



# 基于 PSO-RELM 的绞吸挖泥船产量预测 及其可视化辅助决策

郭博臻, 白一鸣, 赵永生

(大连海事大学船舶电气工程学院, 辽宁 大连 116026)

**摘要:** 为保证绞吸挖泥船的疏浚效率, 泥浆产量预测是一种有效的辅助手段。根据绞吸挖泥船的实际作业数据, 进行数据预处理与主成分分析(PCA), 从而简化了预测模型的复杂程度。然后, 采用粒子群优化的正则化极限学习机(PSO-RELM)建立挖泥船瞬时产量预测模型。预测结果表明: PSO-RELM 相较于常规极限学习机有更好泛化性能, 能够提高挖泥船瞬时产量的预测精度。从而生成可视化图表, 辅助挖泥船操纵人员调整疏浚策略。

**关键词:** 绞吸挖泥船; 主成分分析; 正则化极限学习机; 产量预测; 可视化

中图分类号: TP 273; U 616

文献标志码: A

文章编号: 1002-4972(2021)09-0147-05

## Production prediction and visual decision support of cutter suction dredger based on PSO-RELM

GUO Bo-zhen, BAI Yi-ming, ZHAO Yong-sheng

(School of Marine Electrical Engineering, Dalian Maritime University, Dalian 116026, China)

**Abstract:** To ensure the dredging efficiency of the cutter suction dredger, the slurry production prediction of the dredger is an effective auxiliary method. Based on the data obtained from the actual operation of a cutter suction dredger in a certain place, this paper carries out data preprocessing and principal component analysis(PCA) to simplify the complexity of the prediction model. The particle swarm optimization regularized extreme learning machine(PSO-RELM) is used to establish the instantaneous production prediction model of the dredger. The prediction results show that PSO-RELM has better generalization performance than conventional extreme learning machine, and can improve the production prediction accuracy of the cutter suction dredger. Thus, visual charts is generated to assist the dredger operator to adjust the dredging strategy.

**Keywords:** cutter suction dredger; principal component analysis; PSO-RELM (particle swarm optimization regularized extreme learning machine); production prediction; visualization

绞吸挖泥船因其用途广泛、经济性好的优点, 成为疏浚作业中使用最为广泛的一种挖泥船。为保证绞吸挖泥船实际疏浚的工程效益、确定影响产量的主要因素, 进行产量预测具有重要意义。

在产量的影响因素方面, 刘永杰等<sup>[1]</sup>研究了绞刀转速的变化对产量的影响, 并使用 CFX 软件进行数值模拟, 结果显示较进口泥沙量而言, 绞

刀转速对泥沙吸入量的影响更明显; 唐建中<sup>[2]</sup>分析绞刀系统对产量的影响, 提到了影响泥浆浓度和泥浆流速的因素有切削深度、切削控制量、绞刀转速; 周禹扬等<sup>[3]</sup>研究了运转效率、停工、管线、船机故障、生产停歇等对绞吸挖泥船施工效率的影响, 并提出了提高时间利用率的方法; 刘璟等<sup>[4]</sup>研究了重型绞吸挖泥船挖掘中风化岩石

收稿日期: 2021-01-03

作者简介: 郭博臻(1996—), 男, 硕士研究生, 研究方向为船舶控制系统设计。

性质，并分析了此开挖土质下施工效率的影响因素，主要有绞刀切片厚度、绞刀转速、绞刀功率等。上述对产量影响因素分析都是基于某个部件或运行状态的，具有一定的局限性。本文在采集特定工况下大量施工数据的基础上，运用数据挖掘算法分析这些因素对产量的影响程度。

在产量优化与预测方面，丁树友等<sup>[5]</sup>采集疏浚信息，利用数学模型对绞吸挖泥船产量进行优化，但是这种优化方式较为传统；王其松<sup>[6]</sup>提出了施工经验与施工数据结合的产量优化方法，但是这种方法过于依赖数据集，采集整理的难度较大；王柳艳等<sup>[7]</sup>用 RBF 神经网络预测产量，但得到的决定系数较低；陈秀静等<sup>[8]</sup>建立了带有时滞的 BP 神经网络预测模型；杨金宝等<sup>[9]</sup>采用 Levenberg Marquardt 算法的双层 BP 神经网络预测绞吸挖泥船的产量。然而，文献[8-9]使用的 BP 神经网络容易陷入局部极小点。

本文基于主成分分析与 PSO-RELM 建立了绞吸挖泥船产量的预测模型，并对预测产量做可视化。首先，选取两组工况较好的完整横移周期数据进行主成分分析，减少了影响产量的变量的个数。然后，建立绞吸挖泥船的 PSO-RELM 产量预测模型，达到了良好的预测精度。最后，利用 Echarts 实现预测产量的可视化分析及其辅助决策。

1 数据预处理与主成分分析

1.1 实际数据的预处理

数据选自某绞吸挖泥船的实际疏浚数据，通过横移周期，挑选两组工况较好的数据，相关信息见表 1。

表 1 浓度、流速、产量数据

组名	平均浓度/ %	平均流速/ (m·s <sup>-1</sup> )	平均产量/ (m <sup>3</sup> ·h <sup>-1</sup> )
1	28.546	5.536 3	3 244.3
2	30.542	5.414 5	3 378.7

挖泥船实际疏浚得到的数据并不都是良好的数据，需要进一步的处理：

- 1)插值。使用邻近值插值方式，补充未记录的数据。
- 2)清除无效数据和异常值处理。为了能反映绞吸挖泥船正常工作时的特性，成组删掉数据中包含 0 值以及明显不符合实际作业状态的数据。
- 3)数据平滑处理。采用中值滤波方式对后面的数据进行滤波。
- 4)数据标准化。本文进行的是 z-score 标准化处理，经过处理后的数据均值为 0，标准差为 1，变换函数是：

$$x^*=(x-\mu)/\sigma$$

(1)

式中： $x$  为原始数据； $x^*$  为标准化后的数据； $\mu$  为原始数据的均值； $\sigma$  为原始数据的标准差。部分处理数据见表 2。

表 2 部分处理数据

序号	转速/(r·min <sup>-1</sup> )			转速/(m·min <sup>-1</sup> )		绞刀深度/m
	绞刀	1#甲板泵	2#甲板泵	左横移绞车	右横移绞车	
1	0.372 4	-0.776 3	0.217 1	0.170 2	0.135 6	-0.791 7
2	0.054 4	-1	0.289 5	-0.472 3	-0.368 6	0.500 0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
2302	0.640 2	0.080 8	0.427 6	-0.403 4	-0.097 5	0.666 7

1.2 主成分分析

影响挖泥船产量的直接因素有泥浆流速和泥浆浓度，而影响泥浆流速的控制因素主要是泥泵转速，影响泥浆浓度的控制因素主要有横移速度、绞刀转速、绞刀深度等，因此对 1#甲板泵转速、2#甲板泵转速、左横移绞车转速、右横移绞车转

速、绞刀转速、绞刀深度等产量影响因素进行主成分分析<sup>[10]</sup>，结果见表 3。按方差贡献率由大到小依次排列各产量影响因素，主成分个数根据方差累计贡献率大于 85%时选取。可以得到影响挖泥船产量的主成分为：右横移绞车转速、左横移绞车转速、绞刀深度、1#甲板泵转速、绞刀转速。

表 3 主成分分析结果

影响因素	贡献率/%	累计贡献率/%
右横移绞车转速	33.16	33.16
左横移绞车转速	25.45	58.61
绞刀深度	14.61	73.22
1#甲板泵转速	9.64	82.86
绞刀转速	8.74	91.60
2#甲板泵转速	8.40	100.00

2 产量预测模型

2.1 粒子群优化的正则化极限学习机(PSO-RELM)

因为影响绞吸挖泥船瞬时产量的因素较多,且挖泥船控制系统中大多是非线性部分,难以得到绞吸挖泥船的机理模型。极限学习机(ELM)具有训练速度快<sup>[11]</sup>、泛化性能好的优点,可以将绞吸挖泥船的控制

系统建立一个黑箱模型。极限学习机的结构是一种典型的单隐层前馈神经网络(SLFN)。极限学习机的结构见图 1。

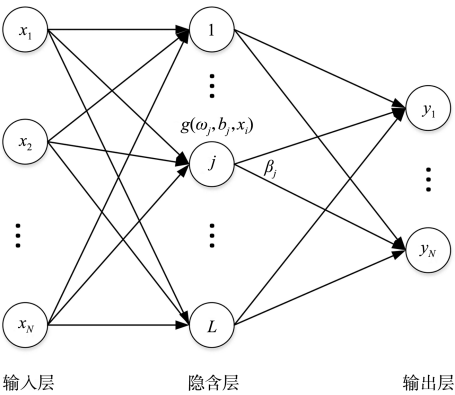


图 1 极限学习机的结构

RELM 算法<sup>[2]</sup>: 若  $N$  个训练样本,  $\mathbf{x}=\{x_1, x_2, \cdots, x_N\}^T$  作为输入,  $\mathbf{T}=\{t_1, t_2, \cdots, t_N\}^T$  作为输出, 设定隐含层节点为  $L$ , 激活函数为  $g(\omega, b, x)$ , 随机选取输入层权值  $\omega=\{\omega_1, \omega_2, \cdots, \omega_L\}^T$  与隐含层偏置  $\mathbf{b}=\{b_1, b_2, \cdots, b_L\}^T$ 。极限学习机的训练目标是令:

$$\mathbf{T}=\mathbf{H}\mathbf{\beta} \tag{2}$$

式中:

$$\mathbf{H}=\begin{pmatrix} g(\omega_1, b_1, x_1) & g(\omega_2, b_2, x_1) & \cdots & g(\omega_L, b_L, x_1) \\ g(\omega_1, b_1, x_2) & g(\omega_2, b_2, x_2) & \cdots & g(\omega_L, b_L, x_2) \\ \vdots & \vdots & \ddots & \vdots \\ g(\omega_1, b_1, x_N) & g(\omega_2, b_2, x_N) & \cdots & g(\omega_L, b_L, x_N) \end{pmatrix}_{N \times L};$$

$\mathbf{\beta}=(\beta_1, \beta_2, \cdots, \beta_L)^T$  为输出层权值。对于式 (2), 由最小二乘法

$$\mathbf{\beta}=(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{T} \tag{3}$$

引入正则化系数可提高 ELM 的结构稳定性, 得到 RELM:

$$\mathbf{\beta}=(\mathbf{H}^T\mathbf{H}+\mathbf{C}\mathbf{I})^{-1}\mathbf{H}^T\mathbf{T} \tag{4}$$

式中:  $\mathbf{C}$  为正则化系数;  $\mathbf{I}$  为单位矩阵。

PSO-RELM 算法即利用 PSO<sup>[13]</sup> 对 RELM 的隐层参数进行迭代选优, 粒子群优化的正则化极限学习机流程见图 2。

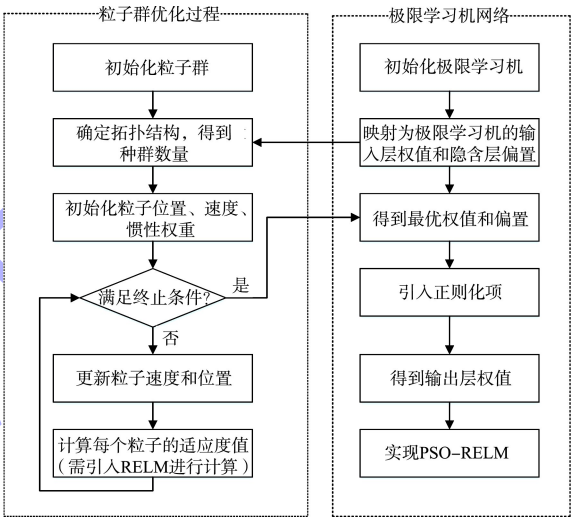


图 2 粒子群优化正则化极限学习机流程

PSO-RELM 具体步骤如下:

步骤 1: 初始化。确定粒子个数  $d=L(n+1)$ , 其中  $L$  为隐含层节点数,  $n$  为输入变量的个数。设定每个粒子的初始速度和位置  $\{(v_i, x_i) \mid i=1, 2, \cdots, d\}$ , 个体学习因子  $c_1$ 、群体学习因子  $c_2$ 、惯性权重  $\omega$ 、终止条件(达到迭代次数或小于目标值)。

步骤 2: 确定适应度值。本文中将训练样本输出和实际输出的均方误差值(MSE)设为适应度, 并在每次迭代中计算。

步骤 3: 根据适应度寻优并更新每个粒子的速度和位置。将适应度值与每个粒子  $v_i$  的个体极值  $p_{best_i}$  比较, 将更好的  $p_{best_i}$  作为本次迭代的适应度值, 并更新对应的  $p_{best_i}$  位置; 同理, 将每个粒子的  $p_{best_i}$  与群体极值  $g_{best_i}$  对比, 更新群体最优位置  $g_{best_i}$ , 并更新  $g_{best_i}$  对应位置。根据式 (5)、(6) 更新速度和位置。

$$v_i = \omega v_i + c_1 \cdot \text{rand}() \cdot (p_{\text{best } i} - x_i) +$$
$$c_2 \cdot \text{rand}() \cdot (g_{\text{best } i} - x_i)$$

(5)

$$x_i = x_i + v_i$$

(6)

步骤 4: 不断迭代, 直到满足终止条件。迭代过程中, 使用式(4)来确保正则化项的引入。

步骤 5: 将步骤 4 得到的输入权值和偏置代入到 RELM 网络结构中, 实现 PSO-RELM。

2.2 产量预测模型建立

将表 3 中得到的主成分作为输入变量, 以瞬时产量作为输出变量, 建立基于极限学习机的产量预测模型。各类型极限学习机的参数设定值见表 4。

表 4 各类型极限学习机参数设定值

极限学习机类型	隐含层节点数	正则化参数	迭代次数	激活函数
ELM	50	-	-	Sigmoid
RELM	50	0.5	-	Sigmoid
PSO-RELM	50	0.5	100	Sigmoid

对表 2 中处理后的数据随机选取 2 250 组作为训练样本, 选取剩下的 52 组作为测试样本。各类型极限学习机的预测结果见图 3。

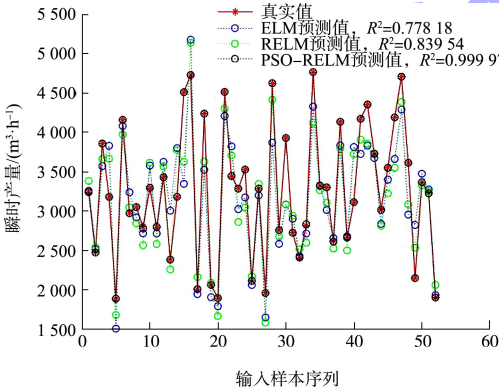


图 3 各类极限学习机的预测结果

通过均方误差 (MSE) 和决定系数 ( $R^2$ ) 来评价预测结果好坏。其中  $R^2$  的计算方式如下:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

(7)

式中:  $SSR = \sum (\hat{y}_i - \bar{y})^2$ ;  $SST = \sum (y_i - \bar{y})^2$ ;  $SSE = \sum (y_i - \hat{y}_i)^2$ ;  $y_i$  为真实值,  $\bar{y}$  为  $y_i$  的平均值;  $\hat{y}_i$  为  $y_i$  的预测值。 $R^2$  的取值在  $[0, 1]$  之间,  $R^2$  值越接近 1, 拟合的效果越好。各类型极限学习机预测结果见表 5。

表 5 各类型极限学习机预测结果

极限学习机类型	均方误差 MSE	决定系数 $R^2$
ELM	252 897.375 2	0.778 18
RELM	173 238.287 9	0.839 54
PSO-RELM	216.563 4	0.999 97

由图 3 和表 5 可以得出以下结论:

- 1) 各类 ELM 算法均可以对绞吸挖泥船的产量进行预测。
- 2) RELM 确实能略微提高 ELM 的泛化能力, 拟合效果比 ELM 提高了 7.88%, MSE 比 ELM 的 MSE 降低了 31.5%, 但是二者仍在同一数量级, 由此也证明了 RELM 对 ELM 预测效果的提升是有限的。
- 3) PSO-RELM 的泛化能力最为出众, 能够达到良好的拟合效果, 其拟合效果比 RELM 提高了 19.1%、比 ELM 提高了 28.5%, PSO-RELM 的 MSE 比 RELM 和 ELM 的 MSE 减小了 3 个数量级。
- 因此, 利用具有良好泛化能力的 PSO-RELM 的产量预测模型, 可进行预测产量的可视化。

3 挖泥船预测产量的可视化及其辅助决策

3.1 挖泥船预测产量的可视化

采用 Echarts 对预测产量进行可视化显示和分析。根据本次采集数据中产量的范围, 设定 5 个产量区间, 分别为低 ( $0 \sim 2\,500\text{ m}^3/\text{h}$ )、较低 ( $>2\,500 \sim 3\,000\text{ m}^3/\text{h}$ )、中 ( $>3\,000 \sim 3\,500\text{ m}^3/\text{h}$ )、较高 ( $>3\,500 \sim 4\,000\text{ m}^3/\text{h}$ )、高 ( $>4\,000\text{ m}^3/\text{h}$ )。并且将整个显示框内连续时间段的最大值、最小值以及平均值进行显示。为反映各横移周期的疏浚情况, 将绞吸挖泥船预测产量按照横移周期进行折线图显示 (图 4); 为反映实时的控制变量与预测产量情况, 将绞吸挖泥船的控制变量雷达图和预测产量折线图动态联动显示 (图 5); 为反映历史的控制变量与预测产量情况, 将绞吸挖泥船预测产量和控制变量大小的历史数据联动显示 (图 6)。图 4 与图 6 中的折线图均设定了数据缩放轴, 缩放过程中数据会相应改变。



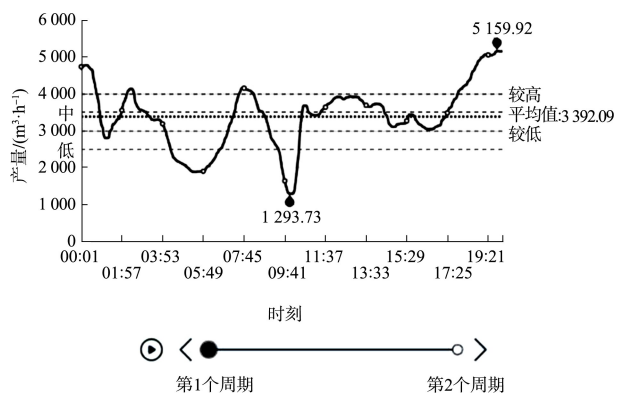


图 4 按横移周期显示预测产量

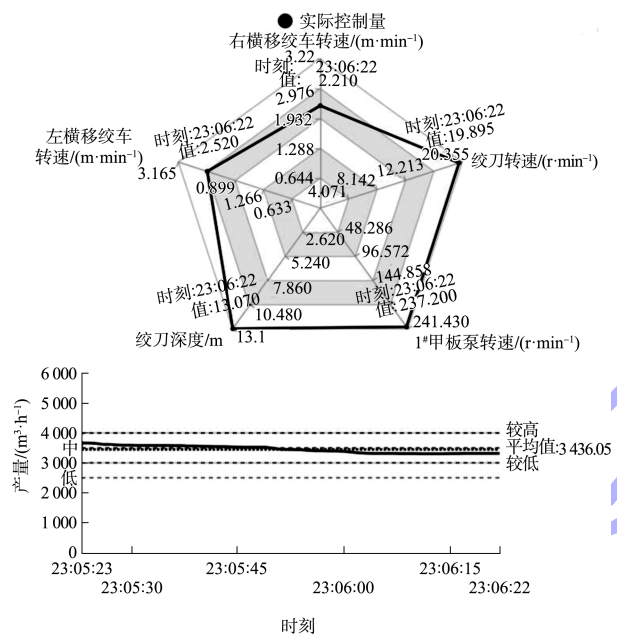


图 5 绞吸挖泥船的控制变量和预测产量的动态联动显示

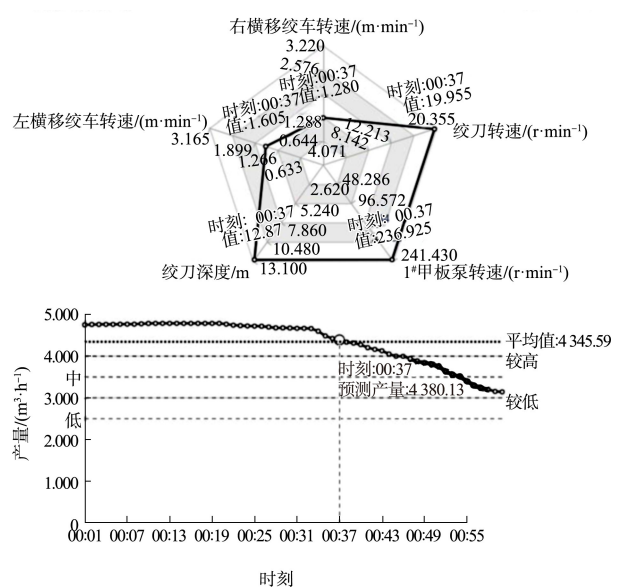


图 6 绞吸挖泥船预测产量和控制变量大小的  
历史数据联动显示

3.2 可视化分析辅助决策

针对本文数据, 每个横移周期内平均产量在 3 000 m<sup>3</sup>/h 以上, 挖泥船的疏浚效益是可观的。

由图 4~6 可提供如下的辅助决策: 1) 根据图 4 按横移周期的平均预测产量, 可以帮助操纵人员判断效益是否可观。若疏浚效益不可观, 可以在下个周期适当改变疏浚策略, 以提高疏浚效益。2) 从图 6 的历史数据中可得到预测产量 > 3 000 m<sup>3</sup>/h 的点所对应控制变量的大小。此控制变量大小可为图 5 中当前产量较低时控制变量调整提供参考, 达到辅助决策。

上述 Echarts 可视化图表相比于第 2 节中直接进行预测的结果, 能够更加清晰直观地显示预测产量的变化情况、控制变量与预测产量之间的对应关系。通过图表的分析为操纵人员提供实时的操纵帮助。

4 结论

1) 对绞吸挖泥船产量影响因素进行主成分分析, 能够减少输入变量的个数、合理选取预测模型的变量。

2) 利用 PSO-RELM 预测产量, 预测精度更高、泛化能力更好。

3) 基于 PSO-RELM 的产量预测模型建立 Echarts 可视化图表, 通过分析图表, 能够辅助操纵人员控制挖泥船, 有助于保证疏浚效益。

参考文献:

[1] 刘永杰, 朱汉华, 吴楹, 等. 绞吸式挖泥船泥沙吸入量影响因素的数值模拟分析[J]. 船海工程, 2016, 45(4): 166-169.

[2] 唐建中. 绞吸式挖泥船疏浚作业优化与控制研究[D]. 杭州: 浙江大学, 2007: 65-68.

[3] 周禹扬, 张蓓. 绞吸船施工效率影响因素分析[J]. 中国水运(下半月), 2017, 17(5): 317-318.

[4] 刘璟, 刁海岛, 张陈浩, 等. 重型绞吸式挖泥船挖掘中风化岩影响因素分析[J]. 水运工程, 2014(4): 182-186.

[5] 丁树友, 田俊峰. 绞吸挖泥船作业产量的现场优化技术[J]. 水运工程, 2007(5) 13-15, 103.