

DOI: 10.19650/j.cnki.cjsi.J2614997

融合概率数据增强与动态相似性的寿命预测*

李延峰, 李思洁, 任维波, 陈忠鑫
(中北大学机械工程学院 太原 030051)

摘要:滚动轴承作为旋转机械的核心零部件,其健康状态直接关系到整台装备的运行安全与生产效率。准确预测滚动轴承的剩余使用寿命,是实现设备预测性维护、降低非计划停机风险的关键技术支撑。在众多预测方法中,基于相似性的寿命预测方法因其原理直观、无需建立复杂物理模型等优势而受到广泛关注。然而,历史全寿命数据匮乏制约了参考轨迹的完备性,而传统全局度量方式难以刻画退化过程中的非平稳局部波动,二者共同限制了预测的泛化能力。针对上述问题,提出一种基于概率数据增强与局部波动感知的动态相似性度量方法。首先,从振动信号中提取均方根特征,并构建两阶段退化模型生成健康指标;进一步,采用高斯、伽马与威布尔混合分布对关键退化参数进行概率建模,生成统计特性与真实过程一致的增强轨迹库,有效扩充退化样本;其次,提出基于滑动窗口局部波动分析的自适应相似性度量机制,通过动态量化序列波动强度并结合自适应时域衰减权重,使匹配更精准聚焦于当前退化阶段的动态演变特征。最后,在 XJTU-SY 和自测数据集上的实验结果表明,所提方法在小样本条件下显著提升了预测精度,性能优于多种现有方法。

关键词:剩余使用寿命预测;退化建模;概率数据增强;退化趋势相似性检测

中图分类号: TH133.33 **文献标识码:** A **国家标准学科分类代码:** 460.10

Remaining useful life prediction via probabilistic augmentation and dynamic similarity

Li Yanfeng, Li Sijie, Ren Weibo, Chen Zhongxin

(School of Mechanical Engineering, North University of China, Taiyuan 030051, China)

Abstract: Rolling bearings are critical in rotating machinery. Their health directly affects operational safety and efficiency. Accurately predicting their remaining useful life is essential for predictive maintenance and reducing unplanned downtime. Similarity-based methods are attractive due to their intuitive nature and avoidance of complex physical models. However, the scarcity of full-lifecycle data limits reference trajectory completeness, and global metrics struggle to capture non-stationary local fluctuations during degradation. To address these challenges, this article proposes a dynamic similarity metric method integrating probabilistic data augmentation and local fluctuation perception. Root mean square features are first extracted from vibration signals, and a two-stage degradation model constructs health indicators. Then, key degradation parameters are then modeled using Gaussian, Gamma, and Weibull mixture distributions to generate an enhanced trajectory library that preserves real statistical characteristics. An adaptive similarity metric, based on sliding window local fluctuation analysis, dynamically quantifies fluctuation intensity and incorporates time decay weights to align with current degradation dynamics. Extensive experiments on the XJTU-SY and self-collected datasets show that the proposed method significantly improves prediction accuracy under small-sample conditions, outperforming several existing approaches, with ablation studies validating the contribution of each module.

Keywords: remaining useful life prediction; degradation modeling; probabilistic digital twin database; similarity detection of degradation trends

0 引言

预后与健康管理技术在提高目标系统可靠性方面具有关键作用。其中,剩余使用寿命(remaining useful life, RUL)预测作为预后与健康管理技术的核心部分,在预测性维修中至关重要,并得到了广泛的研究和关注^[1]。现有的 RUL 预测方法可以分为物理方法和数据驱动方法两种。相较于物理方法,数据驱动方法因不需要彻底了解系统故障的物理机理而被广泛研究^[2]。

在数据驱动方法中,基于相似性的方法是从历史失效设备中寻找与待测设备经历相似的退化行为,并根据这些退化行为的寿命值预测待测设备的剩余寿命。这种方法实现原理简单且具有较高的适用性。许先鑫等^[3]基于 Copula 函数对传感器数据进行相关性建模分析,构建发动机传感器之间的 Copula 结构。赵洪利等^[4]提出一种基于曼哈顿距离和斜率双尺度的改进相似性寿命预测方法。孙秀慧等^[5]提出了一种基于 Copula 熵进行非线性传感器选择的方法,通过精准重构健康因子并定义相似距离。施锦涛等^[6]提出时空相似性度量准则,通过对时序距离和空间距离进行非线性加权,构建时空相似性度量指标。胡聪等^[7]利用余弦相似度判断图像之间的相似度。曲岳哈等^[8]提出基于动态时间规整距离(dynamic time warping, DTW)的自适应阈值建模方法,通过曲线形态相似性构建自适应故障阈值模型并预测故障时间置信区间。早期研究多采用余弦相似性^[9]、皮尔逊相关系数^[10]或动态时间弯曲^[11]等经典度量。随后,研究者们提出了如信息融合^[12]等策略以整合多种度量。近年来,更先进的时频脊估计^[13]、健康指数^[14]考虑了距离相似性和空间方向相似性等。然而,这些度量方法大多侧重于轨迹的整体形态比对,未能对时间序列中每个特定点的贡献进行差异化考量,因而对随机波动较为敏感,容易产生度量误差。

相似性 RUL 预测往往需要大量全生命周期数据。但此类信号在实际中不容易获取,而数据增强技术是解决数据不足问题的有效手段。Yu 等^[15]通过正向扩散和反向去噪过程,并引入信号标签作为条件信息,以指导生成相关类别样本。Tang 等^[16]设计了一种信号增强和压缩方法,通过添加噪声的方法进行数据增强。顾辉等^[17]提出一种将数据重叠采样与随机采样相结合的数据增强算法。火久元等^[18]提出了一种对失衡类别的数据进行裁剪、混合处理生成新的数据样本来增加样本量。然而,在复杂工况下,现有方法生成的数据难以覆盖全工况特征,导致模型泛化能力受限。近年来,基于生成对抗网络(generative adversarial network, GAN)及其变体的技术为数据增强提供了新突破。马同旭等^[19]根据预测风速,基

于最小二乘条件生成对抗网络(least squares conditional generative adversarial networks, LSCGAN)模型生成对应预测风速下的多维时间序列数据。Zhang 等^[20]提出了一种与 GANomaly 融合的方式进行轴承数据增强。王报祥等^[21]提出了改进循环一致生成对抗网络,用于生成高质量的故障数据样本,从而提高数据不平衡下的机械故障智能诊断精度。但是,GAN 及其变体在训练过程中容易出现模式崩溃、梯度消失或爆炸等问题,导致生成的数据质量不稳定,可能生成大量重复或无意义的样本,从而影响数据增强的效果。

为在小样本条件下实现滚动轴承剩余使用寿命的精准预测,本文提出一种融合概率数据增强与动态相似性的寿命预测方法。通过混合分布建模与蒙特卡洛重采样生成概率增强数据库(probabilistic augmented database library, PADL),有效扩充与真实退化统计特性一致的合成轨迹;并设计局部波动感知的动态相似性度量,精准聚焦当前退化阶段的动态特征。实验结果表明,该方法显著提升了非平稳退化过程的预测精度与泛化能力。

1 基于两阶段退化模型的 PADL 构建

1.1 两阶段退化模型建立

轴承的振动信号包含了大量的信息,但是原始信号波动较大,退化趋势不明显,难以直接应用于轴承的 RUL 预测。通过提取特征可以更好地理解和分析振动信号所携带的信息。均方根值(root mean square, RMS)因其在物理意义、故障敏感性方面的优势,故选取 RMS 作为关键特征。

图 1 展示了原始信号的 RMS 值,其退化过程呈现出显著的两阶段特征。

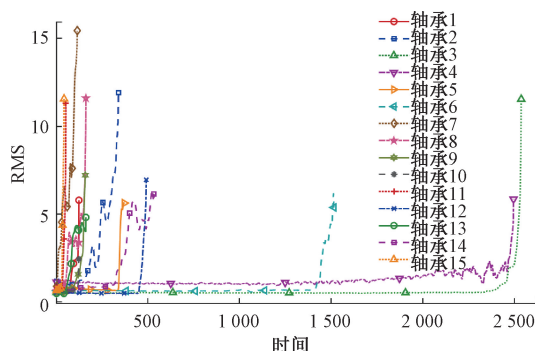


图1 原始信号的 RMS 值

Fig. 1 RMS of the original signal

为更精确地刻画这一物理过程,本文采用两阶段混合模型对 RMS 值进行建模以构建健康指标(health indicators, HI)。其表达式如式(1)所示。

$$HI = \begin{cases} \phi + kt, & t \leq t_c \\ \phi + kt_c + \theta(e^{\beta(t-t_c)} - 1), & t > t_c \end{cases} \quad (1)$$

式中: k 表示缓慢退化阶段的退化速率; θ 与 β 分别表示快速退化阶段的形状参数与退化速率; 而 t_c 则为两阶段之间的变点出现的时刻。

为简化参数辨识过程, 该模型对参数 ϕ 采用固定值 1。此设定确保了所有 HI 均从统一的初始健康状态 (即 0) 开始。不仅消除了不同轨迹间初始状态的差异性, 还显著降低了模型的复杂度与参数识别的复杂度。因此, 实际需要识别的系数为 k, θ, β 与 t_c 。

图 2 展示了基于两阶段退化模型构建的 15 组 HI 轨迹。

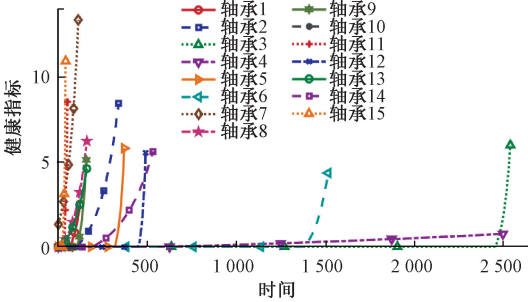


图 2 拟合后的 HI 轨迹

Fig. 2 Fitted HI trajectory

本研究将每条轨迹的退化模式抽象为一个参数向量, 从而构建出刻画退化行为的特征空间。据此通过拟合该参数向量的联合概率分布, 并运用蒙特卡罗采样, 生成大量遵循真实退化统计规律的合成 HI 轨迹。

1.2 模型参数辨识与统计特性分析

基于上述两阶段退化模型, 本文采用最小二乘法对关键参数 k, θ, β 与 t_c 进行辨识。在从历史退化数据中获取该参数集的初始估计值后, 下一步旨在构造更多符合其统计特性的 k, θ, β, t_c 参数集。

该方法的前提在于对原始参数概率分布的准确建模。为此, 首先对历史拟合得到的 k, θ, β 与 t_c 值进行分布分析, 其频率分布直方图如图 3 所示。

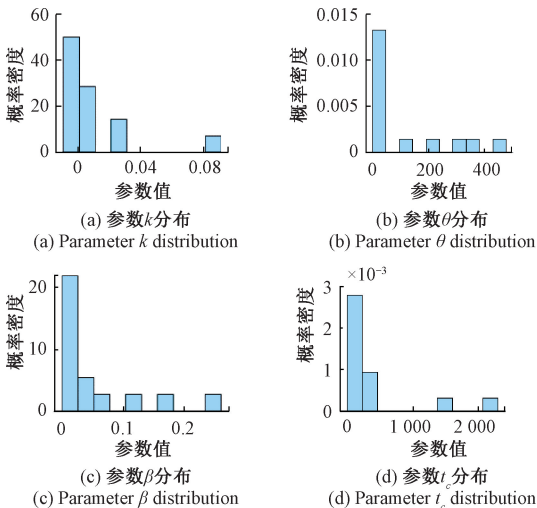


图 3 原始参数分布直方图

Fig. 3 Histogram of raw parameter distribution

在此基础上, 将选取适当的概率分布函数对这些经验分布进行拟合, 以精确刻画其随机特性, 从而为后续的参数生成提供概率依据。本文选取了威布尔分布 (Weibull)、高斯分布 (Gaussian) 和伽马分布 (Gamma) 3 种广泛采用的理论分布, 分别对 4 个参数的分布进行拟合。

拟合结果如图 4 所示, 初步评估不同分布对参数特征的捕捉能力, 以确保生成系数在统计特性上与真实拟合得到的参数组保持一致, 从而保证新生成的 HI 轨迹既具备多样性, 又维持了物理意义上的真实性且符合原始分布规律。

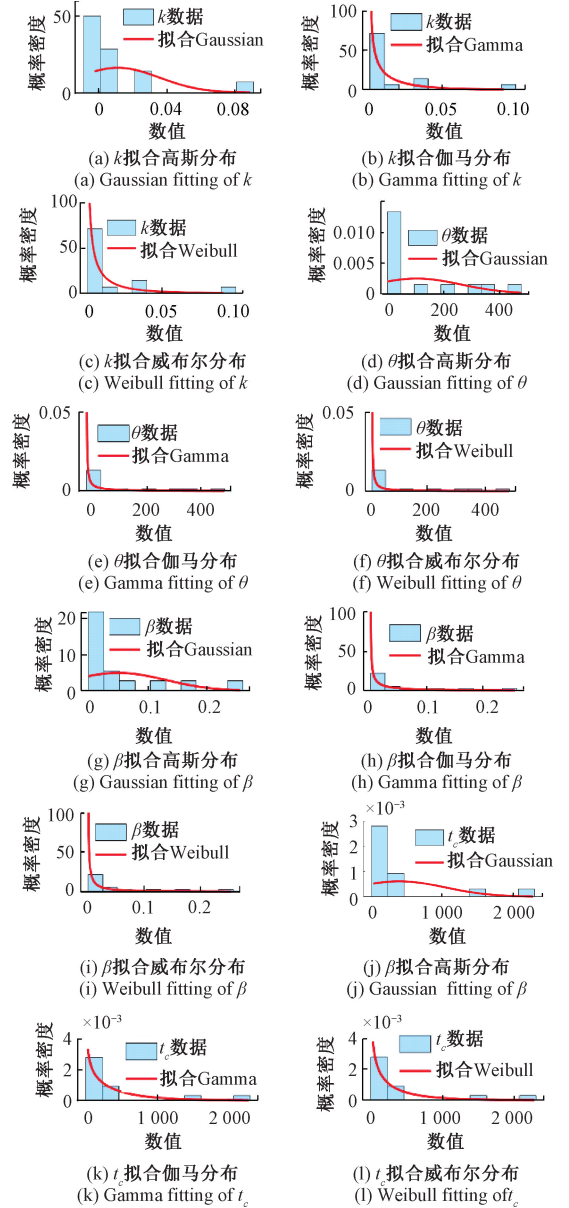


图 4 4 个关键参数对 3 种分布的拟合

Fig. 4 Fitting of three distributions to the four

key parameters

由图4可知,伽马分布与威布尔分布所拟合的概率密度函数曲线形态高度接近,表明这两种分布在刻画参数不确定性方面均具备良好的适用潜力。为进一步从定量角度评估并筛选最优分布,本研究采用 A-D 检

验 (anderson-darling, AD) 统计量对 3 种候选分布分别进行拟合优度检验,AD 统计量越小,表示分布拟合效果越好。

基于 AD 统计量的拟合优度检验结果如表 1 所示。

表 1 基于 AD 统计量的拟合优度检验

Table 1 Goodness-of-fit test based on the Anderson-Darling statistic

候选分布	k		θ		β		t_c	
	AD	P 值	AD	P 值	AD	P 值	AD	P 值
高斯分布	2.492	0.051	1.796	0.119	1.763	0.125	2.467	0.052
伽马分布	0.923	0.398	0.419	0.827	0.263	0.963	1.009	0.351
威布尔分布	0.882	0.423	0.524	0.719	0.286	0.948	0.753	0.513

分析表明,伽马分布在多个参数上的 P 值整体略高于威布尔分布。因此,在后续的参数生成中,本文选用伽马分布作为各参数的生成模型,以确保合成数据在统计特性上与真实退化行为保持高度的相似性,即合成数据更能代表真实退化的趋势。

1.3 基于 PADL 方法的退化数据生成

采用蒙特卡洛采样方法生成与原始分布一致的 θ 、 β 、 k 、 t_c 参数集,步骤如图 5 所示。

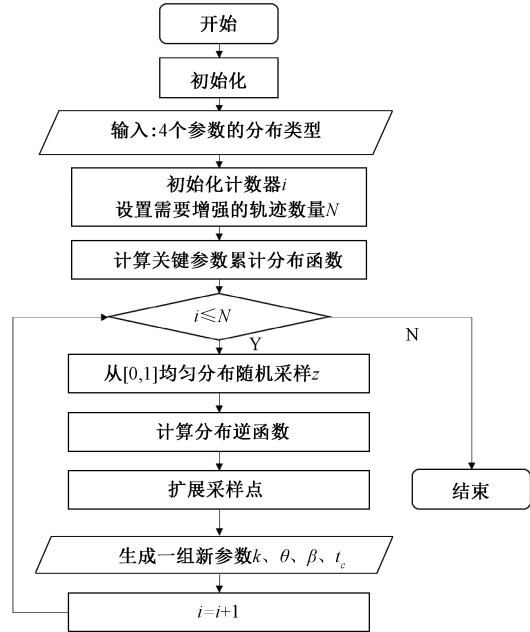


图 5 蒙特卡洛采样方法流程

Fig. 5 Flowchart of the Monte Carlo sampling method

在生成新的 θ 、 β 、 k 、 t_c 参数集后,将这些新生成的系数可代入式(3)中,从而合成全新的退化轨迹。为验证参数集的统计可靠性,图6展示了4个参数在重采样后的分布图,采样后的概率密度分布与重采样前的分布在整体形态上保持高度一致,证明蒙特卡罗采样在保持参数本质统计特性方面的有效性。

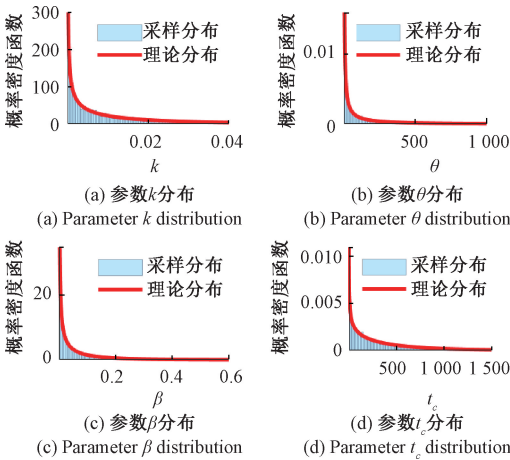


图 6 重采样参数分布

Fig. 6 Resampled parameter distribution

基于上述实验结果与分析,可以得出,基于蒙特卡洛的重采样方法在生成模型参数的同时,有效保持了 θ 、 β 、 k 、 t_c 系数原有的统计分布特征。因此,由该过程所生成的新系数具有与原数据相同的统计属性,能够被可靠地代入退化模型中,合成兼具代表性与多样性的 HI 轨迹。

在本节中,HI 轨迹的数量从原始的 15 条扩充至 5 000 条,构建了一个规模显著增大、统计代表性更强的退化轨迹库。为后续基于数据驱动的剩余使用寿命预测提供充分且高质量的输入和足够的训练数据。

2 融合多尺度时序注意力的动态相似性 RUL 预测方法

2.1 基于多尺度时序注意力机制的动态相似性度量

在基于相似性的 RUL 预测中,最重要的是从历史数据中检索与当前监测序列匹配度最高的相似退化轨迹。然而,传统相似性度量方法强调曲线之间的整体相似性,但忽略了局部趋势和时间关系。因此,本文提

出一种融合多维度时序感知机制的动态加权相似性度量方法,引入了基于序列实时波动特性的自适应衰减调整机制。

通过将数据划分成多个滑动窗口,动态量化序列的波动强度:波动越大,衰减因子相应增大,使相似性度量在快速阶段更集中于近期数据该机制能够依据不同退化阶段的动态特征,实时调节时序权重分配,从而有效提升寿命预测相似性度量的敏感性与准确性。

为了捕捉健康指标序列的局部波动特性,将当前设备运行片段划分为多个重叠的滑动窗口。设每个窗口长度为 L ,步长为 S ,则第 k 个窗口为:

$$W_t = \{x_{t-L+1}, x_{t-L+2}, \dots, x_t\}, t \geq L \quad (2)$$

对于每个窗口内的数据,计算一阶差分序列以捕获局部变化:

$$\Delta X_t = \{x_{t-L+i} - x_{t-L+i-1}, \dots, x_t - x_{t-1}\} = \{\Delta_{x-L+2}, \dots, \Delta_x\} \quad (3)$$

计算滑动窗口内标准差,量化波动分散度,其中序列平均值为 $\overline{\Delta x_t}$:

$$\sigma_t = \sqrt{\frac{1}{L-2} \sum_{i=t-L+2}^t (\Delta x_i - \overline{\Delta x_t})^2} \quad (4)$$

计算滑动窗口内极差变化率,用于量化波动强度:

$$R_t = \frac{\max(\Delta x_t) - \min(\Delta x_t)}{\max(\Delta x_{t-1}) - \min(\Delta x_{t-1})} - 1 \quad (5)$$

极差变化率反映了波动幅度相对于平均水平的相对强度。若极差变化率增大,表示波动增强;反之,则表示波动减弱。将标准差和极差变化率归一化并加权合并,得到窗口 k 的综合波动强度:

$$\sigma'_t = \frac{\sigma_t - \sigma_{\min}}{\sigma_{\max} - \sigma_{\min}} \quad (6)$$

$$R'_t = \frac{R_t - R_{\min}}{R_{\max} - R_{\min}} \quad (7)$$

$$F_t = \theta \sigma'_t + (1 - \theta) R'_t \quad (8)$$

根据综合波动强度,动态调整衰减因子。设定衰减因子的最小值和最大值,记为 $\alpha_{\max}$ 和 $\alpha_{\min}$,则第 k 个窗口对应的衰减因子为:

$$\alpha_t = \alpha_{\min} + (\alpha_{\max} - \alpha_{\min}) \cdot (1 - e^{-k \cdot F_t}) \quad (9)$$

当波动强度大时,衰减因子增大,使得相似性度量更关注近期数据。对于当前轨迹片段,为每个时间点分配权重。假设当前轨迹片段长度为 t ,对于第 i 个时间点,其权重计算为:

$$w_{t,i} = \frac{e^{-\alpha_t(t-j)}}{\sum_{j=t-L+1}^t e^{-\alpha_t(t-j)}} \quad (10)$$

式中: $i \in [t-L+1, t]$,该权重函数使得近期数据获得更高权重;衰减速率由 α_t 动态控制,同时权重进行了归一

化。设当前轨迹片段 $X_t = [x_{t-L+1}, \dots, x_t]$,历史轨迹片段 $Y_t = [y_{t-L+1}, \dots, y_t]$,用加权欧氏距离计算相似性可得:

$$D(X_t, Y_t) = \sqrt{\sum_{i=t-L+1}^t w_{t,i} (x_t - y_k)^2} \quad (11)$$

为了将距离转换为相似性得分,可以使用如下转换:

$$D(X_t, Y_t) = \sqrt{\sum_{i=t-L+1}^t w_{t,i} (x_t - y_k)^2} \quad (12)$$

该数值的大小直接反映了当前序列与历史序列在局部波动特征上的吻合程度:数值越大,说明当前片段与历史片段的动态演变规律越相似,在相似性度量空间中的距离越近。

2.2 基于动态相似性的 RUL 预测

在确定相似性评估度量之后,仍需对轴承寿命终点的 HI 阈值。如式(13)所示,RUL 被定义为从当前时间 t_{current} 到 HI 第 1 次超过其阈值 $t_{\text{threshold}}$ 的时间段,而 HI 阈值则选取原始样本平均 HI 作为识别轴承寿命终止的阈值,如式(14)所示。

$$\text{RUL} = t_{\text{threshold}} - t_{\text{current}} \quad (13)$$

$$\text{HI} = \frac{\sum_{i=1}^{N_t} \text{HI}_i}{N_t} \quad (14)$$

测量样本的 RUL 由所选相似轨迹的 RUL 决定,根据相似度从历史轨迹库中检索出与当前退化状态最相似的若干轨迹并对这些轨迹在当前时刻对应的 RUL 值进行加权融合,从而得到最终的 RUL 预测值。

2.3 方法流程

针对传统相似性方法忽略时序动态特性及由于样本数量不足导致预测精度不足的问题,本文提出一种融合概率数据增强与动态相似性度量的改进预测方法。具体算法流程图如图 7 所示。该方法的实施步骤为:

1) 基于两阶段退化模型建立健康指标序列,并提取模型中的关键参数;

2) 对原始参数进行概率密度拟合,在此基础上利用蒙特卡洛重采样生成多组符合原始分布的新参数;

3) 将生成参数代入退化模型,生成扩展轨迹,构建增强数据集;

4) 通过动态相似性度量方法,计算测试信号与数据集中各轨迹的相似性,按式(12)计算相似度;

5) 依据相似度对所选轨迹的剩余寿命进行加权融合,得到测试轴承的剩余寿命预测值。

3 试验研究及案例分析

为了验证本文提出方法的有效性,使用公开的轴承全寿命周期加速性能退化实验数据 XJTU-SY^[22] 和自测的滚动轴承加速试验台数据集进行分析。

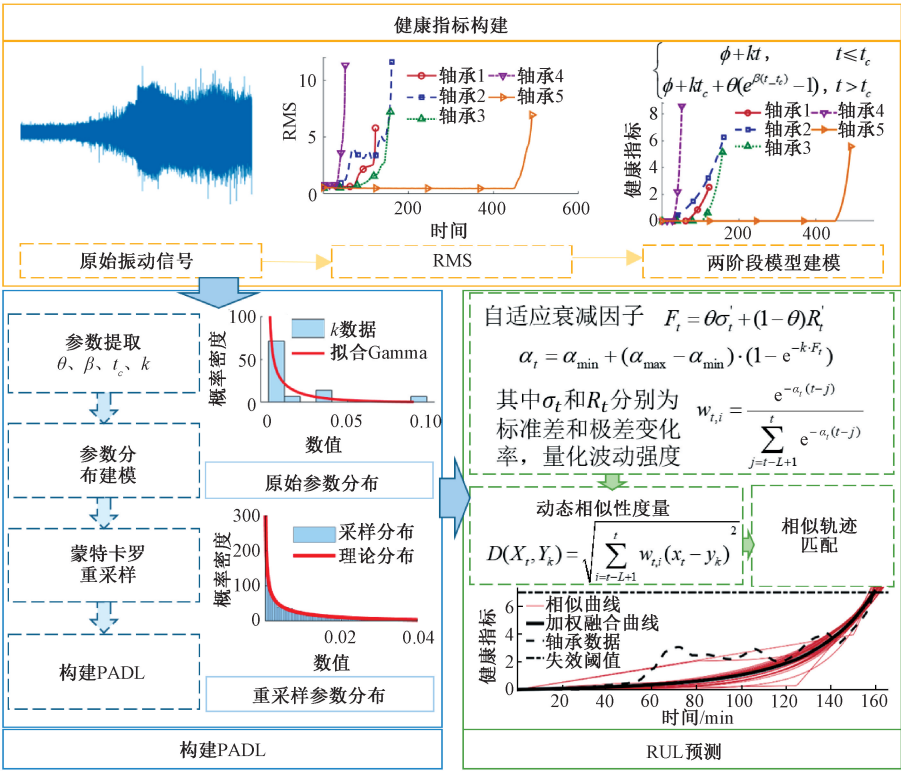


图7 基于PADL的RUL预测方法

Fig.7 PADL-based RUL prediction method

自测的滚动轴承加速试验台将1个新的滚动轴承加速运行至失效,两个加速度传感器分别安装在测试轴承的水平、竖直方向上,测试轴承用的是 BFK-1 滚动轴承,轴承的工作条件:转速为 2 100 r/min,径向力为 10 kN。试验台搭建如图 8 所示,实验设置采样频率为 25.6 kHz,采样间隔为 1 min,每次采样时长为 2 s。

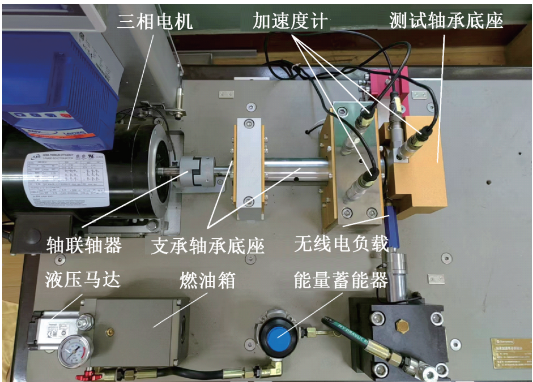


图8 BFK-1 滚动轴承实验台

Fig.8 BFK-1 Rolling bearing test bench

基于相似性的 RUL 预测中,测量样本的 RUL 值由所选相似轨迹的 RUL 值决定,相似轨迹的数量直接影响预测的稳定性和准确性:数量过少可能导致预测结果受个

别异常轨迹干扰,鲁棒性不足;数量过多则可能引入不相关或噪声信息,降低对局部特征的敏感度。

为探讨相似轨迹数量对预测性能的影响,实验分别采用 5 组 HI 曲线(用 XJTU-SY 数据集的 Bearing1_1 至 Bearing1_5 数据)进行分析。比较结果如图 9 所示,随着相似轨迹数量从 2 增至 20,模型预测误差显著降低。然而,当 K 值继续增大至 50 以上,预测误差不降反升,这可能源于过多相似轨迹引入噪声干扰所导致的。因此,本研究最终将相似轨迹数量设定为 20。

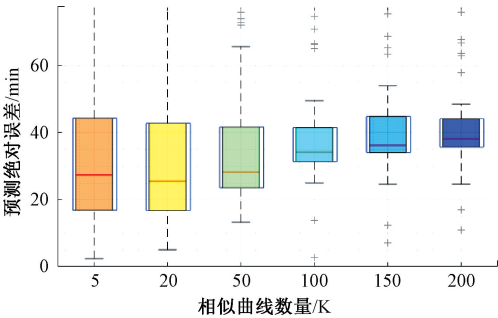


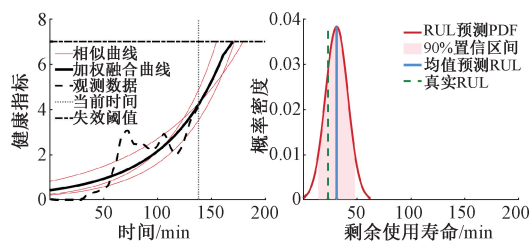
图9 HI 轨迹数对 RUL 预测的影响

Fig.9 Influence of the number of HI trajectories on RUL

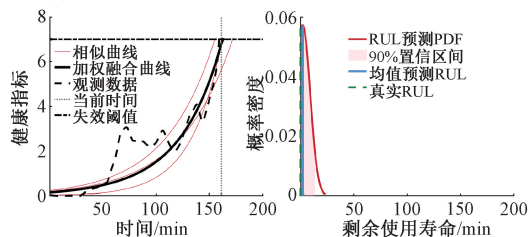
为验证所提方法的有效性,本研究将融合 PADL 的 RUL 预测结果与未融合 PADL 的预测结果进行对比分

析。分别用 XJTU-SY 数据集的 Bearing2_1 数据及自主测得的滚动轴承数据集的 B2_3 进行实验。

如图 10 和 11 所示,图 10(a)~(c)和图 11(a)~(c)为引入 PADL 的预测结果,该模型在预测精度与不确定性控制方面均体现出显著优势。增强后轨迹基于 PADL 的预测方法通过集成多源、多工况下的退化轨迹数据,能够更全面地覆盖设备退化的潜在演化模式,进而在基于相似性的 RUL 预测过程中提供更具代表性的参考轨迹。更丰富的轨迹样本使模型能够更准确地识别与当前状态匹配的历史退化模式,从而得到更精确的寿命估计。相比之下,图 10(d)~(f)和图 11(d)~(f)未采用 PADL 的预测方法由于 HI 库中轨迹数量有限,可供参考的退化模式覆盖不足,其预测结果表现出较高的预测偏差。



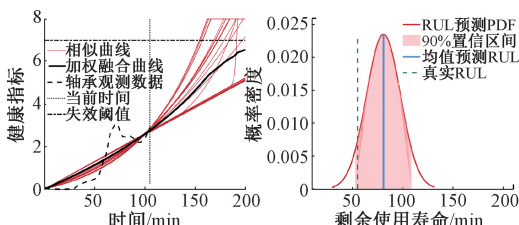
(e) 未融合PADL的80%时间点预测结果
(e) Prediction results at 80% life without PADL integration



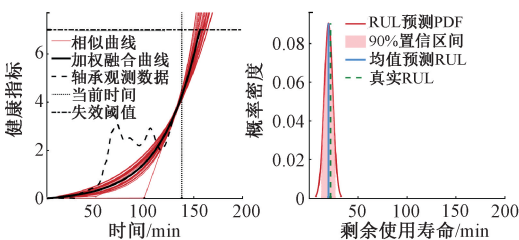
(f) 未融合PADL的最终时间点预测结果
(f) Prediction results at the final time point without PADL

图 10 XJTU-SY 数据集 RUL 预测结果

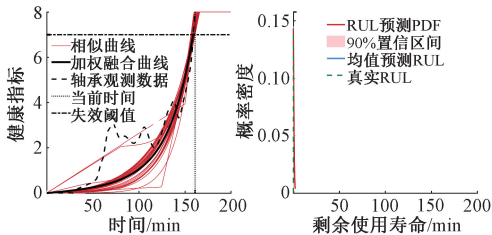
Fig. 10 RUL prediction results on the XJTU-SY dataset



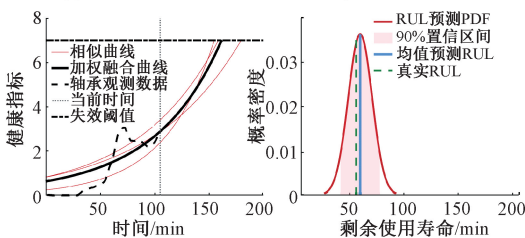
(a) 融合PADL的50%时间点预测结果
(a) Prediction results at 50% life with PADL integration



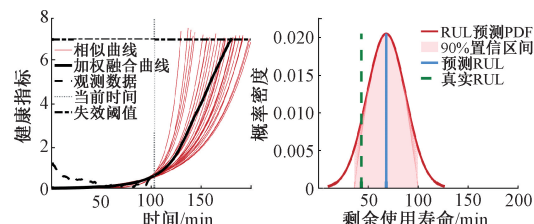
(b) 融合PADL的80%时间点预测结果
(b) Prediction results at 80% life with PADL integration



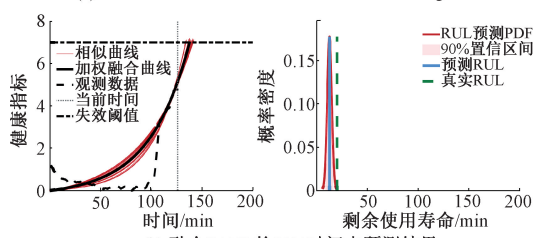
(c) 融合PADL的最终时刻预测结果
(c) Prediction results at the final time point with PADL



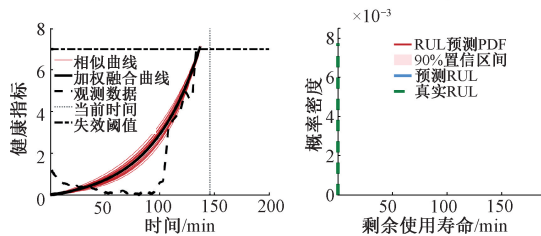
(d) 未融合PADL的50%时间点预测结果
(d) Prediction results at 50% life without PADL integration



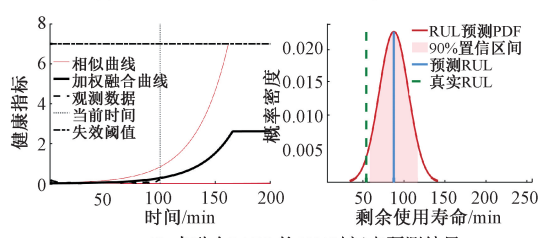
(a) 融合PADL的50%时间点预测结果
(a) Prediction results at 50% life with PADL integration



(b) 融合PADL的80%时间点预测结果
(b) Prediction results at 80% life with PADL integration



(c) 融合PADL的最终时刻预测结果
(c) Prediction results at the final time point with PADL



(d) 未融合PADL的50%时间点预测结果
(d) Prediction results at 50% life without PADL integration

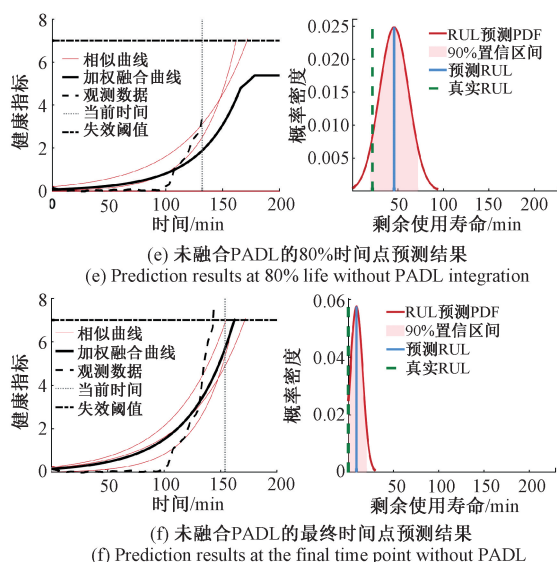


图11 自测数据集的RUL预测结果

Fig. 11 RUL prediction results of the self-test dataset

此外,为验证本文所提出方法在应对小样本与局部相似性度量问题上的优越性,选取了3种具有代表性的先进方法与本文的方法做对比:双动态时间规整的相似性预测方法(dual dynamic time warping, DDTW)^[23]、融合相似性分析的预测方法(Similarity)^[24]以及通过将GAN融入时空长短期记忆网络(spatial-temporal long-short-term memory network, ST-LSTM)的GAN-ST-LSTM方法^[25]。

这些方法分别从数据生成、不同相似性度量等不同角度应对RUL预测中的挑战。通过与上述方法的系统比较,可更清晰地凸显本文方法的综合优势。几种方法的RUL预测对比结果如图12所示。

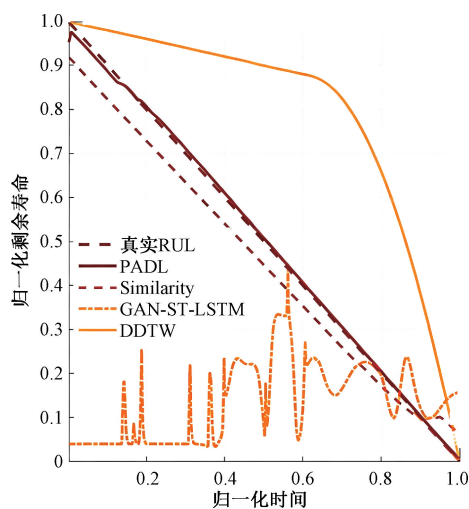


图12 各方法RUL预测结果对比

Fig. 12 Comparison of RUL prediction results among different methods

为保障比较的公平性,各对比方法的参数均遵循其原文献设定。从图中可以看出,本文所提方法的整体预测曲线平滑性较好,且其RUL估计值通常更接近真实退化轨迹。且该方法在多个运行阶段均表现出较好的适应性,能够有效跟踪设备退化趋势,从而取得较高的预测精度。综合而言,本文提出的RUL预测方法在保持输出稳定性的同时,具有更强的工况适应性与更优越的预测准确性。

4 结论

本研究提出了一种融合概率数据增强与局部波动感知相似性度量的滚动轴承剩余寿命预测框架,系统性地应对了小样本条件下的预测难题。首先,通过两阶段退化建模与蒙特卡洛概率重采样,生成了物理意义明确且覆盖多退化模式的合成健康指标轨迹,有效扩充了训练数据;其次,设计了一种局部波动感知的动态相似性度量机制,通过滑动窗口分析与自适应权重调整,使相似性匹配能够精准聚焦于退化过程的非平稳动态演变特征。实验表明,该框架显著提升了数据稀缺场景下的预测精度与鲁棒性。

所提出的方法能够显著提升小样本场景下的预测精度与稳定性,为突破数据稀缺条件下的寿命预测难题提供了新的技术路径。然而,这种方法有一些限制,前提条件是有一些设备从健康运行到失效的数据。当无设备运行到失效的全寿命周期的数据时,可以用迁移学习结合来自动力学模型的仿真数据。

参考文献

- [1] 雷亚国, 许学方, 蔡潇, 等. 面向机械装备健康监测的数据质量保障方法研究[J]. 机械工程学报, 2021, 57(4): 1-9.
LEI Y G, XU X F, CAI X, et al. Research on data quality assurance for health condition monitoring of machinery [J]. Journal of Mechanical Engineering, 2021, 57(4): 1-9.
- [2] WANG ZH J, LI Y J, DONG L, et al. A RUL prediction of bearing using fusion network through feature cross weighting [J]. Measurement Science and Technology, 2023, 34(10): 105908.
- [3] 许先鑫, 李娟, 孙秀慧, 等. 基于Copula相似性的航空发动机RUL预测[J]. 航空动力学报, 2024, 39(8): 512-519.
XU X X, LI J, SUN X H, et al. RUL prediction for aero-engines based on Copula similarity [J]. Journal of Aerospace Power, 2024, 39(8): 512-519.
- [4] 赵洪利, 陈天铭. 基于双尺度相似性的发动机寿命预

- 测[J]. 推进技术, 2022, 43(11): 355-362.
- ZHAO H L, CHEN T M. Engine life prediction based on two-scale similarity[J]. Journal of Propulsion Technology, 2022, 43(11): 355-362.
- [5] 孙秀慧, 李娟, 戴洪德. 基于 Copula 熵传感器选择的发动机相似性寿命预测方法[J]. 航空发动机, 2024, 50(5): 113-121.
- SUN X H, LI J, DAI H D. Engine similarity life prediction based on Copula entropy for sensor selection[J]. Aeroengine, 2024, 50(5): 113-121.
- [6] 施锦涛, 陈磊, 秦凯, 等. 基于时空相似性的即时学习在线建模[J]. 仪器仪表学报, 2022, 43(6): 185-193.
- SHI J T, CHEN L, QIN K, et al. Just-in-time learning online modeling based on spatio-temporal similarity[J]. Chinese Journal of Scientific Instrument, 2022, 43(6): 185-193.
- [7] 胡聪, 李超, 周甜, 等. 基于改进深度相对距离学习框架的车辆再识别算法[J]. 仪器仪表学报, 2020, 41(12): 245-252.
- HU C, LI CH, ZHOU T, et al. Vehicle re-identification algorithm based on improved deep relative distance learning framework [J]. Chinese Journal of Scientific Instrument, 2020, 41(12): 245-252.
- [8] 曲岳晗, 何林, 杨冬锋, 等. 基于多源监测数据和自适应故障阈值的变压器绕组绝缘故障预测方法[J]. 仪器仪表学报, 2026, 47(3): 372-389.
- QU Y H, HE L, YANG D F, et al. Fault prediction method for transformer winding insulation based on multi-source monitoring data and adaptive fault threshold[J]. Chinese Journal of Scientific Instrument, 2026, 47(3): 372-389.
- [9] LAM J, SANKARARAMAN S, STEWART B. Enhanced trajectory based similarity prediction with uncertainty quantification [C]. Annual Conference of the PHM Society, 2014, 6(1): 623-634.
- [10] WANG ZH Y, TANG W, PI D CH. Trajectory similarity-based prediction with information fusion for remaining useful life [C]. International Conference on Intelligent Data Engineering and Automated Learning, 2017: 270-278.
- [11] LI Y F, YANG Y CH, FENG K, et al. Automated and adaptive ridge extraction for rotating machinery fault detection [J]. IEEE/ASME Transactions on Mechatronics, 2023, 28(5): 2565-2575.
- [12] LI Y F, ZHANG X, CHEN Z G, et al. Time frequency ridge estimation: An effective tool for gear and bearing fault diagnosis at time-varying speeds [J]. Mechanical Systems and Signal Processing, 2023, 189: 110108.
- [13] ZHOU ZH, LI T F, ZHAO ZH B, et al. Time-varying trajectory modeling via dynamic governing network for remaining useful life prediction [J]. Mechanical Systems and Signal Processing, 2023, 182: 109610.
- [14] LIU Y CH, HU X F, ZHANG W J. Remaining useful life prediction based on health index similarity [J]. Reliability Engineering & System Safety, 2019, 185: 502-510.
- [15] YU T, LI CH SH, HUANG J, et al. ReF-DDPM: A novel DDPM-based data augmentation method for imbalanced rolling bearing fault diagnosis [J]. Reliability Engineering & System Safety, 2024, 251: 110343.
- [16] TANG H S, LU S L, QIAN G, et al. IoT-based signal enhancement and compression method for efficient motor bearing fault diagnosis [J]. IEEE Sensors Journal, 2021, 21(2): 1820-1828.
- [17] 顾辉, 邵星, 王翠香, 等. 卷积深度森林滚动轴承故障诊断方法 [J]. 机械强度, 2024, 46(6): 1279-1286.
- GU H, SHAO X, WANG C X, et al. Rolling bearing fault diagnosis method based on convolutional deep forest [J]. Journal of Mechanical Strength, 2024, 46(6): 1279-1286.
- [18] 火久元, 李宇峰, 常琛, 等. 基于混合裁剪失衡数据增强与 SwinNet 网络的滚动轴承故障诊断 [J]. 振动与冲击, 2024, 43(6): 64-74.
- HUO J Y, LI Y F, CHANG CH, et al. Fault diagnosis of rolling bearings based on mixed-cutout imbalance data augmentation and the SwinNet network [J]. Journal of Vibration and Shock, 2024, 43(6): 64-74.
- [19] 马同旭, 刘帅, 刘卫亮, 等. 基于两阶段多维数据生成与实时健康指数的风机齿轮箱故障预警 [J]. 仪器仪表学报, 2024, 45(11): 266-276.
- MA T X, LIU SH, LIU W L, et al. Fault warning of wind turbine gearbox based on two-stage multidimensional data generation and real-time health index [J]. Chinese Journal of Scientific Instrument, 2024, 45(11): 266-276.
- [20] ZHANG Y, HAO H J, ZHANG T X. Abnormal detection and fault diagnosis method of bearing based on deep convolutional generative adversarial network [C]. 2023 IEEE 11th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), 2023, 11:

1604-1608.

[21] 王报祥, 丁传仓, 居森, 等. 面向数据不均衡下机械智能故障诊断的改进循环一致生成对抗网络[J]. 仪器仪表学报, 2025, 46(10): 96-106.

WANG B X, DING CH C, JU M, et al. Improved cycle-consistent generative adversarial network for mechanical intelligent fault diagnosis under data imbalance condition[J]. Chinese Journal of Scientific Instrument, 2025, 46(10): 96-106.

[22] 雷亚国, 韩天宇, 王彪, 等. XJTU-SY 滚动轴承加速寿命试验数据集解读[J]. 机械工程学报, 2019, 55(16): 1-6.

LEI Y G, HAN T Y, WANG B, et al. XJTU-SY Rolling element bearing accelerated life test datasets: A tutorial[J]. Journal of Mechanical Engineering, 2019, 55(16): 1-6.

[23] ZHANG J, ZHANG CH, WU Y B, et al. Bearing life prediction method based on novel health indicators and improved similarity curve matching[J]. Measurement Science and Technology, 2025, 36(1): 016189.

[24] XUE B, LI Y, XU ZH B, et al. Remaining useful life prediction for rolling bearing based on a hybrid data-driven method integrating similarity analysis [J]. Measurement Science and Technology, 2025, 36(9): 096217.

[25] BAO W G, MIAO X D, WANG H, et al. Remaining useful life assessment of slewing bearing based on spatial-temporal sequence[J]. IEEE Access, 2020, 8: 9739-9750.

作者简介



李延峰, 2017 年于太原理工大学获得硕士学位, 2021 于太原理工大学年获得博士学位, 现为中北大学机械工程学院副教授, 主要研究方向为复杂设备可靠性建模和数字孪生技术、机械设备的剩余寿命预测和健康管理、数据驱动的智能诊断以及机械信号处理与分析。

E-mail: yanfengli0002@163.com

Li Yanfeng received his M. Sc. degree from Taiyuan University of Technology in 2017, and his Ph.D. degree from Taiyuan University of Technology in 2021. He is currently an associate professor at the School of Mechanical Engineering, North University of China. His main research interests include reliability modeling and digital twin technology of complex equipment, remaining useful life prediction and health management of mechanical equipment, data-driven intelligent diagnosis, as well as mechanical signal processing and analysis.



李思洁 (通信作者), 2022 年于山西农业大学获得学士学位, 目前在中北大学攻读硕士学位, 主要研究方向为机械设备的剩余寿命预测和健康管理。

E-mail: 13383460327@163.com

Li Sijie (Corresponding author) received her B.Sc. degree from Shanxi Agricultural University in 2022. She is currently pursuing her master's degree at North University of China. Her main research interests include remaining useful life prediction and health management of mechanical equipment.