

DOI: 10.20189/j.cnki.CN/61-1527/E.202603007

数模联动框架下考虑样本贡献度差异的 剩余寿命预测方法

罗正杨, 裴 洪*, 郑建飞, 韩其辉, 陈 茜, 司小胜
(火箭军工程大学, 西安 710025)

摘 要: 针对现有数模联动预测框架下设备后期退化特征拟合精度不足、剩余寿命预测误差偏大等问题, 提出一种考虑样本贡献度差异的剩余寿命预测方法。通过引入多种形式的权重函数改进传统目标函数, 确保后期预测效果对优化目标的影响高于早期预测效果。在离线训练阶段, 首先通过多维融合系数将多传感器监测数据融合为复合健康指标; 其次, 考虑到退化过程的非线性特性, 基于指数退化模型构建复合健康指标的演变过程模型, 并结合首达时间概念推导剩余寿命的解析表达式; 最后, 通过构建包含预测偏差与预测方差的目标函数, 优化确定最优融合系数与失效阈值。在线预测阶段, 采用序贯贝叶斯估计对模型参数进行迭代更新, 实现剩余寿命的在线预测。通过航空发动机案例开展对比验证实验, 结果表明: 相较于传统数模联动剩余寿命预测方法, 所提方法的寿命预测均方根误差降低 32.8%, 全周期剩余寿命预测平均相对误差控制在 30% 以内, 且设备运行后期的剩余寿命预测结果更为精确, 可为航空发动机等工业设备的故障诊断与寿命预测提供参考。

关键词: 剩余寿命预测; 非线性退化模型; 数模联动; 样本贡献度

中图分类号: V215; TB114

文献标志码: A

开放科学标识码(OSID):

文章编号: 2097-2741(2026)03-0089-10



听语音
与作者
聊科研

航空发动机、风力发电机、高速列车、导弹武器等大型装备受内外部因素影响不可避免地会发生性能退化, 进而演变为失效^[1-2]。一旦发生失效, 将会造成重大的经济损失和安全事故。为避免此类情况发生, 对设备进行剩余寿命预测并据此开展维修替换决策已成为迫切的现实需要^[3-4]。

剩余寿命预测经过国内外学者的长期研究, 已取得长足发展^[5-7]。当前剩余寿命预测方法主要包括机理建模法、数据驱动法以及机理与数据融合建模法。机理建模法以设备失效机理为核心, 通过构建力学、化学、热学等领域的偏微分方程来描述退化过程^[8-9], 物理意义明确、可靠性强。然而, 实际工程设备结构复杂、运行环境多样, 刻

画实际失效机理模型存在困难, 导致机理建模法及融合建模法难以被推广应用^[10-12]。数据驱动法通过挖掘历史监测信息, 自动提取监测数据与剩余寿命之间的映射关系, 避免了复杂的机理建模过程, 易于推广且准确率高, 得到了广泛研究^[13-14]。

为保障数据驱动法的预测精度, 以反馈思想为核心的数模联动预测框架已在各类工程领域广泛应用^[15], 其基本思路是基于设备运行的多维监测数据构建复合健康指标并建立随机退化模型, 通过最小化目标函数使预测退化指标与实际退化指标形成反馈联动, 从而优化模型参数。例如, 李天梅等^[16]通过加权融合获取复合健康指标, 构建了考虑寿命预测值与设备实际寿命均方误差的优

收稿日期: 2026-03-31

修回日期: 2026-04-27

录用日期: 2026-05-12

基金项目: 国家自然科学基金 (62227814, 62233017, 62373368, 62373369); 中国博士后科学基金 (2023M734286); 陕西省科协青年人才托举计划 (20230127); 陕西省军民融合英才支持项目青年拔尖人才计划 (2025034)

第一作者: 罗正杨 (2005—), 男, 本科生, 主要从事预测与健康管理研究。

***通信作者:** 裴洪 (1992—), 男, 副教授, 主要从事预测与健康管理研究。

引用格式: 罗正杨, 裴洪, 郑建飞, 等. 数模联动框架下考虑样本贡献度差异的剩余寿命预测方法[J]. 火箭军工程大学学报, 2026, 40 (3): 89-98. LUO Zhengyang, PEI Hong, ZHENG Jianfei, et al. Remaining useful life prediction method incorporating sample contribution difference under data-model linkage framework[J]. Journal of Rocket Force University of Engineering, 2026, 40 (3): 89-98.

化目标函数,并基于线性 Wiener 过程模型实现剩余寿命预测;Li 等^[17]通过神经网络融合复合健康指标,构建了综合单调性、凹凸性、失效阈值方差最小等特性的目标函数,并采用随机系数模型实现了剩余寿命预测;Li 等^[18]构建了一个带正则化项、以预测精度为导向的目标函数,通过加权融合获取复合健康指标,并基于线性 Wiener 过程模型实现剩余寿命预测;Li 等^[19]采用堆叠去噪自编码器,以无监督方式融合多维监测数据,构建了考虑预测寿命与实际寿命之间均方误差的优化目标函数,采用 Wiener 过程模型实现剩余寿命预测。

现有研究在构建损失函数时,假定所有样本具有同等贡献度,忽略了不同时刻样本贡献度的差异性。在实际工程中,更加关注的是设备在接近失效时的预测准确度,而同等贡献度的假设会弱化设备运行后期寿命预测效果。此外,现有研究中构建目标函数时采用总的预测寿命与实际寿命之差,而并非各个监测时刻的预测剩余寿命与实际剩余寿命之差,导致训练样本过少,易丢失部分有用信息,直接影响预测精度。

针对上述问题,在数模联动框架下提出一种考虑不同时刻样本贡献度差异的剩余寿命预测方法,通过设计指数函数、幂函数、反比例函数等权重函数以改进传统目标函数,确保后期预测效果对优化目标的影响高于早期预测效果。同时,在目标函数中引入剩余寿命的相关指标作为优化对象,增加了样本数量,提升了剩余寿命预测精度。

1 基于历史设备的寿命预测离线训练

1.1 多传感器数据描述及预处理

假设有 N 个历史随机退化设备,每个随机退化设备有 S 个传感器,故第 i 个设备的多传感器监测数据可表示为

$$U_i = \begin{bmatrix} U_{i,0}^1 & U_{i,1}^1 & \cdots & U_{i,K_i}^1 \\ U_{i,0}^2 & U_{i,1}^2 & \cdots & U_{i,K_i}^2 \\ \vdots & \vdots & \ddots & \vdots \\ U_{i,0}^S & U_{i,1}^S & \cdots & U_{i,K_i}^S \end{bmatrix} \quad (1)$$

式中: $U_{i,k}^j$ 为第 i 个随机退化设备第 j 个传感器的第 k 次监测数据; K_i 为第 i 个随机退化设备总的监测次数。

考虑到不同传感器监测数据的量纲存在差异,可能导致数据间相互干扰,采用最大最小归一化方法对数据进行预处理,将所有数据归一化到

$[0, 1]$ 区间。归一化后第 i 个设备的多传感器监测数据可表示为

$$X_i = \begin{bmatrix} X_{i,0}^1 & X_{i,1}^1 & \cdots & X_{i,K_i}^1 \\ X_{i,0}^2 & X_{i,1}^2 & \cdots & X_{i,K_i}^2 \\ \vdots & \vdots & \ddots & \vdots \\ X_{i,0}^S & X_{i,1}^S & \cdots & X_{i,K_i}^S \end{bmatrix} \quad (2)$$

式中: $X_{i,k}^j$ 表示归一化后第 i 个随机退化设备第 j 个传感器第 k 次监测数据。其具体表达式为

$$X_{i,k}^j = \frac{U_{i,k}^j - \min(U^j)}{\max(U^j) - \min(U^j)} \quad (3)$$

式中: $\max(U^j)$ 、 $\min(U^j)$ 分别为第 j 个传感器所有监测数据的最大值和最小值。

此外,由于传感监测数据不可避免地存在噪声,采用平滑滤波对数据进行处理,以减弱数据随机性对后续建模的影响。

1.2 基于融合系数构建复合健康指标

为将传感器的多维数据转化为一维数据,采用融合系数加权平均构建复合健康指标。第 i 个设备在第 k 个监测周期的复合健康指标为

$$Z_{i,k} = \sum_{j=1}^S X_{i,k}^j w_j \quad (4)$$

式中: w_j 为第 j 个传感器的融合系数。为便于研究,仅考虑各传感器的融合系数在随机退化设备运行全周期内为恒定值的情况,定义融合系数矩阵 $W = [w_1, w_2, \cdots, w_j]$ 。

1.3 指数随机退化过程建模

指数退化模型在电子、机械等领域的可靠性分析与寿命预测中应用广泛,其对非线性退化过程拟合的有效性已得到验证,因此采用该模型对随机退化过程进行建模。

在此情况下,第 i 个随机退化设备在 t ($0 \leq t \leq t_{K_i}$) 时刻退化水平可表示为

$$Z_i(t) = \theta_i' \exp \left\{ \beta_i' t + \sigma_i B(t) - \frac{\sigma_i^2}{2} t \right\} \quad (5)$$

式中: $B(t)$ 为标准布朗运动; θ_i' 、 β_i' 、 σ_i 为模型参数。

式 (5) 所示的退化模型为非线性模型,通过对数变换将其线性化,线性化后第 i 个随机退化设备的复合健康指标及退化水平为

$$S_{i,k} = \ln Z_i(Z_{i,k}) = \ln \left(\sum_{j=1}^S X_{i,k}^j w_j \right) \quad (6)$$

$$S_i(t) = \ln(Z_i(t)) = \ln \theta_i' + \left(\beta_i' - \frac{\sigma_i^2}{2} \right) t + \sigma_i B(t) \quad (7)$$

令 $\theta_i = \ln \theta'_i$, $\beta_i = \beta'_i - \frac{\sigma_i^2}{2}$, 可得

$$S_i(t) = \theta_i + \beta_i t + \sigma_i B(t) \quad (8)$$

通过上述变换过程, 健康指标的退化过程已由非线性转化为线性, 因此, 可基于参数 θ_i 、 β_i 、 σ_i 实现对剩余寿命的预测, 根据文献[20], 采用最大似然估计方法, 可得3个参数的估计值分别为

$$\hat{\theta}_i = S_{i,0} - \hat{\beta}_i t_{i,0} \quad (9)$$

$$\hat{\beta}_i = \frac{S_{i,K_i} - S_{i,0}}{t_{i,K_i} - t_{i,0}} \quad (10)$$

$$\hat{\sigma}_i^2 = \frac{1}{K_i} \sum_{k=1}^{K_i} \frac{\left((S_{i,k} - S_{i,k-1}) - \hat{\beta}_i (t_{i,k} - t_{i,k-1}) \right)^2}{(t_{i,k} - t_{i,k-1})} \quad (11)$$

式中: $t_{i,k}$ 为第 i 个随机退化设备第 k 次监测所对应的监测时间。

基于前述模型, 依据首达时间的概念, 第 i 个随机退化设备的寿命与剩余使用寿命分别可定义为

$$T_i = \inf \{ t_i : S_i(t_i) \geq w \mid S_{i,0} < w \} \quad (12)$$

$$L_{i,k} = \inf \{ l_{i,k} : S_i(t_{i,k} + l_{i,k}) \geq w \mid S_{i,k} < w \} \quad (13)$$

根据文献[20], 结合参数估计结果, 式(12)和(13)定义的 T_i 和 $L_{i,k}$ 服从逆高斯分布, 即

$$f_{T_i}(t_i) = \frac{w - \hat{\theta}_i}{\sqrt{2\pi t_i^3 \hat{\sigma}_i^2}} \exp \left[-\frac{(w - \hat{\theta}_i - \hat{\beta}_i t_i)^2}{2\hat{\sigma}_i^2 t_i} \right] \quad (14)$$

$$f_{L_{i,k}}(l_{i,k}) = \frac{w - S_{i,k}}{\sqrt{2\pi l_{i,k}^3 \hat{\sigma}_i^2}} \exp \left[-\frac{(w - S_{i,k} - \hat{\beta}_i l_{i,k})^2}{2\hat{\sigma}_i^2 l_{i,k}} \right] \quad (15)$$

此外, T_i 和 $L_{i,k}$ 的期望和方差分别为

$$E[T_i] = \frac{w - \hat{\theta}_i}{\hat{\beta}_i} \quad (16)$$

$$\text{var}[T_i] = E[(T_i - \hat{T}_i)^2] = \frac{(w - \hat{\theta}_i) \hat{\sigma}_i^2}{\hat{\beta}_i^3} \quad (17)$$

$$E[L_{i,k}] = \frac{w - S_{i,k}}{\hat{\beta}_i} \quad (18)$$

$$\text{var}[L_{i,k}] = E[(L_{i,k} - \hat{L}_{i,k})^2] = \frac{(w - S_{i,k}) \hat{\sigma}_i^2}{\hat{\beta}_i^3} \quad (19)$$

1.4 考虑样本贡献度差异的复合健康指标构建与随机退化建模交互联动

由式(16)、(17)可知, 预测寿命的期望及方差仅由复合健康指标的初始值及终值决定, 而非全周期监测数据, 可能导致部分隐藏信息丢失, 影响预测精度。同时, 鉴于指数退化模型的马尔可夫性, 由式(18)、(19)得到预测剩余寿命的期望及方差包含当前健康状态信息, 因此在目标函数中选用剩余寿命而非寿命作为优化目标更为合理。

优化目标函数中的优化对象将反馈到退化设备的在线预测, 通过设计目标函数即可获取更为理想的预测效果。在剩余寿命预测的多数场景下更加关注预测结果的准确性和不确定度, 但是在目标函数中引入过多评价指标易使优化目标相互冲突, 因此, 本文的优化目标仅为表征预测准确度的剩余寿命预测偏差和表征预测不确定度的剩余寿命预测方差。

实际在进行寿命预测时, 随机退化设备越接近失效时刻, 预测的准确程度就越重要。若此时设备剩余寿命预测出现偏差, 未及时检修可能造成极为惨重的后果。因此, 需要为不同时刻的预测结果赋予不同权重, 并设计多种形式的权重函数。优化目标函数可表示为

$$J(\mathbf{W}, \mu_w, \sigma_w^2) = \frac{1-\gamma}{N} \sum_{i=1}^N \sum_{k=0}^{K_i} Q(k) (\hat{L}_{i,k} - L_{i,k})^2 + \frac{\gamma}{N} \sum_{i=1}^N \sum_{k=0}^{K_i} Q(k) \hat{L}_{i,k, \text{var}}^2 \quad (20)$$

式中: $\hat{L}_{i,k} = E[L_{i,k}]$ 、 $\hat{L}_{i,k, \text{var}} = \text{var}[L_{i,k}]$ 分别为第 i 个随机退化设备在第 k 个监测周期的剩余寿命及剩余寿命方差; $(\hat{L}_{i,k} - L_{i,k})^2$ 为第 i 个随机退化设备在第 k 个监测周期预测剩余寿命与实际剩余寿命的均方误差, 表征剩余寿命预测的准确程度; γ 为调节因子, 表征决策者对剩余寿命预测准确程度与预测不确定性的重视程度, $\gamma = 0$ 表示只关注预测的准确程度, $\gamma = 1$ 表示只关注预测的不确定性; $Q(k)$ 为权重函数, 由于随着随机退化设备越接近失效时刻, 预测的准确程度越重要, 因此, $Q(k)$ 应为递增函数。

为保障全周期剩余寿命预测准确程度, 设计3种形式权重函数, 具体如下。

①指数函数:

$$Q(k) = ae^{bk} + c \quad (a > 0, b > 0, c \geq 0) \quad (21)$$

该权重函数具有非线性的快速增长趋势, 能在随机退化设备失效前为其赋予极高权重, 适合航空发动机、导弹武器等末期误差敏感的高危装备。但是需对 a 、 b 、 c 3个参数进行反复调节, 成本较高, 且由于弱化了前期权重, 忽略前期预测误差, 可能造成模型对随机退化设备的前期退化趋势拟合严重不足, 影响后期的剩余寿命预测。

②幂函数:

$$Q(k) = \left(\frac{k}{K_i} \right)^n \quad (n \geq 1) \quad (22)$$

增长速率由 n 进行调控, 相比于指数函数更为平缓, 在保证末期权重较大的同时也保留了一定

初期权重, 适合通用机械、工业泵阀等对初期和末期预测误差均较为敏感的设备。

③反比例函数:

$$Q(k) = \frac{1}{K_i - k + \varepsilon} \quad (\varepsilon > 0) \quad (23)$$

权重直接由剩余寿命 $K_i - k$ 决定, 物理可解释性最强, 且不需额外参数, 只需引入一个常数 ε , 避免分母为 0 即可, 数学形式最为简单。同时, 由于引入常数 ε , 使得 $0 \leq k \leq K_i$ 时, 权重增长趋势较为平缓, 避免了指数函数后期增长过快的缺点。

综上所述, 3 种权重函数各有优劣, 可根据随机退化设备的实际特征进行选择, 也可根据 3 种权重函数分别构建 3 种优化目标函数, 选取其中最优的预测效果作为最终目标函数。

2 服役设备的在线剩余寿命预测

通过历史退役设备可确定各传感器监测数据融合系数矩阵的估计值 $\mathbf{W}^*$ 及失效阈值的估计值 w^* , 但对在役随机退化设备进行在线剩余寿命预测时, 由于设备服役时期不同, 且在生产及运行过程中存在各种不确定因素, 需对式 (8) 中 3 个参数进行迭代更新。为便于研究, 将指数退化模型中的 β 作为随机变量, 设 $\beta \sim N(\mu_\beta, \sigma_\beta^2)$, μ_β 与 σ_β^2 分别为均值及方差。

采用序贯贝叶斯估计对参数 β 进行更新, 将上一时刻的后验估计作为下一时刻的先验信息, 利用全周期监测数据对参数进行更新, 假设已运行 k 个监测周期的随机退化设备的对数化复合健康指标为 $\mathbf{S}_{0:k} = [S_0, S_1, \dots, S_k]$ 。

根据序贯贝叶斯估计方法, 在 t_k 时刻 β 的后验分布为

$$\begin{aligned} p(\beta | \mathbf{S}_{0:k}) &= \frac{p(\beta, \mathbf{S}_{0:k})}{p(\mathbf{S}_{0:k})} = \\ &= \frac{p(S_k | \mathbf{S}_{0:k-1}, \beta) p(\beta | \mathbf{S}_{0:k-1}) p(\mathbf{S}_{0:k-1})}{p(\mathbf{S}_{0:k})} = \\ &= \frac{p(S_k | \mathbf{S}_{0:k-1}, \beta) p(\beta | \mathbf{S}_{0:k-1})}{p(S_k | \mathbf{S}_{0:k-1})} \propto \\ &= p(S_k | \mathbf{S}_{0:k-1}, \beta) p(\beta | \mathbf{S}_{0:k-1}) \end{aligned} \quad (24)$$

式中: $p(\beta | \mathbf{S}_{0:k-1})$ 为 t_{k-1} 时刻 β 的后验分布; $p(S_k | \mathbf{S}_{0:k-1}, \beta)$ 为 β 已知时 t_k 时刻复合健康指标的概率密度函数。

由于式 (8) 表示的随机过程是高斯的, 因此, $p(S_k | \mathbf{S}_{0:k-1}, \beta)$ 也是高斯分布, 具体可表示为

$$\begin{aligned} p(S_k | \mathbf{S}_{0:k-1}, \beta) &= \frac{1}{\sqrt{2\pi\sigma^2(t_k - t_{k-1})}} \cdot \\ &\cdot \exp\left\{-\frac{[S_k - S_{k-1} - \beta(t_k - t_{k-1})]^2}{2\sigma^2(t_k - t_{k-1})}\right\} \end{aligned} \quad (25)$$

由于假设 $\beta \sim N(\mu_\beta, \sigma_\beta^2)$, 因此, $p(\beta | \mathbf{S}_{0:k-1})$ 也是高斯的, 故 $p(S_k | \mathbf{S}_{0:k-1}, \beta)$ 与 $p(\beta | \mathbf{S}_{0:k-1})$ 共轭, 由式 (24) 可知 $p(\beta | \mathbf{S}_{0:k})$ 也是高斯的, 设 $\beta | \mathbf{S}_{0:k} \sim N(\mu_{\beta,k}, \sigma_{\beta,k}^2)$, $\mu_{\beta,k}$ 及 $\sigma_{\beta,k}^2$ 的值分别为

$$\mu_{\beta,k} = \frac{\sigma^2}{\sigma_{\beta,k-1}^2(t_k - t_{k-1}) + \sigma^2} \mu_{\beta,k-1} + \frac{\sigma_{\beta,k-1}^2(S_k - S_{k-1})}{\sigma_{\beta,k-1}^2(t_k - t_{k-1}) + \sigma^2} \quad (26)$$

$$\sigma_{\beta,k}^2 = \frac{\sigma^2 \sigma_{\beta,k-1}^2}{\sigma_{\beta,k-1}^2(t_k - t_{k-1}) + \sigma^2} \quad (27)$$

由历史退役设备的监测数据可得式 (26) 和 (27) 中 $\mu_{\beta,0}$ 及 $\sigma_{\beta,0}^2$ 的具体值为

$$\mu_{\beta,0} = \frac{1}{N} \sum_{i=1}^N \hat{\beta}_i \quad (28)$$

$$\sigma_{\beta,0}^2 = \frac{1}{N-1} \sum_{i=1}^N (\hat{\beta}_i - \mu_{\beta,0})^2 \quad (29)$$

进行在线剩余寿命预测时, 可根据历史退役设备的监测数据确定 θ 与 σ 的估计值分别为

$$\hat{\theta} = \frac{1}{N} \sum_{i=1}^N \hat{\theta}_i \quad (30)$$

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N \hat{\sigma}_i^2 \quad (31)$$

考虑到参数 β 的随机性, 可得 t_k 时刻随机退化设备剩余寿命的概率密度函数为

$$\begin{aligned} f_{L_k | \mathbf{S}_{0:k}}(l_k | \mathbf{S}_{0:k}) &= \int_{-\infty}^{+\infty} f_{L_k | \beta, \mathbf{S}_{0:k}}(l_k | \beta, \mathbf{S}_{0:k}) p(\beta | \mathbf{S}_{0:k}) d\beta = \\ &= \frac{w - S_k}{\sqrt{2\pi l_k^3 (\sigma_{\beta,k}^2 + \sigma^2)}} \cdot \exp\left\{-\frac{(w - S_k - \mu_{\beta,k} l_k)^2}{2l_k (\sigma_{\beta,k}^2 + \sigma^2)}\right\} \end{aligned} \quad (32)$$

为直观展现所提方法的总体实现路径, 给出剩余寿命预测原理如图 1 所示。

3 应用验证

3.1 传感器选择及数据预处理

为验证所提方法的有效性, 采用 C-MAPSS 商用航空发动机退化数据集集中的 FD001 数据集进行验证^[21]。由于 FD001 数据集中 100 台测试发动机的监测数据都并非全寿命周期监测数据, 选用 100 台训练发动机进行实验验证。其中前 60 台发动机作为历史退役设备进行离线训练, 后 40 台发

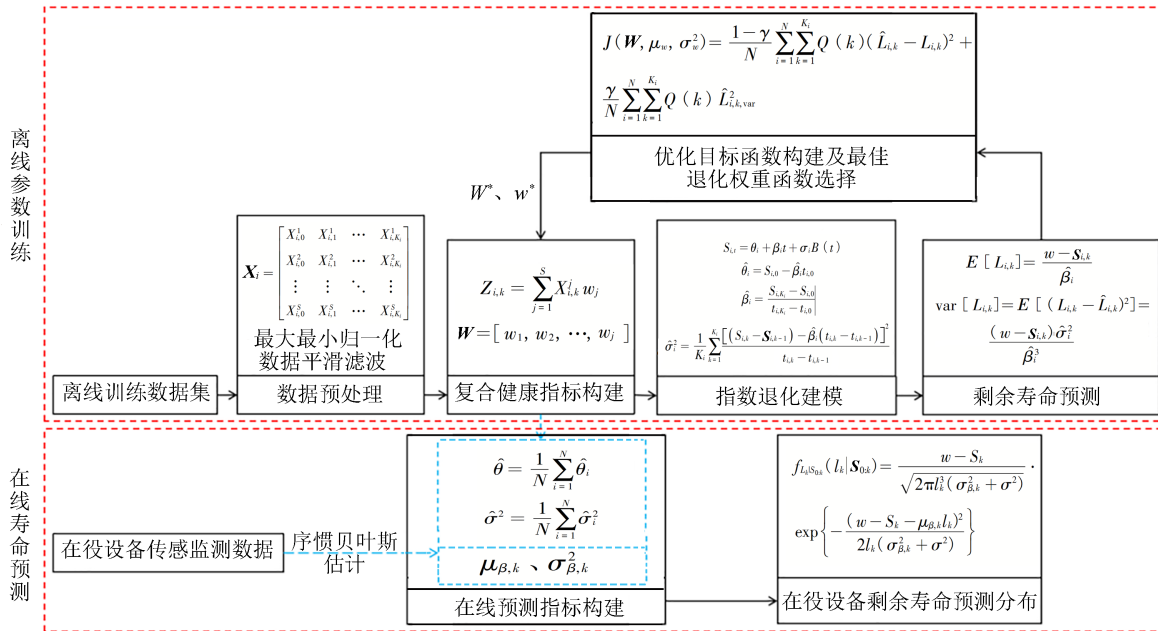


图1 剩余寿命预测原理

Fig. 1 Principle of RUL prediction

动机作为在役设备进行在线寿命预测。

每台发动机的监测数据都包括21个传感器的全寿命周期监测数据,由于部分传感器的监测数据并不具有时变趋势,无法反映随机退化设备的退化趋势,因此,需要对传感器进行筛选。

皮尔逊相关系数^[22]可反映两个变量之间的相关性,传感器监测数据的皮尔逊相关系数绝对值越大,表明其与监测时刻的相关性越强;同时,由于各台发动机的生产及运行工况相似,同一传感器在不同发动机上的皮尔逊相关系数方差应较小,否则,说明该传感器在运行过程中易受噪声干扰,反映的退化规律不可靠。表1给出了由100台发动机监测数据得到的各个传感器的皮尔逊相关系数平均值及方差,筛选出其中均值较大且一致性较好的10个传感器分别为T24、T30、T50、P30、Ps30、phi、BPR、htBleed、W31、W32。

对上述10个传感器的监测数据进行最大最小归一化及平滑时间窗口长度为10的平滑滤波处理,即可得到用于离线寿命预测的训练数据。

3.2 引入权重函数后的寿命预测离线训练

式(4)将10个传感器的监测数据融合为复合健康指标,将对数化的复合健康指标代入式(9)~(11)可得航空发动机指数退化模型参数 $\hat{\theta}_i$ 、 $\hat{\beta}_i$ 、 $\hat{\sigma}_i$,利用式(18)、(19)得到每台发动机在各个监测时刻的剩余寿命预测值及预测方差,采用五重交叉验证方法得到调节因子 γ 为0.25时皮尔逊相关系数平均值最高,设置调节因子 $\gamma = 0.25$ 。

表1 21个传感器信息

Table 1 Detailed description of 21 sensors

序号	标号	皮尔逊 相关系 数均值	皮尔逊 相关系 数方差	说明
1	T2	0	0	风机进口总温度
2	T24	0.675 3	0.003 5	低压压气机出口总温度
3	T30	0.644 0	0.003 3	高压压气机出口总温度
4	T50	0.781 6	0.001 4	低压涡轮出口总温度
5	P2	0	0	风机入口压力
6	P15	0	0	旁路管道总压力
7	P30	-0.761 5	0.003 5	高压压气机出口总压力
8	Nf	0.683 5	0.036 9	发动机风扇速度
9	Nc	0.386 0	0.321 7	发动机核心转速
10	epr	0	0	发动机压力比
11	Ps30	0.810 6	0.001 6	高压压气机出口静压
12	phi	-0.789 7	0.002 8	燃料流量与Ps30的比率
13	NRf	0.687 3	0.035 8	校正风扇转速
14	NRc	0.223 5	0.469 7	校正核心转速
15	BPR	0.724 8	0.002 4	旁路比率
16	farB	0	0	燃烧器燃料空气比
17	htBleed	0.673 1	0.003 1	出血焓
18	Nf.dmd	0	0	要求风扇速度
19	PCNfR	0	0	要求校正风扇速度
20	W31	-0.714 1	0.002 7	高压涡轮冷却剂排放
21	W32	-0.716 7	0.002 7	低压涡轮冷却剂排放

通过式 (20) 可得到最优的融合系数矩阵 W^* 与失效阈值 w^* 。对于此类优化问题, 已有较多成熟高效的求解算法, 通过 Matlab 中的 “fminsearch” 函数进行多维搜索求解。

图 2 为引入 3 种权重函数时预测寿命与实际寿命的对比结果, 可以看出, 引入反比例函数的寿命预测值与实际值吻合度较高, 这是因为反比例函数中并不包含额外参数, 使用 “fminsearch” 求解效率最高。

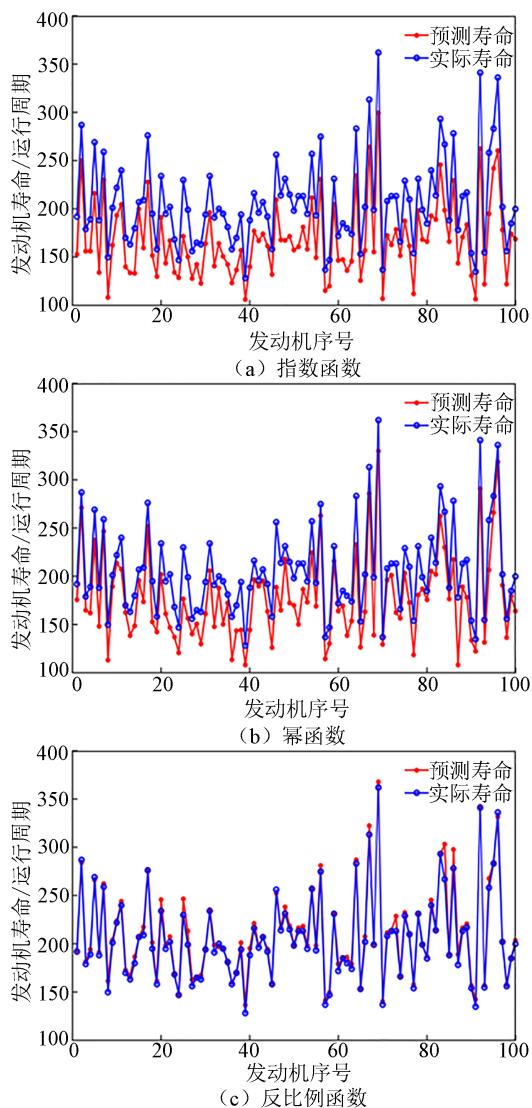


图 2 引入 3 种权重函数后的寿命离线预测效果

Fig. 2 Offline life prediction results with three weight functions

鉴于引入反比例函数后的寿命预测效果最好, 且其形式最为简单, 物理可解释性最强, 因此, 选择反比例函数 $Q(k) = \frac{1}{K_i - k + 1}$ 作为权重函数。随着监测次数的增加, $Q(k)$ 的值逐渐增大, 当 $k = K_i$ 时, $Q(k) = 1$ 。求解后的多维融合系数矩阵

$W^* = [0.91828, 0.25080, 0.72939, -0.33513, 0.33056, -0.39531, 0.46049, 0.08045, -0.44333, 0.48665]$, 失效阈值 $w^* = 0.78671$ 。为直观反映不同传感器数据对复合健康指标的影响, 给出复合健康指标的具体表达式为

$$Z_k = 0.91828X_k^2 + 0.2508X_k^3 + 0.72939X_k^4 - 0.33513X_k^7 + 0.33056X_k^{11} - 0.39531X_k^{12} + 0.46049X_k^{15} + 0.08045X_k^{17} - 0.44333X_k^{20} + 0.48665X_k^{21} \quad (33)$$

式中: X_k 为第 j 个传感器在第 k 个监测时刻归一化后的监测数据。

3.3 在役设备剩余寿命预测

运用序贯贝叶斯估计方法, 将第 61~100 台航空发动机的监测数据代入式 (32), 可得所有监测时刻剩余寿命的概率密度分布。为验证所提方法对在役设备进行寿命预测的效果, 随机选取第 69、88 台发动机作为在役设备。图 3 给出了两台在役设备预测剩余寿命与实际剩余寿命的对比结果。

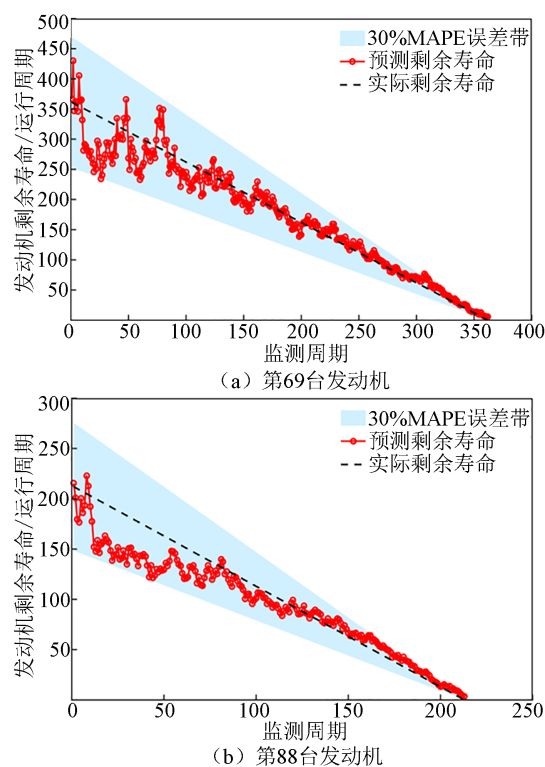


图 3 第 69、第 88 台发动机的剩余寿命在线预测效果

Fig. 3 Online RUL prediction results for Engines 69 and 88

可以看出, 随着监测周期的增多, 预测剩余寿命与实际剩余寿命的拟合程度逐渐提升。以剩余寿命真值为基准构建 30% 平均绝对百分比误差 (MAPE) 误差带, 用以界定相对误差 $\pm 30\%$ 的合理区间。两台发动机的全周期预测结果均稳定落

在该误差范围内，充分验证了所提方法对剩余寿命预测的准确度。

第88台发动机共有213组监测数据，图4给出了第88台发动机全周期预测剩余寿命的95%置信区间，可以看出，随着设备运行时长增加，预测置信区间持续收窄。在此基础上，图5给出了随机选取的4个监测周期的剩余寿命预测分布，4个概率密度函数峰值分别为0.005 4、0.007 7、0.010 1、0.136 5。可以看出，随着监测次数的增多，概率密度函数曲线呈现出愈发尖锐的趋势，表征预测的不确定性逐渐减小。

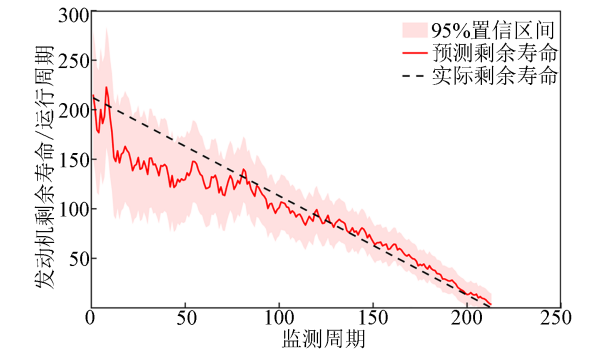


图4 第88台发动机预测剩余寿命的95%置信区间
Fig. 4 95% confidence interval of predicted RUL for Engine 88

为进一步验证所提方法的有效性，采用均方误差（mean squared error, MSE）作为性能比较指标以量化预测效果，其表达式如式（34）所示。

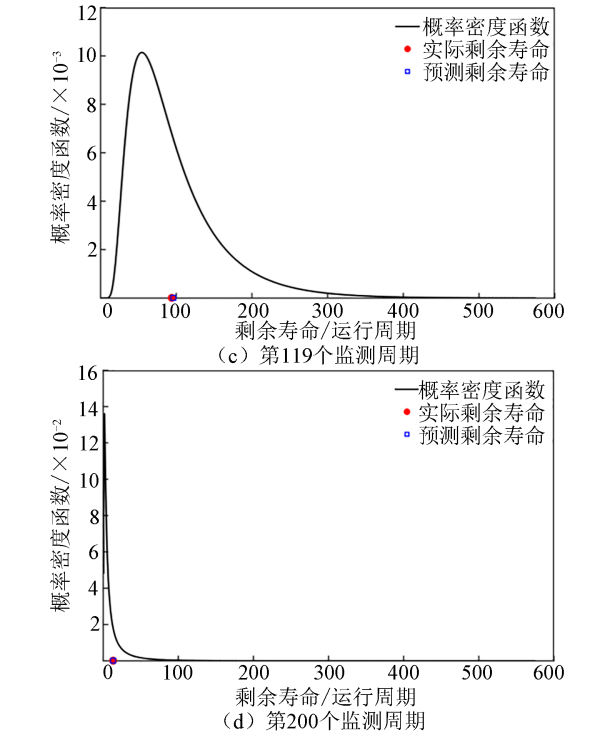
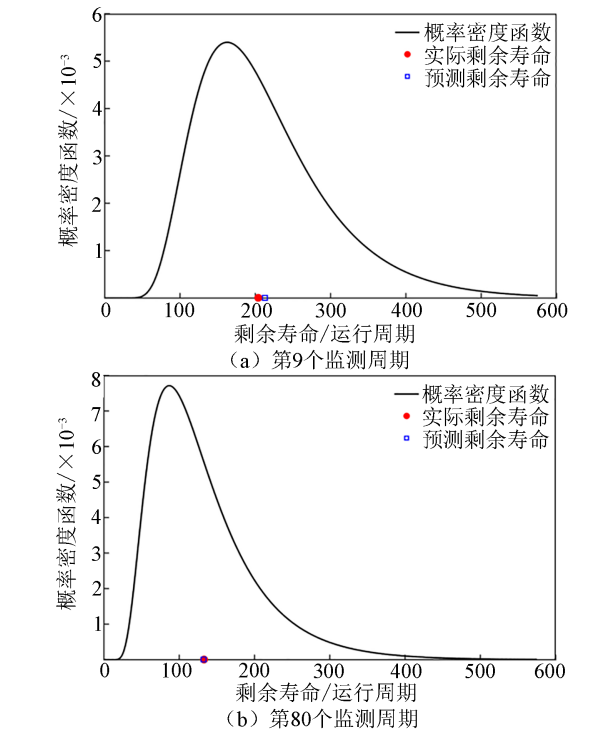


图5 第88台发动机在4个监测周期的剩余寿命预测分布
Fig. 5 Prediction distribution of RUL for Engine 88 across four monitoring cycles

$$M_{\text{MSE}} = \frac{1}{M} \sum_{i=1}^M (\hat{T}_i - T_i)^2 \quad (34)$$

式中： $\hat{T}_i$ 为第*i*个随机退化设备的预测寿命； T_i 为第*i*个随机退化设备的实际寿命； M 为随机退化设备的数量。

基于给出的预测指标，表2给出了本文方法与基于深度置信网络+后向传播神经网络+改进粒子滤波算法的方法^[23]、基于卷积神经网络+线性维纳过程的方法^[24]、基于加权融合+线性维纳过程+数据模型交互的方法^[25]的对比结果。可以看出，依据所提方法得到的寿命预测结果的MSE值最小，说明寿命预测的精确度最高。

表2 不同方法预测性能比较 Table 2 Comparison of prediction performance of different methods	
方法	MSE
基于深度置信网络+后向传播神经网络+改进粒子滤波算法的方法 ^[23]	283.00
基于卷积神经网络+线性维纳过程的方法 ^[24]	136.42
基于加权融合+线性维纳过程+数据模型交互的方法 ^[25]	78.27
本文方法	52.60

特别说明的是，文献[25]所示的方法同样为

数据与模型交互的方法,但并未考虑样本贡献度差异,本文方法的MSE值较之降低32.8%。由于本文方法为后期预测结果赋予了更高权重,随着监测周期不断接近失效时刻,剩余寿命的均方误差值将进一步降低。

4 结论

为进一步提升寿命预测精度,提出一种考虑样本贡献度差异的剩余寿命预测方法。相较于现有的基于机器学习和基于统计数据驱动的经典剩余寿命预测方法,文中所提方法可有效降低设备退化后期预测误差与预测方差,在时序退化趋势拟合与不确定性量化层面表现出更为突出的优势,其寿命预测均方根误差相较于传统数模联动方法降低32.8%,全周期剩余寿命预测平均相对误差控制在30%以内,能够更好适配航空发动机、高速列车等复杂装备的退化演化特性,有效达到预设的高精度寿命预测目标。

尽管所提方法通过考虑样本贡献度有效降低了后期预测的偏差及方差,仍有以下问题值得进一步研究。首先,不同随机退化设备的退化机理不尽相同,如何高效地从形式多样的函数中选取满足预测需求的权重函数有待进一步研究;其次,构建目标函数时主要考虑了预测准确性和预测不确定性,并为两者选取了相同的权重函数,若引入更多表征预测效果的指标,并根据需求为各个指标选取不同的权重函数,有望进一步提升预测精度;再次,现有方法在处理强噪声监测数据时,无法有效识别偏差数据,仍为其赋予较高贡献度,需要研究如何优化偏差数据权重分配方法;最后,所提方法仅是针对数模联动框架下单一失效模式的随机退化设备剩余寿命预测方法,如何将其推广到多工况、多失效模式的情形有待进一步研究。

参考文献

- [1] 胡姚刚,李辉,廖兴林,等.风电轴承性能退化建模及其实时剩余寿命预测[J].中国电机工程学报,2016,36(6):1643-1649.
Hu Yaogang, Li Hui, Liao Xinglin, et al. Performance degradation model and prediction method of real-time remaining life for wind turbine bearings[J]. Proceedings of the CSEE, 2016, 36(6): 1643-1649.
- [2] 田贵双,王少萍,石健.考虑多元件退化相关的列车牵引系统可靠性评估与寿命预测[J/OL].北京航空航天大学学报,(2024-01-10)[2026-02-15].
<https://doi.org/10.13700/j.bh.1001-5965.2023.0797>.
- [3] 陆宁云,陈闯,姜斌,等.复杂系统维护策略最新研究进展:从视情维护到预测性维护[J].自动化学报,2021,47(1):1-17.
Lu Ningyun, Chen Chuang, Jiang Bin, et al. Latest progress on maintenance strategy of complex system: from condition-based maintenance to predictive maintenance[J]. Acta Automatica Sinica, 2021, 47(1): 1-17.
- [4] Liu Xinwei, Zhang Zongzhen, Li Zhouli, et al. Advancements in bearing health monitoring and remaining useful life prediction: techniques, challenges, and future directions[J]. Measurement Science and Technology, 2025, 36(3): 032003.
- [5] 蔡志强,胡昌华,王兆强,等.改进Transformer架构的混合时频域增强分解模型的重大装备剩余寿命预测[J].火箭军工程大学学报,2025,39(3):76-89.
Cai Zhiqiang, Hu Changhua, Wang Zhaoqiang, et al. Improved transformer architecture-based hybrid time-frequency domain enhanced decomposition model of useful life prediction of major equipment[J]. Journal of Rocket Force University of Engineering, 2025, 39(3): 76-89.
- [6] 成心怡,郑建飞,张琪,等.面向多传感器耦合数据健康指标构建的装备剩余寿命预测方法[J].火箭军工程大学学报,2025,39(4):11-21.
Cheng Xinyi, Zheng Jianfei, Zhang Qi, et al. Remaining useful life prediction of equipment based on health index construction from multi-sensor coupled data[J]. Journal of Rocket Force University of Engineering, 2025, 39(4): 11-21.
- [7] 韩其辉,杨立浩,郑建飞,等.基于空间重构与并行化建模的滚动轴承剩余寿命预测方法[J].火箭军工程大学学报,2025,39(5):127-136.
Han Qihui, Yang Lihao, Zheng Jianfei, et al. Spatial reconstruction and parallelizing modeling-based remaining useful life prediction method of rolling bearings[J]. Journal of Rocket Force University of Engineering, 2025, 39(5): 127-136.
- [8] Ben Ali J, Chebel-Morello B, Saidi L, et al. Accurate bearing remaining useful life prediction based

- on Weibull distribution and artificial neural network[J]. *Mechanical Systems and Signal Processing*, 2015, 56/57: 150-172.
- [9] Benkedjouh T, Medjaher K, Zerhouni N, et al. Remaining useful life estimation based on nonlinear feature reduction and support vector regression[J]. *Engineering Applications of Artificial Intelligence*, 2013, 26(7): 1751-1760.
- [10] Ferreira C, Calves G. Remaining useful life prediction and challenges: a literature review on the use of machine learning methods[J]. *Journal of Manufacturing Systems*, 2022, 63: 550-562.
- [11] El-Dalahmeh M, Al-Greer M, El-Dalahmeh M, et al. Physics-based model informed smooth particle filter for remaining useful life prediction of lithium-ion battery[J]. *Measurement*, 2023, 214: 112838.
- [12] Lui Yuhui, Li Meng, Downey A, et al. Physics-based prognostics of implantable-grade lithium-ion battery for remaining useful life prediction[J]. *Journal of Power Sources*, 2021, 485: 229327.
- [13] Xu Pengcheng, Lei Yaguo, Wang Zidong, et al. A self-data-driven approach for online remaining useful life prediction of machinery using a recursive update strategy[J]. *Mechanical Systems and Signal Processing*, 2025, 229: 112541.
- [14] 雷亚国, 贾峰, 孔德同, 等. 大数据下机械智能故障诊断的机遇与挑战[J]. *机械工程学报*, 2018, 54(5): 94-104.
- Lei Yaguo, Jia Feng, Kong Detong, et al. Opportunities and challenges of machinery intelligent fault diagnosis in big data era[J]. *Journal of Mechanical Engineering*, 2018, 54(5): 94-104.
- [15] 冯磊, 张正新, 李天梅, 等. 数据驱动的闭环控制系统剩余寿命预测方法综述[J]. *中北大学学报(自然科学版)*, 2024, 45(1): 1-11.
- Feng Lei, Zhang Zhengxin, Li Tianmei, et al. A review of data-driven remaining useful life prediction methods for closed-loop control system[J]. *Journal of North University of China(Natural Science Edition)*, 2024, 45(1): 1-11.
- [16] 李天梅, 司小胜, 刘翔, 等. 大数据下数模联动的随机退化设备剩余寿命预测技术[J]. *自动化学报*, 2022, 48(9): 2119-2141.
- Li Tianmei, Si Xiaosheng, Liu Xiang, et al. Data-model interactive remaining useful life prediction technologies for stochastic degrading devices with big data[J]. *Acta Automatica Sinica*, 2022, 48(9): 2119-2141.
- [17] Li Zhen, Wu Jianguo, Yue Xiaowei. A shape-constrained neural data fusion network for health index construction and residual life prediction[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(11): 5022-5033.
- [18] Li Huiqin, Zhang Zhengxin, Si Xiaosheng. A dual-purpose data-model interactive framework for multi-sensor selection and prognosis[J]. *Reliability Engineering & System Safety*, 2025, 258: 110904.
- [19] Li Tianmei, Pei Hong, Si Xiaosheng, et al. Prognosis for stochastic degrading systems with massive data: a data-model interactive perspective[J]. *Reliability Engineering & System Safety*, 2023, 237: 109344.
- [20] Pei Hong, Si Xiaosheng, Li Tianmei, et al. Interactive prognosis framework between deep learning and a stochastic process model for remaining useful life prediction[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 35(12): 18000-18012.
- [21] Wang Cunsong, Lu Ningyun, Cheng Yuehua, et al. A data-driven aero-engine degradation prognostic strategy[J]. *IEEE Transactions on Cybernetics*, 2021, 51(3): 1531-1541.
- [22] 纪德洋, 金锋, 冬雷, 等. 基于皮尔逊相关系数的光伏电站数据修复[J]. *中国电机工程学报*, 2022, 42(4): 1514-1523.
- Ji Deyang, Jin Feng, Dong Lei, et al. Data repairing of photovoltaic power plant based on pearson correlation coefficient[J]. *Proceedings of the CSEE*, 2022, 42(4): 1514-1523.
- [23] Peng Kaixiang, Jiao Ruihua, Dong Jie, et al. A deep belief network based health indicator construction and remaining useful life prediction using improved particle filter[J]. *Neurocomputing*, 2019, 361: 19-28.
- [24] Wen Long, Su Shaoquan, Wang Bin, et al. A new multi-sensor fusion with hybrid convolutional neural network with wiener model for remaining useful life estimation[J]. *Engineering Applications of Artificial Intelligence*, 2023, 126: 106934.
- [25] Li Tianmei, Si Xiaosheng, Pei Hong, et al. Data-model interactive prognosis for multi-sensor monitored stochastic degrading devices[J]. *Mechanical Systems and Signal Processing*, 2022, 167: 108526.

Remaining Useful Life Prediction Method Incorporating Sample Contribution Difference under Data-Model Linkage Framework

LUO Zhengyang, PEI Hong*, ZHENG Jianfei,

HAN Qihui, CHEN Xi, SI Xiaosheng

(Rocket Force University of Engineering, Xi'an 710025, China)

ABSTRACT: To address the insufficient fitting accuracy of late-stage equipment degradation features and excessive prediction errors of remaining useful life (RUL) under existing data-model linkage frameworks, a novel improved RUL prediction method incorporating sample contribution differences was proposed. Multiple weight functions were embedded to revise the conventional objective function, assigning a higher priority to late-stage prediction performance during model optimization. In the offline training phase, multi-sensor monitoring data were first integrated into composite health indicators using multidimensional fusion coefficients. Second, given the nonlinear characteristics of equipment degradation, an evolution model of composite health indicators was established, and the analytical formula of the RUL was derived based on the first-passage time theory. Finally, an objective function combining the prediction bias and variance was constructed to obtain the optimal fusion coefficients and failure thresholds. In the online prediction phase, sequential Bayesian estimation was utilized for iterative model parameter updating, thereby enabling real-time RUL prediction. Comparative validation experiments conducted on aircraft engine cases demonstrate that, compared to traditional data-model linkage RUL prediction methods, the proposed approach reduces the root mean square error by 32.8% and maintains an average relative error within 30% across the entire life cycle. It achieves higher accuracy in late-stage RUL predictions, providing a valuable reference for fault diagnosis and life prediction of industrial equipment, such as aircraft engines.

KEYWORDS: remaining useful life prediction; nonlinear degradation model; data-model linkage; sample contribution

(编辑: 于福兴)