



基于 PCA 和 RBF 神经网络的 绞吸挖泥船实时产量预测

王伟¹, 戴文伯¹, 王柳艳¹, 沈彦超¹, 何俊²

(1. 中交疏浚技术装备国家工程研究中心有限公司, 上海 200120; 2. 中港疏浚有限公司, 上海 200120)

摘要: 绞吸挖泥船在实际作业过程中的动态特性非常复杂, 影响产量的控制因素众多。若这些控制因素全部参与产量预测比较耗时。为了实时训练网络及预测产量, 先对影响绞吸挖泥船产量的控制因素进行主成分分析(PCA), 再根据分析结果约减控制因素; 在系统仿真建模中, 分别以全部因素和约减后因素作为径向基(RBF)神经网络的输入变量, 以产量作为输出变量来建立绞吸挖泥船产量预测模型。结果表明, 减少输入变量, 不仅降低产量预测模型的复杂程度, 减少神经网络计算耗时, 而且能保持模型良好的预测精度, 从而为施工现场的操作人员提供实时的产量参考。

关键词: 绞吸挖泥船; 主成分分析; 径向基神经网络; 产量预测

中图分类号: U 616

文献标志码: A

文章编号: 1002-4972(2021)04-0206-05

Real-time production prediction of CSD based on PCA and RBF neural network

WANG Wei¹, DAI Wen-bo¹, WANG Liu-yan¹, SHEN Yan-chao¹, HE Jun²

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

2.CHEC Dredging Co., Ltd., Shanghai 200120, China)

Abstract: In the actual operation process, the dynamic characteristics of the cutter suction dredger (CSD) are very complicated, and many factors affect the production. It will consume much time if all these control factors participate in the output prediction. To train the network and predict the production in real-time, the PCA is utilized to these control factors. Then, the number of control factors is reduced according to the results. In the system simulation modeling, the whole control factors and reduced control factors are used separately as the input variables of the RBF neural network, and the production is used as the output variable to establish the prediction model of CSD's production. The results show that the prediction model of production can be simplified and the computation time can be reduced while the input variables are reduced, and the prediction accuracy of the model can be maintained, providing real-time production reference for operators on the construction site.

Keywords: cutter suction dredger (CSD); principal criteria analysis (PCA); RBF neural network; production prediction

绞吸挖泥船是目前疏浚工程中运用比较广泛的一种疏浚船舶, 其对土壤适用范围广, 主要用于风浪小、流速低的内河湖区和沿海港口的吹填施工。在疏浚作业过程中, 为保证工程效益, 有必要分析各种关联变量对产量的影响程度。目前已有一些研究成果: 刘永杰等^[1]研究了绞刀转速的变化对产量

的影响; 唐建中^[2]在“疏浚作业系统动态特性的学习和预测”一节提到影响泥浆浓度和流速的因素有切削深度、切削控制量、绞刀转速; 周禹扬等^[3]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。以上是从绞刀系统、时间利用率等方面分析绞吸挖泥船施工效率的

影响因素,而本文从实际操作过程中可控制因素出发,使用主成分分析(PCA)方法对影响产量的控制因素进行影响因子分析。产量预测对施工操作人员挖泥有指导意义。目前,用于产量预测的方法有数理统计分析、拟合曲线、模糊规则、神经网络^[4-9]等。本文使用径向基神经网络进行产量预测,但可控制因素全部参与产量预测比较耗时,须根据控制因素的影响贡献率约减因素,降低计算耗时,从而实现产量的实时预测。

1 数据来源及预处理

1.1 数据来源

本文所采用的数据来自多船多地的施工数据,各施工地点的土质类型不一样,因此对这些施工数据进行分析更具有代表性,见表1。

表1 施工工况土质类型

绞吸挖泥船	施工地点	土质类型
S1	P1	细粉砂
S2	P2	中粗砂
S2	P3	铁板砂

1.2 数据预处理

在实际工况中绞吸挖泥船传感器得到的数据并不全是良好的数据,数据中可能存在对分析结果有影响的无效数据、异常数据、噪声数据等,不同数据列之间可能存在量纲,因此数据分析之

前有必要进行预处理。数据预处理流程见图1。

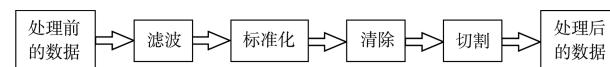


图1 数据预处理流程

步骤如下:

1) 对原始数据进行滤波,滤掉噪声信号,由于滤波器存在群延迟,故需要对各种控制因素的数据使用相同特性的滤波器,保证群延迟的一致性。

2) 对滤波后数据进行标准化处理,常用方法是Z-SCORE标准化,经过处理后的数据均值为0,标准差为1,变换函数是:

$$x^* = (x - \mu) / \sigma \quad (1)$$

式中: x^* 为标准化后的数据, x 为原始数据, μ 为原始数据的均值, σ 为原始数据的标准差。

3) 从标准化后数据中清除无效数据和异常数据。

4) 由于CSD横移到两边线时速度很小,因此需要切掉这部分数据,保留中间数据,即PCA和产量预测时使用的是稳定产量的数据。根据横移角度变化判断横移方向,选择相邻相同横移方向的2组施工数据作为分析对象。本文从3地施工数据中挑选6组有代表性的工况数据,使用第1、3、5组数据做PCA,使用第2、4、6组数据做产量预测,见表2。

表2 施工数据信息

施工地点	组别	横移方向	横移时长/ s	产量均值/ (m ³ ·h ⁻¹)	流速/ (m·s ⁻¹)	排压/ MPa	泥泵转速/ (r·min ⁻¹)	浓度/ %
P1	1	左	350	2 550	4.5	0.30	230	38
	2	左	482	2 435	4.5	0.30	230	35
P2	3	左	433	1 886	4.5	0.28	243	43
	4	左	285	1 853	4.5	0.28	243	42
P3	5	右	517	722	4.3	0.28	240	26
	6	右	444	670	4.3	0.28	240	24

2艘绞吸挖泥船上密度计和流量计都安装在船尾处,距绞刀头约100多米,因此产量数据存在时滞问题。为了得到真实的产量数据,通过计算吸泥口至密度计之间的管道长度与流速的比值,从而得到时滞时间为23、23、24 s,PCA和产量预测中的产量数据都已进行滑移处理。

2 影响产量的控制因素及PCA

2.1 影响产量的控制因素

通常,绞吸挖泥船的产量 $R(m^3/h)$ 计算公式如下:

$$R = 3 600 Q C_w \quad (2)$$

式中: Q 为泥浆流量(m³/s); C_w 为泥浆浓度(%)。

泥浆流量是泥浆流速和输泥管横截面积的乘积，即

$$Q = v \cdot \pi D^2 / 4 \quad (3)$$

式中： D 为输泥管内径(m)， v 为输泥管中的流速(m/s)。

由式(2)和(3)可知，影响产量的直接因素有泥浆流量和浓度，而影响泥浆流量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泥泵转速、横移速度、绞刀转速、入土深度、进关量等，因此在绞吸挖泥船 S1 上影响产量的控制因素主要有：水下泵转速、1# 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；在绞吸挖泥船 S2 上影响产量的控制因素主要有：水下泵转速、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。由于横移速度可以由横移绞车转速计算得到，因此没有横移速度数据时，可使用左、右横移绞车转速来代替；由于入土深度是由潮位、耳轴吃水和绞刀深度决定的，而在一个横移周期内，潮位和耳轴吃水的观测值变化很小，因此使用绞刀深度代替入土深度；由于在一个横移周期内进关量数值几乎不变，与产量实时变化的相关性很小，因此分析时进关量不作考虑。绞吸挖泥船挖泥过程中，泥泵、横移绞车、绞刀、桥架这 4 部分是联动的，与它们对应的上述控制因素共同影响产量的变化，故将这些控制因素作为分析对象。

2.2 PCA

一般分析前需要做 KMO 和巴特利特球形检验。KMO 检验可以检查各变量之间的偏相关性，取值范围是 0~1。KMO 的结果越接近 1，表示变量之间的偏相关性越好。分析时 KMO 值大于 0.5 时，可以做 PCA。巴特利特球形检验是用来判断相关矩阵是否为单位矩阵。在 PCA 中，若拒绝各变量独立的原假设，则说明可以做 PCA；如果不拒绝原假设，则说明这些变量可能独立提供一些信息，不适合做 PCA。分析时巴特利特球形检验结果小于 0.05 时，球形假设被拒绝，原始变量之间存在相关性，适合做 PCA。根据上面所述，1.2 节里所

选 3 组数据检验结果均满足条件，见表 3。

表 3 KMO 和巴特利特球形检验结果

组别	KMO 指标	显著性指标
1	0.54	2.8×10^{-6}
3	0.66	0.0
5	0.52	7.1×10^{-3}

接下来对 1.2 节预处理后的第 1、2、3 组数据进行 PCA，主成分个数根据 Kaiser 经验法则选取，选择特征值大于 1 的成分作为主成分，分别为 3、3、2 个。

根据 PCA 计算得到的贡献率和成分矩阵，计算控制因素的影响贡献率，选取与主成分数量相对应的控制因素作为约减后的因素，见表 4。

表 4 影响产量的控制因素

施工地点	全部控制因素	影响贡献率/%	约减后控制因素
P1	左横移绞车转速	38	
	绞刀深度	27	左横移绞车转速、绞刀深度、绞刀转速
	绞刀转速	22	
	水下泵转速	7	
	甲板泵转速	6	
P2	左横移绞车转速	45	
	1# 甲板泵转速	15	左横移绞车转速、1# 甲板泵转速、2# 甲板泵转速
	2# 甲板泵转速	14	
	水下泵转速	10	
	绞刀深度	9	
	绞刀转速	7	
P3	右横移绞车转速	50	
	1# 甲板泵转速	17	
	水下泵转速	15	右横移绞车转速、1# 甲板泵转速
	绞刀深度	11	
	绞刀转速	7	

以上是从 3 地施工数据里挑选 3 组数据的 PCA 结果，其结果应用于产量预测。

3 产量预测模型及结果分析

由于影响绞吸挖泥船产量的因素较多，且挖泥船控制系统中大多是非线性的，难以得到绞吸挖泥船的机理模型，所以本文采用径向基(RBF)神经网络来构建控制因素与产量的预测模型。

3.1 RBF 神经网络

RBF 神经网络的结构是一种三层前向网络。此网络包括输入层信号源节点、输出层节点和一

层隐含层节点。其基本思想是:用RBF作为隐单元的“基”构成隐含层空间,这样就可将输入矢量直接映射到隐空间。当RBF的中心点确定以后,这种映射关系就确定了。而隐含层空间到输出空间的映射是线性的,即网络的输出是隐单元输出的线性加权。此处的权即为网络可调参数。因此,从总体上看,网络由输入到输出的映射是非线性的,而网络输出对可调参数而言却又是线性的。因而,RBF神经网络具有以任意精度逼近任意连续函数的优点,可以将绞吸挖泥船的控制系统构建成一个黑箱模型。

3.2 产量预测模型

根据表3,构建全部控制因素作为输入变量的RBF神经网络模型,其模型的输出变量是产量(图2)。约减后因素的RBF神经网络模型在全部因素的RBF神经网络模型的基础上,根据表3内

容保留约减后的因素作为输入变量即可。在系统仿真模型中,模拟产量预测RBF神经网络的实时训练,将训练得到的最优平滑因子代入预测函数中,计算下一秒的预测产量。

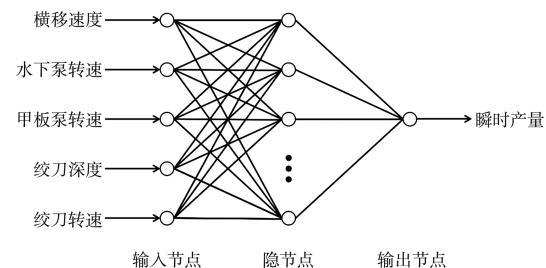
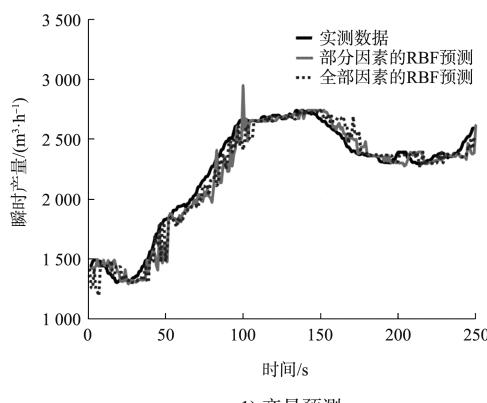


图2 全部控制因素的RBF神经网络模型

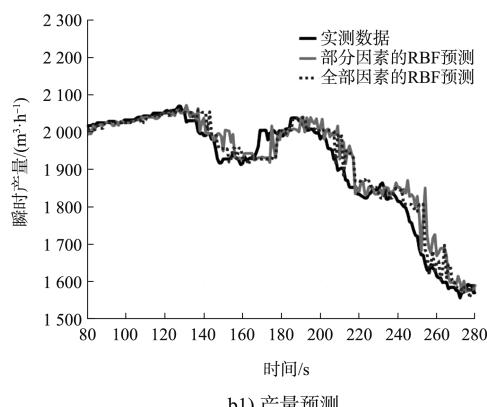
3.3 产量预测结果

采用1.2节中预处理后的第2、4、6组数据作为训练和预测数据集。绞吸挖泥船的产量预测结果及相对误差见图3。



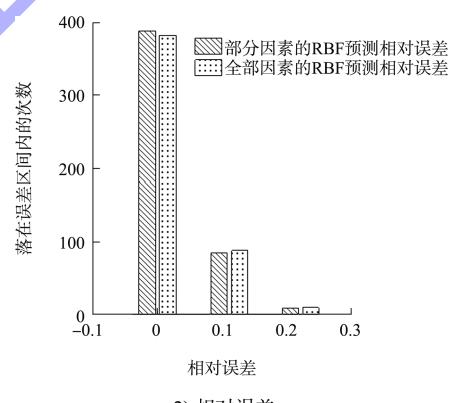
a1) 产量预测

a) 施工地点 P1

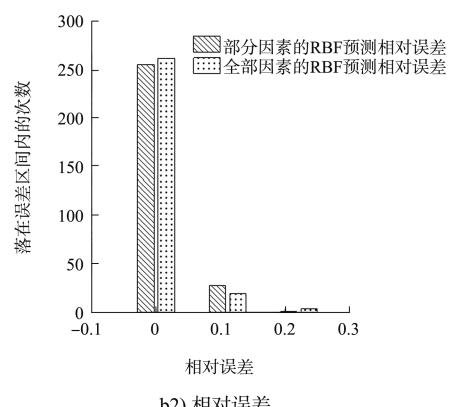


b1) 产量预测

b) 施工地点 P2



a2) 相对误差



b2) 相对误差

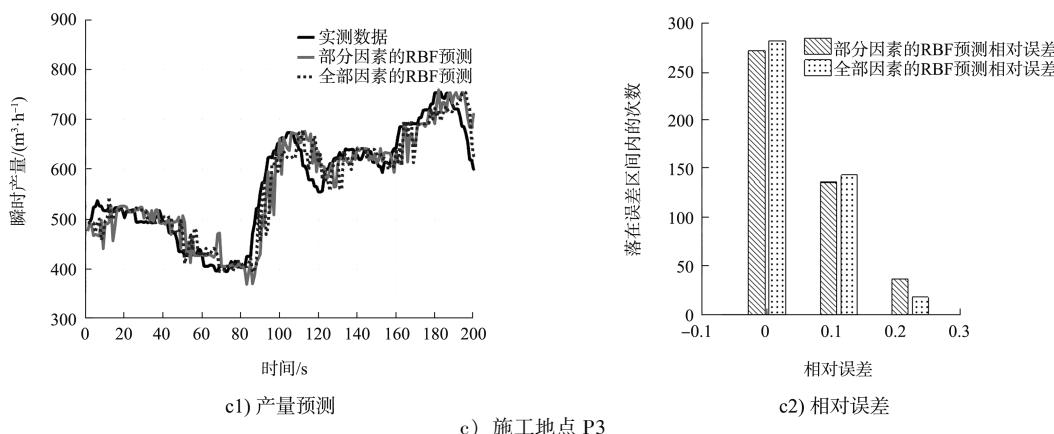


图3 不同地点的产量预测及相对误差

由图3可知,2种模型产量的预测值和实际值的曲线拟合度都较高,2种模型的相对误差都能控制在20%以内,绝大多数相对误差分布在10%以内,预测精度都可以达到较高水平。产量预测结果分析见表5。

表5 2个神经网络模型预测结果分析

施工地点	预测结果分析指标	计算耗时/s	相对误差		拟合度 R^2
			均值/%	偏差/%	
P1	部分因素 RBF 分析	529	3.52	0.955	0.950
	全部因素 RBF 分析	670	3.74	0.950	
P2	部分因素 RBF 分析	225	2.56	0.903	0.928
	部分因素 RBF 分析	288	2.11	0.928	
P3	部分因素 RBF 分析	492	6.06	0.904	0.916
	全部因素 RBF 分析	704	5.47	0.916	

表5中计算耗时是一个横移周期内产量预测所需计算时间;相对误差是预测值与观测值得误差百分比;拟合度 R^2 是指回归直线对观测值的拟合程度, R^2 最大值为1,其值越接近1,说明回归直线对观测值的拟合程度越好。由表5可以看出,2种模型的相对误差均值和拟合度 R^2 变化不大,但是约减后因素的神经网络计算耗时明显比全部因素的神经网络少。

4 结语

1)在3种土质的工况条件下,经过PCA得出横移速度对产量的影响较大,影响贡献率分别为38%、45%、50%,故施工操作过程中,为了提高产量,应优先控制横移速度,这与疏浚施工现场操作人员的经验相符合。

2)通过比较2种RBF神经网络模型的计算耗时,约减控制因素后的3地产量预测耗时分别缩

短141、63、212 s,经计算,每次预测产量的平均耗时为1.03 s,故实际应用时可以通过改变控制信号的采样周期(大于1.03 s),从而实现产量的实时预测。

3)本文是在仿真模型中使用实船历史数据进行产量影响因子分析和产量预测,下一步须在实船上进行试验及应用,结合现场实际情况,进一步研究产量预测的实时性和可靠性。

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].船舶工程,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] 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.
- [5] 王其松,严军,邓家泉.绞吸式挖泥船产量优化研究[J].水运工程,2013(3):48-55.
- [6] 张德新,蒋爽,李志刚.一种改进Smith预估绞吸挖泥船产量控制方法研究[J].机械与电子,2015(7):57-60.
- [7] 王柳艳,陈新华,王伟.基于RBF神经网络的绞吸挖泥船施工产量预测研究及分析[J].中国港湾建设,2019,39(12):64-68.
- [8] 杨金宝,倪福生,魏长赞,等.基于双隐含层BP神经网络的绞吸挖泥船产量预测[J].计算机与数字工程,2016,44(7):1234-1237.
- [9] 闭治跃,王庆丰,唐建中.基于径向基函数神经网络的挖泥船排泥管道泥浆浓度软测量模型研究[J].传感技术学报,2007,20(7):1630-1634.