

# 基于 LSTM 神经网络的船舶主机油耗实时预报

王子铭<sup>1,2)</sup> 陈顺怀<sup>1,2)</sup>

(武汉理工大学交通学院<sup>1)</sup> 武汉 430063) (高性能船舶技术教育部重点实验室<sup>2)</sup> 武汉 430063)

**摘要:**通过分析船舶航行过程中的实时主机运行参数有助于准确地评估主机能效状态,以某江海直达船实际航行数据为样本,提出一种基于深度学习的船舶主机油耗实时预报方法。对船舶实时航行数据进行预处理和数据降维处理,并构造基于 LSTM 神经网络的主机油耗实时预报模型。利用处理后的数据对神经网络超参数进行调优并对模型进行测试。实验结果表明:模型预测结果可靠,误差较小。

**关键词:**深度学习;LSTM 神经网络;油耗预报

**中图分类号:**U676.4

**doi:**10.3963/j.issn.2095-3844.2020.05.030

## 0 引言

船舶传统的能效管理模式依赖于船舶操纵人员的经验,有很大不确定性,并且降低了船舶的能源利用效率<sup>[1]</sup>。寻找更有效的船舶实时数据分析方法有助于更准确的评估船舶能效状态并提高船舶能效管理效率。孙峰等<sup>[2]</sup>基于数据挖掘技术对船舶能效状态进行了评估,周根明等<sup>[3]</sup>利用层次分析法提出了一种评估船舶主机能耗方法。这些研究虽然提供了船舶主机能耗的评估方法,但并没有对船舶的能效数据进行实时处理。叶睿等<sup>[4]</sup>利用神经网络分析了客滚船的实时数据,构建了主机油耗模型,但其输入数据仅为船体姿态数据,这影响了模型的准确性与鲁棒性。陈维南等<sup>[5]</sup>利用 BP 神经网络预测主机油耗,从而对船舶主机能耗状况进行评估,但由于 BP 神经网络未考虑船舶不同时序上能效数据之间的关联,这影响了模型的可靠性。

文中以某江海直达船实际航行的主机运行数据和船舶航行姿态数据为样本,结合循环-长短期记忆(recurrent neural networks-long short-term memory, RNN-LSTM)神经网络算法,构建了船舶主机油耗实时预报模型。将预处理及数据降维

后的数据作为网络的输入,从而降低数据噪声和网络复杂性。通过大量实验对神经网络的超参数进行优化,从而提高模型的准确性,神经网络输出得到船舶主机下一时刻油耗的预测值。

## 1 构建模型

本船舶主机油耗实时预报模型通过输入当前时刻的航行数据得到下一时刻的船舶主机油耗预测值,主要的工作过程见图 1。

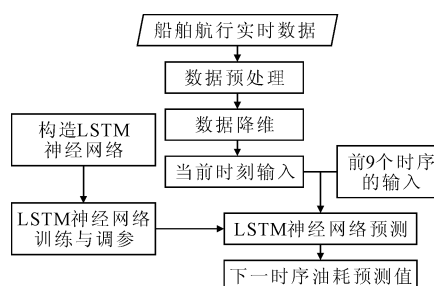


图1 船舶主机油耗实时预报流程

在不考虑海况骤变的情况下,训练好的 LSTM 神经网络对于固定航域可重复使用。并且需要指出的是,不同船舶与主机匹配关系有所不同,并且航行环境有所差异,所以针对其他船舶及航域,还需要重新训练神经网络,时间成本在可接

收稿日期:2020-07-14

王子铭(1996—):男,硕士生,主要研究领域为船舶设计

受范围内. 本模型并未考虑航域中风、浪、流等信息的影响, 故为简化处理, 假设船舶在该固定航域航行时, 环境状况稳定.

### 1.1 目标船及传感器

船舶实时数据均来源于某江海直达船, 该船总长 154 m, 型宽 24 m, 型深 11.8 m, 设计吃水 9.1 m, 空载排水 4 801.8 t, 满载排水 27 300.7 t, 主机型号为 5S35MC-9.2, 额定功率为 3 700 kW, 额定转速为 142 r/min.

船速( $V_s$ )、瞬时油耗( $q$ )、首吃水( $d_f$ )、尾吃水( $d_a$ )、左吃水( $d_l$ )、右吃水( $d_r$ )、主机转速( $n_c$ )和轴功率( $P_c$ )均由传感器采集得到. 其中, 根据船舶主机的设备和型号, 目标船传感器选择如下:

1) 主机流量计考虑到目标船主机工作时具有很高的温度, 需要耐高温性好的流量计. 故本文选择 BELZ 系列流量计中带电子显示屏的椭圆齿轮流量计. 该流量计仪表可实时读取主机的瞬时油耗量, 并通过 4~20 mA 输出信号. 流量计图片见图 2.

2) 四角吃水传感器分别在目标船船首、船中左右舷和船尾安装压力式液位传感器. 这些压力式液位传感器测量范围为 0~10 m, 并将测量到的四角吃水以 4~20 mA 输出. 压力式液位传感器见图 3.



图 2 BELZ 流量计



图 3 压力式液位传感器

3) 轴功率仪传感器 目标船采用 TT Revolution 系统采集主机曲轴转速及实时功率. 该系统为单通道、不接触、感应供电的系统, 并通过 4~20 mA 的输出信号输出采集到的数据.

### 1.2 数据预处理

选取目标船由舟山前往张家港某航次的 93 242 条记录作为样本. 由于传感器将采集到的数据传输到 PC 端过程中受到船上电器设备电磁干扰等因素, 使得最终接收到的数据有异常值. 图 4 为样本数据预处理前的情况.

由图 4 可知, 采集到的很多数据超出了数据的正常范围或不符合正常变化趋势. 神经网络预测前需对数据进行滤波及去除异常值处理, 以避免神经网络欠拟合或预测异常等问题. 限幅滤波法及箱线图法去除异常值具有计算效率高的优点<sup>[6]</sup>, 符合实时处理的要求. 表 1 为以发动机额定

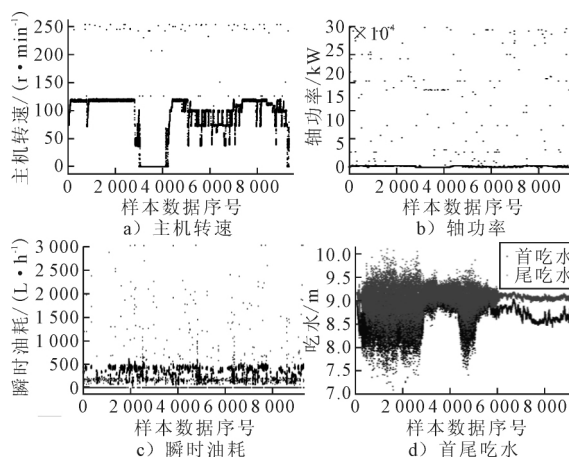


图 4 样本原始数据

转速为依据, 根据实际情况设置各状态参数理论最大值 max 与理论最小值 min. 若采集到的数据大于理论最大值或小于理论最小值则认为该数据异常, 在分析前应剔除.

表 1 目标船状态参数

| 名称                          | max   | min | 符号    | 名称    | max | min | 符号    |
|-----------------------------|-------|-----|-------|-------|-----|-----|-------|
| 船速/km                       | 15    | 0   | $V_s$ | 首吃水/m | 10  | 0   | $d_f$ |
| 瞬时油耗/( $L \cdot h^{-1}$ )   | 1 000 | 0   | $q$   | 尾吃水/m | 10  | 0   | $d_a$ |
| 主机转速/( $r \cdot min^{-1}$ ) | 142   | 0   | $n_c$ | 左吃水/m | 10  | 0   | $d_l$ |
| 轴功率/kW                      | 4 500 | 0   | $P_c$ | 右吃水/m | 10  | 0   | $d_r$ |

数据预处理的过程见图 5, 处理后的数据见图 6. 航速由自带数据预处理装置的 GPS 传感器采集, 可直接使用. 航速数据见图 7.

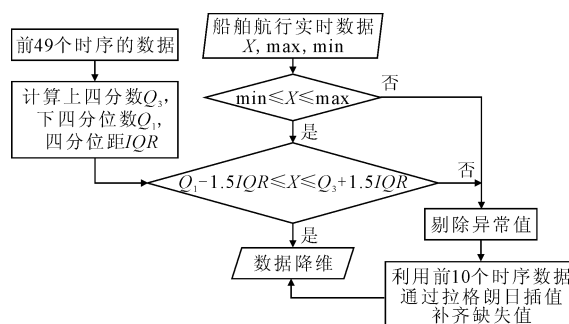


图 5 数据预处理流程图

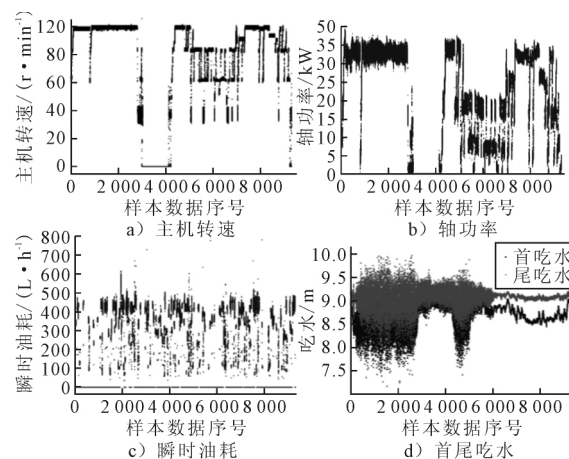


图 6 样本处理后数据

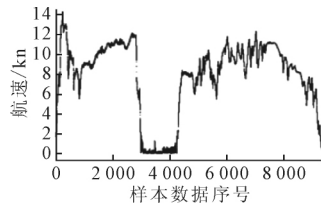


图 7 航速

### 1.3 数据降维

网络输入数据的维度过大不利于网络的实时计算效率,并且船舶各航行参数反映的信息在一定程度上有重叠.为消除这种影响并提高网络计算效率,本文使用主成分分析法(principal component analysis,PCA)对预处理后的数据进行降维处理.

主成分分析法可将高维的输入数据合成为低维的相互无关的主成分,各主成分能够反映原始

高维数据的大量信息,并且各主成分所包含的信息不重叠.主成分分析法的操作流程见图 8,经计算得到表 3 的成分矩阵.由表 3 可知,第一主成分  $Z_1$ 、第二主成分  $Z_2$ 、第三主成分  $Z_3$  累计方差贡献率达到 99.99%,可以很好反映原数据所包含的信息.故将  $Z_1$ 、 $Z_2$ 、 $Z_3$  代替原数据进行后续的计算.

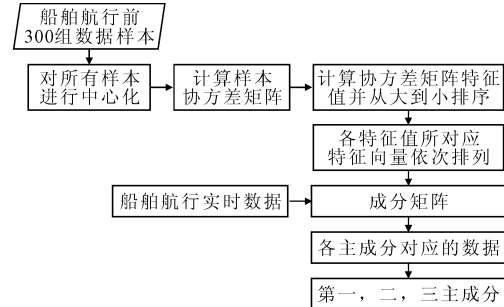


图 8 主成分分析法计算流程图

表 3 成分矩阵

| 主成分   | $Z_1$                    | $Z_2$                   | $Z_3$                   | $Z_4$                   | $Z_5$                   | $Z_6$                   | $Z_7$                   | $Z_8$                   |
|-------|--------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| $n_c$ | $-2.756 \times 10^{-2}$  | $-8.331 \times 10^{-3}$ | $9.942 \times 10^{-1}$  | $1.039 \times 10^{-1}$  | $9.514 \times 10^{-5}$  | $1.059 \times 10^{-3}$  | $2.286 \times 10^{-3}$  | $-1.098 \times 10^{-8}$ |
| $P_c$ | $-9.996 \times 10^{-1}$  | $-9.981 \times 10^{-3}$ | $-2.769 \times 10^{-2}$ | $-9.449 \times 10^{-4}$ | $-9.354 \times 10^{-5}$ | $-1.193 \times 10^{-4}$ | $-4.967 \times 10^{-5}$ | $4.326 \times 10^{-10}$ |
| $D_a$ | $1.468 \times 10^{-4}$   | $2.102 \times 10^{-5}$  | $-9.696 \times 10^{-4}$ | $1.438 \times 10^{-2}$  | $-3.908 \times 10^{-1}$ | $-8.980 \times 10^{-1}$ | $2.016 \times 10^{-1}$  | $-6.490 \times 10^{-8}$ |
| $D_b$ | $3.874 \times 10^{-5}$   | $8.138 \times 10^{-7}$  | $-5.249 \times 10^{-4}$ | $2.671 \times 10^{-3}$  | $-6.481 \times 10^{-1}$ | $2.828 \times 10^{-1}$  | $3.285 \times 10^{-3}$  | $-7.071 \times 10^{-1}$ |
| $D_c$ | $1.703 \times 10^{-5}$   | $1.847 \times 10^{-5}$  | $2.091 \times 10^{-3}$  | $3.488 \times 10^{-3}$  | $-8.479 \times 10^{-2}$ | $-1.830 \times 10^{-1}$ | $-9.794 \times 10^{-1}$ | $-4.219 \times 10^{-9}$ |
| $D_d$ | $3.874 \times 10^{-5}$   | $8.115 \times 10^{-7}$  | $-5.249 \times 10^{-4}$ | $2.672 \times 10^{-3}$  | $-6.481 \times 10^{-1}$ | $2.828 \times 10^{-1}$  | $3.285 \times 10^{-3}$  | $7.071 \times 10^{-1}$  |
| $V_s$ | $-1.926 \times 10^{-3}$  | $-1.024 \times 10^{-3}$ | $1.039 \times 10^{-1}$  | $-9.945 \times 10^{-1}$ | $-9.421 \times 10^{-3}$ | $-1.200 \times 10^{-2}$ | $-2.623 \times 10^{-4}$ | $-3.470 \times 10^{-9}$ |
| $q$   | $-1.0209 \times 10^{-2}$ | $9.999 \times 10^{-1}$  | $8.113 \times 10^{-3}$  | $-1.619 \times 10^{-4}$ | $1.031 \times 10^{-6}$  | $1.719 \times 10^{-5}$  | $3.212 \times 10^{-5}$  | $-1.269 \times 10^{-8}$ |
| 方差贡献率 | $9.813 \times 10^{-1}$   | $1.852 \times 10^{-2}$  | $1.686 \times 10^{-4}$  | $2.132 \times 10^{-6}$  | $3.595 \times 10^{-8}$  | $2.155 \times 10^{-8}$  | $3.673 \times 10^{-9}$  | $3.354 \times 10^{-23}$ |

## 2 基于 LSTM 神经网络的主机油耗实时预报模型

### 2.1 RNN-LSTM 循环神经网络

RNN 神经网络也称循环神经网络,它很适合解决输入在时序上存在关系的问题,它很好的考虑到了不同时序的输入之间隐含的信息.图 9 为 RNNs 按时序展开的网络结构示意图.

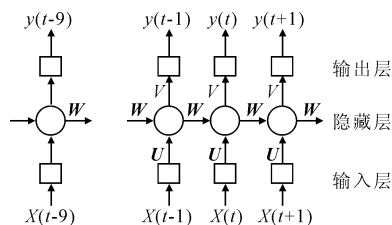


图 9 循环神经网络结构展开示意图

输入层到隐藏层的激活过程为

$$s_t = f_1(Ux_t + Ws_{t-1}) \quad (1)$$

式中: $s_t$  为隐藏层神经元的输出; $U$  为输入层或上一层隐藏层到当前隐藏层的连接权值矩阵; $W$  为上一时序隐藏层与当前时序隐藏层连接权值矩阵; $f_1(x)$  为隐藏层激活函数.正因为各时序隐

藏层之间存在连接权值矩阵  $W$ ,使 RNN 网络考虑到了不同时序输入之间的联系.隐藏层到输出层的激活过程为

$$y_t = f_2(Vs_t) \quad (2)$$

式中: $V_{ij}$  为隐藏层到输出层连接权值矩阵; $f_2(x)$  为输出层激活函数.普通的 RNN 网络随着学习的数据量加大存在结合点记忆快速衰弱的缺点,容易出现梯度爆炸或梯度消失等问题.长短期记忆(long short-term memory,LSTM)是一种特殊的 RNN,主要是为了解决长序列训练过程中的梯度消失和梯度爆炸问题,对于长序列的输入有更好的表现<sup>[7]</sup>.LSTM 神经网络除外部 RNN 循环外还有内部“LSTM 细胞”循环,其在 RNN 的基础上加入记忆单元  $C$  来记忆信息的长期状态,并通过在隐藏层各单元之间加入遗忘门、输入门、候选门、输出门,使时间序列上的记忆信息可控,从而使 RNN 网络具备了长期记忆功能<sup>[8]</sup>.LSTM 隐藏单元完整逻辑结构图见图 10.

遗忘门控制上一时刻记忆单元状态  $C_{t-1}$  将多少信息保留到当前时刻的记忆单元状态  $C_t$ ,输入门和候选门控制当前时刻的输入  $x_t$  有多少信

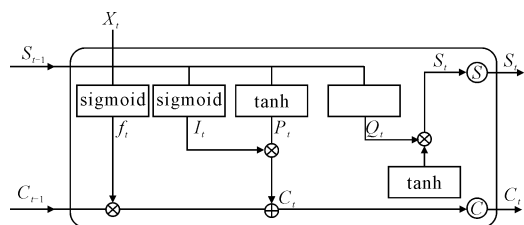


图 10 LSTM 隐藏单元完整逻辑结构图

息导入到当前时刻记忆单元状态  $C_t$ , 输出门则控制当前记忆单元状态将多少信息输出到下一时刻隐藏层。遗忘门数学模型为

$$f_t = \text{sigmoid}(\mathbf{W}_f^T \times s_{t-1} + \mathbf{U}_f^T \times x_t + b_f) \quad (3)$$

式中:  $\mathbf{W}_f, \mathbf{U}_f, b_f$  分别为遗忘门循环权重、输入权重、偏置。输入门数学模型为

$$i_t = \text{sigmoid}(\mathbf{W}_i^T \times s_{t-1} + \mathbf{U}_i^T \times x_t + b_i) \quad (4)$$

式中:  $\mathbf{W}_i, \mathbf{U}_i, b_i$  分别为输入门循环权重、输入权重、偏置。候选门数学模型为

$$p_t = \tanh(\mathbf{W}_c^T \times s_{t-1} + \mathbf{U}_c^T \times x_t + b_c) \quad (5)$$

式中:  $\mathbf{W}_c, \mathbf{U}_c, b_c$  分别为候选门循环权重、输入权重、偏置。输出门数学模型为

$$o_t = \text{sigmoid}(\mathbf{W}_o^T \times s_{t-1} + \mathbf{U}_o^T \times x_t + b_o) \quad (6)$$

式中:  $\mathbf{W}_o, \mathbf{U}_o, b_o$  分别为输出门循环权重、输入权重、偏置。记忆单元数学模型及 LSTM 细胞最终输出为

$$C_t = f_t \times C_{t-1} + i_t \times C'_t \quad (7)$$

$$S_t = O_t \times \tanh C_t \quad (8)$$

LSTM 神经网络在输入大量数据后, 通过向前传递与反向传播机制对网络内的权值进行修正, 直到网络收敛. 本文利用 python 的 kares 框架构建 4 层的 LSTM 神经网络, 其中输入层与输出层为全连接层, 两层隐藏层为 LSTM 层. 并选取 sigmoid 作为神经网络的激活函数, 性能函数采用绝对值误差 (mea) 函数. 选取 10 为时间步, 即 LSTM 网络当前时刻输入数据与前 9 个时序的输入存在联系.

## 2.2 超参数调优

LSTM 神经网络有许多超参数影响着网络的拟合准确度, 针对不同特点的数据, 各超参数均有其相对更合适的值. 本文针对 batch size (批尺寸)、epochs (世代)、优化算法、网络权值初始化方法和一二层隐藏层神经元数量进行调优.

由于目前并没有成熟的理论解决超参数取值问题, 本文通过遍历的方法, 在 3 次交叉验证的基础上以测试集的 RMSE (均方根误差) 为依据确定各超参数的最优设定. RMSE 计算公式为

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_{obs,i} - y_{model,i})^2}{n}} \quad (9)$$

式中:  $n$  为样本数目;  $y_{obs,i}$  为预测值;  $y_{model,i}$  为正确值. 图 11 为超参数取不同值时的 RMSE. 故而经过计算, 各超参数最优设定见表 4.

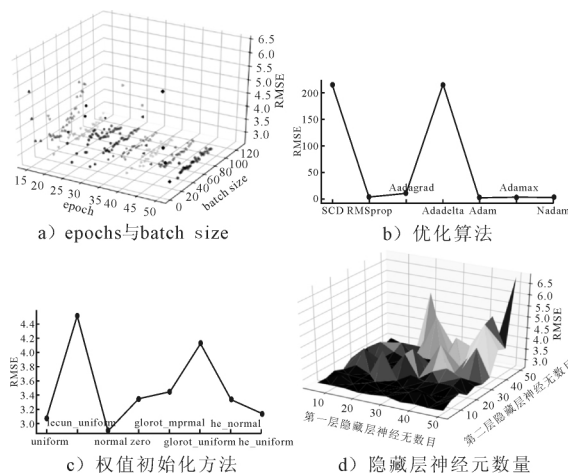


图 11 超参数取不同值时的 RMSE

表 4 神经网络超参数最优取值

| 超参数 | epochs | batch size | 优化算法  | 权值初始化方法 | 第一隐藏层神经元数 | 第二隐藏层神经元数 |
|-----|--------|------------|-------|---------|-----------|-----------|
| 最优值 | 50     | 30         | adam  | normal  | 40        | 15        |
|     | 2.897  | 2.897      | 2.820 | 2.804   | 2.709     | 2.709     |

## 3 模型验证及分析

为符合分别将数据集中 80% 和 20% 的数据作为训练集和测试集的训练策略, 并提高预测精度, 将某航次采集到的数据样本前 83 242 个时序的数据作为训练集, 后 10 000 个时序的数据作为测试集. 为了避免神经网络拟合过程中不必要的数据问题, 将数据样本中每个时序数据的第一主成分、第二主成分、第三主成分归一化后作为网络的输入. 在 GPU 上对网络进行训练, 其中计算机显卡型号为 Nvidia GeForce GTX 1060, CUDA 版本为 9.0. 每 epochs 平均训练时间为 23 s, 测试集测试结果见图 12.

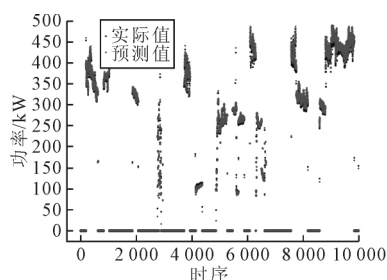


图 12 实时油耗预测结果

由图 12 可知,数据集中区域预测效果较好,当某时序数据相比上时序数据有较大变化时,预测值相对小于实际值.这是因为网络输入数据并未考虑到航行过程中风、浪、流等信息,降低了网络的鲁棒性,导致实时数据相对上一时序数据有较大改变时,网络未能及时响应,使网络存在一定的拟合误差,但误差在可接受范围之内.

预测结果显示,测试集的  $RMES$  值为 2.714.考虑到实时数据采集存在一定的不确定性,并且计算时存在一定舍入误差,这里认为神经网络预测值与实际值之间绝对误差不超过 1 可认为预测值与实际值相同,这时得到测试集相对误差百分数绝对值为 3.71%,可见网络预测结果较好.

#### 4 结 束 语

文中结合某江海直达船实际航行数据,经过数据预处理及数据降维,将高维的航行数据剔除异常值并降为 3 维,从而降低了数据噪声并提高了后续运算效率.基于 LSTM 神经网络建立主机油耗实时预报模型,以目标船舶开往张家港某航次的数据作为样本对网络进行训练及超参数调优,从而对主机油耗进行预测,测试集均方根误差

为 2.714.运用该模型可对正常状态下航行的船舶进行主机油耗实时预报,为船舶主机能效状态评估提供参考.

#### 参 考 文 献

- [1] 董建华,吴雨华,郑士军.船舶状态评估及预算系统[J].中国航海,2009,32(1):26-29.
- [2] 孙峰,黄连中,刘伊凡,等.一种应用数据挖掘技术的评估柴油机性能方法[J].大连海事大学学报,2017,43(3):83-88.
- [3] 周根明,唐晓霞,郭霆,等.基于 AHP 方法的船舶柴油机健康状态评估[J].柴油机,2013,35(6):9-13.
- [4] 叶睿,许劲松.基于人工神经网络的船舶油耗模型[J].船舶工程,2016,38(3):85-88.
- [5] 陈维南,黄连忠,等.基于 BP 神经网络的船舶主机能耗状态评估[J].中国航海,2018,13(4):127-133.
- [6] 李辉,杨增辉,周海洋,等.快速滤波算法用于在线监测数据预处理[J].高电压技术,2002,28(7):30-31.
- [7] 杨帆,何正伟,何帆.基于 LSTM 神经网络的船舶异常行为检测方法[J].武汉理工大学学报(交通科学与工程版),2019,43(5):886-892.
- [8] 高梓博,杜太利,张勇,等.聚类算法在船舶能效数据挖掘中的应用[J].武汉理工大学学报(交通科学与工程版),2019,43(2):286-290.

## Real-time Forecast of Fuel Consumption of Ship Main Engine Based on LSTM Neural Network

WANG Ziming<sup>1,2)</sup> CHEN Shunhuai<sup>1,2)</sup>

(School of Transportation, Wuhan University of Technology, Wuhan 430063, China)<sup>1)</sup>

(Key Laboratory of High Performance Marine Technology,  
Ministry of Education, Wuhan University of Technology, Wuhan 430063, China)<sup>2)</sup>

**Abstract:** It is helpful to accurately evaluate the energy efficiency status of the main engine by analyzing the real-time operating parameters of the main engine during the ship sailing. Taking the actual sailing data of a river-sea direct ship as a sample, a real-time prediction method of fuel consumption of ship main engine based on deep learning was proposed. The ship's real-time navigation data was pre-processed and data dimension reduced, and a real-time prediction model of main engine fuel consumption based on LSTM neural network was constructed. Using the processed data, the neural network hyperparameters were optimized and the model was tested. The experimental results show that the prediction results of the model are reliable and the error is small.

**Key words:** deep learning; LSTM neural network; fuel consumption forecast