

# 基于改进 RNN 和 VAR 的船舶设备故障预测方法



曾友渝 谢强

南京航空航天大学计算机科学与技术学院 南京 211106

(1056532596@qq.com)

**摘要** 针对现有的多变量时间序列预测方法不能适用于船舶多设备故障预测的问题,提出一种基于改进的循环神经网络和向量自回归的船舶设备故障预测方法。该方法既能够学习多个变量之间的相互依赖关系和时间序列的长期依赖关系,又有助于减轻传统神经网络对预测时间序列的输入尺度不敏感性。首先,从船舶历史数据库中提取出正常状态数据和故障状态数据,将其多变量时间序列转化为监督学习问题的输入;然后,通过注意力机制捕获船舶多变量之间复杂的相关性;接着,将注意力机制的输出同时作为循环神经网络和向量自回归的输入,分别捕获船舶时间信号的非线性关系和线性关系;最后,将循环神经网络组件和向量自回归组件的输出进行处理后作为最终预测的结果。实验结果表明,提出的预测方法在船舶设备故障预测中训练过程的稳定性高,测试结果的均方根误差低于 1.2,从而能更精确地预测船舶设备属性的趋势并避免故障的发生。

**关键词:** 注意力机制;循环神经网络;向量自回归;故障预测;船舶设备

**中图分类号** TP206

## Fault Prediction Method Based on Improved RNN and VAR for Ship Equipment

ZENG You-yu and XIE Qiang

College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

**Abstract** Aiming at the problem that the existing multivariable time series prediction methods cannot be applied to the multi-sensor fault prediction of ships, an improved recurrent neural network and vector autoregressive fault prediction method for ships equipment is proposed. This method can not only learn the interdependence of multiple variables and the long-term dependence of time series, but also help to reduce the insensitivity of traditional neural network to the input scale of time series prediction. Firstly, the data of normal state and fault state are extracted from the ship history database and converted into the input of the supervised learning problem. Then, the complex correlation between ship variables is captured by the attention mechanism. The nonlinear and linear relationship of ship time signals are captured by inputting the output of attention mechanism into recurrent neural network and vector autoregression. Finally, the outputs of recurrent neural network components and the outputs of vector autoregressive components are processed as the final prediction results. The experimental results show that the proposed method is more stable in the training process of ship equipment fault prediction, and the root-mean-square error of the test results below 1.2. It can more accurately predict the trend of ship equipment properties and fault occurrence.

**Keywords** Attention mechanism, Recurrent neural network, Vector autoregression, Fault prediction, Ship equipment

### 1 引言

近年来,随着船舶工业和科技的飞速发展,船舶系统日趋复杂化和智能化,导致船舶主要设备(如柴油机、锅炉等)发生故障和失效的概率逐渐增大<sup>[1-2]</sup>。目前,对船舶设备的维修主要采用定期维修和事后维修的方式,定期维修的方式不仅效率低还耗费资源,而事后维修无法避免事故的发生<sup>[3]</sup>。随着对船舶安全和可靠行驶的要求不断提高,维修策略需要从传统的事后维修和定期维修转向视情维修。船舶设备故障预测技术是基于船舶设备当前的运行状态,通过监测对应健康指

标的变化规律,预测其未来发展的趋势,进而判断设备性能的退化时刻或关键部件的失效时刻,及时进行干预或维护,避免重大事故的发生。

船舶设备运行状态数据实质上是多变量时间序列,多个传感器采集不同的设备属性,而每一个设备属性的数据就是一个时间序列。针对时间序列的预测,研究人员已经做了很多工作。自回归综合移动平均(Autoregression Integrated Moving Averagemodel, ARIMA)<sup>[4]</sup>和线性支持向量回归(Support Vector Regression, SVR)<sup>[5]</sup>在时间序列预测上都表现很优秀,但它们局限于线性单变量时间序列,不能很好地适

应船舶的多变量时间序列数据。为了解决这一问题,基于 AR 模型的向量自回归(Vector Autoregression, VAR)模型被提出,用于预测多变量时间序列。然而,VAR 模型不能捕获到船舶多变量时间序列数据的非线性。为了捕获多变量时间序列的非线性,大量学者使用神经网络来预测多变量时间序列。Zhang 等<sup>[6]</sup>从特征工程着手,提出了一种特征工程和 Dropout 深度前馈网络相结合的预测方法,并通过刀具磨损实验验证了此预测方法的训练过程稳定和预测准确率高。Bai 等<sup>[7]</sup>提出了一种新型网络——基于扩张卷积和残差连接设计的时间卷积网络(Temporal Convolutional Network, TCN),并通过多个实验验证了该方法在复杂的序列建模任务中的表现明显优于通用的循环架构。Taylor 等<sup>[8]</sup>提出了一种具有可解释参数的模块化回归模型,其思想类似于 STL (Seasonal and Trend Decomposition Using Loess)分解思路,他们通过实验证明此方法在控制程度和可解释性上比传统的时序模型更有优势。Lai 等<sup>[9]</sup>设计出第一个专门为 MFTS (Multivariate Financial Time Series)预测的模型——长期和短期时间序列网络(LSTNet),该方法结合了 CNN, RNN 以及自回归组件的优点,在多个数据集上验证了该方法对强周期模式数据集的有效性。Shih 等<sup>[10]</sup>提出了 TPA-LSTM 模型,并引入注意力(attention)机制选择相关变量加权,通过实验证明了该方法结构简单,并且适用于各种数据集,甚至是非周期性和非线性的数据集。由于用于训练船舶故障预测模型的数据来自船舶历史数据集,历史数据集中包含有正常状态数据和故障状态数据。针对这些多变量时间序列预测的模型只适用于船舶单个设备的多个属性或某个器件的预测,无法满足船舶多设备的故障预测。

针对现有方法无法满足船舶设备故障预测的问题,本文引入改进的注意力机制,自适应地选择与预测相关的船舶设备性能参数并加权,然后利用循环神经网络捕获船舶设备属性时间序列的长期依赖关系,利用向量自回归模型捕获船舶多变量时间序列的线性关系,结合线性与非线性模型对船舶设备实时状态数据进行稳健的预测。本文提出的方法既能够学习船舶多个设备属性之间的相互依赖关系<sup>[11]</sup>和时间序列的长期依赖关系,又有助于减轻传统神经网络对预测时间序列的输入尺度不敏感性。

## 2 船舶设备故障预测模型

本文提出的船舶设备故障预测模型如图 1 所示。第 1 步,从船舶历史数据库中获取船舶设备历史运行数据;第 2 步,将船舶历史数据进行归一化处理,然后添加滑动窗口转为监督学习问题;第 3 步,将处理后的船舶数据作为模型的输入,并将其划分为训练集和测试集;第 4 步,在训练集部分判断训练次数是否达到要求,若没有达到则进入第 5 步,继续训练网络,若达到则进入第 8 步,生成预测模型;第 5 步,将改进的注意力机制(attention mechanism)作用于训练集,将生成的输出同时输入循环神经网络和向量自回归;第 6 步,同时使用循环神经网络和向量自回归预测船舶设备属性的未来值得

到相应的预测结果,然后将两者的预测结果处理后作为训练集的最终预测结果;第 7 步,计算训练损失,调参并优化网络,然后转到第 4 步;第 8 步,生成预测模型;第 9 步,将预测模型作用于测试集,进行网络测试并生成预测结果;第 10 步,根据预测结果评估船舶设备故障预测模型。

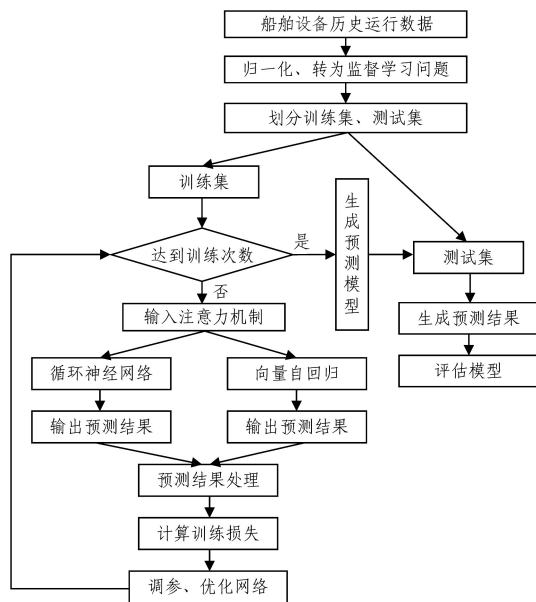


图 1 基于改进的 RNN 和 VAR 的船舶设备故障预测模型

Fig. 1 Ship equipment fault prediction model based on improved RNN and VAR

## 3 基于改进 RNN 和 VAR 的船舶设备故障预测的设计

### 3.1 船舶设备原始数据的注意力处理

在时间序列预测<sup>[12]</sup>中,典型的注意力机制常被用在时间维度上选择相关时间步骤来帮助预测结果的输出,但是它无法捕获每个时刻的多个特征之间的相互关系<sup>[13]</sup>。由于船舶系统收集的数据是多变量时间序列(Multivariate Time Series, MTS)数据,不同的序列之间可能存在复杂的动态相关性,如柴油机供油量不足可能会导致船舶柴油机排气温度升高,也可能导致船舶柴油机转速下降或转速不急,这些相关性非常重要,但是难于捕获和分析。因此,本文提出了一种改进的注意力机制,自适应地选择与预测相关的时间序列,而不是选择相关的时间步骤。

在船舶设备故障预测中,给定窗口内船舶所采集的原始数据作为输入  $X = \{x_1, x_2, \dots, x_n\}$  或  $X = \{x^1, x^2, \dots, x^n\}$ , 其中,  $x_t (t=1, 2, \dots, T)$  表示第  $t$  时刻的观测值,  $x^i (i=1, 2, \dots, n)$  表示第  $i$  个时间序列在窗口内的观测值,  $T$  表示滑动窗口的大小,  $n$  表示船舶所监测的设备属性个数, 时刻  $t$  的观测值  $x_t = \{x_t^1, x_t^2, \dots, x_t^n\}$ ,  $x_t^i (i=1, 2, \dots, n)$  表示第  $i$  个时间序列在  $t$  时刻的观测值。

注意力机制的成功很大程度上源于它在大量信息中快速捕获并筛选出高价值的信息,但是在 RNN 上应用典型的注意机制进行 MTS 预测时存在一个严重的缺陷,即典型的注意机制选择与当前时间步相关的信息。该设计适合每个时间

步包含一段信息的任务,如 NLP 任务,其中每个时间步对应一个单词。如果每个时间步中存在多个变量,则无法捕获变量之间复杂的相关性。本文所提出的注意力机制只关注其时间序列(行向量  $x^i$ ),行上的注意力权重选择有助于预测的变量<sup>[14]</sup>。本文提出的注意力机制如图 2 所示,具体步骤如下:

$$e_i^t = v_m^T \tanh(W_m [h_{t-1}; s_{t-1}] + U_m x^i) \quad (1)$$

$$a_i^t = \text{softmax}(e_i^t) \quad (2)$$

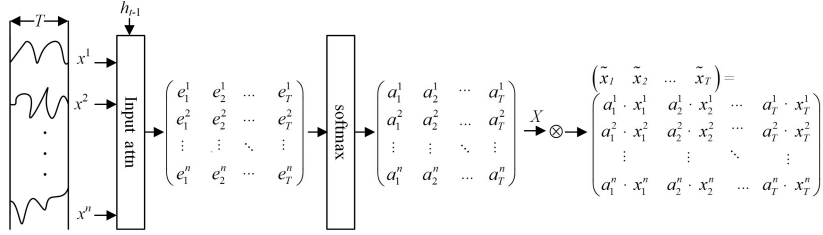


图 2 船舶设备原始数据的注意力处理

Fig. 2 Attention processing of ship equipment raw data

隐藏状态  $h_t$  由上一时刻  $h_{t-1}$  和新的  $t$  时刻的观测值  $\tilde{x}_t$  共同决定:

$$h_t = o(h_{t-1}, \tilde{x}_t) \quad (4)$$

其中,  $o(\cdot)$  是 GRU 单元。根据所提出的输入注意力机制,每个时间序列获得不同的权值,代表不同的重要程度,因此模型可以有选择地聚焦于某些时间序列。

### 3.2 基于改进的循环神经网络的船舶设备故障预测

通过输入注意力组件得到的多变量时间序列  $\tilde{X}$  同时被送入循环神经网络和向量自回归,分别得出船舶故障预测结果,然后将两者预测结果的平均值作为船舶故障预测的结果。预测结果的处理过程如图 3 所示。

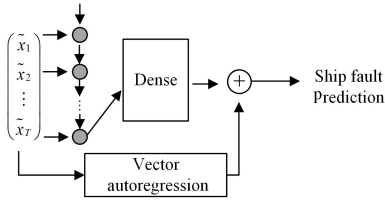


图 3 船舶设备预测结果的处理

Fig. 3 Architecture of ship equipment fault prediction

本文使用 GRU(Gated Recurrent Unit)作为循环组件,捕获表示船舶运行状态的多变量时间序列的时间依赖性以及长期模式<sup>[15]</sup>。GRU 的内部结构如图 4 所示。

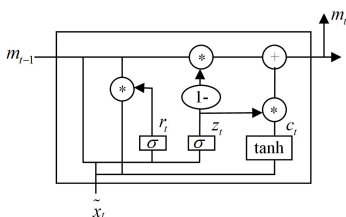


图 4 GRU 内部结构

Fig. 4 Internal structure of GRU

具体公式如下:

$$z_t = \sigma(W_z \cdot [m_{t-1}, \tilde{x}_t]) \quad (5)$$

其中,  $v_m^T \in R^T$ ,  $W_m \in R^{T \times 2q}$ ,  $U_m \in R^{T \times T}$  都是需要学习的参数,  $q$  是隐藏层的维度,  $h_{t-1}$  是上一时刻的隐藏状态,  $s_{t-1}$  是上一个时刻的单元状态。  $e_i^t$  表示在  $t$  时刻第  $i$  个时间序列的重要性。通过一个 softmax 函数确保所有注意力之和为 1。利用这些注意力权值,可以得到新的多变量时间序列  $\tilde{X} = \{\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_T\}$  如下:

$$\tilde{x}_t = (a_1^t x_1^t, a_2^t x_2^t, \dots, a_n^t x_n^t)^T \quad (3)$$

$$r_t = \sigma(W_r \cdot [m_{t-1}, \tilde{x}_t]) \quad (6)$$

$$c_t = \tanh(W \cdot [r_t * m_{t-1}, x_t]) \quad (7)$$

$$m_t = (1 - z_t) * h_{t-1} + z_t * c_t \quad (8)$$

其中,  $\tilde{x}_t$  是经过注意力组件所得到的第  $t$  时刻的多变量时间序列,  $m_t$  表示  $t$  时刻 GRU 单元的隐藏状态,  $c_t$  表示  $t$  时刻 GRU 单元的单元状态,  $W_z, W_r, W$  都为权重参数,  $\sigma(\cdot)$  为 sigmoid 激活函数,  $\cdot$  表示矩阵相乘,  $*$  为点积<sup>[16]</sup>。然后使用一个 Dense 层来连接 GRU 的输出,得到船舶设备未来一段时间的预测值:

$$O_t^R = \text{Dense}(m_t) \quad (9)$$

### 3.3 基于向量自回归的船舶设备故障预测

循环神经网络虽然能够捕获船舶多变量时间序列的时间尺度,但是输出的规模对输入的规模并不敏感。为了解决这一问题,本文引入向量自回归组件,使用输入注意力组件所得到的船舶多变量时间序列  $\tilde{X}$  作为输入。给定输入  $\tilde{X} \in R^{N \times T}$ ,其在  $t$  时刻的观测值可以表示如下:

$$\tilde{x}_t = \sum_{k=1}^d \mathbf{A}_k \tilde{x}_{t-k} + \epsilon_t \quad (10)$$

其中,  $t = d + 1, \dots, T$ ;  $\mathbf{A}_k \in R^{N \times N}$  表示向量自回归模型的系数矩阵,  $k = 1, 2, \dots, d$ ;  $\epsilon_t$  可视为高斯噪声<sup>[17]</sup>。采用最小二乘法求解最优的系数矩阵  $\tilde{\mathbf{A}}$ , 向量自回归模型的输出为:

$$O_t^V = \sum_{k=1}^d \tilde{\mathbf{A}}_k O_{t-k}^V + \epsilon_t \quad (11)$$

### 3.4 预测结果的处理

本文将循环神经网络与向量自回归模型的输出平均值作为船舶设备的最终预测结果:

$$\tilde{O}_t = (O_t^R + O_t^V) / 2 \quad (12)$$

$\tilde{O}_t$  代表本文提出的模型在  $t$  时刻的最终预测结果。本文的输出是船舶设备属性变量的未来窗口时间段内的预测值,然后判断预测值是否处于该设备属性的正常运行范围,若不处于正常范围,则提示预警信息,可以避免重大事故的发生。

本文使用均方误差 MSE 作为目标函数,即船舶设备某

属性的真实值与预测值差平方的期望:

$$MSE = \frac{1}{N} \sum_{n=1}^N (O_n - \tilde{O}_n)^2 \quad (13)$$

其中,  $O_n$  代表第  $n$  个样本的真实值,  $\tilde{O}_n$  代表第  $n$  个样本的预测值,  $N$  表示样本数目。MSE 值越小, 说明模型的预测效果越好。

## 4 实验与结果分析

### 4.1 样本构建

对船舶的故障预测实质上是预测多个振动信号的运动趋势, 预测在未来一段时间内是否处于正常范围。但由于船舶设备之间复杂的关联性, 一个设备属性的变化可能会引起其他设备属性发生改变, 因此不能将对船舶的故障预测当作多个独立的单变量时间序列预测。实验数据集为从某船舶历史数据库中选取监测船舶 7 种主要设备(如船舶主柴油机、辅柴油机、主辅锅炉、空压机、热油加热器等)运行状态的 28 个属性(如燃油进击压力、重燃油油温、冷却水的压力、冷却水的出口温度等)的历史数据, 并同时获取船舶正常运行时该 28 个属性的正常运行范围。每个船舶设备是否处于正常状态是由多个属性值决定的, 设备之间相互影响。共选取 200000 个数数据点, 每 10000 个数数据点表示不同的运行期间。首先船舶原始数据的输入维度为 (10000, 28), 第一维表示数据点的个数, 第二维表示属性的个数, 将原始数据归一化后转化为监督学习问题的输入, 窗口大小设置为 10, 则原始数据的维度变为 (199990, 560)。然后将第二维展开转为 (199990, 10, 28), 表示共有 199990 个样本, 每个样本表示 28 个属性在连续 10 个时间点的观测值。最后将 70% 的样本作为训练集, 30% 的样本作为测试集。

### 4.2 实验参数

为确定滑动窗口大小, 对滑动窗口大小  $T$  进行了 {5, 10, 15, 20} 的网格搜索(即  $T=5, 10, 15, 20$ ), 分别进行 20 次训练, 在训练集上  $T=10$  时达到最佳性能, 因此设  $T$  为 10 进行测试, 本文方法的训练损失如图 5 所示。由于船舶数据量巨大且船舶数据具有非稳态性, 本次实验使用 Adam 优化器作为优化方法。

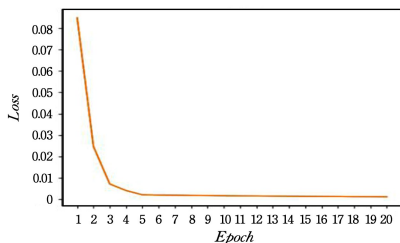


图 5 本文方法的训练损失

Fig. 5 Training loss of the proposed method

### 4.3 实验结果与分析

图 6 是使用本文方法预测船舶冷却水出口温度的结果与真实值的趋势图, 可以看到预测值与真实值的运动趋势完全一致。图 7 是船舶冷却水出口温度的预测结果与真实值的对

比, 可以看出两者趋势几乎一致。当 Time 为 5000 左右时振幅波动大, 是因为此时获取的船舶数据处于船舶冷却水出口温度异常状态。可以看出, 在故障状态下, 本文模型也能很好地预测出船舶属性的趋势。在 Python3.6 环境下, 将均方根误差(Root Mean Square Error, RMSE)作为评价指标, 使用本文方法对船舶冷却水出口温度进行预测。由图 7 可知, 本文方法的  $RMSE=1.19$ , 预测值与真实值几乎重合, 可以说明本文方法适用于船舶属性未来状态的预测。

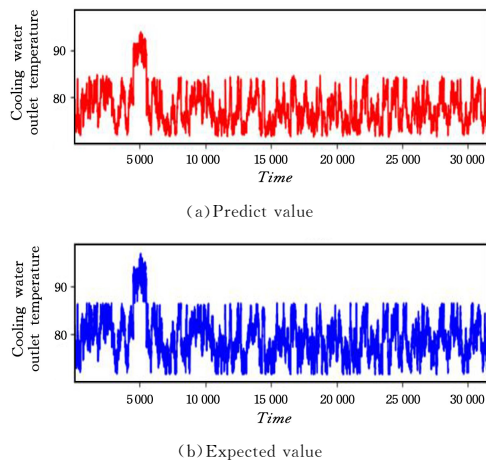


图 6 船舶冷却水出口温度预测值和真实值

Fig. 6 Predicted and real value of outlet temperature of ship cooling water

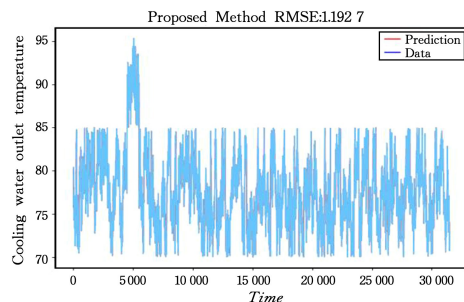


图 7 船舶冷却水出口温度预测值和真实值对比(电子版为彩色)

Fig. 7 Comparison between the predicted and real value of the outlet temperature of ship cooling water

为了验证本文方法对船舶故障预测的有效性和先进性, 本文将其与 4 种方法(门控循环神经网络(GRU)、向量自回归模型(VAR)、Seq2Seq 模型以及 LSTNet 模型)预测船舶冷却水出口温度的结果进行对比, 并用 RMSE 作为评估指标, 对比结果如表 1 所列。

表 1 5 种方法的 RMSE

Table 1 RMSE of five methods

Model	RMSE
VAR	3.22
GRU	1.855
Seq2Seq	1.81
LSTNet	1.46
Our Method	1.19

VAR 是 MTS 预测最著名的模型, 但由于船舶数据存在

大量非线性特征,且船舶历史数据库中包含大量故障数据,VAR对故障数据并不敏感,因此VAR并不能很好地预测船舶数据。GRU是时间序列处理领域著名的神经网络,解决了循环神经网络的长范围依赖和梯度消失的问题,但是由于它只捕获船舶时间序列的长期时间依赖模式,忽略了船舶属性之间的复杂相关性,当数据处于转折点时,不能很好地预测,因此GRU在船舶故障预测中的效果也不是很好。Seq2Seq

实质上是一个编码和解码的过程,编码过程中会将船舶信息进行压缩,由于船舶历史数据库中的信息量大,将Seq2Seq用于船舶状态预测会损失较多的信息,因此影响船舶故障预测的准确性。虽然基于LSTNet方法预测得到的船舶冷却水出口温度结果和真实值的RMSE较小,但LSTNet中的递归跳跃层的跳跃长度需要人工参与手动调整。这4种方法对船舶冷却水出口温度预测的实验结果如图8所示。

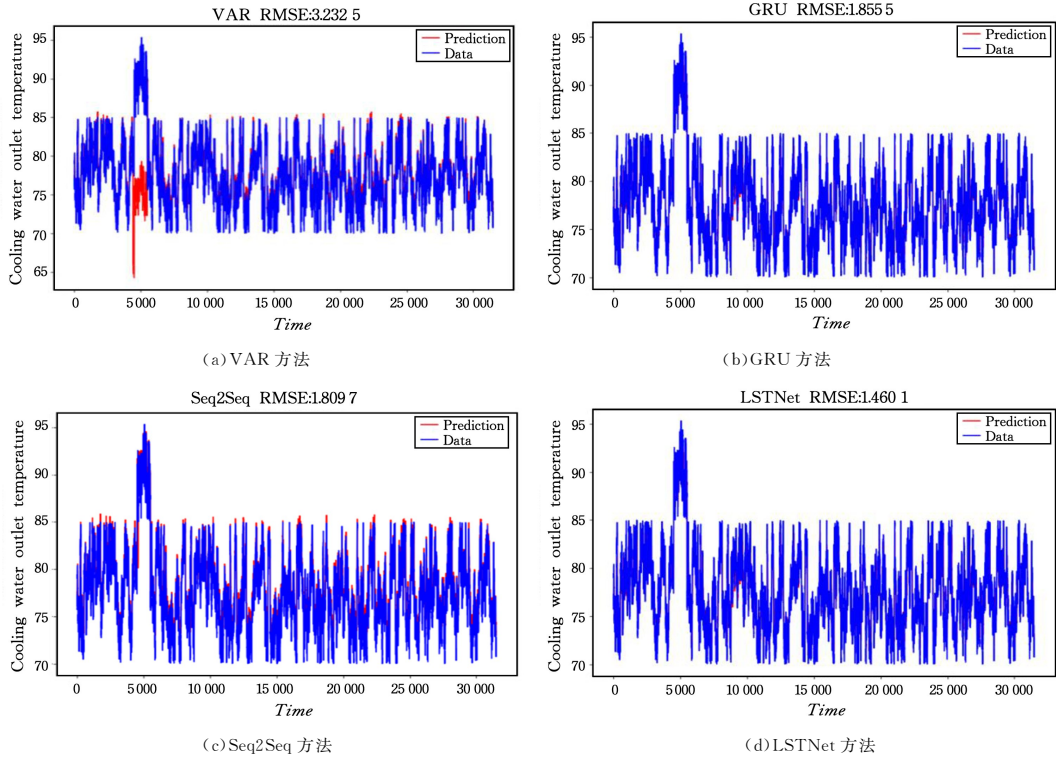


图8 基于4种方法预测船舶冷却水出口温度(电子版为彩色)

Fig. 8 Outlet temperature of ship cooling water predicted based on four methods

本文提出的方法简单且不需要人工参与,通过输入注意力自适应地选择相关的船舶属性,同时通过线性与非线性相结合的方法对船舶属性进行预测。经实验证明,本文方法对船舶故障预测的效果更好。

**结束语** 本文提出了一种输入注意力机制和循环神经网络以及向量自回归组件相结合的船舶设备故障预测方法。新引入的输入注意力机制可以自适应地选择与预测结果相关的船舶属性,并根据不同的相关性分配不同的权重。该方法将注意力机制的输出同时作用于循环神经网络和向量自回归模型,通过循环神经网络捕获时间序列的时间尺度,通过向量自回归模型捕获船舶历史数据的线性关系,从而使其不仅能自适应地选择最相关的输入特征,还能准确地预测船舶设备故障。经实验证实,本文方法在船舶主要设备的状态预测中具有良好的效果。该方法能够预测设备属性的状态趋势,可以满足船舶设备故障预测的需求。

就船舶设备故障预测而言,未来亦可以加入关联规则来预测故障,当某一设备即将或者已经出现故障时,利用关联规则可以预测相关的可能会出现故障,亦可以帮助预测船舶其他设备故障。

## 参考文献

- [1] KE Y, SONG E Z, YAO C, et al. Review of Marine diesel engine fault prediction and health management technology [J]. Journal of Harbin Engineering University, 2020, 41(1): 125-131.
- [2] JI P, LIU W T, ZHANG Y, et al. Research on PID Controller of Ship Power Station Based on RBF Neural Network [J]. Journal of Chongqing Institute of Technology, 2020, 32(2): 203-209.
- [3] MIAO J G, WANG J Y, ZHANG H, et al. Research progress of Fault diagnosis technology of unmanned aerial vehicle [J]. Journal of Instrumentation, 2020, 07: 1-15.
- [4] MALLIKA I L, RATNAM V, RAMAN S, et al. Machine learning algorithm to forecast ionospheric time delays using Global Navigation satellite system observations [J]. Acta Astronautica, 2020, 173: 221-231.
- [5] WANG C A. Taxi Demand Forecast Based on SVR Space-time grid Model [J]. Electronic World, 2020(3): 51-52.
- [6] ZHANG C B, SUN Y M. Prediction of tool wear in combination of feature engineering and deep feed forward Network [J]. Mechanical Design and Manufacturing, 2020(6): 190-193.

- [7] BAI S, KOLTER J Z, KOLTUN V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling[EB/OL]. <https://arxiv.org/abs/1803.01271>, 2018.
- [8] TAYLOR S J, LETHAM B. 2017: Forecasting at scale [EB/OL]. <https://doi.org/10.7287/peerj.preprints.3190v2>.
- [9] LAIG K, CHANG W C, YANG Y M, et al. Modeling Long- and Short-Term Temporal Patterns with Deep Neural Networks [C]// The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR'18). 2018:95-104.
- [10] SHIH S Y, SUN F K, LEE H Y. Temporal pattern attention for multivariate time series forecasting [J]. Machine Learning, 2019, 108(8/9):1421-1441.
- [11] YI L R, WANG S Y, YIN L L, et al. Prediction of industrial Sensor timing Data based on multivariable LSTM [J]. Intelligent Computer and Applications, 2018, 8(5):13-16.
- [12] CHEN Y P, YU L, CHEN H. Traffic Anomaly Detection Based on Wavelet Neural and ARMA Model in Big Data Environment [J]. Journal of Chongqing Institute of Technology, 2020, 33(10):149-154.
- [13] GHOSH S, DAS N, DAS I, et al. Understanding Deep Learning Techniques for Image Segmentation[J]. ACM Computing Surveys, 2019, 52(4):1-35.
- [14] LAI G, CHANG W C, YANG Y, et al. Modeling long-and short-term temporal patterns with deep neural networks [C]// 41st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2018). Ann Arbor, MI, United States 2018:95-104.
- [15] WANG J J, YAN J X, LI C, et al. Deep heterogeneous GRU model for predictive analytics in smart manufacturing: Application to tool wear prediction [J]. Computers in Industry, 2019, 111:1-14.
- [16] SUN Y S, JIANG Q, HU J, et al. Generation model of pedestrian trajectory prediction based on attention mechanism [J]. Computer Applications, 2019, 39(3):668-674.
- [17] SARA M, DAVID L. Life-Cycle Modeling of Structural Defects via Computational Geometry and Time-Series Forecasting [J]. Sensors (Basel, Switzerland), 2019, 19(20):4571-4589.



**ZENG You-yu**, born in 1996, postgraduate. Her main research interests include fault diagnosis and prediction, knowledge engineering.



**XIE Qiang**, born in 1972, associate professor, master supervisor, is a member of China Computer Federation. His main research interests include knowledge engineering and data mining.