

T/GDHYXH

广东海洋协会团体标准

T/GDHYXH XXXXX—XXXX

高频地波雷达数据质量评估与异常处理 规范

Technical Protocol for Quality Assessment and Anomaly Handling of High
Frequency Ground Wave Radar Data

(征求意见稿)

2026 - XX - XX 发布

2026 - XX - XX 实施

广东海洋协会团体标准 发布

目 次

前言	I
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 要求	2
4.1 数据质量要求	2
4.2 异常数据分类与处理要求	5
4.3 数据填补要求	6
4.4 数据质量等级要求	7
5 方法	7
5.1 数据质量评估方法	7
5.2 异常检测与处理方法	10
5.3 数据填补方法	11
5.4 数据质量等级评定方法	12
附录 A（资料性） 高频地波雷达数据质量评估表	14
附录 B（规范性） 高频地波雷达数据异常分类、代码及检测依据	15
附录 C（资料性） 高频地波雷达异常数据质控记录表	16
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。

本文件由中山大学提出。

本文件由广东海洋协会归口。

本文件起草单位：.....。

本文件主要起草人：.....。

高频地波雷达数据质量评估与异常处理规范

1 范围

本文件规定了高频地波雷达（HFGWR）数据质量评估指标、方法、流程及异常分类与处理要求。

本文件适用于HFGWR在海表流、海面风、海浪等参数探测中的数据质量控制，涵盖实时监测与延时资料处理场景，为雷达运营单位、海洋观测站、数据处理中心、监管机构及科研院所等相关方提供数据质量控制依据。

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 12763.2 海洋调查规范 第2部分：海洋水文观测
- GB/T 14914.4—2021 海洋观测规范 第4部分：岸基雷达观测
- HY/T 0279—2019 高频地波雷达电性能检验方法
- HY/T 0280—2019 高频地波雷达现场比测试验规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

高频地波雷达 high frequency ground wave radar

一种利用高频电磁波沿地表传播，并基于海洋表面相互作用的基本原理，获取海表流（场）、波浪（场）、海表风（场）等海洋动力学参数的雷达。

[来源：HY/T 0279—2019, 3.1, HY/T 0280—2019, 3.1, 有修改]

3.2

Bragg峰 Bragg Peak

海洋表面对雷达电磁波的一阶布拉格共振散射产生的特征谱峰，其频率偏移量用于反演海面径向流速度，是高频地波雷达数据质量判定的基础依据。。

3.3

核心区 Core Area

指雷达波束主瓣有效覆盖、信噪比稳定高于阈值的区域，通常位于雷达探测范围的中近场，数据质量较高且稳定。

3.4

边缘区 Marginal Area

指雷达探测范围的远场边界区域，信噪比自然衰减，数据获取率和质量低于核心区。

3.5

海表流 surface ocean current

海洋表面海水的宏观流动，以流速和流向表征。

注：流速单位为cm/s。流向指海水流去的方向，单位为度（°），正北为零，顺时针计量。

[来源：GB/T 12763.2—2007，3.5，有修改]

3.6

径向流 radial flow

流体运动速度在雷达观测视线方向上的投影分量。

3.7

矢量流场 vector flow field

由至少两个非共线雷达站的径向流合成得到的海流，以矢量形式呈现，包含速度大小和运动方向的完整流场描述。

4 要求

4.1 数据质量要求

4.1.1 数据完整性

4.1.1.1 空间数据获取率

空间数据获取率应满足HY/T 0280—2019规定的以下要求：

a) 核心区

数据获取率应不小于90%。

b) 边缘区

数据获取率应不小于80%。

4.1.1.2 时间连续性检验

数据时间序列应连续。若同一观测要素的连续有效数据点间实际时间间隔超出雷达系统设定的标准观测周期时，应判定该时段存在数据缺失，并记录数据缺失起止时间及缺失间隔时长。

4.1.1.3 无效值占比

a) 单次数据

数据中的无效值数量占数据总量的比例应不大于30%。

b) 关键时段

数据中的无效值数量占数据总量的比例应不大于20%。关键时段应包括：

1) 重大气象海洋过程期

根据权威气象海洋部门发布的预报或实况，在雷达探测覆盖区域内出现或预计出现以下情况时：

——台风预警生效期间。

——风暴潮警报生效期间。

——寒潮、大风等恶劣天气过程。

2) 特定任务保障期

为满足重大科研、国防或公共服务等特定任务而设定的专项观测期。

3) 重要现象观测窗口期

根据科研或业务需求，预先确定的用于捕捉特定海洋现象的观测期。

4.1.2 数据准确性

4.1.2.1 现场观测

高频地波雷达数据应通过现场观测验证其准确性。高频地波雷达数据与现场观测数据的均方根误差（RMSE）应不大于针对该海洋要素及应用场景所确定的预设阈值。现场观测应符合以下要求：

- c) 观测频次
每年至少开展 1 次，单次连续观测时间不少于 25 小时。
- d) 环境条件
观测区域应满足 HY/T 0280—2019 规定的以下环境条件：
 - 1) 观测区域内不应有影响高频地波雷达测量的海上构筑物。
 - 2) 高频地波雷达测量区域连续无遮挡角度应不小于 50°。
 - 3) 观测区域水深应不小于 15m。
- e) 站位布设
观测站位或观测断面应合理布设，站点布设应覆盖高、中、低数据密度及易干扰区，站点布设位置和数量应满足观测工作需要。
- f) 观测设备
观测设备应经计量校准合格，其精度至少优于雷达系统标称精度一个量级。
- g) 准确度要求
在此基础上测得的数据应符合 GB/T 14914.4—2021 规定的准确度：
 - 1) 海表流速度
单位为厘米每秒（cm/s），准确度为 20cm/s；
 - 2) 海表流方向
单位为度（°），准确度为 30°。

4.1.2.2 物理阈值验证

高频地波雷达数据应在合理的物理阈值范围内：

- a) 全局范围阈值
 - 径向流：径向速度应在 0~300cm/s 范围内。
 - 矢量流场：合成的矢量流场流速应在 0~300cm/s 范围内，流向不应出现 180°突变。
- b) 局部范围阈值
对于受特殊海洋动力过程影响的局部海域，其径向速度阈值范围应根据该海域的自然地理条件与海洋动力特征进行差异化设定。典型区域类型的阈值设定应考虑以下因素：
 - 1) 热带气旋影响区域
阈值应能适配台风、热带低压等天气系统影响下可能出现的极端流场变化范围。
 - 2) 近岸浅水区域
阈值应符合观测海域地形和近岸潮汐、风生流主导的日常和季节性流场运动规律。
 - 3) 河口盐淡水交汇区域
阈值应能匹配河流下泄径流、盐度锋面及河口径潮动力叠加等特殊动力环境的影响。
 - 4) 其他特殊区域
应根据观测海域具体特征（如半封闭海湾、潟湖等），设定适用的阈值。
- c) 智能化辅助验证

对于受多因素耦合影响的复杂场景，或传统固定阈值难以准确判定的情况，应采用 4.1.4 智能化评估方法进行辅助验证。智能化评估结果应与物理阈值验证结果综合判断，作为数据准确性评定的参考依据。

4.1.3 数据一致性

4.1.3.1 内部一致性

高频地波雷达数据应满足以下内部一致性要求：

a) 时间序列突变检查

——径向流：3 小时内径向速度变化应不大于 100cm/s，若变化大于此值，应判定为时间序列异常；3 小时内径向速度变化应不小于雷达系统在该海域的径向流测量精度，若变化小于此值，应判定为传感器卡滞，该数据应标记为无效数据。

——矢量流场：除满足径向流突变检查要求外，其 3 小时内相邻时段合成流场的整体流向偏差均值应不大于 30°，若大于此值，应判定为时间序列异常。

b) 空间连续性检验

——径向流：对单测站数据，其相邻网格点之间的流速差异应不大于 50%，若大于此值，应判定为空间异常。

——矢量流场：除满足径向流空间连续性要求外，其在多测站合成区域内，相邻网格点流向差异大于 45°的网格对占比应不大于 5%，若大于此比例，应判定为空间异常。

注：具体阈值可根据观测海域动力环境调整。

4.1.3.2 外部一致性

高频地波雷达数据应满足以下外部一致性要求：

a) 多雷达站重叠区数据比对

应选取多雷达站重叠观测区域，比对该区域内相同时段多组雷达的观测数据，计算 RMSE 作为一致性指标。RMSE 应符合根据雷达性能及观测场景所预先设定的指标阈值。若计算结果偏离该阈值，应进一步分析排查。

b) 遥感数据交叉验证

应将高频地波雷达数据与同期、同区域的卫星遥感反演数据进行交叉验证，计算相关系数(R)、RMSE 作为一致性指标。这些指标应符合根据卫星遥感反演精度及观测场景所预先设定的指标阈值，若指标计算结果偏离该阈值，应检查数据处理过程。

c) 智能化交叉验证

当传统点对点比对方法难以发现复杂的一致性偏差时，应参照智能化评估的输出结果进行综合判断。

4.1.4 智能化评估

高频地波雷达数据宜采用智能化评估方法作为传统评估方法的补充与增强。智能化评估应符合以下要求：

a) 评估能力要求

系统应具备基于机器学习模型对数据质量进行自动评估的能力。

b) 评估结果应用

智能化评估结果应作为数据质量评定的参考依据，与传统方法评估结果综合判断后，形成最终质量结论。

c) 模型管理要求

- 1) 用于质量评估的模型应基于历史符合质量要求的数据与典型异常数据进行训练,并定期更新。
- 2) 当雷达设备参数调整或观测海域环境发生显著变化时,应重新训练或更新模型。
- 3) 模型更新后应通过现场观测数据进行验证,验证误差应符合行业技术要求。

4.2 异常数据分类与处理要求

4.2.1 异常数据分类要求

高频地波雷达异常数据应依据异常成因分为系统故障类、环境干扰类、算法异常类。每类异常的认定,应符合其对应的判定条件与特征描述。各类异常的具体代码、检测指标与量化阈值见附录B。

4.2.1.1 系统故障类

当异常数据由雷达系统故障直接导致,并符合以下任一情况时,应认定为系统故障类异常:

- a) 天线系统故障
天线驻波比异常、天线角度偏移导致信号接收偏差。
- b) 接收/发射设备故障
接收机无信号输入、发射机功率过低。
- c) 伺服系统故障
雷达观测位置漂移,导致数据空间定位错误。
- d) 市政供电中断
市政电网供应异常,导致雷达系统电力供应完全丧失,设备停机、数据采集中断。

4.2.1.2 环境干扰类

当异常数据由外部环境因素干扰直接导致,并符合以下任一情况时,应认定为环境干扰类异常:

- a) 射频干扰
工业电磁辐射、船舶雷达信号叠加等因素导致噪底抬升率 $A_n > 10$ 。计算公式如下:

$$A_n = \frac{N_c}{N_b} \quad (1)$$

其中, A_n 为噪底抬升率, N_c 为当前噪声基底功率, N_b 为基准噪声基底功率。

- b) 电离层杂波
电离层扰动导致电磁波异常传播,产生超视距、宽频带的非海面回波。当超视距回波强度占比 $R_b > 20\%$ 时,判定数据受电离层杂波严重污染。计算公式如下:

$$R_b = \frac{\sum p_b}{\sum p_a} \times 100\% \quad (2)$$

其中, p_b 为超视距区域总功率, p_a 为全距离门总功率。

- c) 极端天气干扰
台风、强雷暴等极端天气导致数据采集中断。

4.2.1.3 算法异常类

当异常数据由数据处理或合成算法错误直接导致，并符合以下任一情况时，应认定为算法异常类异常：

- a) 反演算法异常
一阶 Bragg 峰提取错误，导致计算的径向流速度超出物理阈值。
- b) 合成算法异常
矢量流场合成过程中，因单站数据协同错误导致流向突变、相邻网格流向差异过大。

4.2.2 异常数据处理要求

4.2.2.1 处理原则

高频地波雷达异常数据的处理应遵循以下原则：

- a) 应根据 4.2.1 确定的异常类型，采取针对性处理措施。
- b) 处理应以恢复或保障数据整体质量为目标，确保数据符合数据质量要求。
- c) 处理过程不应引入不可追溯的系统性新误差，所有操作须具备可复现性，相关操作细节应完整记录归档，确保追溯与核验。

4.2.2.2 处理结果要求

高频地波雷达异常数据经处理后，应满足以下要求：

- a) 数据可用性恢复
数据应符合数据质量要求规定的完整性、准确性、一致性要求。若无法恢复，应明确标识为“不可用”。
- b) 系统状态恢复
因设备或环境导致的异常，处理后系统运行状态或观测环境应恢复至可保障数据正常采集与处理的水平。
- c) 过程可追溯
异常识别、处理决策、实施过程及效果验证的全过程应完整记录。

4.3 数据填补要求

4.3.1 填补触发条件

满足以下任一条件时，应触发数据填补流程：

- a) 单次数据
数据输出无效值占比 $> 30\%$ 。
- b) 关键时段
数据缺失占比 $> 20\%$ 。关键时段应满足 4.1.1.3 的相关情景。

4.3.2 填补可行性及基本要求

数据填补应满足以下要求：

- a) 应在存在有效参考数据（如缺失前后有效数据、邻近站点数据、数值模型输出等）的前提下进行。
- b) 填补后的数据应重新验证其是否满足数据质量要求。
- c) 通过验证的填补数据，应在元数据中记录填补方式、填补时段、参考数据来源及验证结论。
- d) 无有效参考数据或填补后验证不达标的数据，应标注“未填补”或“未采用”，并说明原因。

4.4 数据质量等级要求

4.4.1 等级划分

高频地波雷达数据质量等级应划分为“优秀、合格、不合格”三个等级。

4.4.2 等级要求

数据质量等级应满足表1的要求。

表1 数据质量等级要求

质量等级	评估要求	适用场景
优秀	1. 原始数据符合数据质量要求； 2. 经过数据异常处理后，整体数据中异常数据占比 $\leq 5\%$ 。	1. 海洋灾害预警业务； 2. 高精度海洋科学研究； 3. 海上交通保障核心决策。
合格	1. 原始数据符合数据质量要求； 2. 经过数据异常处理后，整体数据中异常数据占比 $> 5\%$ ，且 $\leq 10\%$ 。	1. 常规海洋环境监测； 2. 中低精度科学研究； 3. 海上作业辅助参考。
不合格	不满足优秀和合格等级的数据。	1. 仅可用于设备状态分析； 2. 不可用于业务预警、科研分析及海上作业决策。

5 方法

5.1 数据质量评估方法

高频地波雷达数据质量评估通过以下方法执行，评估结果记录于《高频地波雷达数据质量评估表》（附录A 表A.1）中。

5.1.1 完整性评估

完整性通过以下步骤评估：

a) 空间数据获取率统计

分别统计评估时段内核心区和边缘区的有效网格数与总网格数，并计算空间数据获取率。将计算得到的空间数据获取率与4.1.1.1规定的阈值进行比对，判定其是否符合要求。

b) 时间连续性检验

基于预设观测时间间隔，检测数据时间序列，识别并记录超出标准间隔的数据缺失时段及其时长。按照4.1.1.2的规定判定其是否符合要求。

c) 无效值占比计算

统计数据中被标记为无效的数据数量，计算其占总数据量的比值。将计算得到的无效值占比与4.1.1.3规定的阈值进行比对，判定其是否符合要求。

5.1.2 准确性评估

准确性通过以下步骤评估：

a) 现场观测验证

按 4.1.2.1 的要求, 根据同步采集的雷达数据与现场观测数据, 计算 RMSE。将计算得到的 RMSE 值与确定的预设阈值进行比对, 判定其是否符合要求。

注: 执行验证前, 应已根据具体海洋要素与应用场景确定 RMSE 的预设阈值。

b) 物理阈值验证

将观测数据与 4.1.2.2 规定的物理阈值范围进行逐项比对, 识别并记录所有超出阈值的数据。

c) 智能化辅助验证

- 1) 对于受多因素耦合影响的复杂场景, 或传统固定阈值难以准确判定的情况, 参照 5.1.4 智能化评估方法执行辅助验证。
- 2) 将智能化评估输出的数据质量结果与物理阈值验证结果进行综合比对。
- 3) 当智能化评估结果与物理阈值验证结果一致时, 按一致结果进行判定; 当存在分歧时, 应以物理阈值验证结果为基础, 同时记录智能化评估的参考意见, 并将分歧样本纳入模型优化样本库。

5.1.3 一致性评估

一致性通过以下步骤评估:

a) 内部一致性分析

1) 时间序列突变检查

计算 3 小时滑动窗口内的流速变化, 根据数据类型, 对照 4.1.3.1 a) 中规定的相应阈值进行检验, 识别时间序列异常数据。

2) 空间连续性检验

计算单站相邻网格点间的流速或流向差异, 根据数据类型, 对照 4.1.3.1 b) 中规定的相应阈值, 识别空间异常数据。

b) 外部一致性比对

按 4.1.3.2 的要求, 执行以下比对与判定:

1) 多雷达站重叠区数据比对

提取重叠区同期数据, 计算 RMSE。将计算得到的 RMSE 值与确定的预设阈值进行比对, 判定其是否符合要求。

2) 遥感数据交叉验证

获取同期卫星遥感数据, 计算 R 和 RMSE。将计算得到的 R 和 RMSE 值与确定的预设阈值进行比对, 判定其是否符合要求。

注: 执行比对前, 应已根据雷达性能、观测场景及所用卫星产品精度, 确定符合 4.1.3.2 中规定的预设指标阈值。

c) 智能化交叉验证

- 1) 对于传统点对点比对难以发现的复杂一致性偏差, 参照 5.1.4 智能化评估方法执行多源数据融合评估。
- 2) 将智能化评估输出的数据质量结果与外部一致性比对结果进行综合比对。
- 3) 当智能化评估结果与外部一致性比对结果一致时, 按一致结果进行判定; 当存在分歧时, 应以外部一致性比对结果为基础, 同时记录智能化评估的参考意见, 并将分歧样本纳入模型优化样本库。

5.1.4 智能化评估

智能化评估通过以下步骤执行, 作为传统评估方法的补充与增强:

a) 模型构建

- 1) 数据获取

收集历史雷达数据，包含合格数据与典型异常数据。合格数据应覆盖我国近岸主要观测区域及常见海况；异常数据应包含信号干扰、设备故障、数据异常等常见类型，并确保各类异常样本占比均衡。
 - 2) 数据标注

对数据进行标注，明确每个样本的质量类别（如正常、射频干扰、电离层杂波、设备故障等）。
 - 3) 特征工程

从雷达回波原始数据中提取与数据质量强相关的特征，对特征进行归一化、标准化等预处理，形成模型输入向量。提取特征包括但不限于：

 - (1) 原始回波特征

信噪比、噪底抬升率、频谱峰度、一阶布拉格峰功率与位置、超视距回波强度占比等。
 - (2) 衍生特征

径向流速时间序列统计特征（均值、方差、变化率）、相邻网格点空间梯度、有效数据获取率等。
 - 4) 模型选型与训练
 - (1) 模型选型

选用经国内行业实践验证的算法模型（如随机森林、梯度提升树、神经网络等），优先适配海流数据时序特性的模型。
 - (2) 模型训练

使用标注好的训练数据集对模型进行训练，通过交叉验证等方式优化模型参数。
- b) 评估指标选定
- 模型训练完成后，须从均方根误差（RMSE）、平均绝对误差（MAE）、相关系数（ R^2 ）等性能指标中选取合适的指标进行评估，确保模型满足业务应用要求。
- c) 模型输入与输出
- 1) 输入准备

提取待评估时段内的雷达回波原始数据及衍生特征。按模型要求进行数据预处理，形成标准化的输入向量。
 - 2) 模型推理

将输入数据载入已训练好的质量评估模型，执行推理计算。
 - 3) 输出结果

异常类别：输出具体的异常类别。

质量评分：百分制或 0-1 区间评分，反映数据的综合质量水平。

异常定位：记录异常数据对应的时间、空间位置及置信度。
- d) 模型更新与验证
- 1) 定期更新

每 6 个月基于新增数据对模型进行迭代优化。
 - 2) 触发更新

当雷达设备参数调整、观测海域环境显著变化，或模型准确率持续下降，降幅大于 5% 时，立即重新训练模型。
 - 3) 验证要求

更新后的模型须通过国内典型海域现场观测数据验证，验证误差应符合行业技术要求，方可重新投入使用。

5.2 异常检测与处理方法

高频地波雷达数据异常检测与处理通过以下方法执行，检测和处理结果记录于《高频地波雷达异常数据质控记录表》（附录C 表C.1）。

5.2.1 异常检测与标注

高频地波雷达数据的异常检测与标注按以下方法进行：

5.2.1.1 异常数据识别

根据4.2.1规定的异常分类要求与附录B中的检测依据，结合相关辅助信息，对待评估数据进行异常筛查，识别出疑似异常数据。识别出的疑似异常数据由专业人员进行复核。

5.2.1.2 异常数据标注

对于经复核确认的异常数据，依据附录B赋予相应的异常代码。异常代码记录于《高频地波雷达异常数据质控记录表》（附录C 表C.1）。

5.2.2 异常数据分类处理

高频地波雷达异常数据处理通过以下步骤进行：

a) 处理措施实施

根据已确认的异常类型，按照相应条款采取具体处理措施。

b) 处理效果验证

处理完成后，按 5.1 数据质量评估方法对已处理的数据重新进行评估，验证其是否已满足 4.1 数据质量要求。

c) 处理记录

将所采取的处理措施、处理时间、验证结果等信息，记录于《高频地波雷达异常数据质控记录表》（附录 C 表 C.1）。

各类型异常数据的处理措施如下：

5.2.2.1 系统故障类

a) 天线系统故障

- 1) 应对该时段、该区域的数据进行异常标记，并暂停在后续业务化产品中输出。
- 2) 重新校准天线角度、修复驻波比。
- 3) 试采集 3 小时，验证信噪比是否达标。

b) 接收/发射设备故障

- 1) 应对该时段、该区域的数据进行异常标记，并暂停在后续业务化产品中输出。
- 2) 检修或更换接收机/发射机。
- 3) 恢复后验证数据获取率是否恢复正常水平。

c) 伺服系统故障

- 1) 应对该时段、该区域的数据进行异常标记，并暂停在后续业务化产品中输出。
- 2) 调整伺服系统定位，通过 GNSS 重新标定位置，确保偏差不大于 5km。
- 3) 恢复后验证相邻网格数据有无定位偏差。

d) 市政供电故障

- 1) 确认故障，执行供电保障或安全关机。联系市政供电部门，确认故障范围与预计恢复时间。

- 若配备备用电源：启动、切换至备用电源供电。
- 若未配备备用电源：执行设备安全关机流程并上报。
- 2) 市电供电恢复后，执行安全启动或供电切换
 - 若系统已关机，执行雷达系统安全启动流程。
 - 若系统由备用电源供电，安全切换至市政供电后关闭备用电源。
- 3) 验证系统运行状态与数据获取率是否恢复正常水平。

5.2.2.2 环境干扰类

- a) 射频干扰

启用预设滤波算法抑制干扰噪声，处理后重新计算 A_n ，同时核查无效值占比，确保其降至合格阈值以下。
- b) 电离层杂波

启用预设滤波算法抑制干扰噪声，处理后重新计算 R_n ，同时核查无效值占比，确保降至合格阈值以下。
- c) 极端天气干扰
 - 1) 暂停户外设备运行。
 - 2) 天气好转后重启设备。
 - 3) 重启后连续采集 3 个预设标准时间间隔的数据，验证数据连续性，无缺失时段后方可恢复正常观测。

5.2.2.3 算法异常类

- a) 反演算法异常
 - 1) 重新标定 Bragg 峰提取参数。
 - 2) 用现场观测数据验证。
 - 3) 剔除超阈值数据。
- b) 合成算法异常
 - 1) 优化矢量流场合成权重参数。
 - 2) 重新合成数据，确保流向无突变。
 - 3) 验证空间连续性。

5.3 数据填补方法

高频地波雷达数据填补通过以下方法执行：

5.3.1 填补实施

当触发 4.3.1 规定的填补条件时，按以下步骤实施数据填补：

- a) 方法选择

根据数据缺失的时长、空间特征及可用参考数据情况，按照 4.3.2 填补可行性的原则，选择适用的填补方法。
- b) 执行填补

运行所选方法对应的填补算法，生成填补数据。
- c) 效果验证

填补完成后，重新对已填补的数据进行评估，以验证其是否已满足 4.1 数据质量要求。
- d) 过程记录

将本次填补的详细信息，包括触发条件、采用方法、填补时段、参考数据、填补时间、效果验证等，记录于《高频地波雷达异常数据质控记录表》（附录 C 表 C.1）。

5.3.2 分类填补方法

5.3.2.1 短期缺失

- a) 适用场景
短时干扰（ ≤ 6 小时），缺失时段前后的数据变化平缓，未出现剧烈变化或非物理跳变，满足线性假设。
- b) 填补方法
时序插值法。
- c) 操作步骤
 - 1) 提取缺失时段前 3 个、后 3 个采样间隔与原观测一致的有效数据点。
 - 2) 采用线性插值法计算缺失时段数据，确保插值结果与前后数据格式统一。

5.3.2.2 中期缺失

- a) 适用场景
局部区域缺失（6 小时~24 小时），周边有 ≥ 2 个正常邻近站点。
- b) 填补方法
空间插值法。
- c) 操作步骤
 - 1) 筛选 3 个邻近站点，站点须满足距离缺失区域 $\leq 50\text{km}$ 、历史同期流速差异 $\leq 20\%$ 。
 - 2) 采用克里金插值法，以邻近站点数据为基础计算缺失区域数据，优先保留缺失区域核心网格的插值结果。

5.3.2.3 长期缺失

- a) 适用场景
长时段故障（ > 24 小时）、极端天气缺失，无有效邻近站点。
- b) 填补方法
数值模型同化填补法。
- c) 操作步骤
 - 1) 选用覆盖缺失海域的海流数值模型（如 ROMS、FVCOM、HYCOM、GLORYS 等），且模型须经前期现场观测验证。
 - 2) 提取缺失时段前后 6 小时的有效雷达数据，作为模型同化的约束条件。
 - 3) 运行数值模型，输出缺失时段的填补数据，确保数据时间步长与雷达观测一致。

5.4 数据质量等级评定方法

5.4.1 评定准备

高频地波雷达数据等级评定基于 5.1 至 5.3 的执行结果。开始评定时，须确认以下工作已完成，并已取得相应结果：

- a) 数据质量评估
已按 5.1 数据质量评估方法完成对数据完整性、准确性、一致性的评估，并获得各项指标的明确评估结论。

- b) 数据异常检测与处理
已按 5.2 异常检测与处理方法完成异常数据的识别与确认。
- c) 数据填补
若已触发过 4.3.1 规定的的数据填补条件,则须确认已按 5.3 数据填补方法完成数据填补及效果验证,并且已对包含填补数据在内的整体数据,重新进行数据质量评估与异常检测。

5.4.2 等级判定方法

数据质量等级的判定按以下流程执行:

5.4.2.1 提取评定参数

从5.4.1确认的结果信息中,提取以下用于等级判定的核心参数:

- a) 原始数据质量达标情况
原始数据是否符合 4.1 数据质量要求。
- b) 异常数据占比
最终数据集中异常数据所占比例。
- c) 整体数据质量达标情况(若适用)
数据经异常处理与填补后,是否符合 4.1 数据质量要求。

5.4.2.2 等级判定

依据5.4.2.1提取的评定参数,对照4.4.2(表1)规定的各质量等级要求进行符合性判定。满足相应等级全部条件的,即判定为该等级。

5.4.2.3 记录判定结论

将最终判定的数据质量等级记录于《高频地波雷达数据质量评估表》(附录A 表A.1)中。

附录 A

(资料性)

高频地波雷达数据质量评估表

表A.1给出了高频地波雷达数据的评估类型、指标及数据质量等级判定的表格范例，可用于开展高频地波雷达数据质量的规范化评估。

表 A.1 高频地波雷达数据质量评估表

评估类型	评估指标	数据质量等级			备注
		优秀	合格	不合格	
完整性	空间数据获取率				
	时间连续性检验				
	无效值占比				
准确性	现场观测				
	物理阈值验证				
一致性	时间序列突变				
	空间连续性检验				
	多雷达站重叠区数据比对				
	遥感数据交叉验证				
数据质量评估结果					

附录 B

(规范性)

高频地波雷达数据异常分类、代码及检测依据

高频地波雷达数据异常的分类、代码、检测指标及量化阈值应符合表B.1的规定。

表 B.1 高频地波雷达数据异常分类、代码及检测依据

异常类型	异常情景	异常代码	检测依据
系统故障类	天线系统故障	H1	天线驻波比超出设备标称值 $\pm 20\%$ ；单站数据获取率在 1 小时内下降超过 30%
	接收/发射设备故障	H2	接收机信号强度低于标称值 30%以上；发射机输出功率低于标称值 20%以上
	伺服系统故障	H3	伺服系统定位监测偏差 $> 5\text{km}$
	市政供电故障	H4	雷达系统电力供应完全丧失，站内所有用电设备同步失电，市电输入接口电压低于标称电压的 20%。
环境干扰类	射频干扰	E1	噪底抬升率 $A_n > 10$
	电离层杂波	E2	超视距回波强度占比 $R_i > 20\%$
	极端天气干扰	E3	同步气象数据显示风力 ≥ 8 级或闪电频次 > 5 次/小时，且数据采集中断
算法异常类	反演算法异常	A1	径向流速度超出 0~300cm/s；与现场观测数据相对误差 $>$ 预设阈值
	合成算法异常	A2	矢量流场流向 180° 突变且相邻网格流向差异 $> 45\%$ 的网格对占比 $> 5\%$

附 录 C

(资料性)

高频地波雷达异常数据质控记录表

表C.1规定了高频地波雷达异常数据的识别类型、质控措施、处理结果及复核记录等内容，用于规范开展异常数据的全流程质量控制与溯源管理。

表 C.1 高频地波雷达异常数据质控记录表

评估类型	评估指标	异常代码	质控措施	处理结果	复核记录
完整性	空间数据获取率				
	时间连续性检验				
	无效值占比				
准确性	现场观测				
	物理阈值验证				
一致性	时间序列突变				
	空间连续性检验				
	多雷达站重叠区数据比对				
	遥感数据交叉验证				

参 考 文 献

- [1] Zhou C, Wei C, Yang F, Wei J. A Quality Control Method for High Frequency Radar Data Based on Machine Learning Neural Networks[J]. Applied Sciences, 2023(13):11826.
- [2] Cosoli S, Bolzon G, Mazzoldi A. A Real-Time and Offline Quality Control Methodology for SeaSonde High-Frequency Radar Currents[J]. Journal of Atmospheric and Oceanic Technology, 2012, 29(9):1313–1328.
- [3] Roarty H, Smith M, Kerfoot J, Kohut J, Glenn S. Automated quality control of High Frequency radar data[C]. In OCEANS 2012 MTS/IEEE, Virginia Beach, VA, USA, 2012.
- [4] Roarty H, Palamara L, Kohut J, Glenn S. Automated quality control of high frequency radar data II[C]. OCEANS 2016 MTS/IEEE Monterey, Monterey, CA, USA, 2016.
- [5] 刘战存. 多普勒和多普勒效应的起源[J]. 物理, 2003, (07):488–491.
- [6] 张苗苗. 高频地波雷达海洋信息反演及质量评估技术研究[D]. 哈尔滨工业大学, 2023.
- [7] 东松林. 基于深度学习的高频地波雷达回波分类与处理技术研究[D]. 武汉大学, 2022.
- [8] 胡余阳, 汪小平, 田甜, 等. 基于人机协同的对空情报雷达受干扰等级评估方法[J]. 现代雷达, 1-14.
-