

智慧水利背景下空间数据标准化治理体系构建分析

董剑锋 崔佳琪 苏晶 王辉

常州市水资源服务中心, 江苏 常州 213000

[摘要]智慧水利平台持续接入河湖岸线、水利工程、监测站点、三维模型及跨部门共享数据, 空间基准差异和对象编码分散逐步影响专题研判。面向一体化平台数据资源中心, 研究构建由标准框架、目录归集、语义映射、质量校验及反馈更新组成的治理体系, 并引入来源权重估计、稳健配准、对象匹配置信度计算及规则参数更新等方法。工程试验结果表明, 该体系能够降低空间残差, 减少对象误配及拓扑异常, 保持来源记录可核验, 提高专题图查询和编制效率, 可为河湖管控、水旱灾害防御、水资源监管及工程运行提供稳定的数据基础。

[关键词]智慧水利; 空间数据; 标准化治理; 语义映射; 质量校验

DOI: 10.64635/ja.2026.1424

中图分类号: TV213.4

文献标识码: A

Analysis of the Construction of a Standardized Spatial Data Governance System in the Context of Smart Water Conservancy

Dong Jianfeng, Cui Jiaqi, Su Jing, Wang Hui

Changzhou Water Resources Service Center, Changzhou, Jiangsu 213000, China

Abstract: Smart water conservancy platforms continuously access data on river and lake shorelines, water conservancy projects, monitoring stations, three-dimensional models, and cross-departmental shared resources. Differences in spatial datums and scattered object coding have gradually affected thematic analysis and judgment. Oriented toward the data resource center of an integrated platform, this paper studies the construction of a governance system composed of a standard framework, catalog aggregation, semantic mapping, quality verification, and feedback updating. Methods such as source weight estimation, robust registration, object matching confidence calculation, and rule parameter updating are introduced. Engineering test results show that the system can reduce spatial residuals, decrease object mismatching and topological anomalies, maintain verifiable source records, and improve the efficiency of thematic map query and compilation. It can provide a stable data foundation for river and lake management, flood and drought disaster prevention, water resources supervision, and project operation.

Keywords: smart water conservancy; spatial data; standardized governance; semantic mapping; quality verification

引言

智慧水利建设正由单项业务系统汇集转向数据资源中心支持的综合应用。智慧水利一体化平台已形成一个数据中心、一个门户、一张图及多个业务应用的总体架构, 数据资源池涵盖基础数据、业务数据、监测数据、三维数据、空间数据和视频数据, 后续建设还包括河湖管理、防汛物资、节点工程三维及工程监测等专题数据库。不同系统形成于不同建设时期, 坐标参考、字段口径、对象编码和更新时间存在差异, 同一水利对象进入专题图后容易产生位置偏移、名称异写、属性缺项及状态不一致。空间数据标准化需贯穿数据接入、对象识别、质量修复、专题服务和规则更新全过程, 为跨业务查询及空间分析提供可信

依据。

1 既有智慧水利空间数据标准化治理的运行局限

现有平台虽能汇聚多类数据, 但治理仍依赖格式转换、字段清洗和固定阈值检查, 难以准确判定来源记录是否指向同一河段、监测站或工程。河湖长制、防汛、水资源等系统在对象边界、状态字段和时间粒度方面存在差异, 固定距离阈值在密集河网易引发相邻支流误配, 在丘陵区域又可能遗漏同源对象。部分质量规则仅在首次入库时运行, 巡查纠错、专题图核验和工程改造信息未能反馈至标准库; 来源版本保存不足还会造成旧成果覆盖新数据^[1]。现有研究虽涉及数据安全、多源融合和对象编码, 但专题数据仓仍缺少标准配置、对象映射、质量校验及应

用反馈贯通的运行机制。为此,需将标准版本、主对象编码、时空质量和业务反馈纳入统一治理过程,使修订结果能够追溯来源,并反馈至下一轮数据接入。

2 智慧水利空间数据标准化治理体系构建

2.1 空间数据标准框架分层配置

标准框架以水利对象作为基本单元,为河湖、闸泵站、监测站、取排水口、涉河项目及工程构件设置主编码。框架划分为标识、基准、语义、质量和服务五层,分别约束对象身份、时空基准、字段含义、质量条件及服务范围。规则记录版本号、生效时间和适用对象,可支持历史成果回溯计算。标准配置采用公共规则叠加对象规则的方式,统一坐标系统、时间格式、计量单位及来源信息的规范,校核河道连通、湖泊闭合、站点归属、工程边界和三维投影。规则发布前开展样本验证,运行后记录命中、误判及修订结果,使目录归集、对象匹配和质量校验使用统一版本规则^[2]。局部规则更新时,系统依据引用关系定位受影响数据,减少全库重复处理,并为目录归集提供约束。

2.2 多源数据资源目录统一归集

数据目录承接前述标准框架,作为多源数据进入治理的统一入口。相关研究指出,水资源监管平台有必要将不同渠道的资料进行集成,其中包括监测数据、气象数据和地形数据,以便对水资源状况有一个综合的认识^[3]。据此,目录记录数据名称、责任单位、对象类型、覆盖范围、坐标参考、更新方式、字段版本、质量状态和服务去向。文件、数据库表及接口报文接入后,系统读取元数据并生成来源记录;坐标缺失、版本不明或责任主体无法确认的数据进入隔离区,信息补齐后再转入标准区。

归集区划分为原始区、标准区、待审区和服务区,分别承担源数据留存、坐标字段转换、异常复核及专题产品生成。目录维护主编码和来源标识码,增量数据仅处理变更记录。静态边界依据版本更新,动态监测数据先关联稳定对象再入库,三维模型保存二维范围及坐标转换参数。每次转换均保留输入版本、转换规则和处理时间,为语义映射及质量追溯提供依据。

2.3 空间对象语义模型规范映射

目录限定候选范围后,语义模型把来源记录映射至主对象。模型整合基础属性、空间特征、业务关系、时间状态及来源信息,并保留正式名称、曾用名和系统简称。河段选用缓冲相交检索,站点选用邻近检索,工程面选用包围范围检索,再依据多类证据完成匹配,减少全库逐项比对造成的计算开销。来源记录*i*和主对象*j*的匹配置信

度定义为:

$$C_{ij} = w_n s_{n,ij} + w_a s_{a,ij} + w_g \exp\left(-\frac{d_{ij}}{\sigma}\right) + w_t s_{t,ij} + w_p s_{p,ij},$$

$$\sum_{r \in \{n,a,g,t,p\}} w_r = 1 \quad (1)$$

式中, C_{ij} 为综合置信度; $s_{n,ij}$ 为名称相似度; $s_{a,ij}$ 为属性相容度; d_{ij} 为空间距离; σ 为距离衰减尺度; $s_{t,ij}$ 为时间状态一致度; $s_{p,ij}$ 为业务关系一致度; w_r 为对应证据权重。公式(1)用于开展候选排序及对象归并。置信度高于确认阈值时,来源记录自动挂接至主对象;处于复核区间时,系统输出名称差异、空间距离、属性差异及来源依据;低于复核阈值时,该记录继续保留为独立候选,避免证据不足的对象进入正式库。

距离项选用衰减函数,可缓解固定距离阈值存在的区域适应问题。主城区河网较密集时,减少业务关系及拓扑位置能够限制相邻支流误配;丘陵区对象较分散时,名称、属性和所属流域仍可支持同源判断^[4]。主对象编码发布后不随名称变化,来源记录仍保留原编码,使平台能够依据历史版本查找对象变化。该模型还可处理监测站迁移、工程改造及河段更名,保证对象状态变化不会直接形成重复档案。

2.4 时空数据质量规则分级校验

对象身份确定后,质量校验依据阻断、复核和提示三个等级开展。阻断规则用于处理坐标缺失、几何无效、主编码重复及时间倒置;复核规则用于处理边界重叠、站点偏离所属河段、工程范围跨越管理边界等问题;提示规则则关注字段缺项、更新延迟及精度下降^[5]。点对象校核重复位置和归属关系,线对象校核断裂、悬挂及流向,面对象校核闭合、自相交和缝隙,三维对象校核高程基准及二维投影范围,监测记录校核采集间隔和计量口径。源坐标校正采用基于来源权重的稳健估计方法。设原始控制点为,统一基准下的参考点为,采集时间差为,空间变换为,则:

$$\hat{T} = \arg \min_T \sum_{i=1}^m w_i \exp(-\lambda_t |d_{ti}|) \rho\left(\frac{1}{T} \|p_i - q_i\|_2\right) \quad (2)$$

式中, $\hat{T}$ 为最优变换参数; m 为控制点数量; w_i 为来源可信权重; ρ 为Huber损失函数; λ_t 为时间衰减系数; $\|\cdot\|_2$ 为空间残差的二范数。公式(2)用于开展河湖岸线、工程边界及三维节点投影校正。异常控制点对整体参数的影响受到稳健损失约束,采集时间差异较大的历史点位还会赋予较低权重,避免旧控制点主导现状成果。

坐标校正完成后,系统再次执行对象质量校验。线对

象的局部偏差选用分段参数处理,防止全局变换改变已完成对齐的河段。面对象修复前需判断权属属性及来源等级,较小缝隙可依据容差进行吸附,边界重叠则进入复核区。自动修复需保留原始几何、修复后几何、规则编号和处理时刻。涉及权属范围、工程安全范围或行政边界的数据不作自动覆盖,核验结论确认后再写入标准区。

2.5 应用反馈驱动标准动态迭代

专题图投入运行后,巡河定位偏差、站点归属错误、工程范围更新及字段解释差异均会形成反馈样本。反馈记录包括对象编码、专题名称、问题位置、核验结论、责任单位和修订时间。系统把有效反馈纳入核验样本,重新估计匹配权重、距离尺度、校验阈值及吸附容差。参数更新目标为:

$$\theta^{(k+1)} = \operatorname{argmin}_{\theta} \left[L_m(\theta; F_k) + \lambda_q L_q(\theta; E_k) + \lambda_s \|\theta - \theta^{(k)}\|_2^2 \right] \quad (3)$$

式中, $\theta^{(k)}$ 为第 k 轮规则参数; F_k 为已核验对象样本; E_k 为质量问题样本; L_m 为匹配损失; L_q 为质量损失; λ_q 控制质量反馈的影响; λ_s 限制相邻版本参数突变。公式(3)用于周期标定。连续出现的对象误配会改变相关证据权重,样本量不足时,参数变化受到稳定项限制。新参数需在独立核验集上达到发布条件,随后生成新版本并反馈至目录归集环节。

该闭环把数据接入与业务使用连接起来。专题图发现的问题能够用于修正规则,规则变化又会作用于后续增量

数据。历史版本、原始记录及修复证据均予以保留,业务人员可以说明对象归并和边界修订依据,也可在规则运行异常时恢复前一版本。治理结果不再局限于一次性入库,标准库能够随工程改造、河湖变化及业务核验持续更新。

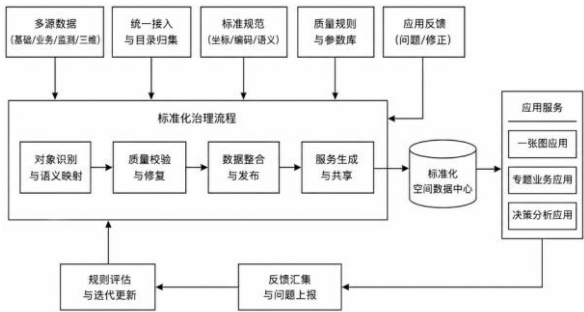


图 1 空间数据标准化治理闭环

3 试验验证

3.1 试验对象参数设置

项目资料未披露可直接复算的生产样本,试验选取主城区密集河网及丘陵区水系构建脱敏工程样本。试验数据涵盖河湖岸线、水利工程、监测站点、涉河项目、节点工程三维模型及气象栅格。人工标注 1200 组对象对应关系,并设置 420 处几何核验点。试验平台配置 16 核处理器、64GB 内存和固态存储,软件环境选用 PostgreSQL15、PostGIS3.4、GDAL3.8 和 Python3.11。匹配确认阈值设为 0.86,复核阈值设为 0.72,吸附容差设为 0.5~6.0m,参数更新步长设为 0.08。

表 1 试验数据参数

数据类型	样本规模	比例尺分母	空间精度/m	更新间/min	坐标参考	数据格式
河湖岸线	4862	2000	1.00	43200	CGCS2000 高斯投影	Shapefile
水利工程	1276	500	0.25	43200	CGCS2000 地理坐标	FileGDB
监测站点	842	5000	2.00	5	WGS84	JSON
涉河项目	638	1000	0.50	1440	WebMercator	PostGIS
节点工程三维模型	164	500	0.10	43200	独立工程坐标	3DTiles
气象栅格	144	1000000	1000	10	WGS84	GeoTIFF

试验设置直接叠加、固定规则治理和标准化治理体系三种方法。各方法在相同硬件环境、相同数据样本及相同

核验基准下完成一次性试验,记录空间配准、对象关联、属性规范、拓扑质量、来源追溯及专题服务指标。

表 2 不同方法单次试验结果

评价指标	直接叠加	固定规则治理	分层标准驱动的闭环治理
配准均方根误差/m	18.42	6.37	2.18
对象关联 F1 值	0.684	0.851	0.934
属性一致度	0.742	0.887	0.949
拓扑错误数/处	326	104	28
来源追溯度	0.538	0.816	0.963
查询中位时延/s	4.8	3.1	1.7
专题图编制时间/min	84	53	28
全量处理耗时/s	412	538	596

3.2 多源对象匹配可信度提升

直接叠加仅处理格式及坐标投影,来源记录仍保留原编码,对象关联 F1 值为 0.684。固定规则治理选用统一距离阈值, F1 值提高至 0.851,但密集河网中的相邻支流仍会发生误连。标准化治理体系把名称、属性、距离、时间状态及业务关系纳入统一判断过程, F1 值达到 0.934,来源追溯度达到 0.963。试验形成 1107 组自动确认对象, 93 组进入人工复核区。复核记录保留候选得分及来源依据,低置信记录未进入主对象库。

该结果表明,稳定编码能够保持来源记录与主对象之间的长期关系,多证据判断则可降低单一距离阈值造成的区域偏差。名称相同但所属河流不同的站点会因业务关系不一致而降低得分。工程更名后,只要空间位置、功能类型及管理关系保持稳定,来源记录仍可挂接原主对象。因此,对象匹配结果具备可核验条件,并可为后续几何修复限定准确范围。

3.3 空间数据规范化质量改善

标准化治理体系的配准均方根误差为 2.18m,低于直接叠加及固定规则治理。稳健估计剔除 9 个异常控制点,主城区选用分区变换参数,丘陵区域选用统一参数。质量校验识别 312 个异常对象,第 1 至第 4 轮反馈质量得分依次为 0.781、0.864、0.913 和 0.926,末轮剩余 41 个。表 2 显示,属性一致度达到 0.949,拓扑错误降至 28 处。

误差下降源于来源权重及时间衰减共同参与控制点筛选。局部偏移不会牵引全区域变换,历史点位也不会覆盖较新的边界成果。对象身份确认后,几何规则可依据河段、站点及工程范围分别运行,减少不同类型数据共用容差造成的修复偏差。未自动处理的记录主要涉及权属边界和施工期临时范围,系统保留原始几何并提交人工核验,避免自动吸附改变管理边界。

3.4 专题图业务支撑效能增强

统一主对象编码建立后,河湖管控专题图可在同一对象下汇集岸线、巡查记录、监测站点、取排水口及涉河项目,水旱灾害防御专题图也可依据河段关联雨情、水情和工情。查询中位时延由 4.8s 降至 1.7s,专题图编制时间由 84min 降至 28min。标准区已完成对象归并、字段转换、

单位处理及空间索引,业务应用无需在每次查询时重复识别对象。

标准化治理体系的全量处理耗时为 596s,高于两种对比方法,新增时间主要用于来源记录生成、候选证据计算及反馈参数更新。专题库稳定运行后,新增数据可依据主对象编码完成局部更新,已确认对象无需再次进入全量处理。新增计算成本主要集中于发生变化的对象,不会随历史数据持续累加。该运行方式适宜水利一张图周期更新,也能为监测数据连续接入保留可核验依据,使河湖管控、防汛研判及工程调度共享统一的对象基础。

4 结束语

面向智慧水利多源空间数据接入场景,研究构建由标准配置、目录归集、语义映射、质量校验及反馈更新组成的治理体系。试验结果表明,该体系能够降低空间残差、对象误配及拓扑错误,提高专题查询和编制效率。研究成果可为河湖管控、水旱灾害防御、水资源监管及工程运行提供统一的对象基础,减少业务系统重复整理成本,并为后续应用接入保留统一扩展条件。后续将研究实时感知数据增量校验及三维对象跨尺度更新方法。

[参考文献]

- [1]储节旺,张子芳.多源异构空间数据融合的情报挖掘和知识发现研究[J].情报理论与实践,2025,48(11):112-121+179.
- [2]刘佳兴,刘德双,燕青青,等.智慧城市建设中空间数据驱动的规划方法研究[J].住宅与房地产,2025,(15):44-46.
- [3]代斌元.面向智慧水利的水资源监管平台建设研究[J].水上安全,2024,(24):64-66.
- [4]于博文.智慧水利工程建设与管理实施新路径研究[J].中国水运,2024,24(14):105-107.
- [5]曹思林.测绘地理信息中的高精度空间数据采集与处理方法研究[J].城市建设理论研究(电子版),2024,(17):175-177.

作者简介:董剑锋(1972—),男,江苏常州人,学历:本科,生产运行管理类中级工程师,主要研究方向:水资源集约节约利用、节约用水规划设计与咨询、非常规水源配置及利用、信息化管理。