

DOI:10.19651/j.cnki.emt.2210778

# 基于层次散布熵的滚动轴承剩余寿命预测方法<sup>\*</sup>

吴 茜 张守京

(西安工程大学机电工程学院 西安 710600)

**摘要:** 针对滚动轴承寿命预测中提取的特征不准确以及预测精度低等问题,提出一种基于层次散布熵(HDE)和门控循环单元(GRU)的滚动轴承剩余寿命预测方法。首先将振动信号时间序列进行层次分析,计算各个节点的散布熵,将散布熵重构融合得到HDE;其次将相关性、单调性和鲁棒性组合形成综合指标,用来验证HDE的优越性;最后划分训练集和测试集,通过GRU网络进行寿命预测试验。结果表明,HDE的综合指标值最优,所提方法HDE-GRU的预测误差比RMS-GRU、DE-GRU和MDE-GRU分别低42.77%、39.57%和20.24%,且运行时间最短,预测精度更高,为滚动轴承健康管理提供了实际价值。

**关键词:** 层次散布熵;GRU网络;滚动轴承;寿命预测

中图分类号: TH165.3; TH17 文献标识码: A 国家标准学科分类代码: 460.99

## Remaining life prediction of rolling bearings based on hierarchical dispersion entropy

Wu Rui Zhang Shoujing

(College of Mechanical and Electrical Engineering, Xi'an Polytechnic University, Xi'an 710600, China)

**Abstract:** Aiming at the problems of inaccurate features extracted and low prediction accuracy in the life prediction of rolling bearings, a method of remaining life prediction of rolling bearings based on hierarchical dispersion entropy (HDE) and gated recurrent unit (GRU) was proposed. Firstly, the time series of vibration signals were analyzed by hierarchical analysis, and the dispersion entropy of each node was calculated, and the dispersion entropy was reconstructed and fused to obtain the HDE. Secondly, correlation, monotonicity and robustness are combined to form a comprehensive index to verify the superiority of HDE. Finally, the training set and test set of HDE are divided, and the life prediction test is carried out by GRU network. The results show that the comprehensive index value of HDE is the best, and the prediction error of the proposed method HDE-GRU is 42.77%, 39.57% and 20.24% lower than that of RMS-GRU, DE-GRU and MDE-GRU, respectively. It has the shortest running time and higher prediction accuracy, which provides practical value for rolling bearing health management.

**Keywords:** hierarchical dispersion entropy; GRU network; rolling bearing; life prediction

## 0 引言

滚动轴承作为机械设备的核心零部件,在运行过程中易发生故障,从而严重影响机械设备的使用寿命和工作效率<sup>[1]</sup>。若能预知滚动轴承的剩余使用寿命(remaining useful life, RUL),便能及时对其进行更换或维修,保障机械设备正常运行,同时降低生产成本。滚动轴承发生故障时,其振动信号的频带会产生变化,故振动信号中的有效信息能反映轴承的运行状态<sup>[2]</sup>。因此,从振动信号中提取对轴承全寿命周期运行状态更敏感的有效特征在健康管理方

面具有一定的现实意义。

许多学者对振动信号的特征提取进行了深入研究。王付广等<sup>[3]</sup>先对滚动轴承振动信号进行经验模态分解,得到多个能量特征,再对这些特征进行模糊熵分析,最后利用极限学习机完成RUL预测;徐洲常等<sup>[4]</sup>提取滚动轴承的时域和时频域特征,通过改进回归型支持向量机预测RUL;陈彦龙等<sup>[5]</sup>将经验模态分解的分量应用于滚动轴承故障诊断中;周建明等<sup>[6]</sup>提取时域特征和经验模态分解能量熵,从两类特征中筛选出最优特征,实现滚动轴承故障分类;杨小东等<sup>[7]</sup>利用卷积神经网络提取滚动轴承的时域和时频域退化

收稿日期:2022-07-22

\*基金项目:国家重点研发计划项目(2019YFB1707205)资助

特征,并且利用该网络进行 RUL 预测;王勉等<sup>[8]</sup>从振动信号中提取多尺度散布熵(mutil-scale dispersion entropy, MDE)并与时移概念结合,利用支持向量机进行滚动轴承故障诊断;Sharma 等<sup>[9]</sup>提取振动信号的多尺度波动散布熵(multi-scale fluctuation-based dispersion entropy, MFDE)应用于旋转设备的状态检测;杨潇谊等<sup>[10]</sup>计算信号的散布熵(dispersion entropy, DE)同时通过余弦欧氏距离筛选出有效特征作为退化指标,实现了滚动轴承性能退化评估。由此可见,有效特征在滚动轴承寿命预测、故障诊断和性能退化评估方面必不可少。此外,获取有效特征后需要将其输入到预测模型完成 RUL 预测。长短时记忆网络(long short-term memory, LSTM)因其运行速率快、鲁棒性好已被广泛应用于寿命预测这类时间序列问题中。王奉涛等<sup>[11]</sup>利用 LSTM 对轴承进行 RUL 预测;赵广谦等<sup>[12]</sup>将卷积神经网络和双向 LSTM 结合实现滚动轴承 RUL 的智能预测。谭峰等<sup>[13]</sup>将 LSTM 循环神经网络成功应用于数控机床热误差预测;Guo 等<sup>[14]</sup>利用卷积神经网络自适应提取得到敏感特征,再采用长短时记忆网络完成轴承寿命预测。

综上所述,传统特征提取方法仅在振动信号中提取单层次或多尺度特征,未考虑不同层次和不同频段的特征。DE 运行效率高,层次散布熵(hierarchical dispersion entropy, HDE)包含多层次的 DE,相较于其他特征更完整,已成功应用在滚动轴承的故障诊断中,但其在 RUL 预测方面的应用报道较少。门控循环单元(gated recurrent unit, GRU)作为 LSTM 的改进网络具有更强的鲁棒性和运算效率。故本文提取 HDE 特征利用 GRU 网络进行滚动轴承 RUL 预测。

## 1 方法理论

### 1.1 散布熵

DE<sup>[15]</sup>是一种可以衡量时间序列复杂性的非线性动力学评价指标,其值大小反映时间序列的状态,值越小,序列规律性越强,反之,序列规律性越弱。给定时间序列  $\{X = x(i), i=1,2,\dots,N\}$ , DE 的计算公式为:

$$DE(x, m, c, d) = - \sum_{n=1}^{c^m} p(\pi_{v_0 v_1 \dots v_{m-1}}) \ln(p(\pi_{v_0 v_1 \dots v_{m-1}})) \quad (1)$$

式中: $m$  为嵌入式维数,  $c$  为映射类别数,  $d$  为时间延迟常数,  $p(\pi_{v_0 v_1 \dots v_{m-1}})$  为散布模式  $\pi_{v_0 v_1 \dots v_{m-1}}$  的概率,  $v_0 v_1 \dots v_{m-1}$  为映射后的时间序列号。

利用尺度因子  $\tau$  将时间序列粗粒化为  $y(j)$ ,对该序列进行 DE 分析得到 MDE<sup>[16]</sup>。MDE 只包含时间序列的低频部分,忽略了高频频序,很难有效反映轴承振动信号的趋势。MDE 计算公式为:

$$MDE = \frac{1}{\tau} DE(y(j), m, c, d) \quad (2)$$

### 1.2 层次散布熵

HDE 是将 DE 与层次分析结合,同时分析时间序列的高频和低频部分,如图 1 所示,虚线框内为信号的低频部分,其余为高频部分。HDE 的计算过程为:

1)根据轴承的振动信号,定义  $k$  层算子  $Q_j$  为:

$$Q_j = \begin{bmatrix} \frac{1}{2} & (-\frac{1}{2})^j & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \frac{1}{2} & (-\frac{1}{2})^j & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & \frac{1}{2} & (-\frac{1}{2})^j \end{bmatrix}_{2^{n-1} \times 2^n} \quad (3)$$

式中: $j=1$  或  $0, 2^{n-1}$  为算子的长度,  $n$  为整数, 轴承振动信号的长度决定  $Q_j$  的形式。则定义  $Q_0$  为平均算子,  $Q_1$  为差分算子, 其表达式为:

$$Q_0(x) = \frac{x(2i) + x(2i+1)}{2}, i = 0, 1, 2, \dots, 2^{n-1} \quad (4)$$

$$Q_1(x) = \frac{x(2i) - x(2i+1)}{2}, i = 0, 1, 2, \dots, 2^{n-1} \quad (5)$$

式中: $Q_0(x)$  和  $Q_1(x)$  分别表示时间序列首层分解的低频和高频成分。

2)为获取各层的层次节点  $X_{k,e}$ ,需要构造一个  $n$  维向量  $[\gamma_1, \gamma_2, \dots, \gamma_n] \in \{0,1\}$  和整数  $e = \sum_{j=1}^k \gamma_j 2^{k-p}, p = 1, 2, \dots, k$ , 则第  $k$  层的第  $e$  个节点可以表示为:

$$X_{k,e} = Q_{\gamma_n} \cdot Q_{\gamma_{n-1}} \cdot \cdots \cdot Q_{\gamma_1}(x) \quad (6)$$

3)计算各节点的 DE 值,将得到的 DE 融合即为 HDE。运算公式为:

$$HDE = DE(x(i), k, m, c, d) \quad (7)$$

式中: $k$  为分解层数,  $m$  为嵌入式维数,  $c$  为映射类别数,  $d$  为时间延迟常数。

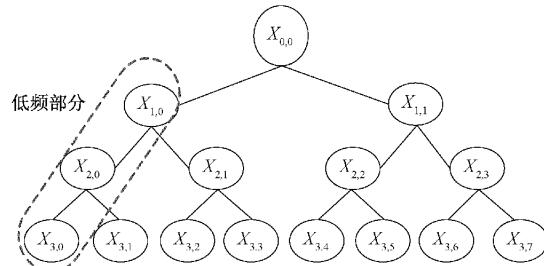


图 1 时间序列层次分析( $k=3$ )

### 1.3 门控循环单元

GRU 将 LSTM 的输入门和遗忘门融合为更新门,同时引入重置门。保留 LSTM 原本功能的基础上减少网络参数,提高训练速率且降低网络复杂度。GRU 的标准结构如图 2 所示。

重置门  $r_t$  动态控制前一时刻特征信息的遗忘程度,阈值越小,遗忘信息越多,更新门  $z_t$  的作用是决定前一时刻的特征信息对当前特征信息的重要性,阈值越大,当前输入对整体的影响程度越大。GRU 网络通过相关参数和偏置

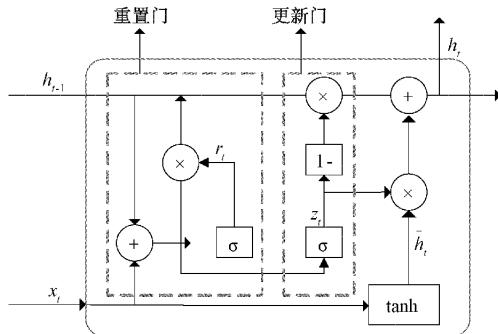


图 2 GRU 网络结构图

来调整激活函数  $\sigma$  的输入输出,从而决定门是否启动。具体计算如下:

$$\begin{cases} r_t = \sigma(W_r[h_{t-1}, x_t]) \\ z_t = \sigma(W_z[h_{t-1}, x_t]) \\ \tilde{h}_t = \tanh(W_h[r_t * h_{t-1}, x_t]) \\ h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \end{cases} \quad (8)$$

其中,  $x_t$  是当前时刻  $t$  的输入;  $h_t$  和  $h_{t-1}$  分别是当前时刻输出和上一时刻输出,  $\tilde{h}_t$  是通过当前输入和重置门叠加运算形成的候选输出集;  $W_r$ ,  $W_z$ ,  $W_h$  分别代表重置门、更新门和候选输出集的权重系数;  $\sigma = 1/(1 + e^{-x})$ ;  $[ ]$  代表向量间的连接;  $*$  代表矩阵间的乘积。

## 2 基于 HDE 的滚动轴承剩余寿命预测流程

首先将滚动轴承振动信号分层,再计算各节点的 DE 值融合为 HDE,得到包含低频和高频时段的可以全面反映滚动轴承运行特征的敏感特征集,将 HDE 输入到 GRU 网络实现轴承剩余寿命的高效精准预测。所提方法的具体流程如图 3 所示。

在计算层次散布熵的过程中,首先要设置相关参数的初始值。若分解层数  $k$  的取值过大,则会导致节点过多,计算效率低,若  $k$  取值过小,则时间序列频带过粗,易忽略重

要的低频或高频信息,故  $k$  的最优值一般为  $3^{[17]}$ 。文献[18]利用鸟群优化算法对层次多尺度散布熵的参数进行优化分析,故本文选取嵌入式维数  $m=2$ ,映射类别  $c=5$  以及时间延迟  $d=1$ 。

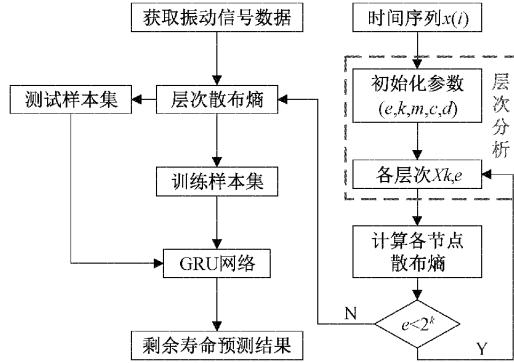


图 3 预测流程

此外,层次散布熵作为敏感特征集分为训练集和测试集,将训练样本输入到 GRU 网络中进行模型的学习训练。而将测试集输入到模型中进行寿命预测之前,需要选取预测点,以提高算法运算效率以及预测结果的精确度。

## 3 试验验证与分析

### 3.1 滚动轴承数据来源与分析

为验证 HDE 与 GRU 网络结合的 RUL 预测模型的有效性,采用辛辛那提大学(University of Cincinnati)的滚动轴承加速寿命试验数据,轴承型号为 Rex-ford ZA-2115,试验中设置轴承转速为 2 000 RPM,DAQCardTM-6062E 信号采集卡采集轴承振动信号,采样频率 20 000 Hz,采样点 20 480。包含 3 次试验,试验 1 中轴承失效结束时,3 号轴承出现内圈缺陷,4 号轴承出现内圈滚子元件缺陷;试验 2 失效结束时,轴承 1 发生外圈故障;试验 3 中,3 号轴承发生外圈失效。该轴承加速寿命试验平台以及试验轴承和对应传感器的布局如图 4 所示。

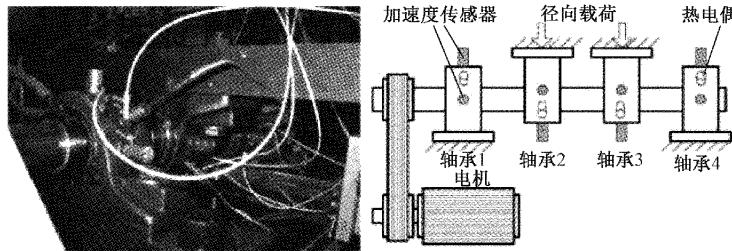


图 4 轴承加速度试验平台

采用试验 2 中 2~1 号轴承数据进行试验验证,将该数据集的全部 984 组样本归一化处理,绘制出轴承性能退化趋势如图 5 所示。前期振动信号相对平稳,随着运行时间增加、工作条件、运行环境等其他因素的影响,轴承振动加剧进入退化阶段,直到后期完全失效。计算退化量均值作

为轴承正常阶段和故障阶段的分界点,从性能退化趋势值与均值线 1 相交的部分可以看出第 544 组样本时退化曲线发生突变,即 544 号样本为早期故障点,故本文将以此点开始完成滚动轴承的 RUL 预测以提高运行效率。再计算故障阶段(第 544~984 组样本)的均值,作为故障中期和严

重故障分界点。第 703 组样本时退化趋势与均值线 2 相交,且后续样本在均值线 2 上下波动,故定义第 703 组样本为中期故障点。第 900 组样本之后的退化趋势急剧上升且均没有低于均值线 2,故定义第 900 组样本为严重故障点。

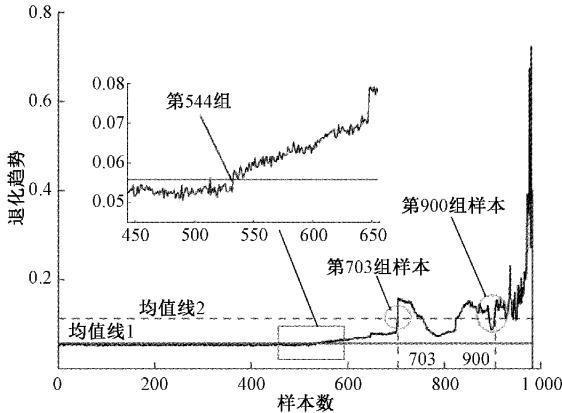


图 5 性能退化趋势

### 3.2 有效特征选择

为验证 HDE 对滚动轴承运行状态的敏感性,从滚动轴承振动信号中提取时域特征(均方根值(root mean square, RMS)、峭度、方差)、频域特征(频谱均值、频率中心、均方根频率)以及 DE 和 MDE,与 HDE 进行对比分析。上述时域和频域特征的计算公式如表 1 所示。

表 1 时域和频域计算公式

特征名称	计算公式
均方根值	$X_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$
峭度	$\beta = \frac{1}{N} \sum_{i=1}^N x_i^4$
方差	$\sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2$
频谱均值	$p_1 = \frac{\sum_{k=1}^K X(k)}{K}$
频率中心	$p_2 = \frac{\sum_{k=1}^K f_k X(k)}{\sum_{k=1}^K X(k)}$
均方根频率	$p_3 = \sqrt{\frac{\sum_{k=1}^K f_k^2 X(k)}{\sum_{k=1}^K X(k)}}$

有效的特征与轴承运行状态信息具有较高的相关性、良好的单调性以及对异常值较强的鲁棒性。对提取特征

计算其相关性、单调性和鲁棒性可以合理选择出最优特征,为滚动轴承剩余寿命预测奠定良好基础。计算过程如下。

1) 对提取的特征进行平滑处理,分解得到均值趋势和随机部分。分解公式为:

$$fea(t) = fea_T(t) + fea_R(t) \quad (9)$$

其中,  $fea(t)$  为  $t$  时刻的状态特征值;  $fea_T(t)$  趋势部分值;  $fea_R(t)$  为随机部分值。

2) 计算相关性  $Corr$ 、单调性  $Mon$  和鲁棒性  $Rob$ 。计算公式为:

$$Corr = \frac{\left| \sum_{t=1}^L (fea(t) - \bar{fea})(t - \bar{t}) \right|}{\sqrt{\sum_{t=1}^L (fea(t) - \bar{fea})^2 \sum_{t=1}^L (t - \bar{t})^2}} \quad (10)$$

$$Mon = \left| \frac{\sum_{t=1}^L diff(fea) - \sum_{t=1}^L -diff(fea)}{L-1} \right| \quad (11)$$

$$Rob = \frac{1}{L} \sum_{t=1}^L \exp\left(-\left|\frac{fea_T(t)}{fea(t)}\right|\right) \quad (12)$$

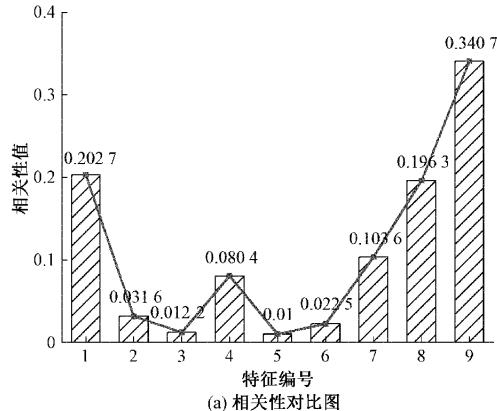
其中,  $L$  为滚动轴承全寿命周期振动信号数据的个数;  $\bar{fea}$  和  $\bar{t}$  分别表示样本均值和时间均值;  $diff(fea)$  是两个连续点之间的差值。

3) 加权线性组合相关性、单调性和鲁棒性。3 个指标的数值均在  $[0, 1]$  之间,且其性能均呈正相关,故将其融合可以综合 3 个指标的优点,同时提高特征选择的准确性。综合指标  $CMR$  的计算公式为:

$$CMR = w_1 Corr + w_2 Mon + w_3 Rob \quad (13)$$

根据文献[19]分析的 3 个指标对轴承退化性能的贡献程度,确定相关性权重  $w_1 = 0.3$ , 单调性权重  $w_2 = 0.5$ , 鲁棒性权重  $w_3 = 0.2$ 。

按表 1 的时域和频域特征顺序,为其编号为 1~6,同时 DE、MDE 和 HDE 的编号分别为 7、8、9。图 6(a)~(c) 是该 9 个特征的相关性、单调性和鲁棒性分布情况。在相关性对比中各特征差异较大,9 号特征 HDE 具有较高的相关性;在单调性和鲁棒性对比中各特征波动不明显,9 号特征 HDE 仍处于较高的优势。



(a) 相关性对比图

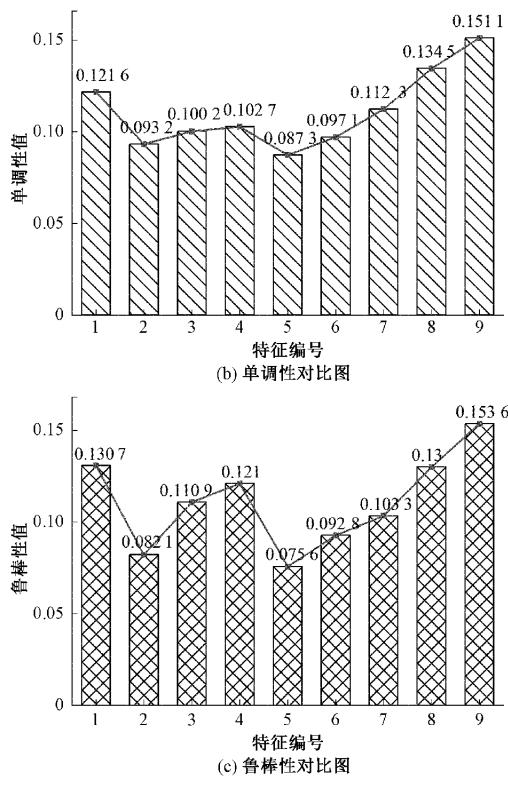


图 6 特征对比分析

综合考虑 3 个指标,计算综合指标 CMR 如表 2 所示。

为筛选最优特征,定义选择特征的综合指标阈值为 0.1,即 HDE 特征最优,其次为 MDE、RMS 和 DE。充分说明 HDE 特征相较于其他特征对轴承敏感性最强,可用 来预测滚动轴承剩余寿命。

### 3.3 基于 HDE 的滚动轴承剩余寿命预测

本文重点研究的是 HDE 特征在滚动轴承剩余寿命预测等时间序列问题上的有效性,为体现该特征的优势,结合前期通过综合指标筛选的较优特征形成 4 组方法进行对比。根据早期故障点设置第 544 组数据文件为预测点,选择前 544 组文件作为训练集,544~984 组文件作为测试集。

从振动信号中提取敏感特征后,输入到 GRU 网络进行 RUL 预测。预测结果如图 7 所示,轴承正常运行阶段和预测阶段的真实寿命值用点表示,预测的轴承寿命结果为曲线。图 7(a)是本文提出的基于 HDE-GRU 的滚动轴承 RUL 预测结果,从早期故障点开始预测,预测曲线与真实寿命值极为接近,且变化趋势一致。图 7(b)是基于 RMS-GRU 的 RUL 预测结果,预测曲线与真实寿命值相差不大,但预测结果波动较大。图 7(c)是基于 DE-GRU 的 RUL 预测结果,预测曲线较为平稳,但与真实寿命值相差较大。图 7(d)是基于 MDE-GRU 的 RUL 预测结果,预测曲线与真实寿命值较为接近。

表 2 综合指标计算结果

特征编号	1	2	3	4	5	6	7	8	9
CMR	0.14775	0.0725	0.07594	0.09967	0.06177	0.07386	0.10789	0.15214	0.20848

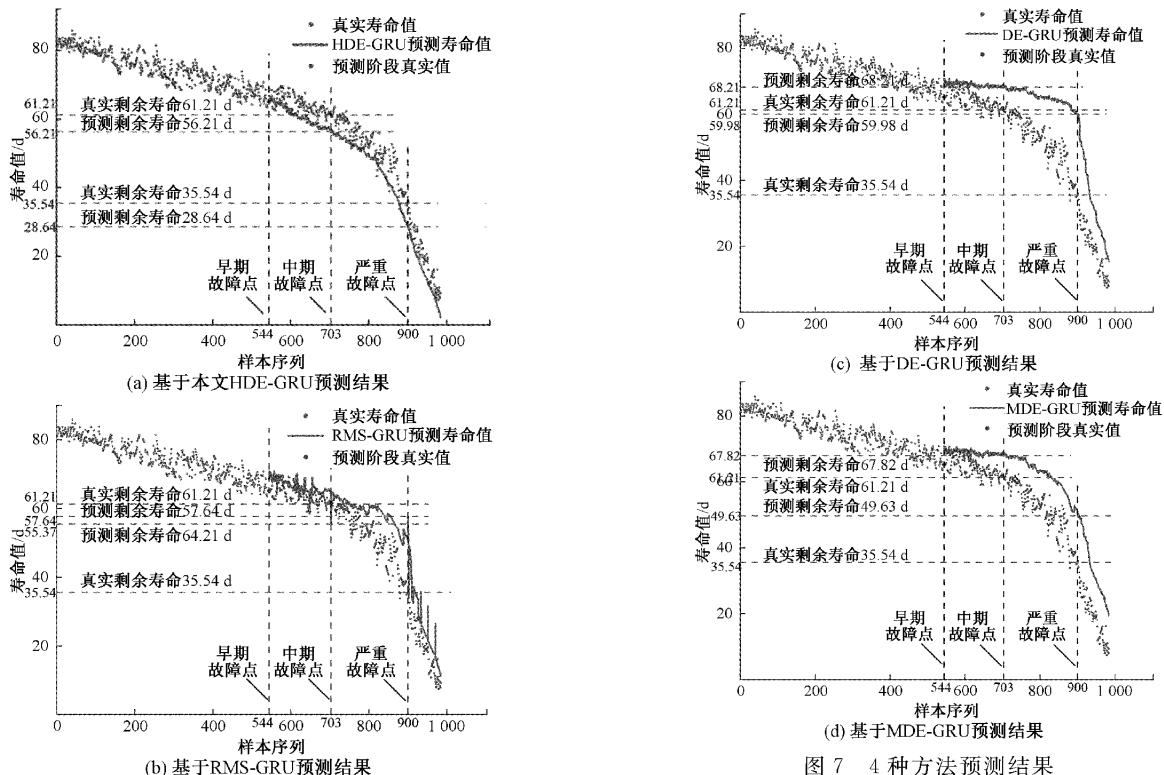


图 7 4 种方法预测结果



- Processing Letters, 2016; 610-614.
- [16] 李从志, 郑近德, 潘海洋, 等. 基于精细复合多尺度散布熵与支持向量机的滚动轴承故障诊断方法[J]. 中国机械工程, 2019, 30(14): 1713-1719,1726.
- [17] ZHONG J, HUANG Y. Time-frequency representation based on an adaptive short-time Fourier transform [J]. IEEE Transactions on Signal Processing, 2010, 58(10): 5118-5128.
- [18] 鄢小安, 贾民平. 基于层次多尺度散布熵的滚动轴承智能故障诊断[J]. 农业工程学报, 2021, 37(11): 67-75.
- [19] WU B, LI W, QIU M. Remaining useful life prediction of bearing with vibration signals based on a novel indicator [J]. Shock and Vibration, 2017, 2017(6): 1-10.

### 作者简介

吴茜,硕士研究生,主要研究方向为机械设备故障预测与健康管理。

E-mail:1126735318@qq.com

张守京(通信作者),博士,主要研究方向为智能制造技术及系统、机械设备信息化及健康管理等。

E-mail:zhangshoujing@xpu.edu.cn