

· 信息技术 ·



大数据技术在航道数字孪生系统中的应用*

张圣丽, 沈 思, 徐 硕
(长江航道测量中心, 湖北 武汉 430010)

摘要: 深入探讨大数据技术在航道数字孪生系统中的应用, 重点研究航道大数据治理技术与多维挖掘算法, 通过构建航道数据资源仓库、设计高效索引技术, 并结合多维挖掘算法, 提出一套创新的航道大数据治理方案, 有效地解决了数据分散、检索效率低等问题, 提升航道数据资源的可检索性、可获取性和可利用性。同时, 通过在长江中游航道示范段构建数字孪生场景、开展船舶交通流分析以及桥区通航指引展示等应用实践, 验证了该技术的实际效果与应用潜力, 并进一步提出大数据与数字孪生技术在未来现代化航道中的发展前景, 为航道智慧化转型提供系统性方案。

关键词: 大数据; 数字孪生; 航道; 数据治理

中图分类号: U61

文献标志码: A

文章编号: 1002-4972(2025)11-0183-07

Application of big data technology in digital twin system of waterways

ZHANG Shengli, SHEN Si, XU Shuo
(Changjiang Waterway Bureau Survey Center, Wuhan 430010, China)

Abstract: This paper delves into the application of big data technology in the digital twin system of waterways, focusing on the governance technology of waterway big data and multi-dimensional mining algorithms. By constructing a waterway data resource warehouse, designing efficient indexing technology, and combining multi-dimensional mining algorithms, an innovative waterway big data governance solution is proposed. This effectively addresses issues such as data dispersion and low retrieval efficiency, and enhancing the retrievability, accessibility, and usability of waterway data resources. Meanwhile, through the construction of digital twin scenarios in the demonstration section of the middle reaches of the Yangtze River waterway, the application practices of ship traffic flow analysis and navigation guidance display in bridge areas, the actual effects and application potential of this technology are verified. Furthermore, the future development prospects of big data and digital twin technology in modern waterways are proposed, providing a systematic solution for the intelligent transformation of waterways.

Keywords: big data; digital twins; waterway; data governance

2019年9月, 中共中央、国务院印发《交通强国建设纲要》, 明确提出“推动大数据、互联网、人工智能、区块链、超级计算等新技术与交通行业深度融合。”的目标任务。交通运输部积极响应, 大力推动以大数据技术为代表的智慧运输服务体系建设, 旨在为交通强国建设提供强劲动力。随着长江航运的快速发展和水资源综合利用的持续深化, 提升航道通航能力和服务水平已成为长江

航运高质量发展的当务之急, 在此背景下, 充分发挥大数据基础支撑和核心驱动作用, 已成为提升航道管理水平和服务能力的有效途径。

近年来, 数字航道发展迅猛, 积累了海量地理信息基础数据以及航道、航标、水深、船舶等动态监测数据, 为航道管理与发展构筑了坚实的数据基础^[1], 然而, 深入剖析发现, 在数据底座构建环节, 仍存在若干关键难题亟需攻克: 1) 数

收稿日期: 2025-02-09

*基金项目: 国家重点研发计划项目“内河航道综合信息服务系统研发及示范”(2018YFB1600405)

作者简介: 张圣丽(1983—), 女, 高级工程师, 从事智慧航道、电子航道图等航道信息化建设管理和科技创新方面的工作。

据资源分散，高效索引技术缺失，导致数据检索与调用效率低，难以支撑航道数据关联监测^[2]；2) 在业务需求方面，航道大数据虽应用场景广泛、潜力巨大，但对其挖掘的广度与深度欠佳，面对不同用户群体持续增长的定制化、个性化诉求，难以构建契合需求的二三维联动服务场景，制约了大数据价值释放^[3]。鉴于此，以航道数字孪生平台为核心，本文聚焦于将航道大数据治理、多维挖掘算法等技术研究成果融入其中，推动数字航道迈向更高层级。

1 航道大数据治理技术

航道大数据治理技术构成了数字孪生系统的数据基础和核心，为提高航道数据资源的可检索性、可获取性和可利用性，从航道信息数据仓库设计与多元航道信息索引技术两方面进行阐述，提出航道大数据高效访问的数据治理创新方案。

1.1 航道数据资源仓库设计

基于航道信息化智能服务的不同信息要求，在对航道大数据相关资源进行全面梳理的基础上，根据航道数字孪生数据要素的应用方向，将航道大数据资源划分为基础业务数据、船舶 AIS 数据、运行结果数据等不同类型。其中，基础业务数据又可以按对象分为航道、航标、水位、流量数据等，按数据存储格式分为电子航道图、结构化数据库和空间数据库等，按作用分为业务数据和数据字典等数据资源；船舶 AIS(自动识别系统，automatic identification system)数据主要分为船

舶静态数据和动态数据，其中船舶静态数据包括船名、编号、吃水、货物种类等，动态数据包括时间、船速、位置、航向等数据^[4]；运行结果数据是系统运行计算过程中所产生的结果数据，按照功能模块划分，又可以分为交通流量分析、重点水域监管、消息定制服务等功能所产生的结果数据。在此基础上，结合各类航道信息应用场景，设计了面向显示、搜索、空间分析和数据服务的航道信息数据仓库，包括基础数据库和专题数据库。数据资源目录包括航道信息、航道维护信息、航道气象信息等多个大类，每个大类都有相应的分类代码和数据字典，确保了数据的标准化而且易于检索。航道数据资源目录及仓库关系配置见图 1、2。

为使长江航道数据资源有效集成并充分发挥其价值，在对仓库数据进行分类基础上，实施了关键 ETL(抽取、转换、加载，extract、transform、load)流程：在数据清洗阶段，通过全面审核基础业务数据、船舶 AIS 数据及运行结果数据，剔除重复、错误及无效信息，并运用算法填补缺失值，确保了数据的完整性和一致性；数据转换阶段，依据数据仓库设计标准，将多源异构数据转化为统一格式，包括电子航道图的矢量化转换、结构化数据的标准化建模及时间戳的统一处理，同时实施了编码标准化以提升处理效率；在数据加载阶段，采用增量加载策略及严格的数据校验机制，将清洗转换后的数据高效准确地加载至基础及专题数据库中，为长江航道的日常运营及特定应用场景提供了坚实的数据支撑^[5]。

| 分类代码 | 大类名称 | |
|----------|-----------|----------------------------|
| HD010101 | 航道信息 | ◆便于人们了解航道概况 |
| HD010102 | 航道维护信息 | ◆便于管理部门和航道维护人员检索信息、业务决策 |
| HD010103 | 航道整治建筑物信息 | ◆便于航道系统治理及工程人员检索信息、工程决策 |
| HD010104 | 航道助航设施信息 | ◆便于港航企业、船舶船东检索信息、船舶装载及航行决策 |
| HD010105 | 航道通航建筑物信息 | ◆跨临河信息的整合 |
| HD010106 | 航道水文地形信息 | ◆与河流紧密相关信息的整合 |
| HD010107 | 航道气象信息 | ◆与气象紧密相关信息的整合 |

图 1 航道数据资源目录

Fig. 1 Catalogue of waterway data resource

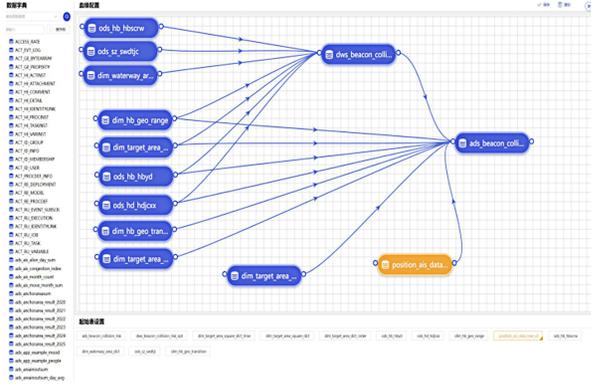


图 2 航道数据资源仓库关系配置

Fig. 2 Relationship configuration of waterway data resource repository

1.2 航道大数据高效索引技术

航道业务数据既有结构化数据, 也有如航道通告类的半结构化数据以及管理文档、监控视频类的非结构化数据。在共享服务这些数据时, 实现半结构化与非结构化数据的快速处理、索引创建

和检索至关重要。

传统检索算法普遍基于顺序存储结构, 对航道大数据集合建立索引时, 顺序存储结构因受到数据文件预置存储规模的限制, 使得存储介质难以支撑单一且庞大的索引文件。为了满足航道大数据快速整合的需求, 在对航道大数据分类分层的基础上, 创新性构建了面向服务的一体化多元航道信息索引结构——分块式链表索引。该索引结构将文档集按块整合, 文档集的索引信息存储在对应的块单元(BlockUnit_File)结构中, 从主索引(Master_Index)中的每个词项(Term)引出一个块索引(Block_Index)链表, 链表中包含该词项在块单元结构中对应的索引信息, 如块号(BlockId)和文档编号(DocId)等信息; 然后主索引、文档索引采用预留位置的定长顺序存储结构, 每个块单元中的文档数量可定义, 具体索引结构见图 3。

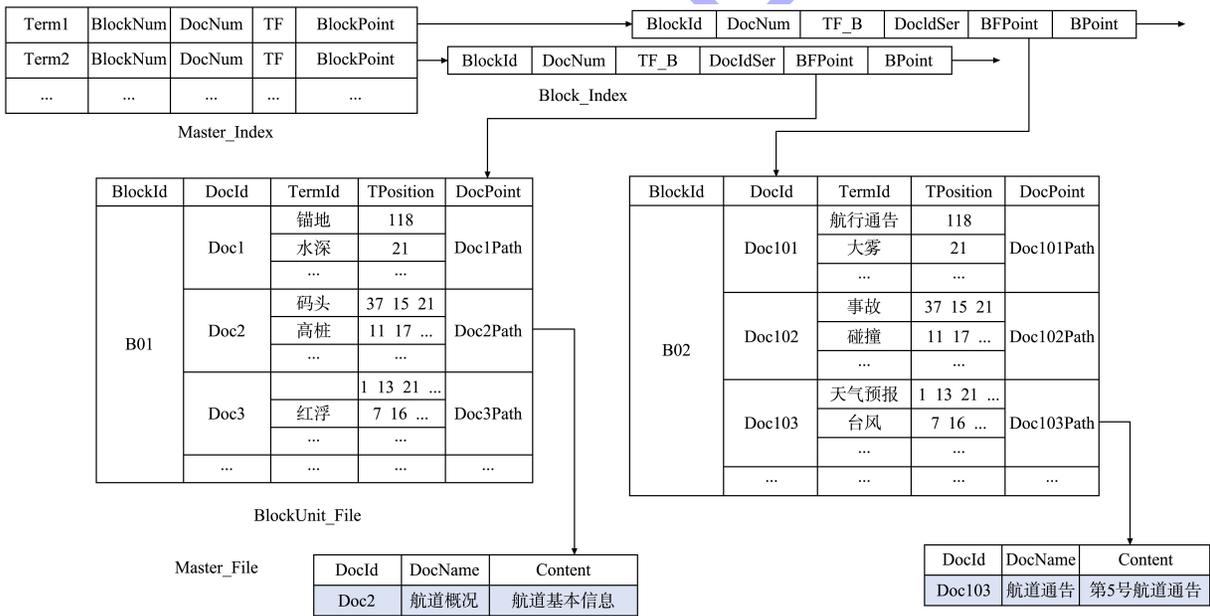


图 3 航道大数据分块式链表索引结构

Fig. 3 Chunked linked list index structure for waterway big data

当系统通过检索语句对索引文件进行检索时, 会对检索语句进行分词处理, 得到多个检索关键词, 并在主索引中进行匹配, 查找涉及相同块单元和文档的检索关键词; 再由检索关键词在主索引

中进行匹配, 查找检索关键字涉及的相同的块单元以及文档, 进而检索读取文档的主要内容, 实现目标数据段的快速索引与访问, 检索过程见图 4。

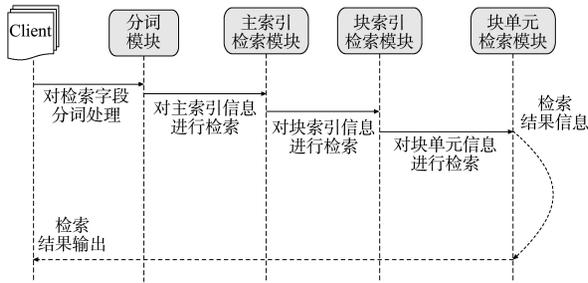


图 4 航道大数据索引检索过程

Fig. 4 Index retrieval process of waterway big data

在航道数据仓库中检索与“某航段 2023 年 9 月船舶 AIS 轨迹、航道维护文档及监控视频”相关的多源异构数据为例，采取分块式链表索引检索流程，分词后，“某航段”定位主索引词汇，直接关联对应的块索引链表，同步并行加载 3 个块单元的文档索引，扫描相关航段 ID 和日期范围，快速定位目标 AIS 轨迹、文档及视频，结合元数据标记，如视频的“航段 ID+时间戳”，通过主索引直接定位文件物理地址，单元结构可减少 60% 以上的冗余数据扫描，块索引加载与关键词过滤效率提升 2 倍以上。

2 大数据多维挖掘算法研究

基于大数据分析和挖掘技术，针对航道养护运行和对外服务需求，建立了航道断面交通流量分析、船舶航行时间预计分析、航线规划导航分析等多种算法模型，为航道业务运行与服务提供了强有力的决策支持。

2.1 航道断面交通流量分析

为准确定量分析周期性航道断面交通流量，基于 AIS 数据构建船舶通过断面识别算法，分析长江上固定断面的船舶流量。该算法具体实现思路为：1) 划定断面；2) 判断船舶是否通过断面；3) 统计每天每个断面的船舶通过数量作为当日断面流量。

该算法核心是判断船舶是否通过断面，即判断属于同一艘船舶的前后相邻的 2 个 AIS 轨迹点的连线是否和断面线相交，如果相交，则认为船舶通过断面。可基于数学上的矢量叉乘法则来判断两线是否相交，也可基于该法则判断船舶通过

断面时的行驶方向。

假设断面左端点是 A，断面右端点是 B，第 1 个轨迹点是 C，第 2 个轨迹点是 D。

1) 当满足如下条件时，两线相交，船舶通过断面： $\vec{AC} \times \vec{AB}$ 与 $\vec{AD} \times \vec{AB}$ 异号；

2) 当满足如下条件时，船舶通过断面并且上行： $\vec{AC} \times \vec{AB} < 0$ 与 $\vec{AD} \times \vec{AB} > 0$ ；

3) 当满足如下条件时，船舶通过断面并且下行： $\vec{AC} \times \vec{AB} > 0$ 与 $\vec{AD} \times \vec{AB} < 0$ 。

以 2021 年 9 月 1 日为例，长江断面船舶流量分析结果见图 5。

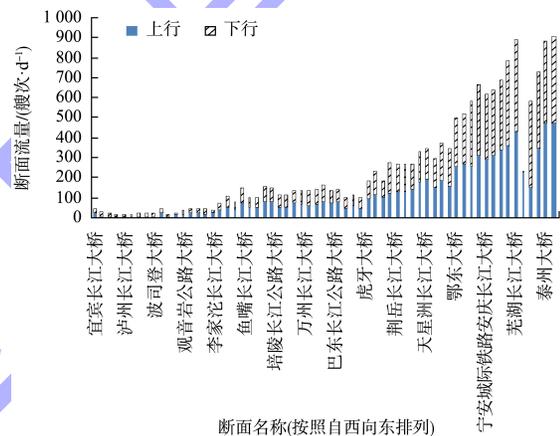


图 5 航道断面日流量

Fig. 5 Daily cross-sectional flow rate of waterway

2.2 船舶航行时间预计分析

为挖掘江上两点之间的航行时间规律，向船方提供承运船舶每航次预计航行时间，给其营运提供参考，同时也为丰富电子航道图的对外服务内容，提升其对外服务水平，研发了基于 AIS 数据的航行时间计算算法，该算法具体实现思路为：1) 对指定航道按照 1 km 进行航段划分；2) 记录船舶到达 1 km 航段起点的时间；3) 记录船舶到达 1 km 航段终点的时间；4) 计算船舶通行 1 km 航段的通行时长；5) 每天对每 1 km 航段统计平均航行时长，并作为该航段的航行预计时间参考值。

航道上两点之间的航行时间等于这两点所跨越的所有 1 km 航段最新的平均航行时长之和，计算方法见图 6。

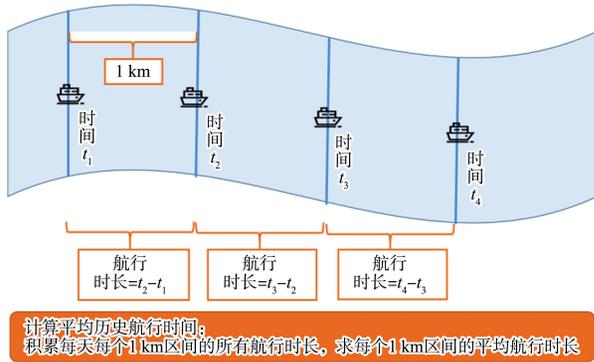


图 6 航段每日平均航行时间计算方法

Fig. 6 Calculation algorithm for daily average voyage time of navigation section

2.3 航线规划导航分析

为辅助用户进行航线规划并提供路径推送服务, 本文提出航线聚类算法框架, 该算法主要分为分割 (partitioning) 和分组 (grouping) 2 个阶段。在分割阶段, 使用最小描述长度原则对航线进行正式的分割, 将航线分割成一系列的线段; 在分组阶段, 则采用基于密度的线段聚类算法, 将相似的线段聚集成簇。此算法的优势在于能够从航线数据库中发现共同的子航线, 而不仅仅是整体上相似的航线。自适应船舶航线规划模型见图 7, 横驶区航线规划模型见图 8。

定义航线之间的距离函数, 包括垂直距离、平行距离和角度距离。该算法具体实现思路为: 1) 轨迹点生成图像; 2) 图像识别出轮廓; 3) 轮廓求航迹中心线; 4) 引进航标按已有的航迹中心线排序后生成蒙版; 5) 蒙版与轨迹点叠加取交集; 6) 重新计算并提取航迹中心线。

对获航线规划中的预计到达时间涉及多个变量和计算步骤: 首先确定出发点和目的地之间的航线距离, 一般通过地理坐标计算两点之间的大圆距离 (即地球表面上两点之间的最短路径)。其次计算船只在整个航线上的平均速度。叠加在速度限制、航行规则、整治施工等某些情况下, 对计算出的结果进行适当调整, 最后具体的计算公式为: 预计到达时间=航线距离/平均速度+其他时间因素。



图 7 自适应船舶航线规划模型

Fig. 7 Route planning model of adaptive ship

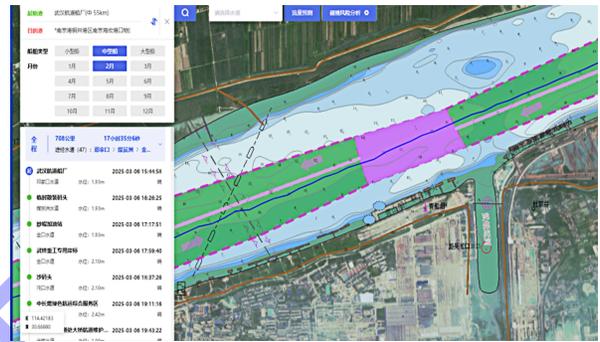


图 8 横驶区航线规划模型

Fig. 8 Route planning model of crossing area

3 航道数字孪生系统中的应用实践

数字孪生平台是面向航道管理需求、支撑航道智慧化服务的重要手段。基于大数据技术, 可为数字孪生航道提供丰富的数据底座和分析决策保障。

3.1 重点示范段数字孪生场景构建

在长江中游 7 km 重点示范航道, 通过多源异构数据融合治理, 构建航道数字孪生场景, 见图 9。利用高保真渲染引擎实现船舶、航标、桥梁、码头、航路、整治建筑物、水下地形、碍航物、沿岸地形与建筑物等要素的三维可视化呈现与交互式查询, 为管理部门提供了直观的航道实时状态信息^[6]。



图 9 两桥段高保真渲染孪生场景

Fig. 9 High-fidelity rendering twin scene of two bridges

3.2 船舶交通流分析与展示

应用航道断面交通流量分析算法，通过船舶AIS数据对船舶断面交通流进行分析，统计1d内上下行通过指定断面的船舶数量，用于航道拥堵预测、控制河段通行信号优化等，为管理部门制定合理的航道通航管理策略提供支持，见图10。

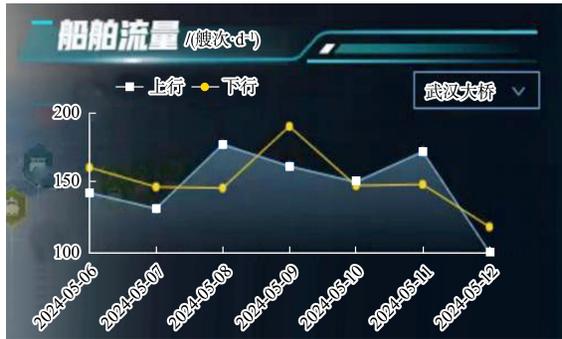


图 10 断面交通流
Fig. 10 Cross-section traffic flow

3.3 桥区通航指引展示

应用航线规划导航及船舶航行时间预计算法，结合现行航路与桥区通航规则，为船舶提供通航桥孔的指示和时间预测服务，引导其安全穿越桥梁，通过重点水域的二三维伴随式导航，提升船舶通行的安全性。桥区通航指引见图11。



图 11 桥区通航指引
Fig. 11 Navigation guidance of bridge area

4 发展展望

利用大数据治理和算法模型技术，目前已成功构建了二三维动静一体化流域实景地图，有效支撑了航道数字孪生平台在航运监管与服务的典型场景应用，为提升航运整体效能提供了有力的技术支撑。展望未来，大数据和数字孪生技术在航道中的融合应用将在以下几个方面进一步深化。

4.1 构建长江数字孪生对象编码体系

统筹规划数据感知与传输网络建设，建立航运全要素数据资源池，涵盖航标、水文、地形、沿岸视频、船舶、港口、通航建筑物、气象等全要素、全生命周期数据。基于地理实体设计唯一标识码，建立孪生对象实体分类与统一编码体系^[7]，从而实现全要素实体分级分类表达与管理，促使多源数据与数字底座融合，实现各类航道数据的标准化管理和高效利用。

4.2 深化数据共享与标准建设的合作

健全内河航道信息交换共享机制，打破数据壁垒，实现内河数据资源的高效整合。这涵盖了明确数据共享的责任主体、共享范围、共享方式及共享流程，确保各部门、各地区能够按照统一的标准和流程进行数据交换与共享，特别是推动航道与港口、物流、船闸等数据的共享合作，构建行业数据联动运行模式。同时，深化大数据标准建设合作机制，与国内外相关标准化组织和机构开展紧密合作，共同制定和完善航运大数据相关标准^[8]，包括数据采集、处理、存储、交换、共享以及应用等方面的标准，推动标准在行业内的广泛应用和落地实施。

4.3 完善航道数字孪生应用体系

进一步优化航道智能模型和专业模型算法，攻关“数据+服务+算法”航道专业模型一体的数据分析及预测模型技术^[9]，运用神经网络、机器学习等AI技术开展三维泥沙动力学模型、航运组织模型、整治建筑物状态分析预测模型^[10]、通航枢纽安全调度模型等研究，从海量航运专业数据中挖掘其复杂规律，拓展大数据技术在航道环境监测、通航风险评估等领域的应用，全面提升航道管理的智能化和精细化水平。

5 结语

1) 数据治理创新：通过引入分块式链表索引结构突破传统检索瓶颈，实现半结构化数据快速处理，解决数据分散与低效检索问题，为数字孪生系统提供标准化、高可用数据底座。

2) 多维算法突破: 断面流量分析、航行时间预测及自适应航线规划算法模型提供了技术体系的工程可行性, 精准挖掘了船舶行为规律, 赋能航道调度与安全导航。

3) 实践验证成效: 长江中游示范段的三维孪生场景与动态服务功能, 验证了大数据技术对提升管理效率、降低通航风险的实际价值。

4) 未来深化方向: 统一数字孪生编码体系、深化数据共享标准、完善智能算法是发展方向, 尤其是充分释放 AI 技术潜能, 可进一步推动航道管理智能化转型, 实现航道管理智能化跃升。

参考文献:

- [1] 胡颖, 付莹, 刘源. 内河航道管理中数字航道技术运用分析[J]. 珠江水运, 2024(8): 48-50.
HU Y, FU Y, LIU Y. Analysis on the application of digital waterway technology in inland waterway management[J]. Pearl River water transport, 2024(8): 48-50.
- [2] 吴静媛, 何卉, 周鹏颖, 等. 内河航道大数据综合利用技术体系研究[J]. 珠江水运, 2021(16): 75-77.
WU J Y, HE H, ZHOU P Y, et al. Research on the technical system for comprehensive utilization of big data in inland waterways [J]. Pearl River water transport, 2021(16): 75-77.
- [3] 严新平. 内河新一代航运系统, 如何构建[J]. 珠江水运, 2023 (4): 22-23.
YAN X P. How to construct a new generation of inland river shipping system [J]. Pearl River water transport, 2023(4): 22-23.
- [4] 夏辉宇, 刘安林, 程佳. 基于 AIS 大数据的南京长江大桥水域规律航迹研究[J]. 水运工程, 2021(1): 189-195.
XIA H Y, LIU A L, CHENG J. Pattern recognition of ship routes in the Nanjing Yangtze River Bridge area based on AIS big data[J]. Port & waterway engineering, 2021(1): 189-195.
- [5] 金琪然, 余杰. 数字孪生和人工智能技术在长江航道大

数据治理中的融合应用研究[J]. 中国水运, 2023(S2): 146-155.

JIN Q R, YU J. Research on the integrated application of digital twin and artificial intelligence technologies in big data governance of Yangtze River waterway [J]. China water transport, 2023, S(2): 146-155.

- [6] 刘静思, 黄巧薇, 张欣. 基于数字孪生航道的长江通航安全保障系统设计与实现[J]. 中国水运, 2024(15): 60-62.

LIU J S, HUANG Q W, ZHANG X. Design and implementation of the navigation safety assurance system for the Yangtze River based on digital twin waterway[J]. China water transport, 2024(15): 60-62.

- [7] 张力, 张航, 刘成堃, 等. 水利数字孪生平台三维模拟仿真技术研究与应用[J]. 人民长江, 2023, 54(8): 9-18.

ZHANG L, ZHANG H, LIU C K, et al. Research and application of key technologies for 3D simulation of water conservancy digital twin platform [J]. Yangtze River, 2023, 54(8): 9-18.

- [8] 洛佳男, 耿雄飞, 文捷, 等. 我国内河智慧航道现状、问题与展望[J]. 水运工程, 2023(2): 123-128, 140.

LUO J N, GENG X F, WEN J, et al. Research status, problems, and prospects of smart channel for inland river in China[J]. Port & waterway engineering, 2023(2): 123-128, 140.

- [9] 翟天放, 张天翼. 浅谈数字孪生水利工程建设框架构建[J]. 吉林水利, 2023(7): 75-78.

ZHAI T F, ZHANG T Y. A brief discussion on the construction framework of digital twin water conservancy project[J]. Jilin water resources, 2023(7): 75-78.

- [10] 张帆, 王平义, 张斌, 等. 长江航道整治建筑物服役状态时空分析与预测 [J]. 水运工程, 2024 (6): 84-90, 121.

ZHANG F, WANG P Y, ZHANG B, et al. Spatio-temporal analysis and prediction of service status of the Yangtze River channel improvement buildings[J]. Port & waterway engineering, 2024(6): 84-90, 121.

(本文编辑 赵娟)