

基于长短期记忆网络时序数据趋势预测及应用^{*}

杨 柯, 范世东

(武汉理工大学 能源与动力工程学院, 湖北 武汉 430063)

摘 要: 为了研究状态监测大数据对设备运行状态的估计和预测, 提出了一种人工经验与主成分分析相结合的长短期记忆网络方法 (AEPCA-LSTM), 利用运行过程中的监测时序数据对设备运行趋势进行预测。通过基于人工经验的主要成分分析方法 (AEPCA) 从状态监测系统中提取与目标变量最相关的状态变量作为输入; 利用长短期记忆网络 (LSTM) 对目标变量趋势变化进行预测, 并考虑运行过程中新数据样本的持续产生, 对模型进行定期更新, 以提高模型的动态适应性。将所提出的方法应用于船舶副机系统的涡轮增压器转速预测中, 结果表明, 该方法相对于传统的 PCA-LSTM 和 LSTM 具有更小的预测平均误差 0.18037, 展现了其在时序数据趋势预测的优势。

关键词: 主成分分析; 长短期记忆网络; 人工经验; 时序数据; 趋势预测

中图分类号: U672 文献标识码: A 文章编号: 1001-4055 (2021) 03-0675-08

DOI: 10.13675/j.cnki.tjjs.200394

Long Short-Term Memory Network Based Method and Its Application in Time-Series Data Trend Prediction

YANG Ke, FAN Shi-dong

(School of Energy and Power Engineering, Wuhan University of Technology, Wuhan 430063, China)

Abstract: In order to study the estimation and prediction of equipment operation state by condition monitoring big data, a novel method of Long-short Memory Network integrating Principal Component Analysis is proposed based on Artificial Experience (AEPCA-LSTM), which uses the monitoring time series data during operation to predict the equipment health trends. Firstly, the Principal Component Analysis method based on Artificial Experience (AEPCA) is used to extract the state parameters most relevant to target variable from the state monitoring system as input. Secondly, the Long short-term Memory Network (LSTM) is used to predict the trend changes of the target variable considering the continuous generation of new data samples during operation, the model is regularly updated to improve the dynamic adaptability of the model. Finally, the proposed method is applied to the turbocharger speed prediction of marine auxiliary engine system. The results show that the method has a lower prediction loss of 0.18037 compared with PCA-LSTM and LSTM, which indicates its advantages in the prediction of trend in time series data.

Key words: Principal component analysis; Long short-term memory network; Artificial experience; Time series data; Trend prediction

^{*} 收稿日期: 2020-05-31; 修订日期: 2020-08-06。

基金项目: 国家自然科学基金 (51679178)。

作者简介: 杨 柯, 博士生, 研究领域为运维与保障。E-mail: ake1231@163.com

通讯作者: 范世东, 博士, 教授, 研究领域为可靠性与维修性。E-mail: sdfan@whut.edu.cn

引用格式: 杨 柯, 范世东. 基于长短期记忆网络时序数据趋势预测及应用[J]. 推进技术, 2021, 42(3): 675-682.
(YANG Ke, FAN Shi-dong. Long Short-Term Memory Network Based Method and Its Application in Time-Series Data Trend Prediction[J]. *Journal of Propulsion Technology*, 2021, 42(3): 675-682.)

1 引 言

随着系统复杂性和运行不确定性的发展,基于长寿命周期设备状态监测数据中的可靠性信息进行趋势预测正在成为 PHM (Prognosis health management) 领域的新挑战^[1]。相对于传统基于数据统计的安全运行理论,高速率和大容量的时序监测数据需要更先进有效的处理方法与工具^[2]。深度学习理论作为机器学习领域的优势研究成果,通过模拟大脑由堆叠构建多层非线性网络结构,对高维复杂输入数据进行特征提取、学习和挖掘。由于强大的建模和表征能力,深度学习在图像和声音识别领域得到了广泛应用,并逐步被应用于设备健康预测管理领域的大数据处理。

Yan 等^[3]针对工业系统大数据的特点,构建了深层 DAE (Denoising auto-encoder) 模型,对工业过程信号进行处理,得出变化趋势,并结合回归分析的方法对输出结果进行融合,实现了设备状态感知和剩余寿命的预测。雷亚国等^[4]通过多个 DAE 堆叠构成深度神经网络的隐层结构,以机械设备的频域信号为输入对 (Deep neural network, DNN) 进行训练,实现了自适应的特征提取和健康状态智能诊断。但是,深度神经网络的参数规模及小批量训练导致了过度拟合和局部最优解问题的出现。于是产生了由多个限制玻尔兹曼机堆叠分类层或回归层改进的深度置信网络 (Deep belief network, DBN), 以减轻设备剩余寿命预测中的这些问题^[5-6]。此外,卷积神经网络 (Convolutional neural network, CNN) 由于其稀疏交互、参数共享和空间池特性,也是有效的预测模型。Babu 等^[7]将深度 CNN 与 MLP (Multilayer perceptron) 相结合,利用两个卷积层和两个池层提取原始信号的特征,实现了健康状态和剩余寿命的预测。Li 等^[8]通过时间窗口获取样本构建了多变量深度 CNN 网络结构,以估计设备的状态和剩余寿命。然而,由于监测数据的时序性及输入变量的相关性,循环神经网络 (Recurrent neural network, RNN) 在设备运行状态预测和寿命管理方面表现出了更强的优势^[9]。Heimes^[10]使用 BPTT (Back-propagation through time) 训练算法验证了 RNN 的可预测性。Malhi 等^[11]使用竞争性学习方法对输入数据进行聚类,并将处理的数据输入到 RNN 结构中,以加快网络训练并提高长期预测性能。但是,由于记忆流在循环层中无法控制,RNN 会遭受“内存衰减”,在处理长期监测序列时可能会导致大量预测偏差。因此,长短期记忆网

络 (Long short-term memory, LSTM) 通过添加门控单元在长时序数据中选择性记忆重要信息,解决了长序列训练过程中的梯度消失和梯度爆炸问题。Wu 等^[12]应用 LSTM 预测工程设备的使用寿命,并通过 NASA 航空发动机数据库进行方案对比证明了 LSTM 的预测效果明显优于传统 RNN。Miao 等^[13]设计并建立了双任务 LSTM 结构,用于联合学习航空发动机的退化评估和剩余使用寿命,从而提高了操作的可靠性和安全性。Elsheikh 等^[14]提出一种适用于 RUL (Remaining useful life) 估计问题的新目标函数以及一种用于训练 LSTM 网络的新目标生成方法,仅需对系统的实际退化进行较少的假设便可得到较准确的预测结果。

然而,LSTM 对大规模监测数据的高维特性缺乏有效处理,因此常与其它方法结合使用。主成分分析法 (Principal components analysis, PCA) 是使用最广泛的数据降维算法之一,它通过重建 k 维特征将高维故障特征转换为低维特征输入^[15]。Wang 等^[16]采用 PCA 对复杂的化工操作变量进行选择并提取深层特征,为新型深残网络提供了最佳输入。Art-Sahalia 等^[17]通过对高维数据的主成分分析,发现了高频数据于低频数据之间的结构有强烈的一致性。Wang 等^[18]考虑了工业过程复杂多变量的特点,使用主成分分析法压缩数据的同时,使数据整体信息损失最小。由于 PCA 是一种数据驱动的方法,仅从数据角度出发,忽略了工程实际中存在关注一些变量影响的同时而忽视某些经验变量的现象。

本文提出一种结合人工经验 PCA 的 LSTM 时序数据处理方法,在传统 PCA 降维的基础上,增加人工经验对特征选择的干预,保证了变量选择的科学性。然后对筛选后的参数变量使用 LSTM 方法实现时序数据预测建模。同时,通过定时更新扩展数据样本并调整模型参数,保证了模型对运行环境的动态适应性和有效性。为了验证所提出的方法有效性,将所提出的方法应用于船舶运行过程,对其状态监测数据进行建模和趋势预测,并通过与其它算法的对比验证了所提出方法的有效性。针对长寿命周期设备运行过程中的大容量监测时序数据处理及趋势预测问题,本文提出基于人工经验的主成分分析与长短期记忆网络结合 (AEPCA-LSTM) 的方法,并以船舶副机系统的涡轮增压器为对象,验证所提出方法的有效性。通过与其它方法的对比实验,体现了所提出方法的优势并指明了未来的研究方向。

2 基于 AEPCA-LSTM 的时序数据趋势预测方法

2.1 长短期记忆网络(LSTM)

LSTM 作为循环神经网络的一种特殊形式,首先由 Hochreiter 等在 1997 年提出^[19],如今已广泛应用于时间序列数据的预测。经典 LSTM 是在原始 RNN 隐藏层中增加一个传输的单元状态(Cell state),并由三个门控单元(Gate)控制,即输入门 I_t 、遗忘门 F_t 和输出门 O_t ,结构如图 1 所示。

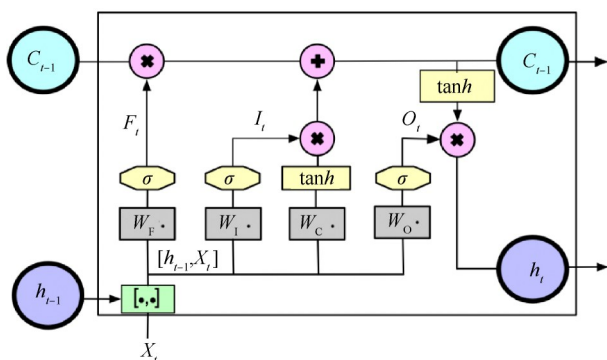


Fig. 1 Structure of classical LSTM

遗忘门 F_t :用于控制信息的遗忘(保留)程度,即上一时刻的单元状态有多少保存到当前时刻的单元状态。

$$F_t = \sigma(W_F \cdot [h_{t-1}, X_t]) + b_F \quad (1)$$

输入门 I_t :用于控制当前时刻的输入信息有多少保存到单元状态。

$$I_t = \sigma(W_I \cdot [h_{t-1}, X_t]) + b_I \quad (2)$$

因此,当前时刻的单元状态更新

$$C_t = F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \quad (3)$$

其中

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, X_t]) + b_C \quad (4)$$

式(4)表示当前时刻的输入产生新信息添加到单元状态中。

输出门 O_t 用于控制当前时刻的输入信息有多少输出到 LSTM 的当前输出,有

$$O_t = \sigma(W_O \cdot [h_{t-1}, X_t]) + b_O \quad (5)$$

最终,LSTM 的输出值为

$$h_t = O_t \odot \tanh(C_t) \quad (6)$$

式中 σ 是 Sigmoid 函数,经处理后为 0~1 的值;0 表示全部不通过;1 表示全部通过; $\odot$ 为 Hadamard 积; W 和 b 为对应的权重矩阵和偏置项。

因此,基于隐藏状态 h_t 和 C_t ,分别定义梯度 $\delta_h^{(t)}$ 和

$\delta_c^{(t)}$,从 $t = \tau$ 开始反向逐层计算误差。

$$\delta_h^{(t)} = \frac{\partial L}{\partial h^{(t)}} \quad (7)$$

$$\delta_c^{(t)} = \frac{\partial L}{\partial c^{(t)}} \quad (8)$$

式中 L 是损失函数。

$$L_{(t)} = \begin{cases} l_{(t)} + L_{(t+1)} & \text{if } t < \tau \\ l_{(t)} & \text{if } t = \tau \end{cases} \quad (9)$$

式中 $l_{(t)}$ 表示 t 时刻的损失, $L_{(t+1)}$ 表示 t 时刻之后的损失。且在最后序列索引位置 τ 的 $\delta_h^{(\tau)}$ 和 $\delta_c^{(\tau)}$ 分别为

$$\delta_h^{(\tau)} = \left(\frac{\partial O_{(\tau)}}{\partial h_{(\tau)}} \right)^T \frac{\partial L_{(\tau)}}{\partial O_{(\tau)}} = V^T (\hat{y}^{(\tau)} - y^{(\tau)}) \quad (10)$$

$$\delta_c^{(\tau)} = \left(\frac{\partial h_{(\tau)}}{\partial C_{(\tau)}} \right)^T \frac{\partial L_{(\tau)}}{\partial h_{(\tau)}} = \delta_h^{(\tau)} \odot O_{(\tau)} \odot (1 - \tanh^2(C_{(\tau)})) \quad (11)$$

因此,对于任一时刻 t ,其梯度 $\delta_h^{(t)}$ 一般表达式由本层 t 时刻的输出梯度误差和大于 t 时刻的误差两部分构成,即

$$\delta_h^{(t)} = \frac{\partial L}{\partial h^{(t)}} = \frac{\partial l_{(t)}}{\partial h_{(t)}} + \left(\frac{\partial h_{(t+1)}}{\partial h_{(t)}} \right)^T \frac{\partial L_{(t+1)}}{\partial h_{(t+1)}} = V^T (\hat{y}^{(t)} - y^{(t)}) + \left(\frac{\partial h_{(t+1)}}{\partial h_{(t)}} \right)^T \delta_h^{(t+1)} \quad (12)$$

而 $\delta_c^{(t)}$ 的反向梯度误差则由前一层 $\delta_c^{(t+1)}$ 和本层 $\delta_h^{(t)}$ 两部分构成

$$\delta_c^{(t)} = \left(\frac{\partial C_{(t+1)}}{\partial C_{(t)}} \right)^T \frac{\partial L}{\partial C_{(t+1)}} + \left(\frac{\partial h_{(t+1)}}{\partial C_{(t)}} \right)^T \frac{\partial L}{\partial h_{(t+1)}} = \left(\frac{\partial C_{(t+1)}}{\partial C_{(t)}} \right)^T \delta_c^{(t+1)} + \delta_h^{(t+1)} \odot O_{(t)} \odot (1 - \tanh^2(C_{(t)})) = \delta_c^{(t+1)} \odot F_{(t+1)} + \delta_h^{(t+1)} \odot O_{(t)} \odot (1 - \tanh^2(C_{(t)})) \quad (13)$$

综上,基于 $\delta_h^{(t)}$ 和 $\delta_c^{(t)}$,通过反向梯度下降便可完成所有参数的迭代更新。

2.2 基于人工经验的主成分分析方法(AEPCA)

主成分分析法(Principal components analysis, PCA)是一种通过正交变换对一组可能存在相关性的多元变量进行线性投影而获得低维主成分的主要数据处理方法。因此,对于设备状态监测系统的复杂、关联性变量数据集,可以采用 PCA 消除噪声并简化计算。但由于 PCA 是一种纯粹的数据驱动方法,无需任何先验知识,这可能会忽略在工程实践中一些未能反映在数据中的重要知识。因此本章节提出基于人工经验的主成分分析方法(AEPCA),在传统 PCA 的基础上增加人工经验(Artificial experience)对

变量的特征选择,在排除无关变量减轻计算负担的同时,重新考虑被忽略的相关变量。

传统 PCA 方法主要通过最近重构性和最大可分性获得所有样本的超平面表示,基于人工经验的 PCA 也满足这两个特征。假设实际系统中获得的传感参数构成数据样本为

$$X_i = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k-m}, \mathbf{x}_{k-m+1}, \dots, \mathbf{x}_k, \dots, \mathbf{x}_i] \quad (14)$$

首先,根据人工经验对所有参数变量进行初步筛选,有

$$X_i' = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k-m}, \mathbf{x}_k, \dots, \mathbf{x}_i] \quad (15)$$

其次,对样本进行中心化,即

$$\sum_i \mathbf{x}_i = 0 \quad (16)$$

然后,对样本 X_i' 投影变换,得到的新坐标系为

$$W = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{k-m}, \mathbf{w}_k, \dots, \mathbf{w}_n) \quad (17)$$

式中 $\mathbf{w}_i$ 是标准正交基向量。

$$\|\mathbf{w}_i\|_2 = 1 \quad (18)$$

$$\mathbf{w}_i^T \mathbf{w}_j = 0 \quad (i \neq j) \quad (19)$$

基于新坐标系舍弃部分坐标,将维度降低到 $n' < n$,则可得到样本 X_i' 在低维坐标系中的投影,即

$$Z_i = \{\mathbf{z}_{i1}; \mathbf{z}_{i2}; \dots; \mathbf{z}_{in'}\} \quad (20)$$

其中

$$\mathbf{z}_{ij} = \mathbf{w}_j^T \mathbf{x}_i \quad (21)$$

式(21)为 $\mathbf{x}_i$ 在低维坐标系下的第 j 维坐标。

因此,可基于低维坐标系下的 Z_i 重构 X_i' ,得到

$$\widehat{\mathbf{x}}_i = \sum_{j=1}^{n'} \mathbf{z}_{ij} \mathbf{w}_j \quad (22)$$

则,原样本 $\mathbf{x}_i$ 与基于投影重构的样本点 $\widehat{\mathbf{x}}_i$ 之间的距离为

$$\sum_{i=1}^k \left\| \sum_{j=1}^{n'} \mathbf{z}_{ij} \mathbf{w}_j - \mathbf{x}_i \right\|_2^2 = \sum_{i=1}^k \mathbf{z}_i^T \mathbf{z}_i - 2 \sum_{i=1}^k \mathbf{z}_i^T W^T \mathbf{x}_i + \quad (23)$$

$$\text{const} \propto -\text{tr} \left(W^T \left(\sum_{i=1}^k \mathbf{x}_i \mathbf{x}_i^T \right) W \right)$$

考虑到 $\mathbf{w}_i$ 是标准正交基,且 $\sum_i \mathbf{x}_i \mathbf{x}_i^T$ 是协方差矩阵,因此将距离目标函数最小化,有

$$\begin{cases} \text{Min } W - \text{tr}(W^T X' X'^T W) \\ \text{s.t. } W^T W = I \end{cases} \quad (24)$$

根据拉格朗日乘子法可得

$$X' X'^T \mathbf{w}_i = \lambda_i \mathbf{w}_i \quad (25)$$

对协方差矩阵 $X' X'^T$ 进行特征值分解,获得特征值

$$\lambda = (\lambda_1, \lambda_2, \dots, \lambda_{k-m}, \lambda_k, \dots, \lambda_n) \quad (26)$$

根据设定重构阈值 μ 满足条件

$$\frac{\sum_{i=1}^{n'} \lambda_i}{\sum_{i=1}^n \lambda_i} \geq \mu \quad (27)$$

选取满足条件的最大 n' 进行特征值排序

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n'} \quad (28)$$

则,特征值对应的特征向量构成主成分的解为

$$W^* = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{k-m}, \mathbf{w}_k, \dots, \mathbf{w}_{n'}) \quad (29)$$

与特征向量相对应的变量,即与目标最相关的变量为

$$X_i'' = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k-m}, \mathbf{x}_k, \dots, \mathbf{x}_{n'}\} \quad (30)$$

由于 PCA 方法可能会对实践中一些重要的变量忽略,因此基于人工经验对相关变量进行增补,有

$$X_i^{*'} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k-m}, \mathbf{x}_k, \dots, \mathbf{x}_{n'}, \mathbf{x}_{n'+1}, \dots, \mathbf{x}_{n'+i}\} \quad (31)$$

2.3 基于 AEPCA-LSTM 的时序数据趋势预测方法

基于 AEPCA-LSTM 的时序数据趋势预测方法主要由两部分组成:AEPCA 部分和 LSTM 部分,具体流程如图 2 所示。

Step 1:使用人工经验来选择相关变量,以减少不相关变量的干扰。

Step 2:使用 PCA 方法,从数据角度分析与目标变量最相关的变量作为输入变量。

Step 3:考虑到某些重要变量可能会基于数据而被忽略,因此使用人工经验增补因 PCA 方法忽略的一些变量作为最终的变量输入。

Step 4:对输入样本数据进行归一化处理。

Step 5:将所有的变量输入 LSTM 网络结构。

Step 6:输入信息通过输入门并获得长期记忆选择,将需要的输入信息保存下来。

Step 7:通过遗忘门对当前输入及上一时刻的记忆选择性保留。

Step 8:更新后的单元状态通过输出门,输出最后的预测值。

Step 9:将预测值与实际值进行对比,误差进行反向传播。整个训练过程反复迭代进行直至误差在可接受范围内。

Step 10:考虑运行过程中新数据样本的持续产生,对模型进行定期更新。

假设模型的定时更新周期为 T ,数据采集频率为 f ,则模型更新时增加的数据样本数量为 $M_s = \frac{T}{f}$ 。因此,获得新样本数据集为 $X_{\text{new},s} = \{X_i^{*'}, X_{M_s}\}$ 。式中 $X_i^{*'}$

为先前的历史数据样本集, X_{M_s} 为更新周期内的新增数据样本。

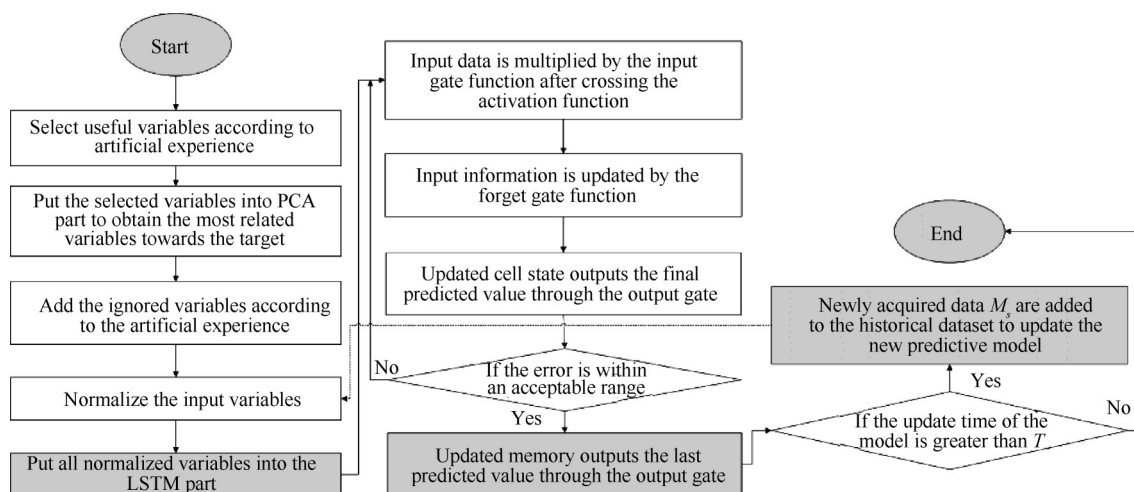


Fig. 2 Flowchart of AEPCA-LSTM

Step 11: 在达到定时更新间隔时,输入新样本数据集 $X_{new,s}$ 对模型进行更新。 $g(X_{new,s}) = \sigma(WX_{new,s} + b)$, $g(\cdot)$ 表示门控函数, W 表示权重矩阵, b 表示偏置项。

3 案例研究

船舶副机在运行过程中通过驱动发电机组为整船系统提供所需动力,它包括燃料系统、滑油系统、冷却系统和压缩空气系统等。其中,压缩空气系统通过发动机废气驱动涡轮并推动叶轮压缩更多的新鲜空气进入气缸(如图3所示),从而提高发动机的燃烧效率且减少废气排放。

作为增压系统的重要组成部分,涡轮增压器长时间在高速高温的恶劣环境中工作,它的性能不仅影响副机的连续可靠运行能力,而且会影响船舶系统的燃油经济性。其中,转速是可以反映空气压缩系统的效率和评估涡轮增压器工作状态的一个重要指标,且同时受到气缸排气和润滑系统的影响。因此,本文选择某船监测系统中与3号副机涡轮增压器

相关的17个运行参数,分别为:增压器转速 N_T , r/min; 柴油机进口滑油压力 p_{O1} , MPa; 柴油机的滑油进口温度 T_O , $^{\circ}\text{C}$; 滑油滤器前压力 p_{O2} , MPa; 增压器进口滑油压力 p_{O3} , MPa; 冷却器后增压空气压力 p_A , MPa; 冷却器后增压空气温度 T_A , $^{\circ}\text{C}$; 增压器前排气温度 T_{E9} , $^{\circ}\text{C}$; 增压器后排气温度 T_{E10} , $^{\circ}\text{C}$; 1~8号气缸出口排气温度分别为 $T_{E1} \sim T_{E8}$, $^{\circ}\text{C}$; 按照80min的采样间隔构造了一个3830采集点的样本数据集,对转速进行趋势预测,来验证所提出的时序数据建模方法。

3.1 特征选择

首先根据运维及机务管理人员的经验,选取与目标变量最相关的17个变量;其次,根据PCA方法对经初步筛选的17个变量进行降维处理,获得当重构阈值满足97%时的10个相关变量(通常情况下当变量对整体的贡献度超95%时便可认为所选择的变量可在很大程度上代表整体的特性),如图4所示。

$$X_i'' = \{P_{O1}, P_{O2}, P_{O3}, T_{E9}, T_{E10}, T_{E1}, T_{E2}, T_{E3}, T_{E4}\} \quad (32)$$

由于滑油的粘度和流动性与温度密切相关,并

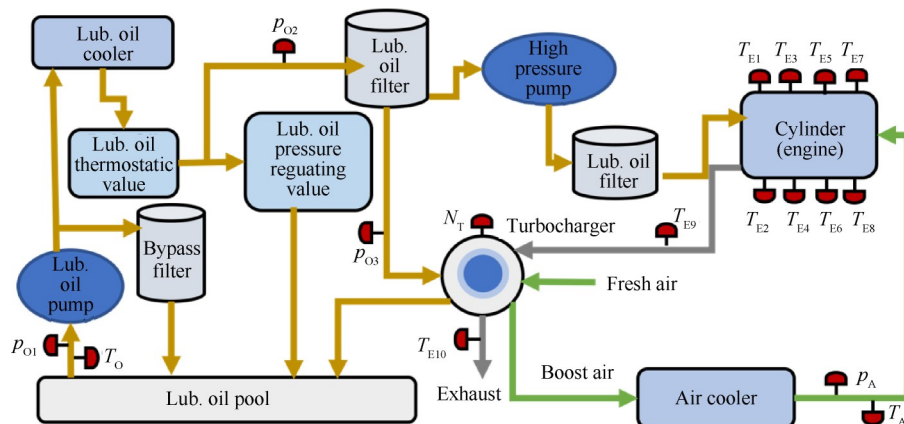


Fig. 3 Diagram of parameters related to operational turbocharger

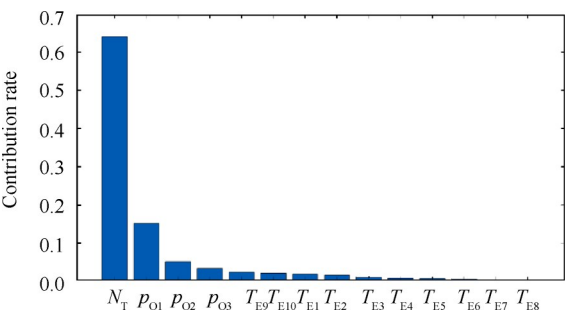


Fig. 4 Contribution rate of variables

可能导致气缸衬套、活塞等其它摩擦副产生异常磨损,进而增加系统积碳并影响进排气效率。因此,根据 2.2 节中基于人工经验的主成分分析方法增加变量 T_0 (柴油机的滑油进口温度),即得到最终的网路输入特征变量为

$$X_i^{*'} = \{p_{01}, p_{02}, p_{03}, T_{E9}, T_{E10}, T_{E1}, T_{E2}, T_{E3}, T_{E4}, T_{E5}, T_{E6}, T_{E7}, T_{E8}, T_0\} \quad (33)$$

3.2 结果分析

基于 3.1 特征选择的结果对样本集进行标准化处理并输入 LSTM 网络。按照遍历法分别抽取不同比例的样本对模型进行迭代训练,并最终选择三组具有代表性的 50%, 70%, 90% 比例的样本测试结果和误差进行分析,如图 5 和表 1 所示。

结合图 5 与表 1 可知,在选择不同数量的历史数据进行训练时,会得到不同的结果。当训练集样本量分别为 50%, 70%, 90% 时,对应的输出误差分别为 0.2526, 0.0509, -0.2509。当训练集样本量为 70% 时,有最小的输出误差,因抽取训练集样本量太少会产

生欠拟合现象,而抽取训练集样本量太多则会产生过拟合现象,这都会导致误差的增大。

采用不同比例的样本量进行模型训练时,首先把全部样本按对应比例分为训练集与测试集,并在训练集中按 7:3 划分数据集,分别作为训练模型时的训练样本与测试样本,进而实现交叉验证。如图 5 所示,测试结果呈现的是使用连续的且未参与过训练的测试集。因此,模型训练及测试有效,且存在可推广性。

Table 1 Errors analysis of the different models

Item	50% data for training	70% data for training	90% data for training
Average error	0.2526	0.0509	-0.2509

3.3 模型更新

考虑到运行环境的变化,对上述模型进行动态定期更新以尽量减少预测误差。在模型更新计算时,沿用 3.2 节中的 7:3 数据量比对模型进行训练学习。设每 100 个采样点为更新周期,得到模型更新后的预测结果和误差分析如图 6 所示。

模型以 100 个采样点为间隔周期进行动态更新,因此分别获得数据规模为 660, 760 和 860 的三次结果对比,即原始模型、首次更新模型和第二次更新模型,如表 2 所示。

基于原始模型,对数据规模为 660, 760 和 860 样本的预测误差分别为 0.04917, 0.10403 和 0.15109,可

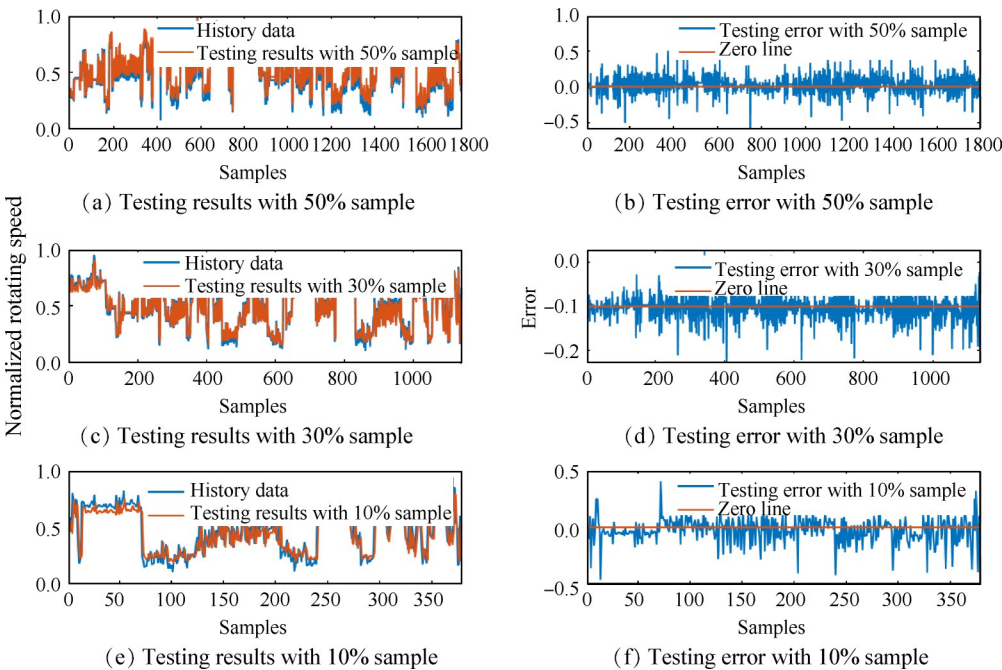


Fig. 5 Testing results of different models

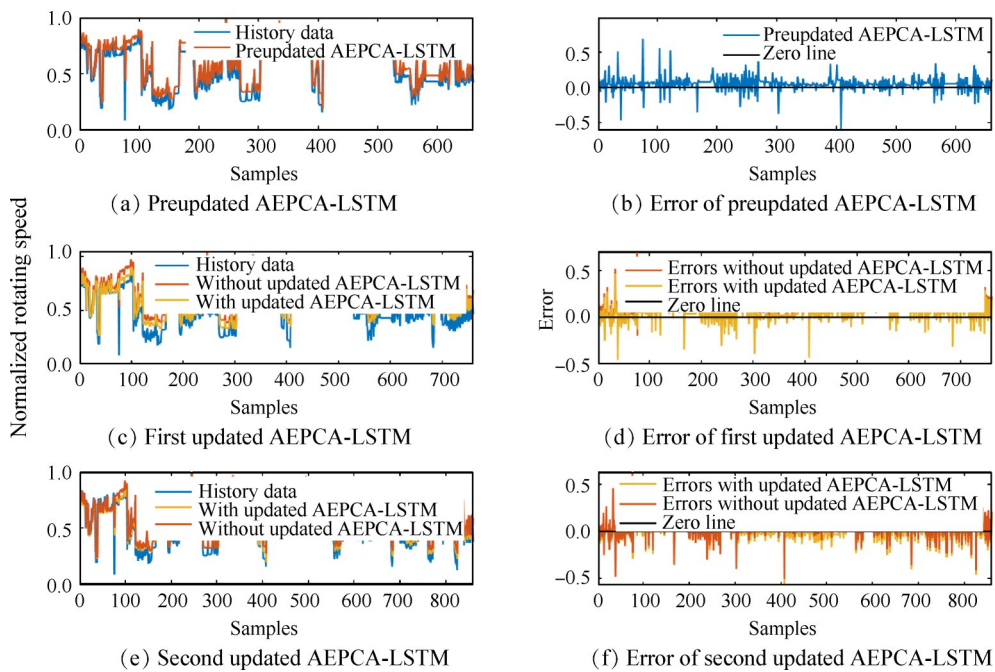


Fig. 6 Results analysis of the updated mode

Table 2 Errors analysis of the updated models

Model	Number of the data	Errors without update	Errors with update
Pre-updated	660	0.04917	—
1st-updated	760	0.10403	0.06179
2nd-updated	860	0.15109	0.05770

见随着新数据到来的样本增大时,原始模型的预测精度逐渐下降。同时,对于同等的数据规模,第一次更新后的模型预测误差由 0.10403 降为 0.06179,第二次更新后的模型预测误差由 0.15109 降为 0.05770,第二次更新相对第一次更新的预测误差则由 0.06179 降为 0.05770。由此可见,随着数据样本的增大,模型进行更新后的预测可靠性保持稳定。

3.4 方法对比

基于更新后的模型,将 AEPCA-LSTM 与 PCA-LSTM 和 LSTM 三种预测方法进行比较,得到结果分析如表 3 所示。

Table 3 Results analysis of the comparative study

Method	AEPCA-LSTM	PCA-LSTM	LSTM
Average error	0.05263	0.18037	0.23659

表中,AEPCA-LSTM 的预测精度最好,平均误差为 0.05263;LSTM 的预测精度相对较低,平均误差为 0.23659。与传统 LSTM 方法相比,PCA-LSTM 引入 PCA 方法,选取与目标变量最相关的变量组合对目标变量进行预测分析,因此可得较高的精度,而 AEPCA-LSTM 方法进一步融入人为因素的考量,增强了变量

选取的科学性与可靠性,进而获得更优的预测结果。

4 结 论

通过本文研究,得到以下结论:

(1) 针对设备监测系统的状态估计和趋势预测问题,提出了一种基于人工经验主成分分析的长短期记忆网络方法(AEPCA-LSTM),结合专家知识对与目标变量相关的变量进行筛选、降维和增补处理,并应用于船舶副机系统增压器转速的预测。通过不同样本数量的遍历选择和交叉验证,获得最优训练误差为 0.0509。

(2) 引入模型定时更新,使得所提出的建模方法能够始终适应变化的工作环境与工作状态。同时,将传统 LSTM 算法与 PCA-LSTM 算法处理得到的预测结果和误差与本文所提出的 AEPCA-LSTM 算法进行对比,所得预测平均误差为 0.23659,0.18037 和 0.05263,表明了所提方法在时序数据趋势预测中的优势。

(3) 本文所提出的 AEPCA-LSTM 方法能够处理一般情况下的时序数据建模问题,具有较强的泛化能力,后期也将进一步考虑多元变量之间的耦合关系及变量间的时延特性对系统建模的影响。

致 谢:感谢国家自然科学基金的资助。

参考文献

[1] Pecht M. Prognostics and Health Management of Electronics[M]. Hoboken: John Wiley & Sons, 2008.

- [2] 胡昌华, 施 权, 司小胜, 等. 数据驱动的使用寿命预测和健康管理技术研究进展[J]. 信息与控制, 2017, 46(1): 72-82.
- [3] Yan H, Wan J, Zhang C, et al. Industrial Big Data Analytics for Prediction of Remaining Useful Life Based on Deep Learning [J]. *IEEE Access*, 2018, 6: 17190-17197.
- [4] 雷亚国, 贾 峰, 周 昕, 等. 基于深度学习理论的机械装备大数据健康监测方法[J]. 机械工程学报, 2015, 51(21): 49-56.
- [5] Deutsch J, He D. Using Deep Learning Based Approaches for Bearing Remaining Useful Life Prediction [C]. *Denver: Annual Conference of the Prognostics and Health Management Society*, 2016.
- [6] Deutsch J, He D. Using Deep Learning-Based Approach to Predict Remaining Useful Life of Rotating Components [J]. *IEEE Transactions on Systems Man & Cybernetics Systems*, 2017, 48(1): 11-20.
- [7] Babu G, Zhao P, Li X. Database Systems for Advanced Applications[M]. *Berlin: Springer*, 2016.
- [8] Li X, Ding Q, Sun J. Remaining Useful Life Estimation in Prognostics Using Deep Convolution Neural Networks [J]. *Reliability Engineering and System Safety*, 2018, 172: 1-11.
- [9] Chow T, Fang Y. A Recurrent Neural-Network-Based Real-Time Learning Control Strategy Applying to Nonlinear Systems with Unknown Dynamics[J]. *IEEE Transactions on Industrial Electronics*, 1998, 45(1): 151-161.
- [10] Heimes F. Recurrent Neural Networks for Remaining Useful Life Estimation[C]. *Denver: International Conference on Prognostics and Health Management*, 2008.
- [11] Malhi A, Yan R, Gao R. Prognosis of Defect Propagation Based on Recurrent Neural Networks [J]. *IEEE Transactions on Instrumentation & Measurement*, 2011, 60(3): 703-711.
- [12] Wu Y, Yuan M, Dong S, et al. Remaining Useful Life Estimation of Engineered Systems Using Vanilla LSTM Neural Networks[J]. *Neurocomputing*, 2017, 275: 167-179.
- [13] Miao H, Li B, Sun C, et al. Joint Learning of Degradation Assessment and RUL Prediction for Aeroengines via Dual-Task Deep LSTM Networks[J]. *IEEE Transactions on Industrial Informatics*, 2019, 15(9): 5023-5032.
- [14] Elsheikh A, Yacout S, Ouali M S. Bidirectional Handshaking LSTM for Remaining Useful Life Prediction [J]. *Neurocomputing*, 2019, 323: 148-156.
- [15] Zhou F, Park J H, Wen C, et al. Average Accumulative Based Time Variant Model for Early Diagnosis and Prognosis of Slowly Varying Faults[J]. *Sensors*, 2018, 18(6).
- [16] Wang Y, Li H. Complex Chemical Process Operation Evaluations Using a Novel Analytic Hierarchy Process Model Integrating Deep Residual Network with Principal Component Analysis [J]. *Chemometrics and Intelligent Laboratory Systems*, 2019, 191: 118-128.
- [17] Aït-Sahalia Y, Xiu D. Principal Component Analysis of High-Frequency Data[J]. *Journal of the American Statistical Association*, 2019, 114(5): 287-303.
- [18] Wang Y, Yang K, Li H. Industrial Time-Series Modeling via Adapted Receptive Field Temporal Convolution Networks Integrating Regularly Updated Multi-Region Operations based on PCA[J]. *Chemical Engineering Science*, 2020, 228(11).
- [19] Hochreiter S, Schmidhuber J. Long Short-Term Memory [J]. *Neural Computation*, 1997, 9(8): 1735-1780.

(编辑:朱立影)