



基于机器学习的排泥管线压降预测

曲嘉铭, 袁超哲, 陶润礼, 孙文博

(中交疏浚技术装备国家工程研究中心有限公司, 上海 201208)

摘要: 福建沿海地区的吹填工程中, 主要的土质颗粒为中粗砂。该类型土质在输送过程中阻力大较易发生堵管, 导致施工进度延缓。在输送环境下施工, 掌握管路内输送阻力的实时信息至关重要。粗颗粒条件下, 常用的经验方法如 Durand 公式法的计算精度较差。基于已有的管路输送研究, 以及福建沿海工程的测试数据, 使用高斯过程回归方法和支持向量机方法建立管路压降预测模型。两种回归模型均能在训练期得到较为理想的效果, 模型 R^2 指标达到 0.80 以上。在模型的预测期, 支持向量机回归模型的 R^2 指标为 0.78, 高斯过程回归预测模型的 R^2 指标达到 0.95。结果表明: 基于高斯过程回归的机器学习模型能够较好地预测中粗砂吹填工程的疏浚参数。

关键词: 机器学习; 排泥管段压降; 支持向量机; 高斯过程回归

中图分类号: U 616

文献标志码: A

文章编号: 1002-4972(2021)01-0196-06

Prediction of pressure drop in mud discharge pipeline based on machine learning method

QU Jia-ming, YUAN Chao-zhe, TAO Run-li, SUN Wen-bo

(CCCC National Engineering Research Center of Dredging Technology and Equipment Co., Ltd., Shanghai 201208, China)

Abstract: In the reclamation project in the coastal areas of Fujian, the main soil particles are medium-coarse sand types. In the process of transportation of this type of soil, large resistance is more likely to cause pipe blockage, which will delay the construction process. Under this working condition, it is very important to grasp the real-time information of the conveying resistance in the pipeline. However, under coarse particle conditions, commonly used empirical methods such as the Durand formula method have poor calculation accuracy. In this paper, based on the existing pipeline transportation research and the test data of the actual coastal engineering in Fujian, the Gaussian process regression method and support vector machine method are used to establish the prediction model of the pressure drop in the pipeline. Both regression models can get more ideal effects during the training period, the R^2 index of the model is above 0.80. In the prediction period of the model, the R^2 index of the support vector machine regression model is 0.78, and the R^2 index of the Gaussian process regression prediction model reaches 0.95. The results show that the machine learning prediction model based on Gaussian process regression can better predict the dredging parameters of the medium-coarse sand dredging project.

Keywords: machine learning; pressure drop; support vector; Gaussian process regression

福建沿海地区的中粗砂吹填工程中, 吹填造地的面积约为 17 km², 施工吹填区域见图 1。对施工工地的砂土进行取样, 中值粒径在 0.3~0.8 mm 变化, 其土颗粒密度为 3 650 kg/m³。施工的船舶为 3 500 m³/h 的绞吸挖泥船, 其泥泵关键指标参

数见表 1。

中粗砂泥浆具有成分复杂、浓度高和颗粒不均的特点^[1]。输送过程中阻力较大、能耗较高, 容易产生堵管和排距受限等问题^[2]。目前, 国内已有学者对排泥管线沿程压力变化情况进行了实

收稿日期: 2020-04-10

作者简介: 曲嘉铭(1994—)女, 硕士, 助理工程师, 从事机器学习应用于疏浚工艺参数优化的研究。

时监测与分析, 但是尚未有学者将机器学习方法应用于排泥管线压降的预测中^[3-7]。



图 1 施工吹填区域(白色框线中的位置)

表 1 泥泵关键指标参数

泥泵	排压/MPa	转速/(r·min ⁻¹)
1#甲板泵	0.977	235
2#甲板泵	1.703	235
水下泵	0.304	235

本文旨在通过建立机器学习预测模型对排泥管道中压降的值进行预测, 从而提前判断管路中泥浆的阻力情况, 及时调整挖泥设备的运行参数, 保证施工持续高效进行。

1 绞吸挖泥船顺直排泥管段阻力影响因素

JTS 181-5—2012《疏浚与吹填工程设计规范》中的 Durand 公式为:

$$\lambda_m = \lambda_w \left[1 + CK_D \cdot \left(\frac{v^2}{gD(\rho_s - 1)} \cdot \frac{\sqrt{gd_s(\rho_s - 1)}}{v_{ss}} \right)^{\frac{3}{2}} \right] \quad (1)$$

式中: λ_m 、 λ_w 分别为输送泥浆及清水时的管路沿程阻力系数; C 为土颗粒体积浓度(%), $C = (\rho_m - \rho_w) / (\rho_s - \rho_w)$, ρ_s 为土颗粒密度(t/m^3), ρ_m 为密度计测量密度(t/m^3), ρ_w 为海水密度(t/m^3); K_D 为试验系数, 取 121; v 为管路泥浆平均流速(m/s); D 为管路内径(m); d_s 为颗粒平均直径(m); v_{ss} 为土颗粒沉降速度(m/s)。

由于 Durand 公式的经验参数由室内试验确定, 参数之间关系复杂。在不同的输送环境下,

需要根据试验进行多次参数校正, 导致最终管路中阻力的计算会产生较多的偏差。

在福建厦门的工程项目中, 利用 Durand 公式计算中值粒径为 0.7 mm 的泥浆输送管内压降, 将计算结果与实测值进行比较, 结果见图 2。显然, Durand 公式的计算结果表现为: 对峰值不敏感, 计算结果在均值附近摆动, 直接导致施工过程中最关注的高浓度和低浓度情况被忽略。

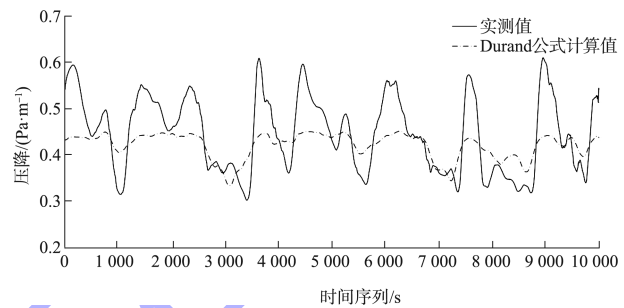


图 2 实测压降与 Durand 公式计算对比

鉴于经验公式法不能准确预测管路中的压降, 本文提出将机器学习预测模型应用于压降的预测中。由 Durand 经验公式可初步判断: 管路中的阻力影响因素较多, 其中流速和浓度作为被动参数, 直接反映了管线内部泥浆的状态。本文建立输入为流速、浓度, 输出为压降的机器学习模型, 根据数据试验设计机器学习网络, 并展开压降预测的研究。

2 机器学习预测模型

机器学习预测模型通过建立映射关系, 选择合适的学习模型, 可以对参数进行预测, 整个过程完全由数据支撑^[7-8]。在不同的管路输送环境下, 只需要利用该管路输送环境产生的数据进行新一轮的学习训练, 则可得到具有较好适应性的预测模型^[9]。本文运用的两种机器学习模型分别为支持向量机回归预测模型和高斯过程回归预测模型, 具体压降预测流程见图 3。

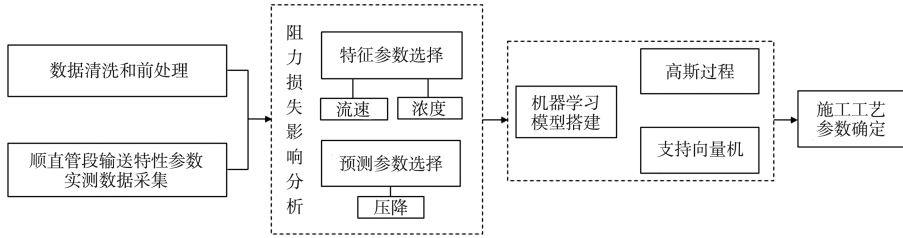


图3 排泥管道参数预测流程

2.1 支持向量机回归预测模型

支持向量机回归是一种有监督学习的方法，其基本思想是：线性可分的情况下寻找两类样本的最优分类超平面；在线性不可分的情况下，加入松弛变量进行分析，通过使用非线性的映射关系将低维样本输入映射到高维空间使变量线性可分，从而实现在高维空间采用线性算法对原本“非线性”的样本进行分析^[10]。

将选为特征变量的施工工艺参数作为模型的输入 $X = \{x_1, x_2, \dots, x_n\}$ ，其对应的需要预测的工艺参数观测值作为模型训练的响应 $Y = \{y_1, y_2, \dots, y_n\}$ ，建立回归模型。在建模前，对工艺参数进行了分析，认为建模样本在原始空间是不可以进行线性分割的样本，故采用非线性回归模型^[11-12]。

对于样本集合 $\{(x_1, y_1), \dots, (x_i, y_i), \dots, (x_n, y_n)\}$ 任意给定的 $\varepsilon > 0$ ，采用非线性映射 ϕ 把样本数据集合映射到高维特征空间获得映射后的样本 $\phi(x_i)$ ，即寻找一个核函数 $K(s, t)$ ，使得 $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ ，在这个高维空间进行线性划分，求解优化问题为求解 $\frac{1}{2} \|w\|^2$ 最小值的问题，求解该问题的约束条件为：假设 $|\langle w, \phi(x_i) \rangle + b - y_i| \leq \varepsilon, i = 1, 2, \dots, l$ 。

构造非线性函数： $f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_j, x) + b$ ， $x_i \in R^n, b \in R$ 作为预测 y 的数学模型，其中， α_i 和 α_i^* 是原目标问题和其对偶问题的拉格朗日算子。

2.2 高斯过程回归预测模型

高斯过程可以看作是多维高斯向无限维的扩展，可以将预测的样本响应 $y = y_1, y_2, \dots, y_n$ 看作是 n 维高斯分布中随机抽取的一个点^[13]。

在建立高斯过程关系时，把选定的疏浚工艺特征值作为模型的训练样本集合，形成模型输入集合 $\{x_1, x_2, \dots, x_n\}$ 与对应的预测响应集合 $\{y_1, y_2, \dots, y_n\}$ ，一般定义响应集合是带噪声的，所以给每个响应建模为某个隐函数 $f(x)$ 加上一个高斯噪声 $N(0, \sigma_n^2)$ ，表达式为(2)：

$$y = f(x) + N(0, \sigma_n^2) \quad (2)$$

$$f(x) \sim GP(0, K) \quad (3)$$

其中： $f(x)$ 为高斯过程的一个先验式， $GP(0, K)$ 为最终训练所得的高斯过程函数。

本文建立高斯过程回归时运用了不同的核函数进行建模，对不同核函数建模预测的结果进行比较分析，最终选择效果较优的“Matern kernel with parameter 5/2”作为高斯过程回归的最终核函数。

3 压降预测与分析

3.1 试验数据

本文基于福建厦门地区中粗砂吹填工程，进行机器学习排泥管线压降预测模型的建立与分析。原始数据来源为施工中某绞吸挖泥船1d监控系统实测数据。其中，输送介质的中值粒径为0.7 mm，实时的浓度和流速分别为船载浓度计和流量计的实时测量值经过时滞校正的结果。绞吸船稳定工作情况下，实时输送浓度见图4，采集样本的整体管段的平均浓度为21.4%；实时输送流速见图5，其平均值为5.1 m/s。平均浓度的计算公式为：

$$y = \frac{M_{\text{pipe}}}{V_{\text{pipe}}} \quad (4)$$

$$C_c = \frac{\rho - \rho_w}{\rho_s - \rho_w} \quad (5)$$

式中： M_{pipe} 为整个测试管段中泥浆总质量； V_{pipe} 为整个测试管段泥浆的总体积； ρ 为管段泥浆的密

度; C_c 为换算后的平均浓度; ρ_s 为原状土密度; ρ_w 为海水密度。

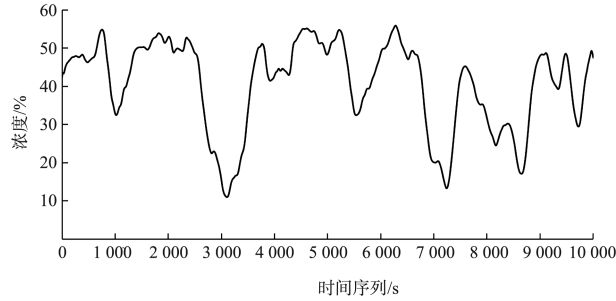


图 4 顺直排泥管段泥浆输送浓度

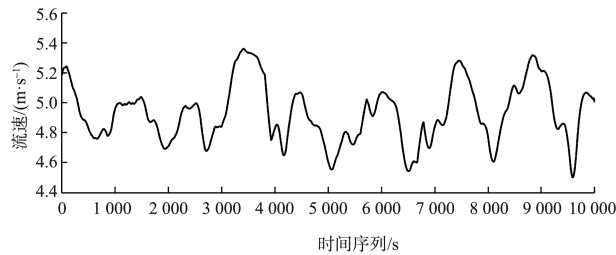


图 5 顺直排泥管段泥浆输送流速

本次研究随机抽选了绞吸挖泥船稳定工作时管段 A 记录的样本数据进行建模, 采用前 6 000 组数据进行模型训练、后 4 000 组数据进行预测, 并对两种机器学习模型的结果进行对比分析。除此之外, 还利用监测管段 B 记录的样本数据进行模型适用性验证, 验证训练结果较好的模型在不同管段中依然具有较强的预测能力。其中, 管段 A 为 214 m 的聚氨酯耐磨管, B 段为 178 m 的普通钢管。

3.2 数据标准化

首先, 对实时测量获得的浓度 C 、流速 v 和对应的压降 ΔP 数据进行标准化处理, 数据的标准化有利于模型初始化的设置, 避免给梯度数值的更新带来问题。同时有利于学习率的调整从而加快模型训练速度、提高计算效率。

实测样本数据标准化方法为:

$$y_i = \frac{x_i - \bar{x}}{\sigma} \quad (6)$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (7)$$

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (8)$$

式中: x_i 为样本数据; y_i 为标准化后的样本数据 (下同); i 为计数标签; n 为样本数量; \bar{x} 为均值; σ 为标准差。

3.3 模型训练结果

最终两种模型的整体预测情况见图 6、7, 误差分布情况见图 8、9。

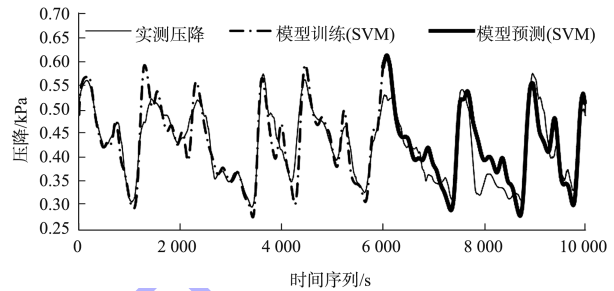


图 6 管段 A 中泥浆压降-支持向量机回归预测结果

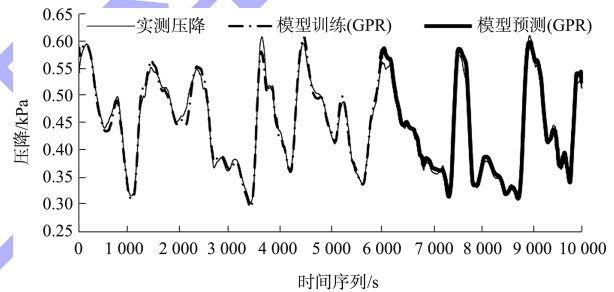


图 7 管段 A 中泥浆压降-高斯过程回归预测结果

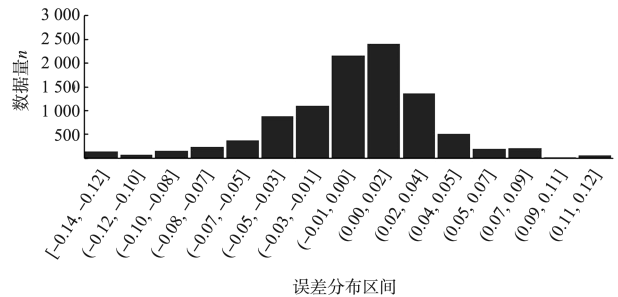


图 8 管段 A 中泥浆压降-支持向量机回归模型整体误差分布

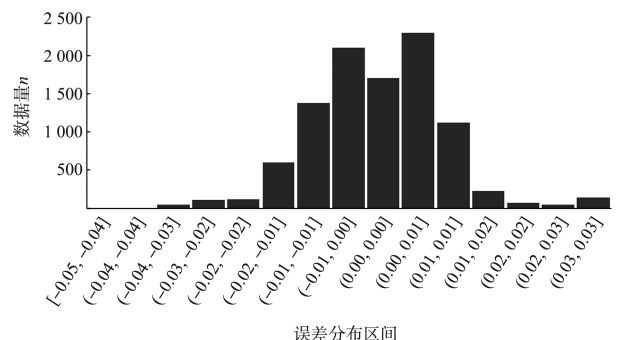


图 9 管段 A 中泥浆压降-高斯过程回归模型整体误差分布

对于机器学习模型的评价一般通过计算 MSE 和 R^2 两个参数进行评价。

MSE 即均方误差,是指模型计算值与真值之间差值的平方的期望。MSE 可以评价数据的变化程度, MSE 的值越小,说明预测模型描述试验数据具有更好的精确度, MSE 的计算公式为:

$$MSE = \frac{1}{m} \sum_{i=1}^m (y_i - y_{\text{pred}})^2 \quad (9)$$

式中: y_{pred} 为模型预测值; m 为计算样本数量(训练期取 6 000, 预测期取 4 000)下同。

R^2 为确定性系数,其通过表征数据的变化来判断模型拟合结果的好坏,它的取值范围为(0~1),当确定性系数越接近于 1 时,则称预测模型对于原始样本的解释能力越强,也就是模型的预测能力越好。确定性系数由两个因素决定,一个是预测数据与原始数据均值之差的平方和 SSR(sum of squares of the regression); 另一个是原始数据与其均值之差的平方和 SST(total sum of squares)。这两个参数的计算公式为:

$$SSR = \sum_{i=1}^m (\hat{y}_i - \bar{y}_i)^2 \quad (10)$$

$$SST = \sum_{i=1}^m (y_i - \bar{y}_i)^2 \quad (11)$$

式中: \bar{y}_i 为 y_i 的均值; \hat{y}_i 为机器学习模型的预测数据。

确定性系数 R^2 的计算公式为:

$$R^2 = 1 - \frac{SSE}{SST} \quad (12)$$

MSE 和 R^2 作为主要评价机器学习模型模拟效果的参数,其结果计算结果见表 2。

表 2 机器学习模型预测成果评价指标

模型	序列	MSE	R^2
SVM-Gaussian	训练期(1~6 000)	0.001 0	0.81
	预测期(6 000~10 000)	0.002 0	0.78
GRP-Matern52	训练期(1~6 000)	0.000 1	0.98
	预测期(6 000~10 000)	0.000 3	0.95

计算评价成果指标显示,支持向量机回归模型对于峰值和谷值的模拟效果较差,导致模型整体的误差较大;且由模型整体误差分布(图 8、9)可见,支持向量机回归模型计算所得的误差分布

范围在-0.14~0.12,误差主要集中于-0.03~0.03 的范围内,说明模型的误差较明显。而高斯过程回归模型的误差分布范围较为集中,仅在-0.05~0.03 的范围内,主要集中于-0.01~0.01 的范围内,说明高斯过程回归模型不仅在模型训练期表现了较优的预测结果,在模型的预测期也表现了较强的数据泛化能力,整体模拟效果较好,其确定性系数比支持向量机回归方法高 0.13,均方误差接近于 0。

3.4 其他管段环境验证

在管路 B 中,直接采用在管路 A 中训练得到的模型进行压降预测,得到的压降预测结果见图 10、11。

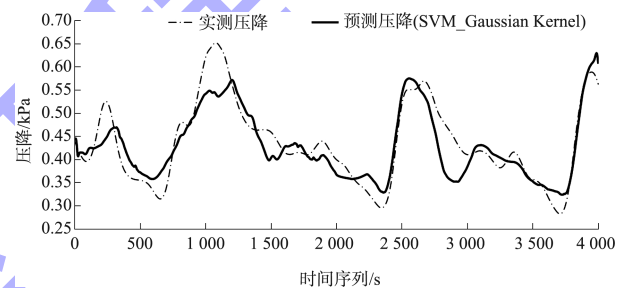


图 10 管段 B 中泥浆压降-支持向量机回归预测结果

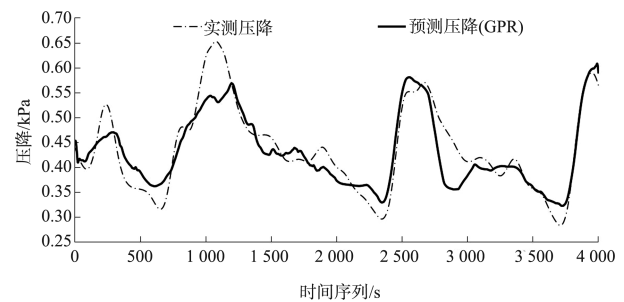


图 11 管段 B 中泥浆压降-高斯过程回归预测结果

管段 B 中泥浆压降预测成果评价指标见表 3。

表 3 机器学习模型在管段 B 的预测成果评价指标

模型	序列	MSE	R^2
SVM-Gaussian	预测期 4 000	0.001 58	0.95
GRP-Matern52	预测期 4 000	0.001 49	0.97

评价指标结果显示,管段 B 中,压降预测模型仍然能够保持其数据泛化性能,得到较出色的压降预测结果。

4 结语

1) 基于支持向量机回归和高斯过程回归的机器学习预测模型均能够得到较为准确的压降模拟模型, 在模型的模拟期, R^2 指标可达到 0.8 及以上; 高斯回归预测模型相较于支持向量机回归模型, 在本文提及的管路输送环境下具有更好的适应性, 能够准确对压降进行预测, 高斯过程回归模型的训练期和模拟期的整体误差控制在 0.01%, 可见机器学习预测模型在疏浚领域具有一定的应用价值。

2) 该方法的探索成果, 可应用于绞吸挖泥船的施工过程。对施工过程进行实时监控, 记录施工参数值, 将参数快速处理后代入参数预测模型, 对未来一段时间内的施工过程参数进行实时预测, 计算预测的阻力变化。当管路中阻力增大至可能堵管的浓度区间, 或降低至产量较低区间的趋势时, 及时调整挖泥船的进泥量, 即绞刀横移速度、绞刀深度、泥泵转速等, 通过改变设备参数, 保持管路系统中泥浆的浓度在输送效率较合理的管路输送环境中施工, 有利于减少施工过程中可能出现的浓度过高的堵管、爆管事故, 或者改善浓度过低、效率过低导致施工成本增加的极端情况, 达到疏浚生产效率最大化的目标。

参考文献:

[1] 陶润礼, 梅志能, 袁超哲, 等. 不同材质吹填管线中粗砂

输送摩阻特性[J]. 中国港湾建设, 2017, 37(12): 11-14.

[2] 牧青, 俞孟蕪. 绞吸挖泥船输泥管路的摩擦阻力损失计算[J]. 船海工程, 2009(4): 17-19, 23.

[3] 田俊峰, 俞孟蕪, 钱献国. 智能疏浚新时代与我国疏浚行业的可持续发展[J]. 中国港湾建设, 2019, 39(01): 1-5.

[4] 闭治跃. 挖泥船泥浆管道输送系统效率优化与控制研究[D]. 杭州: 浙江大学, 2008.

[5] 赵天彪, 袁超哲, 汪金文, 等. 聚氨酯耐磨复合管在中粗砂吹填工况下的应用[J]. 中国港湾建设, 2018, 38(1): 61-65.

[6] 应辉, 李文茂, 张家伟. 管径和管路压力降的计算[J]. 化学工程师, 1998(6): 3-5.

[7] 陈小燕. 机器学习算法在数据挖掘中的应用[J]. 现代电子技术, 2015, 38(20): 11-14.

[8] 陈嘉博. 机器学习算法研究及前景展望[J]. 信息通信, 2017(6): 5-6.

[9] 张燕平. 机器学习理论与算法[M]. 北京: 科学出版社, 2012.

[10] 陶卿, 姚穗, 范劲松, 等. 一种新的机器学习算法: Support Vector Machines[J]. 模式识别与人工智能, 2000, 13(3): 285-290.

[11] 卢虎, 李彦, 肖颖. 支持向量机理论及其应用[J]. 空军工程大学学报(自然科学版), 2003, 4(4): 89-91.

[12] 萧嵘, 王继成, 张福炎. 支持向量机理论综述[J]. 计算机科学, 2000, 27(3): 1-3.

[13] 许雪. 高斯过程回归方法及其预测模型[J]. 科技资讯, 2016, 14(11): 97, 99.

(本文编辑 武亚庆)

参考文献:

[1] 陶润礼, 梅志能, 袁超哲, 等. 不同材质吹填管线中粗砂

(上接第 177 页)

参考文献:

[1] 曹民雄, 蔡国正, 黄海龙, 等. 变动回水区河段的设计最低通航水位确定[J]. 水运工程, 2006(1): 63-67.

[2] 李同西, 马殿光. 龙滩枢纽变动回水区航道设计参数研究[J]. 水运工程, 2008(1): 94-97.

[3] 白金嘉. 水电电站库区回水变动段最低通航水位分析[J]. 水运工程, 2005(6): 57-62.

[4] 徐军辉, 邓伟. 梯级电站变动回水区设计最低通航水位确定方法[J]. 水运工程, 2020(4): 109-114.

[5] 冯宏琳. 西江航道尺度开发潜能研究[D]. 南京: 河海大学, 2006.

[6] 彭钜新. 关于航道整治前后水深改正系数 η 值变化的探讨[J]. 珠江水运, 1994(9): 27-28.

[7] 长江航道规划设计研究院, 中交天津港航勘察设计研究院有限公司. 航道工程设计规范: JTS 181—2016[S]. 北京: 人民交通出版社股份有限公司, 2016.

[8] 四川省交通运输厅交通勘察设计研究院. 嘉陵江川境段航运配套工程二期工程二批次滩险整治工程(亭子口库区)施工图设计[R]. 成都: 四川省交通运输厅交通勘察设计研究院, 2018.

(本文编辑 武亚庆)