、张运林、施坤、张毅博、骆夏杨、崔驰潇、秦伯强、朱广伟、郭宇龙。

高光谱近感水质监测技术规范

1 范围

本文件规定了采用高光谱近感技术开展地表水环境参数反演的总体要求和原则、技术流程、设备选型与布设、同步数据获取与预处理、模型构建与应用、质量控制与评估、安全以及参数扩展要求。

本文件适用于陆基、浮标固定监测站点及船载移动监测平台，利用可见光～近红外波段（400 nm～1000 nm）高光谱近感设备，对内陆淡水湖泊、水库、河流等地表水体开展水环境参数反演。

本文件可作为基于航空或卫星平台的高光谱地表水环境参数遥感反演的部分参考。

2 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。

GB/T 4208-2017 外壳防护等级（IP代码）

GB/T 4883-2008 数据的统计处理和解释 正态样本离群值的判断和处理

GB 11893-1989 水质 总磷的测定 钼酸铵分光光度法

GB 11901-1989 水质 悬浮物的测定 重量法

GB 50093-2013 自动化仪表工程施工及质量验收规范

GB/T 12763.2-2007 海洋调查规范 第2部分：海洋水文观测

HJ 91.2-2022 地表水环境质量监测技术规范

HJ 506-2009 水质 溶解氧的测定 电化学探头法

HJ 636-2012 水质 总氮的测定 碱性过硫酸钾消解紫外分光光度法

HJ 897-2017 水质 叶绿素a的测定 分光光度法

HJ 1075-2019 水质 浊度的测定 浊度计法

HJ 1404-2024 地表水自动监测系统通信协议技术要求

SL 87-1994 透明度的测定（透明度计法、圆盘法）

HJ 212-2025 污染物自动监测监控系统数据传输技术要求

HJ 1446-2026 水质 高锰酸盐指数的测定 草酸钠还原碱性滴定法

HJ 915.1-2024 地表水水质自动监测站选址与基础设施建设技术要求

3 术语和定义

下列术语和定义适用于本文件。

3.1

高光谱近感 hyperspectral proximal sensing

利用高光谱传感器在近距离观测范围内（通常为水面以上2 m～10 m），实时获取目标连续、窄波段光谱信息的技术。

3.2

水环境参数反演 water environmental parameters inversion

利用水体反射率光谱与准同步实测水环境参数之间建立的物理、半分析或数据驱动的定量模型，逆向推算出目标水环境参数浓度的遥感与近感技术过程。

3.3

离水反射率 Water-leaving reflectance

水体离水辐亮度与水面入射辐照度之比，是表征是水体光学特性的核心物理量，也是高光谱近感反演水环境参数的基础输入量。

3.4

训练数据集 training dataset

用于建立和训练水环境参数反演模型的数据子集。用于拟合模型参数和学习输入输出关系的数据子集。

3.5

验证数据集 validation dataset

用于调优超参数、判断水环境参数反演模型是否过拟合或欠拟合，以及模型选择的数据子集。

3.6

测试数据集 testing dataset

在模型训练与调优完成后，用于对模型最终性能进行独立、客观评估的数据子集。在模型结构、特征和超参数全部确定后，仅用于一次性评估最终模型泛化性能的数据子集。

4 总体要求和原则

4.1 总体要求

开展高光谱近感地表水环境参数反演，应满足以下总体要求：

- 实施前应明确监测水体、目标参数、空间范围、时间范围、输出频次和预期用途，并根据水体光学特性及平台条件确定设备和技术方案；
- 系统应具备光谱采集、辅助环境信息记录、数据存储与传输、数据预处理、模型计算、质量控制和结果输出能力，各环节的数据、参数及版本应可追溯；
- 用于建模、验证和测试的实测水环境参数数据应采用现行标准方法获取；光谱观测与水样采集应在观测对象、空间位置和时间上保持一致；
- 反演模型应在规定的水体类型、季节、水文条件、参数浓度范围、设备状态和观测几何范围内使用。输入条件超出模型适用范围时，应标记为待验证数据，经补充验证合格后方可发布；
- 反演结果可用于连续监测、趋势分析和异常筛查；用于地表水环境质量评价、考核或执法判定时，应以国家或行业规定的标准监测方法所得结果为依据。

4.2 总体原则

技术实施应遵循以下原则：

- 科学性原则。参数选择、光谱特征构建和模型方法应有水体光学机理、实测数据或验证试验支撑，不应仅以训练数据集的拟合效果作为模型选择依据；
- 代表性原则。点位和样本应覆盖目标水体的主要空间单元、季节、水文过程及水环境参数变化范围，并避免岸边反射、底部反射、阴影和漂浮物等非目标信号；

- 同步性和可比性原则。实测水环境参数数据与高光谱数据应按统一观测对象和规定时间窗口匹配，建模、验证及业务应用的设备状态、预处理流程和数据单位应保持一致；
- 全过程质量控制原则。应对设备校准、现场采集、实验室测定、数据审核、模型训练、独立检验、部署运行和模型更新实施质量控制，并保留完整记录；
- 适用性和持续改进原则。应依据水体类型、目标参数及应用目的选择模型并定期评估性能；当设备状态、水体光学特性或模型精度发生显著变化时，应重新验证或更新模型。

5 技术流程

高光谱近感地表水环境参数反演技术流程见图1，包含设备选型与布设、同步数据采集与预处理、反演模型构建与应用、质量控制与评估四个阶段。



图 1 高光谱近感地表水环境参数反演技术流程图

6 设备选型与布设

6.1 选型要求

6.1.1 光谱性能要求

应根据目标水体、目标参数、观测频次和部署方式选择高光谱近感观测设备。设备波段设置、光谱分辨率、视场角等应满足表1要求，并应通过到货验收或型式试验确认。

表 1 高光谱近感设备性能参数要求

序号	性能参数	参数值
1	波长范围	400 nm ~ 1000 nm
2	光谱分辨率	≤10 nm
3	时间分辨率	≤5 min
4	视场角	≤10°
5	光谱采样间隔	≤5 nm
6	积分时间	≤100 ms
7	波长重复性	≤0.14 nm
8	波长准确性	±0.5 nm
9	采样方式	支持无人值守连续监测

6.1.2 通讯与防护要求

6.1.2.1 通讯协议要求

通讯协议应符合HJ 1404-2024中相关的系统通讯协议技术要求，并支持 4G/5G、北斗及有线通信中的至少一种方式。系统应具备时钟同步、断点续传、数据校验、状态上报和远程诊断功能。

6.1.2.2 防护和自清洁要求

所选设备应充分考虑长期野外布设需求，外壳材料和内部接线应耐腐蚀，整机防护等级应符合 GB/T 4208-2017 中 IP68 的规定。

此外，设备应具有光学镜片的自清洁功能。

6.1.3 数据存储与传输要求

设备应同时支持本地存储和远程传输。本地存储应至少保存30 d的原始光谱、辅助观测信息、质量标识和运行日志；网络恢复后应自动补传未上传数据，且不得覆盖尚未完成校验的数据。

6.1.4 计算性能要求

设备系统应具备暗电流校正、辐射与波长校正、遥感反射率计算、质量检查、模型调用、结果标识和异常告警能力，并应记录处理软件及算法版本。

6.2 布设环境要求

6.2.1 布设原则

高光谱近感传感器固定站点布设应符合HJ 91.2-2022和HJ 915.1-2024的相关要求，点位应具备监测目标水体的代表性，并兼顾供电、通信、检修及作业安全要求。

6.2.2 安装环境要求

6.2.2.1 高光谱近感传感器距水面安装高度应为 2 m~10 m。

6.2.2.2 高光谱近感传感器安装距离岸边不应小于 2 m，观测区应为光学深水区，应避开浅滩、阴影、底部反射、植物遮挡和漂浮物遮挡的影响；当条件变动致使观测区不再满足要求时，应重新选址或将相应数据标注为无效。

注：光学深水区指水体深度足够大（或水体足够浑浊），使得入射光线在到达水底前被完全吸收和散射，水底反射信号无法到达水面的水体区域。

6.2.2.3 高光谱近感传感器安装环境应视野开阔，观测时段内无高大物体遮挡设备和水面、无强电磁干扰、无腐蚀性气体。

6.2.2.4 安装点周边应有可靠电力供应，包括市电或太阳能等供电方式，供电线路应符合防水及漏电防护要求。

6.2.2.5 安装点位周边应覆盖通讯网络，至少应具备北斗、5G、4G 等无线通讯网络或有线通信网络中的一种。

6.2.2.6 传感器观测方位与太阳主平面的夹角宜为 90°~135°。超出该范围或出现明显太阳耀斑时，应调整观测方位；无法调整的，应将受影响数据标记为无效。

6.2.2.7 浮标和船载应配备稳定装置，陆基设备固定支架应符合 GB 50093-2013 关于防震、稳固性的要求。

7 数据获取与预处理

7.1 数据组成

地表水环境参数反演数据应包含实测水环境参数数据和高光谱反射率数据。地表水环境参数反演数据应包括原始光谱数据、校准和辅助观测数据、遥感反射率、同步实测水环境参数、质量标识以及模型输入输出数据。

每组匹配数据应具有唯一标识，并关联点位、时间、经纬度、设备与校准版本、观测几何、气象水文条件、样品编号、分析方法、数据单位和处理记录。

7.2 数据基本要求

a) 实测地表水环境参数和高光谱反射率数据应满足以下要求：

- 在观测对象、空间位置和时间上保持一致；
- 覆盖目标水体主要季节、水文条件和目标参数变化范围。

b) 数据应满足建模基本要求：

- 每个目标参数的有效匹配样本原则上不应少于 200 组；
- 样本分布应覆盖模型拟应用的参数范围和主要光学水体类型，避免样本集中于少数点位、日期或浓度区间。

7.3 实测水环境参数获取与预处理

7.3.1 水环境参数反演指标

目标参数应按其光学特性分为光学活性参数和非光学活性参数，分别制定验证要求。

- 光学活性参数包括叶绿素 a、总悬浮物浓度、浊度、透明度、漫射衰减系数，具有明确吸收或散射特征的参数；
- 非光学活性参数包括总氮、总磷、溶解氧和高锰酸盐指数，缺乏稳定直接光谱响应的参数可利用其与光学活性组分的稳定协变关系进行光谱反演；输出时应说明其间接性和适用域。

7.3.2 水样采集与测定规范

7.3.2.1 水样应代表传感器观测范围内的水体。水样应在以传感器视场中心为圆心，在 1.5 倍视域半径（根据视场角和安装高度计算）为范围进行采集。当水体空间异质性明显时，应在观测范围内采集至少 3 个子样本均匀组成混合样。船载采样应在船体和水面扰动稳定后进行样品采集，并记录点位、时间、观测足迹、天气和水面状况。

7.3.2.2 水样采集、容器、固定、保存和运输应执行 HJ 91.2-2022 及相应分析方法的规定。不同参数应使用适用的容器和保存条件在规定时间内运送至实验室分析。

7.3.2.3 水环境参数测定方法：

- a) 叶绿素 a 执行 HJ 897-2017 测定；
- b) 总悬浮物浓度执行 GB 11901-1989 测定；
- c) 浊度执行 HJ 1075-2019 测定；
- d) 透明度执行 SL 87-1994 测定；
- e) 漫射衰减系数执行 GB/T 12763.2-2007 测定；
- f) 总氮执行 HJ 636-2012 测定；
- g) 总磷执行 GB 11893-1989 测定；
- h) 溶解氧执行 HJ 506-2009 测定；
- i) 高锰酸盐指数执行 HJ 1446-2026 测定。

7.3.2.4 异常值的判读和处理应执行 GB/T 4883-2008。异常记录应先核查采样、分析、设备状态和环境条件；仅在明确技术原因或预先规定的质量规则时方可剔除。原因不明的极端值应保留原始记录并标记为待确认，不得仅因偏离统计分布而删除。同一参数连续出现异常时，应重新采样或开展设备和分析过程复核。

7.3.3 采集频次

运行期每年应至少开展 4 个批次的同步比对采样，应覆盖丰水期、平水期、枯水期及目标水体的高风险过程。每批样本数量应根据点位差异、参数范围和模型监测需要确定。

7.4 高光谱反射率获取与预处理

7.4.1 获取应满足以下条件

- a) 时间应为北京时间 9:00-16:00（冬季可适当缩短至 10:00-15:00，以太阳高度角 $\geq 30^\circ$ 为基准）；
- b) 环境应为光照稳定的晴天或阴天，且无大雾；
- c) 水面应无结冰、大面积漂浮物、油膜、阴影覆盖。

7.4.2 高光谱反射率采集流程

7.4.2.1 设备首次使用、维修或更换关键光学部件后，以及质量核查发现漂移时，应完成波长和辐射定标和暗电流校正。定标方法见附录 A，定标结果应形成记录并与数据关联

注：光谱定标是确定高光谱遥感器各光谱通道的峰值波长和光谱带宽的过程。

7.4.2.2 每个采样点应连续获取不少于 10 组有效光谱；每组应包括同步的下行辐照度、观测方向辐亮度及环境光信息。

7.4.2.3 应同步记录经纬度、时间、传感器方位与倾角、离水高度、积分时间、增益和设备状态。

7.4.3 高光谱反射率预处理流程

7.4.4 应对饱和、欠曝、突跳、负值、太阳耀斑、漂浮物遮挡和设备状态异常进行检查。可使用光谱角（SAM，见附录 B）辅助筛查，将 $SAM > 0.20 \text{ rad}$ 作为默认预警阈值，表明光谱偏离水体。当原始信号达到设备量程上限的波段应标记为饱和；遥感反射率出现负值或者突跳时，应标记为异常；当全波段同步升高、光谱饱和、近红外异常，标记为耀斑。无效判定应保留原始光谱和原因标识。

7.4.4.1 对 6.4.2.3 中水面上的下行辐照度和水体离水辐亮度记录求平均值，计算得到水体高光谱反射率数据。

7.4.4.2 可根据噪声水平采用 Savitzky-Golay 滤波或其他经验证的方法平滑光谱。窗口宽度、阶数等参数应仅使用训练数据确定，并记录处理前后数据。

7.4.4.3 应对采集的水体高光谱反射率数据进行归一化处理。

7.5 水环境参数与高光谱反射率同步匹配

7.5.1 匹配原则

实测地表水环境参数与高光谱近感水体反射率应在时间和空间上匹配。

7.5.2 匹配方法

实测地表水环境参数与高光谱近感水体反射率匹配应满足以下要求：

- a) 应以样品采集时刻为基准，优先匹配同一观测区内时间最近且质量合格的光谱。一般匹配窗口不应超过 $\pm 5 \text{ min}$ ；
- b) 污染物排放、藻华、强对流、流量突变等导致水体快速变化时，应采用同步采集或将匹配窗口缩短至 $\pm 2 \text{ min}$ ，并记录缩短依据；
- c) 匹配窗口内存在多组有效光谱时，可按统一规则取中位数或质量控制后的平均值；使用的扫描编号、时间差和汇总方法应可追溯；
- d) 超出规定时间或空间范围、观测条件发生明显变化的数据不得组成匹配样本。

8 水环境参数反演模型构建与应用

8.1 数据集划分

8.1.1 数据集划分要求

8.1.1.1 应将有效的同步匹配数据集采用分层随机不放回抽样划分为互不重叠的训练集、验证集和测试集三个独立部分。

8.1.1.2 划分比例应符合以下要求：

- 训练数据集应占同步匹配数据集的 60%~70%，用于拟合预处理参数和模型参数；
- 验证数据集应占同步匹配数据集的 15%~20%，用于模型选择、参数调优和防止过拟合；
- 测试数据集应占同步匹配数据集的 15%~20%，在模型确定后一次性用于最终评价；
- 各集合比例之和应为 100%。

8.1.1.3 当将模型应用于空间或时间外推时，测试集应优先由未参与训练的站点、季节或独立采样批次构成。测试结果未达到要求时，应重新进入模型开发流程并使用新的独立测试数据评价。

8.2 模型构建

8.2.1 模型构建要求

8.2.1.1 应遵循单个指标独立建模要求，每个水环境参数对应一个最优反演模型。

8.2.1.2 测试数据集在模型开发过程中应严格隔离，不应用于任何形式的模型训练或参数调优。

8.2.2 输入输出特征构建与选择

8.2.2.1 输入特征可包括 400 nm~1000 nm 范围内的原始反射率、导数、波段组合、光谱指数。

8.2.2.2 应优先选择对水环境参数具有光学响应或数据证据支持的特征。此外，也可采用相关分析（如 $|r| \geq 0.3$ ）、特征重要性排序、递归特征消除或降维方法优选输入特征。筛选过程应在训练集或交叉验证内部完成，并检查共线性和选择稳定性。

8.2.2.3 应将待反演的水环境参数逐个作为输出特征用于训练目标参数反演模型。

8.2.3 模型类型选择

8.2.3.1 反演模型选择应遵循以下原则：

- a) 适用性原则：应根据目标水环境参数的光谱响应特性、数据量大小、数据复杂度及应用场景选择合适的模型类型；
- b) 性能优先原则：在适用性满足时，应优先选择预测精度高、泛化能力强的模型；
- c) 可解释性原则：在精度满足要求时，应优先选择具有一定可解释性的模型；
- d) 效率原则：应考虑模型训练和推理的计算效率，与业务应用需求相匹配。

8.2.3.2 模型类型与选择依据 应根据目标参数光学特性、样本量和数据特征选择反演模型，各类常见模型适应条件见表 3。

表 2 常用反演模型类型与适用条件

模型类型	常用模型	选择依据
统计模型	多元线性回归模型、偏最小二乘回归模型	样本量<200 组，或目标参数具有光谱响应，存在线性关系。

机器学习模型	支持向量机回归模型、随机森林模型、极端梯度提升树模型、轻量梯度提升树模型	样本量 200~2000 组，目标参数与光谱特征存在非线性关系；或目标参数具有光谱响应，存在线性关系。
深度学习模型	深度神经网络模型、卷积神经网络模型、长短期记忆网络模型	样本量 > 2000 组、数据特征复杂、追求更高精度；目标参数与光谱特征存在线性或非线性关系皆可。

8.2.4 模型训练

- 8.2.4.1 可采用格网搜索、随机搜索和贝叶斯优化等方法调优参数，常用模型参数见表 3。
- 8.2.4.2 参数搜索时应监控损失函数变化情况，可采用均方误差（MSE）或平均绝对误差（MAE）作为损失函数。对于需要迭代优化的模型应选择优化器调整模型权重和偏置以最小化损失函数，包括梯度下降（SGD）或 Adam 或 AdamW 三种优化器。
- 8.2.4.3 参数优化搜索过程中，存在下列任一情况时应停止训练优化：
- 所有的候选参数组合均已穷举评估完毕；
 - 达到预设最大的迭代次数；
 - 模型的损失函数性能提升 < 0.1% 且持续 10 轮或验证集性能开始下降。
- 8.2.4.4 应详细记录搜索范围、方法、最佳参数组合及验证数据集上的性能指标。
- 8.2.4.5 模型训练应采用正则化（如 L2 正则化）或早停或降低模型复杂度等方法控制过拟合。

表 3 可调参数示例

模型	主要参数
线性与正则化回归模	变量选择、正则化类型、正则化强度
偏最小二乘回归	潜变量数、标准化方法
支持向量机回归模型	核函数、正则化参数、核系数、epsilon
随机森林模型	树数量、最大深度、最小分裂样本数、最小叶节点样本数、最大特征数
梯度提升树模型	hidden_layer_sizes（隐藏层结构）：元组（如(100, 50)，(64, 32)）
神经网络	网络结构、激活函数、学习率、批大小、优化器、正则化和早停参数

8.3 模型评估与优选

8.3.1 评估方法

模型开发阶段应采用分组交叉验证评估模型性能。采用k（ $k \geq 5$ ）折交叉验证，将数据集划分为k份，依次使用其中一份作为验证数据集，其余作为训练数据集，重复k次取平均。测试集不得参与交叉验证。

8.3.2 评估指标

回归模型至少应包含决定系数（ R^2 ）、均方根误差（RMSE）、平均绝对误差（MAE）和均方误差（MSE），并给出样本数、参数范围及置信区间。计算公式见附录C。

8.3.3 过拟合与欠拟合检验

应对模型是否过拟合和欠拟合现象进行检查。

- 过拟合：训练集性能远好于交叉验证或验证集。可将训练集与验证集 R^2 差值大于 0.1 作为预警信号，同时结合 RMSE、MAE 和残差分布判定。若出现过拟合，需重新训练模型，可采用正则化、增大训练集、简化模型等措施；
- 欠拟合：训练集和验证集性能均不满足应用需求。可将训练集与验证集 R^2 均小于 0.6（或低于应用场景规定的最低精度要求）作为预警。若出现欠拟合，需重新训练模型，可采用核查目标参数的光学可观测性、数据质量和特征有效性，以及增加模型复杂度等措施。

8.3.4 模型优选

候选模型应依据交叉验证和验证集结果选择，最后基于测试数据集通过计算反演水环境参数结果与实测水环境参数值的评估指标（见 8.3.2），根据误差最小原则遴选出性能最好的水环境参数反演模型。

8.4 模型应用

8.4.1 模型部署

模型在独立测试集达到预先规定的验收要求，并完成版本、输入特征、预处理流程和适用域登记后，方可部署至高光谱近感水环境参数反演系统中。

8.4.2 模型调用

将部署后的模型应用于新采集的高光谱近感数据时，新数据的光谱预处理流程、传感器特性及环境条件应与建模数据保持一致。

8.4.3 模型输出

模型输出应包括目标参数值，并同时提供下列信息：

- 反演的水环境参数值；
- 所使用的模型名称及版本标识；
- 模型不确定性说明或适用性声明，特别是当输入数据条件超出建模数据范围时；
- 光谱数据的采集日期和时间、地理位置和数据处理版本信息。

8.4.4 输出数据格式

输出数据应采用机器可读格式，至少支持 CSV、JSON 或 Excel 中的一种，并明确字符编码、时间基准、参数名称、数值、单位、不确定性、质量标识和模型版本等。

8.5 模型更新与维护

8.5.1 模型更新

8.5.1.1 新的野外水环境参数数据获得之后，应及时评估模型反演效果，为未达到规定要求，应优化模型：

- $R^2 \geq 0.9$ ：高精度模型，无需更新模型；
- $0.6 \leq R^2 < 0.9$ ：可接受模型，针对用户应用精度要求，决定是否更新；如趋势监测可不更新模型，若精准数值监测，则需优化模型；
- $R^2 < 0.6$ ：不可接受模型，需更新模型。

注：上述 R^2 阈值适用于模型测试数据集上的评估结果，不同水质参数可依据其光学可反演性适当调整，但调整应有数据支撑并在模型文档中说明。

8.5.1.2 新增代表性样本达到原训练样本量的 20%时，宜评估重新训练的必要性；仅达到触发条件应更新替换现用模型。

8.5.1.3 模型更新应重新执行数据划分、特征选择、训练、独立测试和验收。新模型未通过验收前，原合格模型及其适用域不得被覆盖

8.5.1.4 应建立模型版本更新管理机制，记录模型版本号、输入特征、超参数组合、训练数据集、验证数据集、测试数据集和性能指标等。

8.5.2 模型维护

应对部署的模型每月进行性能监测，发现性能偏离时应立即启动复核、暂停与更新程序。

9 质量控制与评估

9.1 质量控制

9.1.1 设备管理质量控制

9.1.1.1 设备应按制造商要求和附录 A 进行定期校准与维护，周期不应超过 1 年；维修、搬迁、更换关键部件或质量核查不合格后，应在恢复使用前重新校准。

9.1.1.2 应每日核查光谱设备在线状态、存储、供电和通讯状况；应每 6 个月 1 次定期检查包括光谱范围、分辨率等技术指标，发现异常时应立即进行维护或更换。如指标不满足表 1 要求，应立即进行维护或更换。

9.1.1.3 当出现设备故障、环境异常（如水面结冰）等情况时，应暂停数据采集。恢复采集前，必须进行设备校准和质量核查，验收合格后方可重新联网运行。

9.1.2 数据质量控制

9.1.2.1 实测水环境参数

实测水环境参数数据应执行 HJ 91.2-2022 及相应分析方法的质量保证和质量控制要求。

9.1.2.2 高光谱反射率数据

应按下面要求控制高光谱反射率数据质量：

- 每个观测批次前后宜测量具有有效校准证书漫反射参考板。当实测反射率与参考值的相对偏差绝对值大于 5% 时，应核查设备并重新辐射定标；受影响时段数据应重新处理或标记为无效；
- 应使用经审核的同类水体参考光谱计算 SAM。SAM 大于经验验证阈值时，应结合饱和、耀斑、阴影、漂浮物和设备状态判定；不得仅依据单一统计阈值删除原始记录；
- 标准漫反射参考板连续 10 次测量的变异系数大于 5% 时，应检查照明稳定性、积分时间、污染和设备状态；原因消除并重新测量合格前，该组数据不得作为有效数据。

9.1.2.3 实测水环境参数数据

应执行 HJ 91.2-2022 中的质量保证与质量控制规定。

9.1.2.4 反演水环境参数数据

当高频连续数据超出历史同期数据均值 ± 3 倍标准差范围时，标记为“待确认”，需结合现场照片、风速、水位和排污事件联合研判。若无法合理解释的极值应判定为异常无效数据。

9.1.2.5 模型质量控制

- 应按下面要求控制模型质量：
- 数据划分和交叉验证应按站点、批次或时间分组，防止信息泄漏；
 - 模型应同时满足预先规定的独立测试集误差、偏差和稳定性要求；
 - 运行期应监测模型性能，并设置预警限和更新触发条件。

9.2 质量评估

9.2.1 数据准确性评估

9.2.1.1 将反演结果与同步实测数据对比，并进行误差分析，识别系统误差与随机误差。通过计算两者之间的 R^2 、RMSE、MAE、平均偏差、样本数和实测值范围等指标，评估数据的准确性和一致性。

9.2.2 数据稳定性评估

- 9.2.2.1 应在照明和观测几何稳定条件下，对参考板或稳定水体连续采集不少于 10 次，评价光谱重复性。目标波段或代表性光谱指标的相对标准偏差宜小于 10%；超过限值时应调查环境变化和设备状态。
- 9.2.2.2 使用质量控制图（如均值图、标准差图等）监控采集数据的稳定性，出现异常波动时及时采取补救措施。

10 安全要求

10.1 设备安全

设备外壳防护等级不应低于 IP68。支架、接线和供电系统应具备防水、防腐蚀、防雷、接地和漏电保护；浮标及船载设备还应采取防倾覆、防碰撞和防缠绕措施。

10.2 数据安全

高光谱近感反射率数据和反演结果在互联网传输时，应采用加密传输方式，采用国家密码管理局认可的商用密码算法（如 SM4），密钥管理应符合 HJ 212-2025 的相关规定。应建立数据访问权限管理制度，对高光谱原始数据、反演模型和反演结果实行分级授权管理

10.3 野外作业安全

涉水、船载或高处安装作业应开展风险评估，作业人员应佩戴救生衣、安全带等适用防护装备，并落实双人作业和通信联络。雷电、强风、暴雨、结冰或能见度不足时应停止户外作业。

11 参数拓展与持续改进

11.1 参数拓展

- 新增目标参数前，应证明其具有可利用的直接光谱响应，或在限定水体和时段内与光学活性组分存在稳定关系，并明确预定用途。
- 新增参数应具有现行实测方法、代表性同步样本、独立测试数据和验收指标；
 - 新增设备、跨站点迁移或跨水体应用应重新开展一致性比对和独立验证，不得仅复用原模型参数，应进行更新模型；

11.2 持续改进

- 应定期评审目标参数、设备能力、算法和质量控制要求，并通过受控版本更新保持数据连续性和结果可追溯；
- 应定期进行质量控制与评估，每年应不低于 2 次。当评估结果超过允许误差范围时，仪器处于质量失衡状态，应及时排查和维护仪器，并重新测试。依据评估结果调整反演流程、优化操作步骤，通过建立反馈机制促进持续改进；
- 应定期更新软件和人员培训。当软件功能无法满足需求或出现重大技术革新时，应在 2 个月内进行更新和人员培训。

附 录 A
(资料性)
光谱预处理

- A.1 波长校正方法：采用已知特征波长的标准光源（如汞氙灯），将光谱仪实测的特征峰波长位置与理论值进行比对，通过调整光谱仪的波长校正系数，使全波段波长准确度优于 0.5 nm，波长重复性 ≤ 0.14 nm。
- A.2 辐射定标方法：在相同照明和几何条件下，测量标准漫反射板的反射率光谱，将实测值与参考板已知光谱进行拟合，计算辐射定标系数，用于后续测量数据的校正。当反射率相对偏差 $> \pm 5\%$ 时，立即重新定标。
- A.3 暗电流校正方法：采集前关闭镜头或用标准暗参考进行全波段暗电流采集；每条水体光谱等于原始水体光谱减去同期暗电流光谱，消除传感器热噪声与系统基线偏移。

附 录 B
(资料性)
光谱角计算

B.1 引入 SAM 对光谱的一致性进行评价，SAM 越小，则待测光谱与参考光谱的形状相似性越高。
SAM(x,y)的计算见公式 (B.1)。

$$SAM(x,y) = \cos^{-1} \frac{\sum_{i=1}^n x_i \cdot y_i}{\sqrt{\sum_{i=1}^n x_i^2} \cdot \sqrt{\sum_{i=1}^n y_i^2}} \dots\dots\dots (B.1)$$

式中：
SAM(x,y)——光谱角；
 x_i ——待测光谱第i波段反射率；
 y_i ——参考光谱第i波段反射率。

- B.2 根据 SAM 值对光谱的一致性评价，一般认为：
- $SAM < 0.05 \text{ rad}$ ($\approx 2.9^\circ$)：高度相似，可以认为是同一类；
 - $0.05 \leq SAM < 0.10 \text{ rad}$ ($\approx 2.9^\circ - 5.7^\circ$)：相似，通常仍归为同一类；
 - $0.10 \leq SAM < 0.20 \text{ rad}$ ($\approx 5.7^\circ - 11.5^\circ$)：相似性一般，视具体任务决定；
 - $SAM > 0.20 \text{ rad}$ ($\approx 11.5^\circ$)：通常认为不是同一类。

附录 C
(资料性)
模型精度评价指标

C.1 决定系数 (R²)

用于衡量反演模型对实测水环境参数总变异量的解释程度。R² 值越接近1，表示模型越能够解释实测值变化的规律。计算公式见公式 (C.1)。

$$R^2 = 1 - \frac{\sum_{i=1}^n (a_i - b_i)^2}{\sum_{i=1}^n (a_i - \bar{a})^2} \dots\dots\dots (C.1)$$

式中：
R²——决定系数；
a_i——实验室测量的真实值；
b_i——模型反演值；
 $\bar{a}$ ——真实值的均值；
n——代表样本数。

C.2 均方根误差 (RMSE)

量化模型反演值与真实值之间的平均偏差程度，反映模型反演值的离散程度和稳定性。RMSE值越小，说明模型预测效果越好。计算公式见公式 (C.2)。

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (a_i - b_i)^2}{n}} \dots\dots\dots (C.2)$$

式中：
RMSE——均方根误差；
a_i——实验室测量的真实值；
b_i——模型反演值；
n——代表样本数。

C.3 均方误差 (MSE)

衡量反演值与真实值之间差异的平方平均值。计算公式见公式 (C.3)。

$$MSE = \frac{1}{n} \times \sum_{i=1}^n (a_i - b_i)^2 \dots\dots\dots (C.3)$$

式中：
MSE——均方误差；
a_i——实验室测量的真实值；
b_i——模型反演值；
n——代表样本数。

C.4 平均绝对误差 (MAE)

反演值与真实值之间绝对误差的平均值，反映平均误差幅度，越小表明模型预测结果越精确。计算公式见公式（C.4）。

$$MAE = \frac{1}{n} \times \sum_{i=1}^n |a_i - b_i| \cdots \cdots \cdots (C.4)$$

- 式中：
- MAE——平均绝对误差；
 - a_i ——实验室测量的真实值；
 - b_i ——模型反演值；
 - n——代表样本数。

参 考 文 献

- [1] GB/T 31010-2014 色散型高光谱遥感器实验室光谱定标
 - [2] GB/T 36301—2018 航天高光谱成像数据预处理产品分级
 - [3] GB/T 36540-2018 水体可见光-短波红外光谱反射率测量
-