

基于交通波理论的内河航道拥塞度预测方法研究

余庆^{1,4}, 赖伟鑫¹, 曹德胜², 万程鹏^{2,3}, 沈欣怡¹

(1. 集美大学 航海学院, 福建 厦门 361021;

2. 武汉理工大学 智能交通系统研究中心, 湖北 武汉 430070;

3. 武汉理工大学 水路交通控制全国重点实验室, 湖北 武汉 430070;

4. 武汉理工大学 内河航运技术湖北省重点实验室, 湖北 武汉 430063)

摘要:为实现对航道拥塞度的预测,基于交通波理论提出一种考虑最大排队长度的拥塞度预测方法。模型基于船舶自动识别系统(AIS)数据提取交通流特征参数,结合船舶在不同水域的航行行为差异,提出航路特征区域划分方法。在此基础上,选取交通波理论中排队长度作为拥塞度评价指标,提出基于高斯过程回归的最大排队长度预测方法,实现对航道拥塞程度的预测。针对长江流域裕溪河段开展案例研究,结果表明:该航段2020年7月最大排队长度理论值为0.98 km,建立回归模型的 Adjusted R^2 为0.88,预测最大排队长度1.34 km,与理论值误差0.37 km。该模型具有较高的可解释性,能实现对航道拥塞度的预测,本研究可为海事监管服务水平提升提供理论依据。

关键词:DBSCAN 聚类;交通波理论;最大排队长度;高斯过程回归;交通流饱和度

中图分类号:U698

文献标志码:A

DOI:10.3969/j.issn.1000-4653.2025.01.007

Research on inland waterway congestion prediction method based on traffic wave theory

YU Qing^{1,4}, LAI Weixin¹, CAO Desheng², WAN Chengpeng^{2,3}, SHEN Xinyi¹

(1. Navigation College, Jimei University, Xiamen 361021, China; 2. Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan 430070, China; 3. State Key Laboratory of Maritime Technology and Safety, Wuhan University of Technology, Wuhan 430070, China; 4. Hubei Key Laboratory of Inland Shipping Technology, Wuhan University of Technology, Wuhan 430063, China)

Abstract: Existing studies indicate that the longer the queue length of vessels, the greater the channel saturation. To predict channel congestion, this study proposes a congestion prediction method that considers the maximum queue length based on the fundamental principles of traffic wave theory. The model utilizes Automatic Identification System (AIS) data to extract traffic flow characteristic parameters and, considering the differences in navigation behavior among ships in different waters, proposes a method for dividing the channel into characteristic areas. The queue length in traffic wave theory is selected as the evaluation index for congestion, and a method for predicting the maximum queue length based on Gaussian process regression is proposed to achieve the prediction of waterway congestion levels. A case study is conducted in the Yuxi River section of the Yangtze River Basin. The results show that the theoretical value of the maximum queue length in this section in July 2020 is 0.98 km, and the Adjusted R^2 index of the established regression model is 0.88, predicting a maximum queue length of 1.34 km with an error of 0.37 km compared to the theoretical value. The research results demonstrate that the proposed model has a high degree of interpretability and can effectively predict the maximum

收稿日期:2023-12-18

基金项目:国家自然科学基金(52201412);福建省科技计划自然科学基金(2022J05067)

作者简介:余庆(1989—),男,副教授,博士,研究方向为船舶交通流管理与风险评价。E-mail: qing.yu@jmu.edu.cn

通信作者:万程鹏(1990—),男,副研究员,博士,研究方向为水路交通运输系统韧性。E-mail: cpwan@whut.edu.cn

引用格式:余庆,赖伟鑫,曹德胜,等.基于交通波理论的内河航道拥塞度预测方法研究[J].中国航海,2025,48(1):50-59.

YU Q, LAI W X, CAO D S, et al. Research on inland waterway congestion prediction method based on traffic wave theory[J]. Navigation of China, 2025, 48(1): 50-59. (in Chinese)

queue length, thereby enabling the prediction of channel congestion. This study provides a theoretical basis for improving the level of maritime supervision services.

Key words: DBSCAN clustering; traffic wave theory; maximum queue length; Gaussian process regression; traffic flow saturation

我国正处在经济发展逐步优化的关键时期,内河运输凭借其低能耗高效率的优点占据国民经济中的重要地位。截至2022年底,全国共拥有沿海省际万吨以上干散货船2 427艘、7 982.4万载重吨,吨位同比增加6.5%^[1]。在航船舶数量的增长,使航道拥堵现象时有发生,不仅导致船舶运输成本的增高,还影响水路运输作业效率,威胁船舶航行安全。因此,预测船舶排队长度,研究航道通航能力,有助于船舶调度优化、提高航道利用效率,促进我国内河航道运输业的高质量发展。

交通波起源于GreenShield提出的抛物线型流量—密度方程^[2],并被运用于道路交通领域。例如,马庆禄等^[3]基于交通波理论建立车辆拥堵演化模型,同时分析了干涉措施对出口车道匝道的影响。近几年,该理论被广泛运用于水上交通的研究中。例如,郑枫凡等^[4]构建蜿蜒河流船舶交通波模型,基于格林希尔茨速度-密度关系曲线,计算不同时期的船舶集散波波速。曾石营^[5]构建警戒区内慢速船舶交通波模型,计算船舶在警戒区内堵塞排队的长度和延误的时间,分析船队的集结和消散过程。

基于交通波理论,可结合数理统计或者机器学习方法构建航道拥塞度预测模型。在数理统计方面,CETIN等^[6]提出一种新的自适应短期预测方法,并运用CUSUM算法来实时检测过程平均水平的变化。在此基础上,本文进一步采用自回归积分移动平均模型实现对航道拥塞度的预测;在机器学习方向,郝勇等^[7]将基于遗传算法优化的径向基(Radial Basis Function, RBF)神经网络应用于港口船舶交通流量预测领域,验证RBF神经网络在交通流量预测领域的适用性。为进一步提高预测模型的精确度,部分学者综合考虑交通需求、季节和气候等因素,通过分析船舶流量历史数据,构建考虑周期性波动因素的船舶交通流量预测改进模型,并运用贝叶斯估计和预测方法求解模型,提出新的航道拥塞度预测方法^[8]。

尽管国内外学者针对航道拥塞度预测开展了相关研究工作。然而,对于动态环境下的航道拥塞度预测以及航道服务水平评价方法还缺少研究,有必要开展相关研究工作,从而实现对特殊时段内船舶动态调度的优化。

为解决进一步探讨长江水域船舶航行特征,构

建可用于评估船舶通过能力的量化模型,本研究基于船舶自动识别系统(Automatic Identification System, AIS)数据进行船舶交通流特征参数提取,提出基于船舶特征差异的水域分割方法,构建基于船舶交通波理论的船舶排队长度计算模型。最后选取长江段某水域进行实测验证。

1 基于交通波理论的内河航道拥塞度预测模型

预测模型的技术路线见图1,本模型主要包含特征参数提取、交通波模型构建、回归预测和航道服务水平评价等4个部分。

1) 考虑到AIS数据具有密度分布的特点,使用DBSCAN方法对预处理后的AIS数据进行聚类,从而提取交通流特征参数。

2) 针对提取出的特征参数,结合交通波理论,对最大排队长度进行计算。

3) 采用Spearman相关系数、方差膨胀因子(Variance Inflation Factor, VIF)检验交通波理论中的相关参数进行共线性检验,从而构建回归模型,实现对最大排队长度的预测。

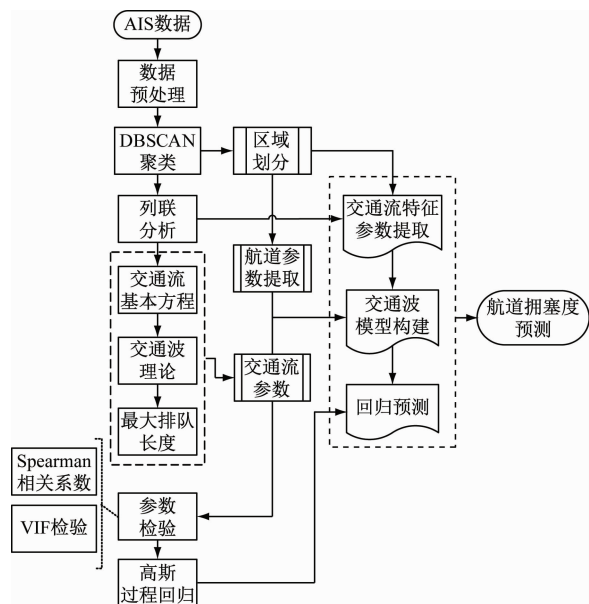


图1 预测模型的技术路线

Fig. 1 Technical route chart for predictive models

4) 结合船舶最大排队长度进行航道拥塞度预测。

1.1 交通流特征参数提取

1.1.1 数据处理

在使用 AIS 数据前,需对其进行转译。转译后的船舶 AIS 数据包含了船舶某时刻的状态特征,需要进一步进行过滤。

1) 根据船舶速度进行筛选,规定 $v=0$ kn 为停泊船舶, $v \neq 0$ kn 为在航船舶,在过滤时仅保留在航船舶的 AIS 数据。

2) 鉴于长江水域航行船舶正常航行速度一般不超过 20 kn,因此对速度大于 20 kn 的 AIS 数据一并进行过滤。

在对速度参数进行过滤后,对位置与时间进行进一步过滤,以便选取研究水域的 AIS 数据。

1.1.2 DBSCAN 聚类

DBSCAN 聚类是一种基于密度的无监督聚类算法^[9],在聚类时不需要事先指定簇的个数,而是通过控制 Eps(半径)与 MinPts(形成簇的最小点)进行聚类,其优点在于可形成任意形状的聚类簇。

1.1.3 提取特征参数

进行区域划分后,使用列联分析方法提取各区域内船舶速度特征,并对航道宽度、航道长度等参数一并进行提取。在提取船舶速度特征时,着重提取、计算各月 AIS 数据的最大速度、平均速度。其中,采用统计方法提取最大速度 v_{\max} ;船舶平均速度 v' 的计算为

$$v' = \frac{\sum_{i=1}^n v_i}{n}, i = 1, 2, \dots, n \quad (1)$$

式(1)中: v_i 为 AIS 数据第 i 艘船舶的速度; n 为数据量。

为减小误差,将船舶各月数据占比作为权重,并结合区域内在航船舶速度特征,确定划分区域的船舶自由速为

$$v_f = \sum_{i=1}^6 v_i \times w_i \quad (2)$$

式(2)中: v_f 为船舶自由速度; v_i 为各月份最大速度; w_i 为各月份的权重。

1.2 交通波模型构建

1.2.1 船舶交通波理论

交通流理论是应用数学或物理学原理,对交通流各参数之间的关系进行定性或定量分析的理论。交通波理论^[10]是基于流体动力学建立起来的一种交通流理论,其描述了同向运动不同状态的两股交通流相遇时的交通状态转移过程,广泛应用于交通流的波动特征和影响分析^[11]。

船舶交通流基本方程为

$$q = \rho v w_1 \quad (3)$$

式(3)中: q 为目标水域流量; ρ 为船舶密度; v 为船舶速度; w 为航道宽度。根据藤井理论^[12],船舶领域的长轴为 $8.0L$,短轴为 $3.2L$,因此,区域船舶密度的计算为

$$\rho_{\max} = \frac{1}{\left[8 \times \left(\frac{L}{1000} \right) \times 3.2 \times \left(\frac{L}{1000} \right) \right]} \quad (4)$$

式(4)中: ρ_{\max} 为区域船舶最大密度; L 为标准船舶长度。

假设航道内 2 个相邻的交通波示意图 2,分别记为区域 A 和区域 B,密度分别为 ρ_1 和 ρ_2 ,交通流内船舶的平均速度分别为 v_1 和 v_2 , S 为波阵面(即密度出现突变的点连成的面),其速度为 v_w 。

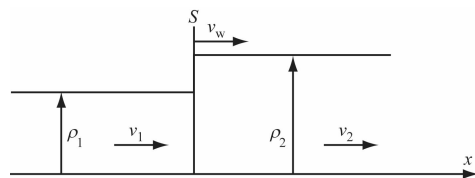


图2 交通波示意

Fig. 2 Diagram of traffic wave

根据质量守恒定律,在 t 时间内,进入波阵面 S 的船舶数量等于离开波阵面 S 的船舶数量,即:

$$\rho_1(v_w - v_1)t = \rho_2(v_w - v_2)t \quad (5)$$

将式(3)代入式(5),有

$$v_w = \frac{q_2 - q_1}{w_1(\rho_2 - \rho_1)} \quad (6)$$

式(6)中: q 为流量。

定义船舶航行方向为正方向,即 x 轴正方向,则有如下情况:

若 $\rho_1 < \rho_2$,即区域 A 的密度小于区域 B。船舶由区域 A 进入区域 B 时,交通流密度增加,船舶减速前进,此时生成的交通波为集结波;若 $\rho_1 > \rho_2$,即区域 A 的密度大于区域 B,船舶由区域 A 进入区域 B 时,交通流密度减小,船舶加速前进,此时生成的交通波为消散波。

假设船舶的速度与密度关系符合格林希尔兹模型^[13]。根据格林希尔兹速度——密度线性关系,有

$$v = v_f \left(1 - \frac{\rho}{\rho_j} \right) \quad (7)$$

式(7)中: ρ_j 为阻塞区域密度。

引入标准交通流 η ,记 $\eta = \frac{\rho_i}{\rho_j}$ 为第 i 股交通流的标准化密度,有

$$\eta_i = 1 - \frac{v_i}{v_f}, \quad i = 1, 2, 3 \quad (8)$$

$$v_i = v_f(1 - \eta_i) \quad (9)$$

将式(8)代入式(9),有

$$v_w = v_f[1 - (\eta_1 + \eta_2)] \quad (10)$$

由式(10)得到集结波 w_{21} 与消散波 w_{32} 的计算为

$$w_{21} = v_f[1 - (\eta_1 + \eta_2)] \quad (11)$$

$$w_{32} = v_f[1 - (\eta_2 + \eta_3)] \quad (12)$$

式(11)和式(12)中: v_1 、 v_2 和 v_3 分别为进入目标水域前、水域内和离开目标水域后的速度。

当集结波与消散波相遇时,船速从 v_3 恢复至 v_1 ,此时,相遇波 w_{31} 的计算为

$$w_{31} = v_f(1 - 2\eta_1) \quad (13)$$

1.2.2 最大排队长度模型

本文定义排队长度为第一艘进入减速状态的船舶至最后一艘处于减速状态的船舶之间的距离。当船舶驶入狭窄航道时,为保障航行安全,会进行减速,船舶处于低速行驶的状态,发生排队现象。

一列船队运行状态示意图3。图3中:船头间距为相邻两列船舶的船头间隔距离,由相邻折线间的垂直距离代表;船头时距为前后2艘船舶前端通过同一地点的时间差,由水平距离代表;斜率则代表船舶速度。假定第一艘船舶的速度变化点为坐标原点,当船舶进入区域II时发生减速现象,速度变为 v_2 ,产生集结波 w_{31} ,方向向前;当船舶由区域II行驶至区域III时,速度变为 v_3 ,产生消散波 w_{32} ,方向向后。

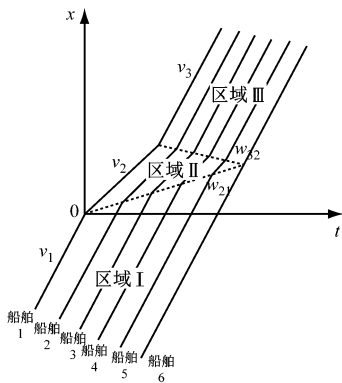


图3 船队运行状态示意

Fig. 3 Diagram of the operating status of the fleet

在此过程中,当两列波传播至船舶6时,由于船舶6位于船队的队尾,受到的影响较小,仍保持速度 v_1 行驶。根据排队的定义,船舶6未被计入排队中。

船舶排队现象见图4。图4中:在0至 t_A 时,船舶在航道内正常航行,未发生排队现象;在 t_A 时刻,

船舶进入狭窄航道,此时通航条件较差,船舶减速前进,发生排队现象,并产生集结波 w_{21} ;在 t_B 时刻,船舶行驶至通航条件较好的区域,并加速前进,产生消散波 w_{32} ;在 t_C 时刻,生成波与消散波相遇,产生相遇波 w_{31} ;随着时间的推移,在 t_D 时刻,船舶疏散完毕,排队现象消失。

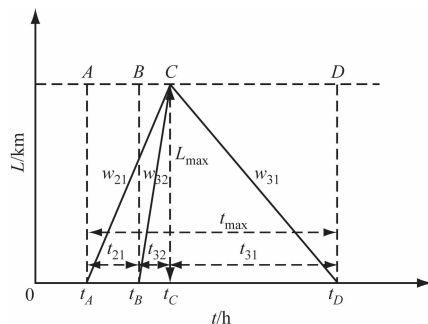


图4 船舶排队现象

Fig. 4 Ship queuing phenomenon

由图4可知:发生排队情况时的最大排队长度为 L_{\max} ,最大排队时间为 t_{\max} 。记集结波、消散波和相遇波的传播时间分别为 t_{21} 、 t_{32} 和 t_{31} ,则有^[14]

$$t_{21} = t_B - t_A = \frac{L_{\max}}{|w_{21}|} \quad (14)$$

$$t_{32} = t_C - t_B = \frac{L_{\max}}{|w_{32}|} \quad (15)$$

$$t_{31} = t_D - t_C = \frac{L_{\max}}{|w_{31}|} \quad (16)$$

$$t_{\max} = t_{21} + t_{32} + t_{31} = t_D - t_A \quad (17)$$

由几何关系,最大排队时间为

$$t_{\max} = \frac{L_{\max}}{|w_{21}|} + \frac{L_{\max}}{|w_{31}|} \quad (18)$$

进一步计算3种波速,可求出最大排队长度为

$$L_{\max} = \frac{|w_{32} \times w_{21}|}{|w_{32}| - |w_{21}|} \times (t_B - t_A) \quad (19)$$

1.3 回归预测

1.3.1 Spearman 相关系数

Spearman 相关系数^[15]可反映2个随机变量的相关程度和强度,其具有如下性质:

1) $|r(X, Y)| \leq 1$,若 $r(X, Y) = 0$,则表示2个变量之间不存在相关性;

2) 若 $X = Y$, $r(X, Y) = 1$,说明2个变量之间具有完全共线性。

1.3.2 共线性检验

在多元回归模型中,多重共线性是指2个或多个自变量之间存在高度相关性,从而导致模型难以解释或预测。VIF^[16]可检验多重共线性,其计算方法为

$$VIF = \frac{1}{1 - R_i^2} \quad (20)$$

式(20)中: R_i^2 为第*i*个自变量与其余自变量进行多元线性回归之后的拟合优度。

1.3.3 高斯过程回归

高斯过程回归适用于处理高维数、小样本和非线性等复杂回归问题。记回归的输入和输出^[17]为 $\{(x_i, y_i) | i = 1, 2, \dots, n\}$, 则他们之间的关系可表示为

$$y_i = f(x_i) + \varepsilon, \varepsilon \sim N(0, \sigma_n^2) \quad (21)$$

式(21)中: x_i 为输入值; $f(x_i)$ 为函数值; ε 为噪声, 其服从均值为0, 方差为 σ_n^2 的标准正态分布。

进一步可得到观测值*y*的先验分布为

$$y \sim N(0, K(X, X) + \sigma_n^2 I_n) \quad (22)$$

观测值*y*与预测值 $f(x^*)$ 的联合先验分布为

$$\begin{bmatrix} y \\ f(x^*) \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(X, X) + \sigma_n^2 I_n & K(x^*, X) \\ K(X, x^*) & K(x^*, x^*) \end{bmatrix}\right) \quad (23)$$

根据式(21)~式(23)计算出预测值 $f(x^*)$ 的后验分布为

$$f(x^*) | X, y, x^* \sim N(f(x^*), \text{Var}[f(x^*)]) \quad (24)$$

式(24)中: $f(x^*)$ 和 $\text{Var}[f(x^*)]$ 为 x^* 对应的均值和方差。

在高斯过程回归中, 使用最广泛的协方差函数为平方指数协方差, 有

$$k(x_i, x_j) = f(x_i)^T f(x_j) = \sigma_f^2 e^{-\frac{1}{2}(x_i - x_j)^T \Lambda (x_i - x_j)} \quad (25)$$

式(25)中: σ_f 和 Λ 为超参数, 采用牛顿法、共轭梯度法可对超参数进行优化。

2 案例分析

为验证本文提出模型的合理性, 选取长江芜湖段部分水域开展模型验证。在该水域内, 一部分船舶沿长江干流行驶, 另一部分船舶从西梁山出口驶出, 然后向着裕溪河的方向前进。最后, 部分船舶会在裕溪河口处驶入, 并继续朝着裕溪闸的方向行驶。研究水域范围见图5。

2.1 数据处理与特征参数提取

2.1.1 船舶 AIS 数据处理

研究水域选取 AIS 数据范围为: 东经 118.29° ~ 118.365°, 北纬 31.409° ~ 31.51°; 选取时间为 2020 年 1—6 月。经统计, 共选取 AIS 数据 141 999 条。研究对所有超过预期 AIS 接收范围(或地理坐标)以及船舶速度等于 0 kn 与大于 20 kn 的异常 AIS 数据进行删除, 共计 79 977 条, 最终保留船舶 AIS 数据 62 022 条。

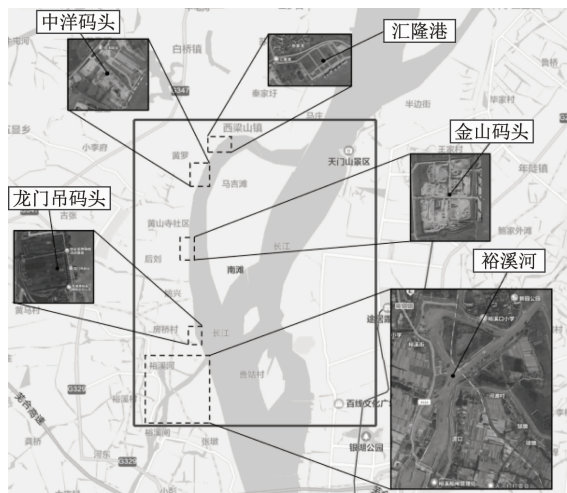


图5 研究水域范围

Fig. 5 Range of research water area

2.1.2 聚类分析

为减少冗余信息, 提高聚类算法的效果, 对主流、支流和岔流交汇处的 AIS 数据进行进一步筛选, 具体范围为: 东经 118.34° ~ 118.35°, 北纬 31.492° ~ 31.497°, 东经 118.31° ~ 118.32°, 北纬 31.428° ~ 31.431°。

对进一步筛选后的 AIS 数据使用 DBSCAN 聚类进行区域划分: 对于主流与支流交汇处, 见图 6a, 控制 Eps 为 0.000 8, MinPts 为 29; 对于支流与岔流交汇处, 见图 6b, 控制 Eps 为 0.000 5, MinPts 为 25; 据以上结果对水域进行划分, 分别记为区域 1、区域 2 和区域 3, 水域划分结果见图 6c。

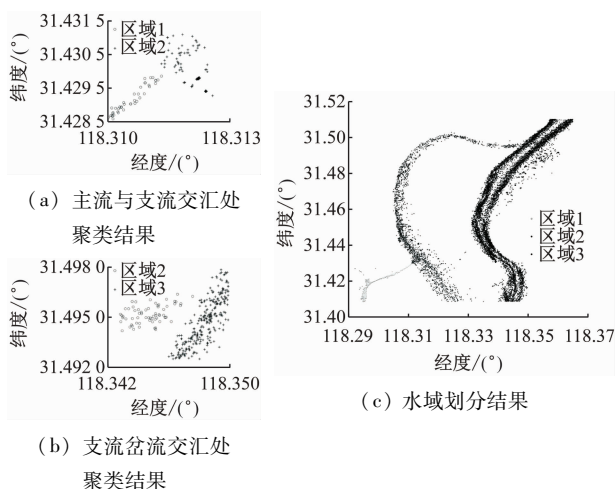


图6 DBSCAN 聚类与划分结果

Fig. 6 DBSCAN clustering and divided results

2.1.3 交通流特征参数提取

分别以月份和区域作为横坐标进行列联分析, 得到的结果见图 7。其中: 不同月份下各区域的数据量与占比情况如图 7a 和 7c 所示; 不同区域中各

月份的占比情况如图7b和图7d所示。

区域船舶速度特征如图7e和图7f所示。其中,图7e表示区域1船舶速度特征,图7f表示区域2船舶速度特征。记区域1船舶平均速度为 v'_1 ,区域2船舶平均速度为 v'_2 ,不难发现 v'_1 与 v'_2 相对接近。然而,区域2的船舶最大速度 v_{2max} 明显高于区域1的船舶最大速度 v_{1max} 。

针对该水域的特点与区域划分结果进一步对各区域内部进行划分和测量,具体结果见图8。其中,航道参数提取结果如图8a~图8d所示;航道进一步划分结果如图8e所示。由图8a~8d可知:裕溪河前段约1.7 km,后段约1.3 km,总长约3 km,河道宽度在100~200 m之间,前段河道宽度为146 m,后段为141 m;该水域中存在多个港口,图8e针对港口的位置进行区域的进一步划分。

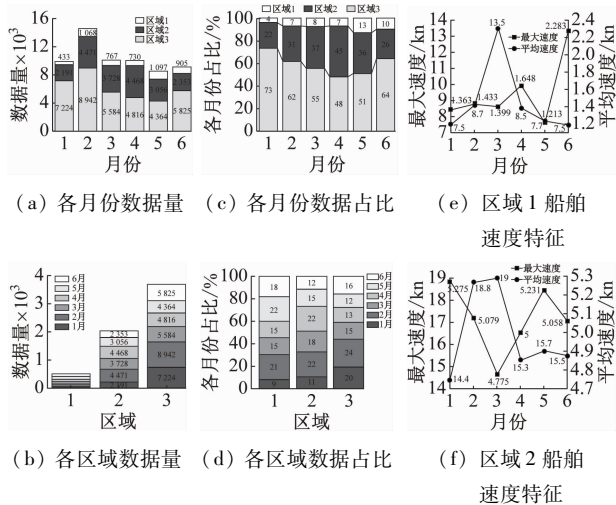


图7 列联分析与船舶速度特征

Fig. 7 Cross-tabulation analysis and ship speed characteristics



图8 航道参数的确定与区域划分

Fig. 8 Determination of river parameters and regional division

2.2 交通波模型构建

在进入裕溪河口前,船舶会经过龙门吊码头,该码头靠近区域2出口,如果将该区域的平均速度为 v_1 ,计算出的 v_1 会偏小,导致最终结果不准确。

当船舶从区域3进入区域2时,无论是前往港口、码头,还是继续向裕溪河前进,非特殊情况下,在抵达目的地之前都不会减速。假设在理想情况下,即水流、风力、水深和航行环境等相对稳定,且船舶密度不变。在这种情况下,船舶驶出区域2出口的速度与刚进入区域2时的速度应该大致相同。统计并计算区域2进入汇隆港前的船舶的平均速度,将其记作 v_1 。

在进入裕溪河前,由于航道狭窄,航行条件恶劣,船舶需减速行驶,发生排队现象,记裕溪河前段船舶平均速度为 v_2 。随着时间的推移,排队现象有所缓解,船舶行驶至裕溪河后段,记裕溪河后段船舶平均速度为 v_3 。

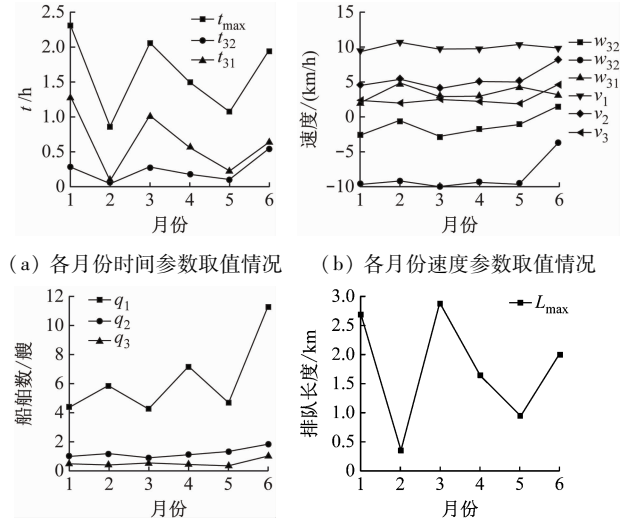
同时,将图6与图7中的统计结果代入式(2),可计算出裕溪河船舶 v_f 为8.85 kn,即16.38 km/h (1 km/h = 5/18 m/s)。

裕溪一线船闸设计通过船型标准为2 000吨级,并且该水域船舶以货船为主。据《内河过闸运输船舶标准船型主尺度系列 第1部分:长江水系》(GB 38030.1—2019),在长江支线船舶过闸的情境下,2 000吨级干散货船对应的船舶尺度为:船长85.0 m,宽度13.8 m。因此,选用该尺度的船舶作为典型船舶。

在实际情况下,船舶的通航时间为随机值。考虑到若采取随机值会导致本研究的场景及计算过程过于复杂。因此,本研究根据类似已建工程资料统计^[18]:当船闸水级2.5 m时,横拉闸门单闸过船时段约50 min,三角闸门单闸过船时段约40 min,假设每艘船舶的过闸时间为45 min,即 $t_b - t_a = 0.75$ 。

由式(11)~式(19)计算出各月份交通流参数

见图9。其中,交通流参数中的时间参数如9a所示;速度参数如图9b所示;流量参数如9c所示;最大排队长度如9d所示。由图9a可知:时间参数之间存在高度相关性。由图9b和图9c可知:各月份的速度参数与流量参数变化较为稳定,没有出现明显起伏。由图9d可知:最大排队长度具有小样本、非线性的特点。



(a) 各月份时间参数取值情况 (b) 各月份速度参数取值情况
(c) 各月份密度参数取值情况 (d) 各月份最大排队长度取值情况
图9 不同月份下各参数取值情况

Fig. 9 The value of each parameter from January to June

2.3 相关性检验与回归

2.3.1 相关性检验

1) Spearman 相关系数。考虑可能存在多重共线性,对变量进行共线性检验。Spearman 相关系数可检查变量之间的相关性强度,通过相关性强度的大小可初步判断是否存在多重共线性问题。对上述变量进行 Spearman 相关性检验,结果见图10。

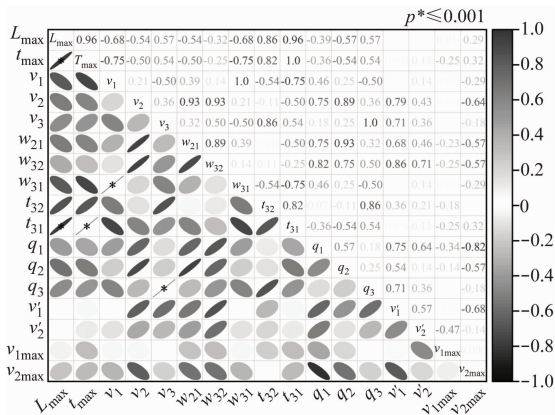


图10 相关系数矩阵

Fig. 10 Correlation matrix

假设检验中的 p 值大小可反映相关系数在统计学意义上是否显著; $p < = 0.001$ 表示有极显著的统计学差异,即得出无效(错误)假设的概率小于等于

0.001,有大于99.999%的可能才得出此结论。

由图10可知:在此次检验中有大于99.999%的可能认为 L_{max} 与 t_{max} 、 v_1 与 w_{31} 、 v_3 与 q_3 等(即图10中标*变量)之间存在完全共线性。其余未标*表示存在相关性,但显著性程度不如 $p \leq 0.001$ 的情况。

2) 共线性检验。图10表明变量之间具有多重共线性,但更具体的情况需要进一步验证。由式(20)使用 VIF 方法进行检验,得到的结果见表1。

表1 VIF 方法检验结果

Tab.1 VIF method test results

变量	VIF
v_1	3.22
t_{32}	6.83
q_2	2.64
v'_2	7.56
v_{1max}	2.39
v_{2max}	3.03

某些变量的 VIF 值过高,因此未在表格中列出,他们的 VIF 值为无穷大。VIF 的大小可反映多重共线性的严重程度:

- ① 当 $0 < VIF < 10$,不存在多重共线性;
- ② 当 $10 \leq VIF < 100$,存在较强的多重共线性;
- ③ 当 $VIF \geq 100$,存在严重多重共线性。

根据表1的结果,表2中变量的 VIF 值均小于10,表明这些变量不存在多重共线性。为保证模型的准确性,使用 VIF 小于5的变量进行回归。

2.3.2 构建高斯过程回归模型

由图9(d)和表1可知:仅保留 v_1 、 q_2 和 v_{1max} 、 v_{2max} 进行高斯过程回归,结果见图11。结果显示,预测值与理论值在数值上的差异较小,基本可起到准确预测的效果。

2.3.3 模型检验与验证

本研究受限于样本量的数量,使用训练数据上的拟合效果来评估模型的性能,结果见表2。表2中:均方误差(Mean Squared Error, MSE)、均方根误差(Root Mean Square Error, RMSE)和平均绝对误差(Mean Absolute Error, MAE)可用来评估模型的效果;MSE、RMSE 和 MAE 的值越低,回归模型的准确性越高; R^2 与 Adjusted R^2 可衡量模型的解释性,他们的取值介于0到1之间,越接近1说明模型的解释性越高。在多元回归中,由于自变量较多,一般采用 Adjusted R^2 对模型的解释性进行评估。

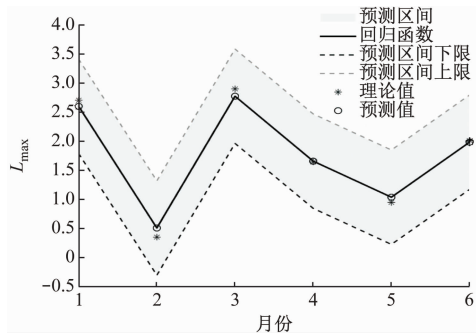


图 11 高斯过程回归结果

Fig. 11 Gaussian process regression results

表 2 模型评估指标

Tab. 2 Model evaluation metrics

指标	数值
MSE	0.000 8
RMSE	0.027 8
MAE	0.129 3
R^2	0.987 7
Adjusted R^2	0.877 7

表 2 中: MSE、RMSE 和 MAE 的值分别为 0.000 8、0.027 8 和 0.129 3,说明该模型的准确性较高; Adjusted R^2 为 0.877 7,其值接近 1,可认为该模型可解释性较高。

为验证预测结果的准确性,本文选取时间范围: 2020 年 1 月 1 日 0 时 0 分 0 秒—2020 年 1 月 10 日 23 时 59 分 59 秒,对 AIS 数据按小时进行划分,得到 240 个时段,并通过统计得到每个时段内最大排队长度的真实值。

最大排队长度实际值与预测值结果见图 12。由图 12 可知:虽然部分时段的实际值与预测值能够较好吻合,但在某些时段仍存在一定差异。

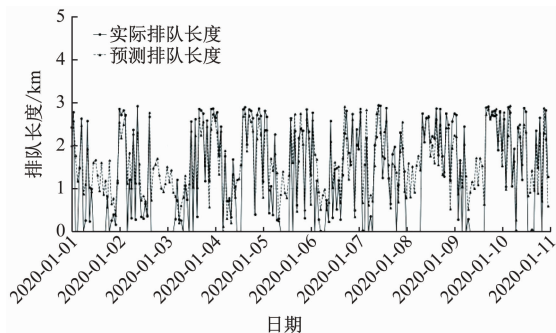


图 12 最大排队长度的实际值与预测值结果

Fig. 12 The true value and the predicted value of the maximum queue length result

计算数据的残差,并绘制残差图,见图 13。由图 13 可知:残差呈现一定的波动性,集中分布在

(-1, 1) km 的范围内。在部分时段内,裕溪河内无船舶驶入,此时最大排队长度为 0 km,但由于预测模型的构建是根据区域内的交通流参数进行的,这就导致最终结果出现真实值为 0 km,而预测值不为 0 km 的情况,导致部分节点的残差较大。

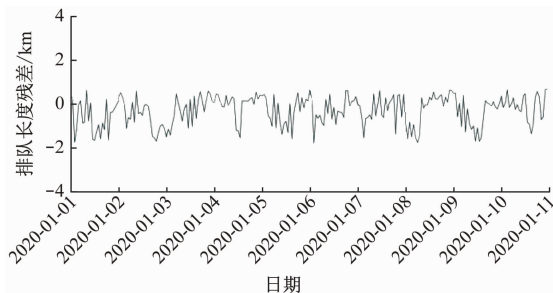


图 13 残差分析结果

Fig. 13 Residual analysis results

对各时段的残差取绝对值后,绘制累计残差图,见图 14。由图 14 可知:10 日的总累计残差约在 145 km,即小时平均总累计残差为 0.60 km。

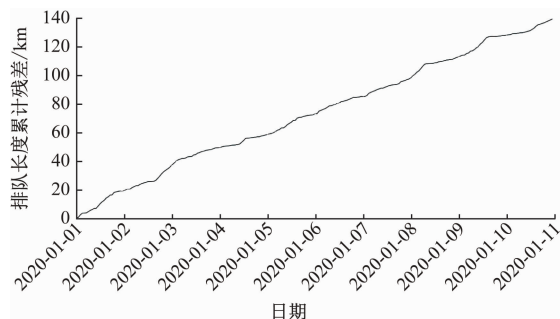


图 14 累计残差结果

Fig. 14 Cumulative residual results

进一步使用箱型图检验结果的合理性,结果见图 15。由图 15 可知:大部分情况下预测值与真实值均位于 25 分位线与 75 分位线内,说明结果是有效的。

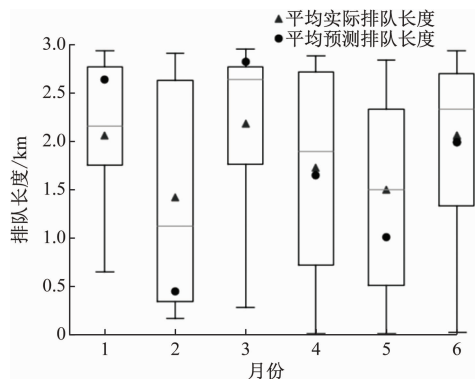


图 15 箱型图检验结果

Fig. 15 Box diagram test results

2.4 对裕溪河航道船舶最大排队长度的预测

利用上述高斯过程回归模型对裕溪河船舶 7 月

最大排队长度进行预测。在进行预测前,对7月船舶AIS数据采用第2.1.1节的方法进行预处理,并根据式(11)~式(19)对各参数进行计算,得到部分变量取值见表3。

表3 7月部分参数取值情况

Tab.3 The value of each parameter in July

变量	数值
v_1'/kn	9.25
v_{1max}/kn	8.90
v_{2max}/kn	14.40
$w_{21}/(km/h)$	1.05
$w_{32}/(km/h)$	-5.40
$w_{31}/(km/h)$	2.12

由式(19)可知:计算出7月的最大排队长度的理论值为0.98 km,回归模型预测7月最大排队长度为1.34 km,其与理论值的误差为0.37 km。

3 结束语

本研究基于交通波理论对船舶最大排队长度进行分析,采用高斯过程回归的方法构建船舶最大排队长度回归模型,并选取长江流域裕溪河段进行验证。本文主要研究结论如下:

1) 从交通波传播的角度,详细分析船队运行状态变化过程中速度的变化情况,并给出船舶排队的定义。

2) 提出一种利用高斯过程回归的船舶最大排队长度预测方法,其 Adjusted R^2 为0.88,具有较高的可解释性,可实现在^[19]小样本、高维度的情况下对船舶最大排队长度的准确预测。

3) 根据已有研究,排队长度与航道通过能力之间存在相关关系,即排队长度越长,航道通过能力越接近饱和,航道服务水平越差。本研究对航道最大排队长度的预测可帮助相关部门更好地进行船舶调度工作,改善航道服务水平,以确保航行安全。

未来可深入研究排队长度与航道通过能力、航道服务水平之间的关系,为海事交通管理提供科学合理的决策依据。

参 考 文 献

[1] 郭佳泰,王琳,李琴. 去年我国水路运输市场整体稳定向好[N]. 中国船舶报,2023-03-24(008).
GUO J T, WANG L, LI Q. Last year, China's waterway transportation market overall stabilized and improved

[N]. China Ship News, 2023-03-24 (008). (in Chinese)

- [2] 管星宇,潘义勇. 基于随机参数线性回归的交通流速度-密度关系模型研究[J]. 森林工程, 2021, 37(5):90-95.
GUAN X Y, PAN Y Y. Research on traffic flow speed-density model based on random parameter linear regression[J]. Forest Engineering, 2021, 37(5):90-95. (in Chinese)
- [3] 马庆禄,牛圣平,曾皓威,等. 网联环境下混合交通流偶发拥堵演化机理研究[J]. 交通运输系统工程与信息, 2022, 22(5): 97-106.
MA Q L, NIU S P, ZENG H W, et al. Mechanism of non-recurring congestion evolution under mixed traffic flow with connected and autonomous vehicles [J]. Journal of Transportation Systems Engineering and Information Technology, 2022, 22(5): 97-106. (in Chinese)
- [4] 郑枫凡,蒋仲廉,余珍,等. 基于船舶AIS数据的蜿蜒河流船舶交通波模型研究[J]. 武汉理工大学学报(交通科学与工程版),2023,47(3):391-395.
ZHENG F F, JIANG Z L, YU Z, et al. Research on the ship traffic wave model for meandering rivers based on AIS data [J]. Journal of Wuhan University of Technology (Transportation Science & Engineering), 2023,47(3):391-395. (in Chinese)
- [5] 曾石营. 基于交通波理论的吴淞口警戒区通航效率分析[J]. 青岛远洋船员职业学院学报, 2019, 40(2):49-53.
ZENG S Y. Analysis of transport efficiency in wusong warning zone based on traffic wave theory[J]. Journal of Qingdao Ocean Shipping Mariners College,2019, 40(2):49-53. (in Chinese)
- [6] CETIN M, COMERT G. Short-term traffic flow prediction with regime switching models [J]. Transportation Research Record, 2006, 1965(1): 23-31.
- [7] 郝勇,王怡. 基于优化RBF网络的港口船舶交通流量预测[J]. 中国航海, 2014(2): 81-84.
HAO Y, WANG Y. Ship traffic volume forecast in port based on optimized RBF neural networks [J]. Navigation of China, 2014(2): 81-84. (in Chinese)
- [8] 张树奎,肖英杰,苏文明. 基于一种改进型线性增长模型的船舶流量预测[J]. 江苏科技大学学报(自然科学版), 2017, 31(4):531-536.
ZHANG S K, XIAO Y J, SU W M. Ship flow prediction based on an improved linear growth model[J]. Journal of Jiangsu University of Science and Technology (Natural Science Edition), 2017, 31(4): 531-536.

- (in Chinese)
- [9] 张朋, 李小林, 王李妍. 基于DBSCAN的动态邻域密度聚类算法[J]. 计算机科学, 2023, 50(增刊1): 599-605.
- ZHANG P, LI X L, WANG L Y. Dynamic neighborhood density clustering algorithm based on DBSCAN [J]. Computer Science, 2023, 50(Suppl.1): 599-605. (in Chinese)
- [10] 余贵珍, 刘玉敏, 金茂菁, 等. 基于交通波的高速公路事故的交通影响分析[J]. 北京航空航天大学学报, 2012, 38(10): 1420-1424.
- YU G Z, LIU Y M, JIN M J, et al. Traffic impact analysis of highway accident based on the shockwave theory[J]. Journal of Beijing University of Aeronautics and Astronautics, 2012, 38(10): 1420-1424. (in Chinese)
- [11] 马江平. 交通事故中的交通波理论分析[J]. 交通世界, 2011(20): 98-100.
- MA J P. Theoretical analysis of traffic waves in traffic accidents[J]. Traffic Safety, 2011(20): 98-100. (in Chinese)
- [12] 马野, 朱景林. 船舶领域及其影响因素研究综述[J]. 青岛远洋船员职业学院学报, 2019, 40(2): 32-38.
- MA Y, ZHU J L. Review of research on ship domain and its influencing factors [J]. Journal of Qingdao Ocean Shipping Mariners College, 2019,40(2):32-38. (in Chinese)
- [13] 白响恩, 肖英杰, 孙玉萍, 等. 基于交通波理论的北槽深水航道通航效率[J]. 中国航海, 2014(1): 61-65.
- BAI X E, XIAO Y J, SUN Y P, et al. Traffic efficiency of north deep-water channel based on traffic wave theory [J]. Navigation of China, 2014(1): 61-65. (in Chinese)
- [14] 马峰, 胡智辉, 潘家财, 等. 海上突发事件下的航道通航效率分析[J]. 集美大学学报(自然科学版), 2018, 23(2): 119-122.
- MA F, HU Z H, PAN J C, et al. Analysis of navigation efficiency under maritime emergencies [J]. Journal of Jimei University (Natural Science Edition), 2018, 23(2): 119-122. (in Chinese)
- [15] 连强. 综合区间数 Spearman 秩相关系数及其应用[J]. 重庆工商大学学报(自然科学版), 2020, 37(6): 71-75.
- LIAN Q. The synthetic spearman rank correlation coefficient of interval numbers and its application [J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2020, 37(6): 71-75. (in Chinese)
- [16] 杨福芹, 李天驰, 冯海宽, 等. 基于无人机数码影像的冬小麦氮素营养诊断研究[J]. 福建农业学报, 2021, 36(3): 369-378.
- YANG F Q, LI T C, FENG H K, et al. UAV digital image-assisted monitoring on nitrogen nutrition of winter wheat in the field [J]. Fujian Journal of Agricultural Sciences, 2021, 36(3): 369-378. (in Chinese)
- [17] 陈刚, 王威, 霍聪. 基于高斯过程回归的船舶动力学模型辨识[J]. 舰船科学技术, 2022, 44(19):1-5.
- CHEN G, WANG W, HUO C. Identification of ship dynamics model based on Gaussian process regression [J]. Ship Science and Technology, 2022, 44(19): 1-5. (in Chinese)
- [18] 陈仁连. 裕溪河复线船闸工作闸门门型选择与设计[J]. 江淮水利科技, 2010(4): 12-13.
- CHEN R L. Working voltage style chooses and design for the yuxihe double-line ship gate [J]. Jianghuai Water Resources Science and Technology, 2010(4): 12-13. (in Chinese)
- [19] 周洪波, 郭国平, 吴斌. 南京长江河桥通航能力基于排队理论[J]. Procedia-Social and Behavioral Sciences, 2013, 96: 2546-2552.