

团 体 标 准

T/CSNAME 186—2026

绿色智能船舶协同研发平台数据质量评估 方法

Data quality assessment method for green intelligent ship collaborative r&d platform

2026 - 10 - 23 发布

2027 - 01 - 23 实施

中国造船工程学会 发 布

目 次

前言 1

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 数据质量评估流程 1

5 数据质量评估 2

 5.1 确定数据评估范围 2

 5.2 选取指标测度 2

 5.3 质量方案 3

 5.4 执行评估 4

 5.5 结果分析与改进 4

附录 A（资料性） 绿色智能船舶协同研发平台数据质量要求示例 错误!未定义书签。

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。

本文件由中国造船工程学会船舶标准化专业委员会提出。

本文件由中国造船工程学会归口。

本文件起草单位：中国船级社、水上载运装备安全研究院（宁波）有限公司、海船舶运输科学研究所有限公司。

本文件主要起草人：邓林义、张旋武、乔继潘、刘桢、陈景昊、郭安康、任东隼、阮佳佳。



绿色智能船舶协同研发平台数据质量评估方法

1 范围

本文件规定了绿色智能船舶协同研发平台数据质量评估流程、评估内容和评估方法。
本文件适用于绿色智能船舶协同研发平台数据质量评估。

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。

ISO 8000-8 数据治理 第8部分 信息和数据质量概念和测量（Data quality — Part 8: Information and data quality: Concepts and measuring）

IACS Recommendation 183 船舶数据质量（Ship Data Quality）

中国船级社，船舶数据质量评估指南，2021

中国船级社，智能船舶规范，2025

3 术语和定义

本文件没有需要界定的术语和定义。

4 数据质量评估流程

数据质量评估流程见图1所示。

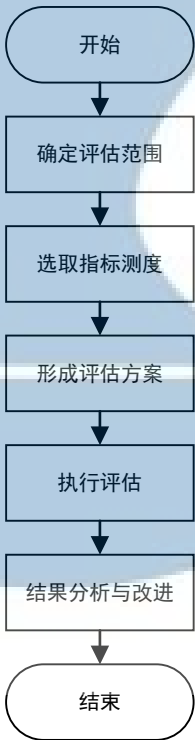


图1 数据质量评估流程

5 数据质量评估

5.1 确定数据评估范围

5.1.1 定义数据范围

为了评估数据质量，应根据绿色智能船舶协同研发平台的功能目标（可参考CCS R012-2025）确定的数据范围。这些数据至少含有以下内容：

- a) 时间/时间戳；
- b) 目标/数据识别；
- c) 协同研发平台功能数据（可参照 CCS GD016 – 2021 附录 1-7 及本文件附录 A）；
- d) 数据源。

5.1.2 元数据要求

绿色智能船舶协同研发平台应建立元数据，对平台内数据内容、数据结构进行定义，包括但不限于数据字段类型、数据格式、值域、精度。元数据应覆盖以下三个维度要求：

- a) 系统特定需求，描述与系统相关的数据特征，这是对系统的统一数据解释和要求；
- b) 技术要求、数据格式、数据源、数据类型等；
- c) 管理要求、数据安全、数据标准等。

5.2 选取指标测度

5.2.1 概述

系统设计阶段，应结合平台业务目标明确数据质量总体要求，包含数据范围、质量维度、质量约束。

示例：开展船舶航行气象数据质量评估时，数据范围应至少包含：风速和风向、有义波高、风浪高度与平均周期、涌高、涌向和平均周期、流速和流向、水深、水温等水文气象条件。

应选取对应的质量测度，包括：

- a) 上述数据采集的完整性；
- b) 数据采集及时性，核查数据是否按规定采集频率向船舶航行相关业务系统提供；
- c) 数据有效性，核查数据是否符合对应数据标准约束。

船舶运行阶段，应依据设计要求开展数据质量评估；可根据评估结果对采集设备实施维护、校准。选取数据质量测度时，应充分考虑当前技术条件下度量工作的可实施性。

5.2.2 数据质量要求

数据质量要求应易于定义，每个数据项的要求都应明确，可能包括数据类型、数据采集频率、格式、范围、计量单位、数据的有效/无效状态、标识等。

数据质量要求应影响数据质量度量维度或特征的选择，数据质量要求的不同测量维度示例见表1。

表1 数据质量要求的不同测量维度示例

质量要求	测量尺寸/特性
数据类型	可以用尺寸来衡量，例如合规性
数据采集频率	可以用连续性、完整性、及时性等维度来衡量。
格式	可以用来测量如合规性、精度等。
范围	可以用来测量如合规性、精度等。
单位	可以用来测量，合规性、精度等。
数据的有效/无效状态	可以用精度、时间、效率等维度来衡量。
标识	可以用来衡量，如合规性、一致性等。

5.2.3 数据质量维度

5.2.3.1 数据标识

数据标识是数据标准化管理的重要基础，应满足以下要求：

- a) 一致性：同一实体或对象在不同数据集中是一致的。

- b) 合规性：数据具有遵守现行标准、惯例或法规以及使用环境中类似规则的属性。例如，数据标识符代码。

5.2.3.2 数据采集

数据收集应满足以下要求：

- a) 完整性：与实体关联的数据具有使用上下文中所有预期属性和相关实体实例的值；
- b) 连续性：根据预定的系统特定要求收集数据，并满足数据收集频率；
- c) 及时性：可以获得符合时间要求的最新数据。

5.2.3.3 数据存储

数据存储应满足以下要求：

- a) 效率：数据可方便地使用。
- b) 精度：数据具有精确的属性，或者在特定的使用环境中提供。例如，数据存储精度可以满足预定要求。
- c) 可用性：数据是否对用户或应用程序可用且有效。
- d) 可恢复性：从数据安全的角度来看，数据可以从备份中恢复。

5.2.3.4 数据集成

针对数据集成过程，应度量数据经过提取、清洗后可用于集成的有效数据情况，数据集成度量维度示例见表 2。

表2 数据集成度量维度示例

尺寸		描述
效率	验证通过率	集成后，验证符合数据定义和规则的数据
	清洁输出率	清理前清理数据与总数据的比率
	质量改进率	清理后的数据质量与清理前的数据质量的百分比
	提取输出速率	提取数据占总源数据的比例

5.2.3.5 数据处理

通过应用系统访问使用情况反映数据利用价值，数据处理度量维度示例见表 3。

表3 数据处理测量维度示例

尺寸		描述
可访问性	数据字段访问率	所有访问字段中的字段比率
	数据集访问率	所有访问数据集中的数据集比率
	平均返回数据	应用程序访问数据集返回的平均数据量

5.3 质量方案

5.3.1 测量周期

应结合数据业务属性与使用目标，设置短期、中期、长期三类数据质量测量周期：

- a) 短期：建议 1 个月；
- b) 中期：建议 4 至 6 个月；
- c) 长期：建议 1 至 2.5 年。

5.3.2 数据质量验证和确认

应开展以下验证确认工作：

- a) 数据值检查：如值域范围校验；
- b) 数据一致性检查：核查数据在网络传输、跨应用流转过程中保持一致性；

c) 数据语义检查：验证数据与现实业务对象、事件、概念的匹配程度。

5.3.3 数据质量报告

数据质量报告应至少包含：评估时间、评估对象、测量方法、评估结果、问题清单和改进建议。

5.4 执行评估

5.4.1 船舶应定期在短期、中期和长期进行数据质量评估。第一次评估应在船舶运行一段时间并积累一定数据后进行。在评估过程中，应首先确定评估数据的范围，选择评估的维度和评估方法，根据质量要求进行评估。此外，不同的数据质量评估措施可能有不同的优先级，应确定优先级以满足数据的系统特定需求。

5.4.2 在每次数据质量评估时，都应对测度进行评估，以确认其是否合适。数据集的评估可通过全面检查或抽样来完成。此外，数据质量的衡量可通过将数据集与预期数据集进行比较来实现，也可通过使用公开可用和广泛接受的数据集来确定数据质量。

示例：对于某些数据质量，如收集频率，首先确定是否收集了足够的数据，然后将收集到的数据与公开可用的数据进行比较，以确定我们使用的数据的质量。

5.4.3 评估方式

评估分为定性评价、定量评价，实际评估可单独使用或组合使用。

5.4.4 评估计算

定量评价可采用以下方法：

- a) 简单比率法，预期结果（E）与总值（T）的比率；
- b) 最小/最大方法，适用于在衡量数据质量时需要聚合多个指标的维度。评价的关键是找到各种指标的最小值或最大值；
- c) 加权平均法，对于复杂的多指标评价，如果可量化每个指标在总体评价中的重要性，可以使用加权平均法。结合需求规格和业务场景需求，按照公式（1）为每一种数据的数据质量需求表中的每一个质量测度制定权重 w_i 。

$$\sum_{i=1}^n w_i = 1 \dots\dots\dots (1)$$

式中：

- w_i ——第*i*个质量测度的权重；
- n ——某一种数据采用的数据质量测度总数。

5.5 结果分析与改进

5.5.1 识别潜在改进项并确定优先级

基于数据质量问题的业务影响程度对问题排序，评估各类改进方案。结合数据分析结果、相关方意见确定修复与改进工作优先级；针对重大质量问题开展根因分析，形成可行改进方案。

5.5.2 制定数据问题处理流程

针对识别的数据质量问题，联合数据使用者、数据质量评估专家开展根因定位。

- a) 非技术类根因处置：开展人员培训，优化业务处理流程；
- b) 技术类根因处置：
 - 1) 直接修正缺陷数据；
 - 2) 优化数据采集、传输、存储环节性能；
 - 3) 改造系统与技术流程，防止同类问题复现。

开展成本-收益权衡，对改进过程实施持续监控。

5.5.3 处置数据质量问题

开展成本效益分析，确定可行处置路径；征求数据使用方、专业专家意见，选择处置方案，分为简单改正措施、根本原因改正两类。

- a) 简单改正措施：直接对记录进行数据修复，例如数据清洗、解析、格式规整；
- b) 根本原因改正：开展系统策略层面变更，制定长期改进计划，改造系统消除问题根源，建立预防机制。预防类措施通常较事后修复成本更低。

编制并落实改正计划；完成修正后，对修正数据集重新开展质量评估，确认整改不会引入新错误，达到预期效果。

5.5.4 监控数据与质量规则符合性

持续监控数据对质量规则的符合状态，输出各级别数据质量评估结果。应完成以下记录：

- a) 数据质量规则、指标、测度、质量得分；
- b) 数据质量随时间变化趋势；
- c) 跟踪整改计划对应的质量问题闭环情况。

按照预设周期重复开展质量评估；设置指标阈值与接收准则。评估结果可采用合格数据占比或异常数据占比表达。

当数据满足质量规则，判定数据适配预期业务用途（如协同研发数据分析）；当数据不满足质量规则时，触发告警，通知相关方，并依据补救计划执行处置措施。



附录 A
(资料性)

绿色智能船舶协同研发平台数据质量要求示例

A.1 绿色智能船舶协同研发平台适用船型

集装箱船、散货船、油船、纸浆船、半潜船、客滚船、化学品船

A.2 绿色智能船舶协同研发平台适用数据范围及频率

- 船舶航行过程中，由船端集成平台采集并发送到岸端协同研发平台的船舶状况、设备运行、周边环境相关的数据，主要包括：
- 通导数据：GPS、计程仪、电罗经、自动舵/舵、倾斜仪、风速风向仪、测深仪等；默认 1 s-2 s/条，可根据船东需求调整存储频率。
- 机舱数据：主机（含推进）、发电机（含轴带）、锅炉、配电板、供油单元、分油机、风机、轴系、轴功率仪、燃油流量计、消防泵、辅助设备 etc；默认 2 s-4 s/条，根据 AMS 数据源和船东需求调整存储频率。
- 甲板数据：四面吃水、压载舱液位容积、燃油舱液位温度容积、货油舱液位温度容积、应力应变数据；默认 2 s-4 s/条，根据 AMS 数据源和船东需求调整存储频率。
- 其他数据：特殊船型数据，如电池动力船舶的电池数据，客滚船的空调数据等。默认 2 s-4 s/条，根据 AMS 数据源和船东需求调整存储频率。