



基于ISCSO-LSTM网络集成的卫星动量轮剩余使用寿命预测

于牧野¹, 徐文青¹, 钱振¹, 符方舟^{1*}, 王大轶²

1. 中山大学航空航天学院, 深圳 518107

2. 北京空间飞行器总体设计部, 北京 100094

* E-mail: fufzh@mail.sysu.edu.cn

收稿日期: 2025-11-30; 接受日期: 2026-02-11; 网络版发表日期: 2026-03-12

国家自然科学基金(批准号: U22B6001, 62473389, 62388101)和广东省自然科学基金(编号: 2025B1515020097, 2024A1515011730)资助项目

摘要 作为姿态控制系统的关键执行器, 动量轮性能退化会直接影响卫星寿命, 有必要对其开展剩余使用寿命(remaining useful life, RUL)预测研究. 针对动量轮RUL预测精度不足、预测性能不稳定的问题, 本文建立了一种基于改进沙猫群优化(improved sand cat swarm optimization, ISCSO)算法与长短期记忆(long short-term memory network, LSTM)网络相结合的集成预测方法. 首先从动量轮原始轴承温度序列中提取多维候选特征, 采用基于单调性与趋势性的综合评估方法筛选敏感特征; 随后构建多个具有不同初始化参数的LSTM网络作为基预测器, 以增强模型泛化能力; 最后提出基于Tent混沌映射的ISCSO算法, 以优化各个基预测器的输出权重. 实验结果表明, 所提方法的多种评价指标结果均优于对比方法, 消融实验进一步验证了改进优化算法与集成策略的有效性.

关键词 剩余使用寿命预测, 动量轮, 长短期记忆网络, 集成策略, 改进沙猫群优化

1 引言

近年来, 国际电信联盟“先登先占、先到先得”的规则加剧了近地轨道及频段资源的竞争, 促使主要航天国家与私营公司加速低轨星座的建设, 商业航天迅速扩张, 低轨星座集群规模空前^[1,2]. 在这种背景下, 若星座卫星发生故障或者出现性能衰退, 不仅会影响任务连续性, 还可能引发在轨碰撞风险.

面对这一挑战, 传统基于固定周期或事后响应的

运维模式已难以满足在轨可靠性需求^[3]. 故障预测与健康维护技术通过综合运用系统的先验知识、状态信息与运行数据, 提前识别退化趋势并对剩余使用寿命(remaining useful life, RUL)进行预测, 能够大幅降低设备的维护成本与运行风险^[4-6]. 作为卫星姿态控制系统的关键执行器, 动量轮长期连续工作在真空和温度波动的复杂环境中, 容易因内部轴承的磨损与润滑劣化而发生性能退化. 此外, 据文献^[7]统计, 动量轮等姿态控制系统部件故障在卫星全寿命周期故障中占有显

引用格式: 于牧野, 徐文青, 钱振, 等. 基于ISCSO-LSTM网络集成的卫星动量轮剩余使用寿命预测. 中国科学: 技术科学, 2026, 56: 666–680
Yu M Y, Xu W Q, Qian Z, et al. Remaining useful life prediction of satellite momentum flywheel based on ISCSO-LSTM network ensemble (in Chinese). Sci Sin Tech, 2026, 56: 666–680, doi: [10.1360/SST-2025-0386](https://doi.org/10.1360/SST-2025-0386)

著比例. 因此, 对卫星动量轮进行RUL预测至关重要.

目前, RUL预测主要分为基于物理模型的方法和数据驱动的方法^[8]. 基于物理模型的RUL预测方法通过精确描述系统部件退化与失效的内在物理规律, 建立反映其工作状态的数学模型(如Paris Law模型^[9]和指数模型^[10]), 以此实现RUL的预测^[11]. 然而, 基于物理模型的方法虽然能取得良好的RUL预测效果, 但其性能严重依赖于高精度数学模型的构建. 建立高精度的动量轮退化模型存在参数获取困难、建模复杂度高等技术瓶颈, 而简化的数学模型又往往难以完整刻画所有潜在的退化模式与随机故障机理, 这使得构建合理且有效的物理模型变得极为困难.

相比之下, 数据驱动的方法无需构建复杂的物理模型, 直接从设备运行的历史监测数据中学习退化规律. 尤其是基于深度学习的方法, 凭借其强大的表征学习能力, 能够自动从高维、非平稳的监测数据中提取与退化相关的特征. 这降低了对先验知识的依赖, 使RUL预测过程更为灵活高效, 因此基于深度学习的RUL预测方法近年来受到了广泛关注与研究. 其具体方法主要可以分为基于卷积神经网络(convolutional neural network, CNN)的方法和基于循环神经网络(recurrent neural network, RNN)的方法^[12,13]. 例如, Chae等^[14]提出了一种融合去噪卷积神经网络(denoising convolutional neural network, DnCNN)与CNN的混合架构, 用于锂离子电池健康状态估计与RUL预测. 该方法通过短时傅里叶变换将传感器信号转换为时频谱图, 并利用DnCNN进行噪声抑制, 在多个实验场景下展现出优于传统去噪方法的估计精度和计算效率, 改善了RUL预测的效果. 然而, CNN卷积核的感受野有限, 导致其对全局时间维度上的动态演变规律建模能力不足, 限制了RUL预测性能的进一步提升. 相比之下, RNN则在处理时序数据方面具有独特的优势. Boujamza和Elhaq^[15]通过引入膨胀卷积结构增强RNN对长序列依赖的建模能力, 同时结合特征选择与超参数优化, 有效提升了航空发动机RUL的预测精度. 门控循环单元(gate recurrent unit, GRU)和长短期记忆神经网络(long short-term memory, LSTM)网络及其各自的变体通过引入门控机制, 有选择地加入新信息与遗忘旧信息, 进一步改善了RNN的长程依赖问题^[16-18]. Baruah和Organero^[19]提出了一种上下文集成循环神经网络, 通过将显式上下文特征以权重调节方式动态融入

GRU, 有效提升了航空发动机在多变工况下的RUL预测精度. 然而, 对于退化过程持续时间更长、时序依赖更复杂的场景, GRU易因记忆衰减而难以充分捕捉长期依赖关系. 相比之下, LSTM网络的门控结构更精细, 在处理长期依赖关系时表现更加稳定^[20]. Wu等^[21]通过构建深度LSTM网络, 利用多传感器时序信号实现了设备RUL的准确预测. 其预测结果在Score指标与RUL误差范围上均优于多种现有先进方法, 验证了深度LSTM网络在多传感器融合与长序列依赖建模中具有显著优势. 然而, 尽管基于单一LSTM网络的预测方法已在RUL任务中展现出良好的性能, 但由于深度学习方法的非凸优化特性, 单一LSTM网络的预测性能受初始参数影响显著, 致使不同的预测结果之间存在较大波动, 稳定性不佳^[22]. 集成策略通过构建多个基预测器并融合其输出, 能够有效平均掉这种随机性, 显著提升预测结果的稳定性, 该方法已在风力发电预测和航天器故障诊断等领域得到了成功验证^[23-25]. 然而, 如何确定各基预测器的最优融合权重, 是提升集成性能的关键. 近年来, 元启发式优化算法因其强大的全局寻优能力, 被广泛应用于神经网络超参数调优与模型融合中, 为求解此类复杂优化问题提供了有效工具^[26,27]. 例如, Chu等^[24]提出了一种集成深度双向GRU网络, 并利用灰狼优化(grey wolf optimizer, GWO)算法搜索多个分类器的最优集成权重, 显著提升了故障分类的精度, 这为利用优化算法增强集成模型性能提供了有力借鉴.

受上述研究启发, 针对卫星动量轮RUL预测精度不足与性能不稳定等问题, 本文提出了一种基于改进沙猫群优化(improved sand cat swarm optimization, ISCSO)-LSTM网络集成的卫星动量轮RUL预测方法, 主要贡献可归纳如下: (1) 构建了基于改进ISCSO-LSTM网络集成框架, 通过多LSTM基预测器集成预测, 有效提升了动量轮RUL预测的准确性与稳定性. (2) 设计了一种结合特征构造与优选的退化特征增强方法, 通过引入多种时域指标, 构造了能突出动量轮性能退化信息的候选特征. 同时基于单调性与趋势性的多指标评估体系, 实现了对退化敏感特征的精准优选, 为LSTM网络集成提供了更具代表性的输入特征. (3) 在标准沙猫群优化(sand cat swarm optimization, SCSO)算法中改进种群初始化操作, 提升了基预测器权重分配的优化效果, 增强了集成的泛化能力.

2 问题描述与预备知识

2.1 问题描述

动量轮是卫星姿态控制系统的核心执行部件, 其性能会随着持续运行逐渐退化. 为避免灾难性后果, 对其RUL进行精确的预测至关重要. 本文旨在通过分析动量轮运行过程中的监测数据, 构建数据驱动模型, 以估计其在发生功能性故障前的剩余运行时长. 剩余使用寿命定义为从当前观测时刻到失效时刻的时间区间:

$$t_R = t_E - t_C, \quad (1)$$

式中, t_R 为剩余使用寿命的值, 单位为天数; t_E 为精确的失效时刻, t_C 为任意当前观测时刻 ($t_C < t_E$). 在工程实践中, t_E 是未知的, 因此RUL是一个需要通过预测模型进行估计的隐变量. 假设 $\mathbf{X}_{1:t_C} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{t_C}\}$ 为从运行初始时刻 ($t=1$)至当前时刻 t_C 所采集到的时间序列数据, 则动量轮的RUL预测问题可表述为一个基于时间序列的回归问题, 其目标是学习一个最优的映射函数, 能够将历史观测数据映射到对应的RUL预测值. 该函数可简化为

$$f(\mathbf{X}_{1:t_C}; \Theta) = \hat{t}_R, \quad (2)$$

式中, f 为映射函数; $\hat{t}_R$ 为RUL预测值; Θ 为预测模型的待定参数集合. 因此, 实现高精度预测的核心在于构建能够有效实现该映射关系的数据驱动模型.

2.2 LSTM网络

LSTM网络是一种专门处理时间序列数据的神经网络结构, 它通过独特的内部门控机制, 有效克服了传统RNN中梯度消失或爆炸的问题, 具备卓越的长期记忆能力, 特别适合处理具有长期依赖关系的时序数据.

如图1所示, LSTM网络的基本单元包括遗忘门 f_t 、输入门 i_t 和输出门 o_t , 每个门控单元通过对过去信息的记忆和过滤来决定如何更新当前时刻的状态. LSTM网络中的三个门的计算方式更新公式如下:

$$\begin{cases} f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \\ i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \\ o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \end{cases} \quad (3)$$

式中, $\sigma(\cdot)$ 为sigmoid函数, 其输出区间为(0, 1); x_t 为当

前单元的输入; h_{t-1} 为上一单元的隐状态; W_* 为输入门、遗忘门和输出门各自对应的权重矩阵, $* \in \{f, i, o\}$; b_* 为输入门、遗忘门和输出门各自对应的偏置.

输入门 i_t 通过与候选状态 $\tilde{C}_t$ 进行元素乘积运算来控制当前单元的信息有多少需要被保存, 而遗忘门 f_t 则通过与上一单元的内部状态进行元素乘积运算来控制上一单元的信息有多少需要被遗忘. 因此, 当前单元内部状态更新的计算方式表示为

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t. \quad (4)$$

而候选状态 $\tilde{C}_t$ 由上一单元的隐状态和当前单元的输入信息构成, 即

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C), \quad (5)$$

式中, $\tanh(\cdot)$ 为双曲正切激活函数, 其输出区间为(-1, 1); W_C 为候选状态的权重矩阵; b_C 为候选状态的偏置.

输出门 o_t 控制 C_t 有多少信息需要输出给隐状态 h_t , 因此LSTM网络当前单元的最终输出为

$$h_t = o_t \odot \tanh(C_t). \quad (6)$$

如式(3)~(6)所示, LSTM网络通过输入门、遗忘门和输出门的精确控制, 自适应地调控信息流动与记忆更新, 从而将输入的历史观测数据 x_t 映射到未来状态. 这一系列门控操作使得LSTM天然适合用于构建式(2)所描述的映射函数 f , 并能有效刻画动量轮退化过程中的非线性与时序依赖特性.

2.3 SCSO算法

SCSO算法是一种模拟沙猫狩猎行为的元启发式算法, 其核心思想是将优化问题视作在搜索空间中寻找最优解的过程, 通过参数自适应调节机制协调全局探索与局部开发之间的平衡^[28]. SCSO算法将优化问题的每个潜在解视为一只沙猫的位置, 通过迭代更新种群位置来寻找最优解, 其流程主要包括种群初始化、搜索猎物、攻击猎物以及搜索-攻击平衡机制四个部分. 下面对SCSO算法的主要步骤进行说明.

在SCSO算法中, 沙猫群体的规模设为 m , 每只沙猫的位置由向量 $\mathbf{P} = (p_1, p_2, \dots, p_e)$ 表示, 其中 e 代表优

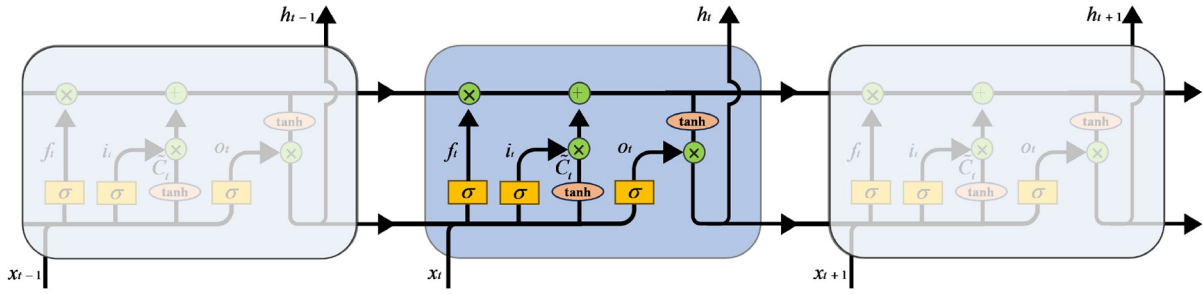


图 1 (网络版彩图) LSTM网络单元基本结构

Figure 1 (Color online) Basic structure of LSTM network unit.

化问题的维度. 算法首先在搜索空间内随机初始化种群位置:

$$p_{i,j} = l_{b_j} + rand \cdot (u_{b_j} - l_{b_j}), \quad (7)$$

式中, $p_{i,j}$ 为第 i 只沙猫的第 j 维变量; $i = 1, 2, \dots, m$; $j = 1, 2, \dots, e$; u_{b_j} 和 l_{b_j} 分别为问题第 j 维的上界和下界; $rand$ 为 $0 \sim 1$ 之间的随机数.

沙猫搜索行为依赖于其对低频信号的感知能力, 其感知范围由参数 r_G 控制. 随着沙猫逐渐接近猎物, r_G 的值从 2 线性递减至 0:

$$r_G = S_M - \left(\frac{S_M \cdot D_c}{D_d} \right), \quad (8)$$

式中, S_M 为沙猫的听觉特性参数, 默认值为 2; D_c 为当前迭代次数; D_d 为最大迭代次数.

在搜索猎物阶段, 沙猫根据以下公式更新位置:

$$\mathbf{P}(t+1) = r \cdot [\mathbf{P}_{bc}(t) - rand \cdot \mathbf{P}_c(t)], \quad (9)$$

式中, $r = r_G \cdot rand$ 为沙猫的灵敏度范围; $\mathbf{P}_{bc}(t)$ 为当前最优位置; $\mathbf{P}_c(t)$ 为当前个体位置.

在攻击猎物阶段, 沙猫在圆形搜索区域内进行局部开发:

$$\begin{cases} \mathbf{P}_{md} = |rand \cdot \mathbf{P}_b(t) - \mathbf{P}_c(t)|, \\ \mathbf{P}(t+1) = \mathbf{P}_b(t) - r \cdot \mathbf{P}_{md} \cdot \cos(\theta), \end{cases} \quad (10)$$

式中, $\mathbf{P}_{md}$ 为沙猫与猎物之间的距离; θ 为 $0^\circ \sim 360^\circ$ 之间的随机角度.

为实现搜索阶段与攻击阶段之间的动态平衡, 算法引入关键控制参数 R :

$$R = 2 \cdot r_G \cdot rand - r_G, \quad (11)$$

$$\mathbf{P}(t+1) = \begin{cases} \mathbf{P}_b(t) - r \cdot \mathbf{P}_{md} \cdot \cos(\theta) & |R| \leq 1, \\ r \cdot [\mathbf{P}_{bc}(t) - rand \cdot \mathbf{P}_c(t)] & |R| > 1. \end{cases} \quad (12)$$

参数 R 能够有效协调两种行为模式的转换, 显著提升算法的收敛效率, 即当 $|R| \leq 1$ 时, 算法自动切换至攻击猎物模式; 当 $|R| > 1$ 时, 保持搜索猎物模式.

ISCSO 算法通过其独特的参数自适应调节与行为切换机制, 实现了在解空间中进行高效的全局探索与局部开发. 这一优化特性使其能够自动、精准地找到多个 LSTM 预测器之间的最优权重分配, 从而为构建高性能的集成预测方法提供关键支撑.

3 基于 ISCSO-LSTM 网络集成的 RUL 预测

3.1 预测框架概述

本文提出的 ISCSO-LSTM 集成预测方法, 其核心框架如图 2 所示, 主要包含四个关键环节. 首先, 对监测参数退化数据进行特征构建与优选, 提取出与退化过程强相关的退化特征, 形成高质量的集成输入序列; 其次, 并行训练多个 LSTM 基预测器, 基于输入序列独立输出未来时刻的监测参数退化值. 训练结束后, 利用 ISCSO 算法对各基预测器进行最优权重分配. 最后, 设定失效阈值, 通过滚动预测机制完成动量轮 RUL 的预测. 下面将依次对特征构造与优选、LSTM 训练、ISCSO 算法优化集成权重及完整 RUL 预测流程进行详细阐述.

3.2 特征构造与优选

直接从原始监测数据中学习映射关系面临噪声干扰与模型复杂度高的挑战. 因此, 本节致力于从原始数

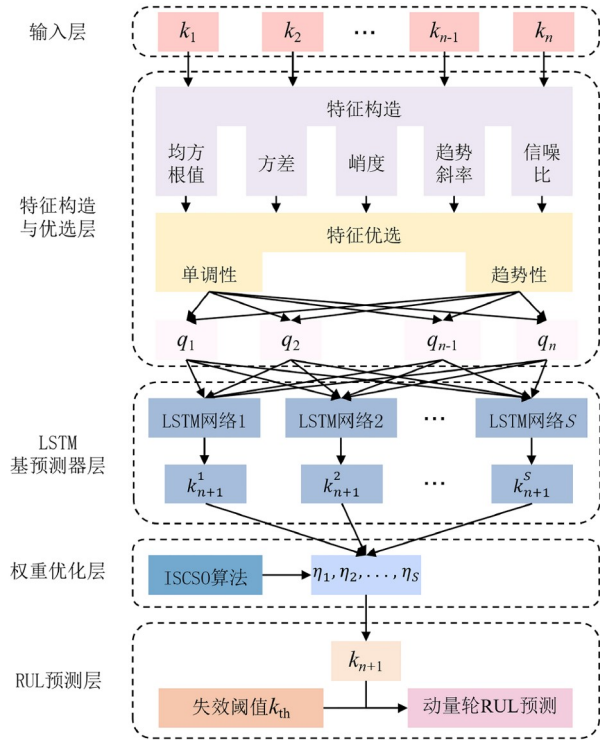


图2 (网络版彩图)基于ISCSO-LSTM网络集成的动量轮RUL预测框架

Figure 2 (Color online) Prediction framework of momentum flywheel RUL based on ISCSO-LSTM network ensemble.

据中构造并筛选出一组能够清晰表征动量轮退化过程的特征, 为后续的LSTM网络提供高质量的输入。

3.2.1 退化特征构造

构造退化特征有助于将与动量轮寿命退化相关的信息显式化, 使预测方法能在有限样本下更快捕获与退化过程直接相关的信息, 从而提高训练效率与泛化能力。已有研究表明, 时域统计特征在健康监测中具有良好的退化敏感性, 被广泛应用于轴承、齿轮及旋转设备的寿命预测任务中^[29-31]。均方根和方差可分别从能量水平和波动强度反映整体退化趋势; 峭度对早期局部损伤引入的冲击成分敏感; 趋势斜率可量化性能衰退的长期单调性; 信噪比则可以表征有效信号成分与背景噪声的强度对比, 反映信号质量的退化程度。这些特征能够从不同角度刻画信号幅值分布、波动程度与时间相关性等信息, 共同构建出对动量轮退化状态的多维度表征。因此, 本节选取这5种时域指标作为刻画退化趋势的候选特征。为将全局时序信息转化为

适用于滚动预测的局部动态特征, 本节所有特征均基于滑动窗口进行局部计算。定义特征构造的滑动窗口长度为 N , k_i 为第 i 个采样点的信号幅值, 5种候选特征的公式如下。

(1) 均方根

$$\text{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N k_i^2}. \quad (13)$$

(2) 方差

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N \left(k_i - \frac{1}{N} \sum_{i=1}^N k_i \right)^2. \quad (14)$$

(3) 峭度

$$K = \frac{\frac{1}{N} \sum_{i=1}^N \left(k_i - \frac{1}{N} \sum_{i=1}^N k_i \right)^4}{\left(\frac{1}{N} \sum_{i=1}^N \left(k_i - \frac{1}{N} \sum_{i=1}^N k_i \right)^2 \right)^2}. \quad (15)$$

(4) 趋势斜率

$$a = \frac{N \sum_{i=1}^N i \cdot k_i - \sum_{i=1}^N i \sum_{i=1}^N k_i}{N \sum_{i=1}^N i^2 - \left(\sum_{i=1}^N i \right)^2}. \quad (16)$$

(5) 信噪比

$$\text{SNR} = 10 \log_{10} \left(\frac{\frac{1}{N} \sum_{i=1}^N k_i^2}{\frac{1}{N} \sum_{i=1}^N (k_i - \bar{k})^2} \right). \quad (17)$$

3.2.2 退化特征优选

然而, 并非所有候选特征都同等地适用于RUL预测。此外, 过多的特征会导致计算量增大, 增加计算时间; 过少的特征则不能完全反映监测参数退化趋势, 导致预测结果较差。为量化每个候选特征与退化进程的相关性, 本节提出一种基于多指标评估的候选特征优选方法。具体而言, 首先对各候选特征的单调性^[32]与趋势性^[33]进行量化, 然后基于上述两种评估指标, 计算每个特征的综合评分, 最终筛选出综合评分靠前的特征子集作为LSTM网络的输入。

单调性能衡量特征值随时间呈现单向变化趋势的程度, 是退化特征的重要属性。其公式为

$$M(Q^j) = \frac{\left| \sum_{i=1}^{n-1} \text{sign}(q_{i+1}^j - q_i^j) \right|}{n-1}, \quad (18)$$

式中, $Q^j = q_1^j, q_2^j, \dots, q_n^j$ 为第 j 种特征的时间序列; i 为时序索引; $\text{sign}(\cdot)$ 为符号函数; n 为序列长度. 单调性 $M(Q^j)$ 的取值范围为 $[0, 1]$, 值越接近 1, 表明该特征随退化进程单调变化的能力越强; 值越接近 0, 则表明该特征变化方向不定, 与退化进程的关联性较弱.

趋势性能够表征特征随时间呈现稳定变化趋势的程度, 其公式为

$$T(Q^j) = \frac{n \sum_{i=1}^n i \cdot q_i^j - \left(\sum_{i=1}^n i \right) \left(\sum_{i=1}^n q_i^j \right)}{\sqrt{\left[n \sum_{i=1}^n i^2 - \left(\sum_{i=1}^n i \right)^2 \right] \left[n \sum_{i=1}^n (q_i^j)^2 - \left(\sum_{i=1}^n q_i^j \right)^2 \right]}}. \quad (19)$$

趋势性 $T(Q^j)$ 的取值范围为 $[-1, 1]$, 绝对值越接近 1, 表明该特征的退化趋势越明显且稳定.

考虑到在动量轮 RUL 预测任务中, 退化特征的单调递增或递减特性对于准确判定失效阈值至关重要, 而趋势性则进一步刻画了特征变化的稳定性与一致性, 有助于识别出具有明显退化规律的特征. 值得注意的是, 尽管熵权法等自适应权重方法能根据训练数据动态分配权重, 但其结果易受训练集统计波动影响, 可能使特征筛选过程过度适应训练数据而损害泛化性能. 固定权重作为一种轻量正则化手段, 可以有效避免这一过拟合风险, 进而确保特征选择准则的稳定性. 因此本文采用固定权重策略, 同时为了突出单调性在退化特征筛选中的主导作用, 将特征综合评分函数的公式定义为

$$\text{Score}(Q^j) = 0.6 \times M(Q^j) + 0.4 \times T(Q^j). \quad (20)$$

最终, 通过上述特征构造与优选流程, 能够得到一个可清晰反映动量轮性能退化规律的低维特征序列, 为后续构建高精度预测模型提供了可靠输入.

3.3 LSTM网络训练

在获得低维特征序列后, 为了从序列中学习退化动态规律, 需要一个能够有效捕捉长期依赖关系的预测模型. 如 2.2 节所述, LSTM 网络在此方面具有天然优势. 因此, 本节将详细阐述 LSTM 网络的训练方法.

3.3.1 训练样本构建

本节采用滑动窗口方法构建训练样本. 设优选后的退化特征序列为 $\mathbf{Q} = [q_1, q_2, \dots, q_n]$, 其中 $q_t \in \mathbb{R}^d$ 为 t 时刻的 d 维特征向量. 给定时间步长 L , 第 i 个训练样本的输入为前 L 个时间步的特征序列 $\mathbf{Q}_i^{\text{train}} = [q_i, q_{i+1}, \dots, q_{i+L-1}]$, 对应的输出为下一时间步的特征值 $\mathbf{y}_i^{\text{train}} = q_{i+L}$. 这种单步预测方法不仅使得训练过程收敛更稳定, 还可以将长序列预测任务分解为可迭代的短期任务, 便于后续递推得到长期演化曲线.

3.3.2 网络训练与优化

为了确保 LSTM 网络能够充分学习动量轮退化过程的时序特征并具备良好的泛化能力, 本文采用系统化的训练策略. 在损失函数选择方面, 考虑到 RUL 预测本质上是一个回归问题, 因此采用均方误差作为网络的优化目标, 其公式为

$$\mathcal{L}(\Theta) = \frac{1}{N} \sum_{i=1}^N (\hat{y}_{t_i} - y_{t_i})^2 + \lambda \|\Theta\|_2^2, \quad (21)$$

式中, Θ 为模型全部参数集合; N 为样本总个数; $\hat{y}_{t_i}$ 为第 i 个样本的预测值; y_{t_i} 为第 i 个样本的真实值; λ 为 L_2 正则项系数. 该函数能够有效衡量预测特征值与真实值之间的偏离程度. 此外, 针对深度学习模型容易过拟合的问题, 本文在 LSTM 层间引入 Dropout 机制, 强制网络学习更具鲁棒性的特征表示. 网络训练则采用自适应矩估计优化 (adaptive moment estimation, Adam) 算法, 通过反向传播和梯度下降迭代更新网络参数 Θ , 直至损失函数收敛. 训练完成的网络具备根据过去 L 个时间步的特征数据预测下一时刻特征值的能力.

通过上述流程, 能够得到一个训练完毕、具备单步预测能力的 LSTM 网络. 该网络是实现动量轮 RUL 预测的核心模块, 其性能也是后续所有预测任务的基础.

3.4 基于 ISCSO 的集成权重优化

尽管单个 LSTM 网络已能实现基本的 RUL 预测功能, 但单一网络的性能易受其初始随机权重的影响, 预测结果随机波动, 导致其泛化能力与稳定性受限. 此外, 相较于需独立调试大量超参数, 且组合训练过程难以稳定的异构网络结构基预测器, 同构 LSTM 网络

通过不同的随机初始化, 在深度神经网络高度非凸的优化过程中足以产生丰富的预测多样性^[22,24], 能够有效保证方法的简洁性. 鉴于此, 本文采用集成策略, 构建一组结构相同但初始参数各异的LSTM基预测器. 由于基预测器集成性能的核心关键在于如何实现各基预测器预测结果的最优融合, 为此, 本节提出利用ISCSO算法为各基预测器分配集成权重, 以此构建一个高性能的集成预测框架, 确保RUL预测的稳定性和泛化能力.

3.4.1 集成输出建模

为了能够将多个LSTM基预测器的输出进行有效融合, 本节建立集成输出的数学模型. 设集成预测框架中包含 S 个LSTM网络, 第 s 个网络在 t 时刻的预测输出记为 $\hat{x}^s(t)$. 集成预测的基本形式为加权求和:

$$x_e(t) = \sum_{s=1}^S \eta_s \hat{x}^s(t), \quad (22)$$

式中, η_s 为第 s 个LSTM网络的融合权重, 满足 $\sum_{s=1}^S \eta_s = 1$ 且 $\eta_s \geq 0$ 的约束条件.

3.4.2 ISCSO算法

由2.3节可知, SCSSO算法通过其独特的搜索、攻击及切换机制, 展现了良好的优化特性. 然而, 与许多群智能算法一样, 其初始种群采用随机生成, 可能会出现多样性不足和收敛速度较慢的问题. 混沌作为一种非线性的自然现象, 以其混沌序列具有遍历性、随机性等优点, 被广泛用于优化搜索问题^[34]. 为了能够让种群的位置均匀分布在整个问题的解空间内, 本文提出了一种改进的SCSSO算法, 即利用Tent混沌映射对种群进行初始化. 令混沌序列 z_k 满足Tent映射:

$$z_{k+1} = \begin{cases} \mu z_k, & 0 \leq z_k < \frac{1}{2}, \\ \mu(1 - z_k), & \frac{1}{2} \leq z_k \leq 1, \end{cases} \quad (23)$$

式中, k 为混沌序列的索引; μ 取2得到满射的混沌序列(初始 $z_0 \in (0, 1)$ 且非特殊周期点). 用Tent序列作为均匀扰动生成初代沙猫个体可表示为

$$p_{i,j}^0 = l_{b_j} + z_k \cdot (u_{b_j} - l_{b_j}), \quad k = (i-1)e + j, \quad (24)$$

式中, $p_{i,j}^0$ 为第 i 只沙猫初始位置的第 j 维分量; z_k 为由式(24)迭代生成的Tent混沌序列; u_{b_j} 和 l_{b_j} 的定义见式(7).

将上式作为SCSSO算法的种群初始化步骤, 替换原始的随机初始化, 能够提高和改善初始种群的多样性和质量, 加快其收敛速度.

3.4.3 基于ISCSO的权重优化模型

本节旨在构建基于ISCSO的权重优化模型, 对基预测器集成的最优权重系数 η_s 进行求解. 该模型将每个沙猫个体的位置编码为一组候选的权重向量, 并通过优化适应度函数来自动寻找最优的权重分配方案. 具体实现过程如下.

(1) 编码策略

将集成权重优化问题映射至ISCSO算法的搜索空间. 对于一个包含 S 个LSTM基预测器的集成框架, 每个沙猫个体的位置向量 $\mathbf{P} = (p_1, p_2, \dots, p_e)$ 直接表征了一组候选的集成权重 $\boldsymbol{\eta} = (\eta_1, \eta_2, \dots, \eta_S)$. 搜索空间的下界 l_b 和上界 u_b 均设置为 S 维的0和1向量, 以确保权重初始值在合理范围内.

(2) 适应度函数设计

适应度函数用于评估一组权重分配的优劣, 其设计直接决定优化方向. 本文以提高集成模型在验证集上的预测精度为目标, 适应度函数选择均方根误差, 其计算方式同式(13), 用于衡量集成预测值与真实值之间的偏差. ISCSO算法的优化目标即最小化该适应度函数的值.

(3) 约束处理

为确保权重满足式(22)的约束条件 $\sum_{s=1}^S \eta_s = 1$ 且 $\eta_s \geq 0$, 在计算适应度前, 需对每个沙猫位置向量进行归一化处理:

$$\eta_s = \frac{\max(x_s, 0)}{\sum_{j=1}^S \max(x_j, 0)}. \quad (25)$$

此处理能有效将任意位置向量映射为合理的权重向量.

3.4.4 权重优化流程

基于ISCSO的权重优化流程如下.

(1) 设置ISCSO算法参数, 包括种群规模 m 、最大迭代次数 D_d , 并利用Tent混沌映射初始化沙猫种群.

(2) 对于种群中的每个个体, 按上述方法进行解码和归一化, 得到权重向量 η .

(3) 使用该权重向量对 S 个LSTM基预测器在验证集上的输出进行加权融合, 并根据适应度函数公式计算适应度值.

(4) 运行ISCSO算法, 根据2.3节与3.4.2节所述的位置更新机制, 迭代优化沙猫种群.

(5) 当达到最大迭代次数时, 输出全局最优位置 X^* , 经解码和归一化后, 即得到最优集成权重向量

$$\eta = (\eta_1^*, \eta_2^*, \dots, \eta_s^*).$$

通过上述流程, 本节将LSTM集成权重的分配问题转化为一个可由ISCSO算法高效求解的数值优化问题, 可确保集成预测框架能够以理论最优的方式融合各基预测器的输出, 为最终实现精确的RUL预测提供核心保障.

3.5 RUL预测流程

综合前文所述的特征重构与优选、LSTM网络训练、ISCSO优化权重与集成策略, 本文所提出的动量轮RUL预测方法形成了一个系统性的解决方案, 其完整工作流程总结如算法1所示.

4 仿真验证

为了全面评估基于ISCSO-LSTM网络集成的动量轮RUL预测方法的性能, 本节首先根据动量轮轴承温度模型获取退化数据, 其次给出算法参数设置与方法评价指标, 并对退化数据进行特征构造与特征优选, 最后通过与其他两种预测方法对比分析以及消融实验, 验证本文方法在预测精度和鲁棒性方面的优势.

4.1 数据准备与实验设置

动量轮轴承在持续运行过程中因摩擦作用会产生显著温升, 由于轴承温度的渐进变化能够有效表征动量轮的退化状态, 因此本文选用轴承温度作为分析动量轮退化特征的依据. 然而, 动量轮作为典型的长寿命产品, 具有小子样、无失效数据的特点, 现有故障数据较少, 故障数据难以获得. 为了获取动量轮轴承

算法1: 基于ISCSO-LSTM集成的动量轮RUL预测算法

输入: 原始监测数据序列 $[k_1, k_2, \dots, k_n]$; 失效阈值 k_{th} ; 滑动窗口长度 L ; LSTM网络数量 S ; ISCSO算法种群规模 m ; 最大迭代次数 D_d

输出: 动量轮的剩余使用寿命 t_R

```

1: // 阶段一: 特征构造与优选
2: 根据3.2.1节, 从 $[k_1, k_2, \dots, k_n]$ 中计算候选特征集
3: 根据3.2.2节, 计算各特征的单调性 $M(Q^j)$ 与趋势性 $T(Q^j)$ 
4: 根据综合评分 $Score(Q^j)$ 筛选最优特征子集, 构成特征序列
    $Q = [q_1, q_2, \dots, q_n]$ 
5: // 阶段二: LSTM网络训练
6: 使用 $Q$ 与滑动窗口 $L$ , 根据3.3.1节构建训练集与验证集
7: for  $s=1$  to  $S$  do 并行训练 $S$ 个LSTM基预测器
8:   随机初始化第 $s$ 个LSTM网络的权重
9:   根据3.3.2节所述策略(损失函数、Dropout机制等)训练网络
10:  在验证集上得到网络的预测序列 $[\hat{x}^s(t)]$ 
11: end for
12: // 阶段三: 基于ISCSO的集成权重优化
13: 初始化ISCSO算法(采用3.4.2节的Tent混沌映射)
14: While  $D_c \leq D_d$  do 开始迭代优化
15:   for 每个沙猫个体 $P = (p_1, \dots, p_e)$  do
16:     根据3.4.3节的编码与归一化策略, 将其解码为权重向量 $\eta$ 
17:     使用 $\eta$ 加权融合 $S$ 个网络的验证集预测结果 $[\hat{x}^s(t)]$ 
18:     计算集成预测的均方根误差作为适应度值
19:   end for
20:   更新ISCSO算法中的全局最优位置 $X^*$ 和个体位置
21:    $D_c = D_c + 1$ 
22: end While
23: 输出最优权重向量 $\eta^*$ 
24: // 阶段四: RUL预测
25: 将当前时刻的优选特征序列 $Q_{current}$ 输入至 $S$ 个已训练的LSTM网络
26: 使用最优权重 $\eta^*$ 对 $S$ 个网络的输出进行加权融合, 得到下一时刻特征的集成预测值 $k_{n+1}$ 
27: 将 $k_{n+1}$ 加入序列, 滚动预测 $k_{n+2}$ , 以此类推, 直至预测值 $\hat{k}_{n+T} \geq k_{th}$ 
28: return  $t_R = T - T$ 为从当前时刻到首次达到阈值的预测时间步长

```

温度数据, 本节采用文献[35]中提出的用于模拟动量轮轴承退化过程的具有线性漂移的维纳过程模型. 该模型能够有效表征轴温的两个关键特征: 一是由长期磨损导致的温度线性上升趋势; 二是由太阳入射角变化、功耗波动等随机因素引起的温度波动. 模型的表

达式为

$$Y(t) = a_0 + a_1 t + \sigma_w W(t), \quad (26)$$

式中, a_0 为动量轮轴温初始值; a_1 为动量轮轴温漂移系数, 反映轴承磨损速率; $W(t)$ 为标准维纳过程; σ_w 为扩散系数, 与 $W(t)$ 共同描述随机环境因素对温度的影响; $Y(t)$ 为动量轮 t 时刻的轴温。

考虑到动量轮个体间的性能差异, 本文将模型参数 a_0 和 a_1 设为随机变量, 分别服从正态分布 $N(\mu_{a_0}, \sigma_{a_0}^2)$ 和 $N(\mu_{a_1}, \sigma_{a_1}^2)$ 。同时, 参考文献[35]中的温度数据参数, 并结合式(26)的温度退化模型, 生成动量轮轴温退化数据, 如图3所示共计3650个样本点。本节将前1200个样本作为建模数据集, 按8:1:1的比例将数据集划分为训练集、验证集和测试集, 剩余样本作为外推动量轮RUL的预测集。

在实验参数设置方面, 本节基于最小预测误差准则[36], 滑动窗口大小设为15, LSTM基预测器个数设为5, 训练过程中则采用Adam优化算法对网络进行训练, 初始学习率设置为0.001。为提高训练效率并保证梯度更新的稳定性, 批处理大小设置为128。同时, 将Drop-out随机失活比例设置为0.2以防止过拟合。

4.2 评价指标

为了评估所提方法的预测性能, 本文采用多维度评价指标体系, 具体包括均方误差(mean square error,

MSE)、平均绝对误差(mean absolute error, MAE)、平均绝对百分比误差(mean absolute percentage error, MAPE)以及决定系数 R^2 。这些指标分别反映了预测值与真实值之间的差异程度、接近程度及拟合优劣情况。其中, MSE通过平方损失函数放大较大误差的影响, 反映预测结果的整体精度, 数值越低表示预测精度越高; MAE量化了预测结果与真实值之间的绝对偏差程度, 该指标数值越小表明网络的拟合优度越高; MAPE通过相对误差的百分比表示形式, 消除了数据尺度的影响, 能够更直观地反映预测精度水平, 其数值降低直接对应着预测准确性的提升; R^2 则能够反映预测值与真实值间的线性关系强度, 取值范围在0~1之间, 接近1时表示网络拟合度高。4种评价指标的计算公式分别如下:

$$\text{MSE} = \frac{1}{n} \sum_{k=1}^n (x_k - \hat{x}_k)^2, \quad (27)$$

$$\text{MAE} = \frac{1}{n} \sum_{k=1}^n |x_k - \hat{x}_k|, \quad (28)$$

$$\text{MAPE} = \frac{1}{n} \sum_{k=1}^n \left| \frac{x_k - \hat{x}_k}{x_k} \right| \times 100\%, \quad (29)$$

$$R^2 = 1 - \frac{\sum_{k=1}^n (x_k - \hat{x}_k)^2}{\sum_{k=1}^n (x_k - \bar{x})^2}, \quad (30)$$

式中, x_k 为序列中第 k 个真实值; $\hat{x}_k$ 为序列中第 k 个预测值; n 为序列长度; $\bar{x}$ 为所有真实值的平均值。

4.3 数据处理

在特征构造阶段, 式(13)~(17)中的特征构造窗口长度 N 直接影响特征的敏感性: 较小的 N 使特征对近期变化敏感, 但抗噪性弱; 较大的 N 能平滑噪声、反映长期趋势, 但会引入响应滞后。本文基于最小预测误差准则[36]将 N 的值定为5, 以在敏感性与稳定性间取得平衡。从原始一维温度时间序列中构造5类候选特征如图4所示, 这些特征能够从不同角度描述动量轮的退化状态, 有效弥补了单一特征带来的局限性。然而, 并非所有特征对RUL预测都很重要。因此, 本文采用基于单调性与趋势性的多指标评估方法对候选特征进行优选, 通过式(18)~(20)计算各特征的综合评分函数值。特征优选的结果如表1所示, 最终选取综合评分最高的均方

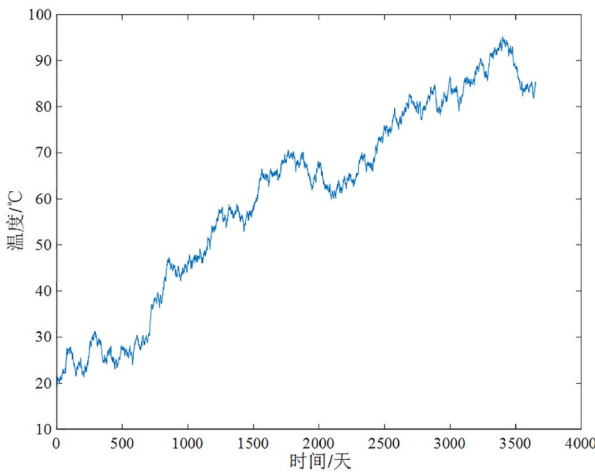


图3 (网络版彩图)动量轮轴温退化仿真数据

Figure 3 (Color online) Simulation data of temperature degradation of momentum flywheel.

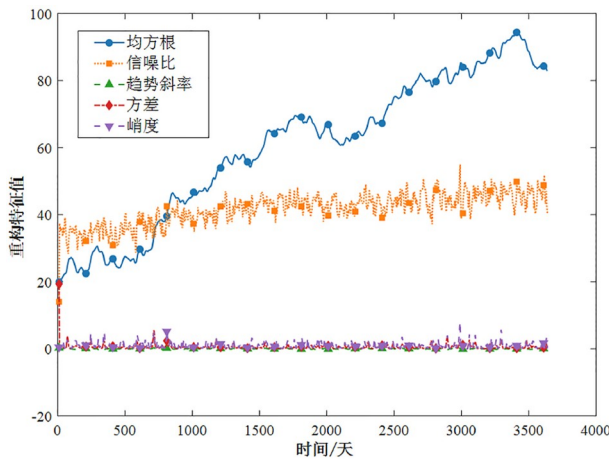


图 4 (网络版彩图) 重构特征值

Figure 4 (Color online) The value of reconstruct feature.

表 1 特征优选结果

Table 1 Results of feature selection

候选特征	单调性	趋势性	综合评分	排名
均方根	0.160	0.945	0.474	1
信噪比	0.008	0.610	0.249	2
趋势斜率	0.021	0.008	0.016	3
方差	0.017	0.005	0.012	4
峭度	0.008	0.000	0.005	5

根、信噪比与趋势斜率三个时域指标, 以其计算出的时序数据作为最终的特征子集。

4.4 对比实验

为验证本文提出的ISCSO-LSTM网络集成方法在卫星动量轮RUL预测中的有效性, 本节选取向量自回归(vector autoregression, VAR)模型、GRU网络、基于粒子群优化(particle swarm optimization, PSO)的LSTM (PSO-LSTM)集成与GWO-LSTM集成预测方法作为对比基准。各方法在测试集上的预测结果如图5所示。

由图5可见, 本文提出的ISCSO-LSTM网络集成方法所得到的预测曲线与真实值最为接近, 在整个测试集上都保持了良好的预测性能。相比之下, VAR模型的预测曲线波动较大, 且存在明显的相位偏差, 表明模型存在欠拟合问题; GRU网络虽然能够捕捉总体趋势, 但对于退化后期的预测也出现了较大的幅度误差, 反映了其在捕捉温度序列动态特性方面的不足。值得

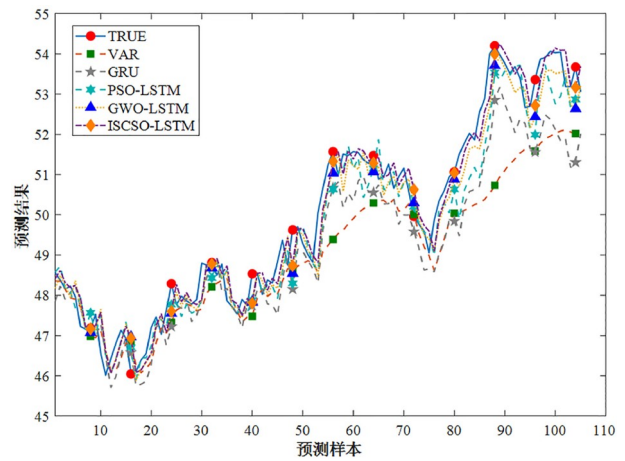


图 5 (网络版彩图) 对比实验预测值与真实值对比

Figure 5 (Color online) Comparison between the predicted and true values in comparative experiment.

注意的是, 同样采用LSTM集成框架的PSO-LSTM网络集成方法与GWO-LSTM网络集成方法, 其预测效果显著优于VAR模型与GRU网络, 证实了“LSTM集成+智能优化权重”策略的有效性。然而, 二者的预测曲线仍与真实值存在可见偏差, 且稳定性不及ISCSO-LSTM网络集成方法, 体现了ISCSO算法在求解权重优化问题上的优越性。

表2展示了各方法的预测性能指标对比结果。可以看出, ISCSO-LSTM网络集成方法在MSE、MAE、MAPE和 R^2 四个评价指标上均取得最优值。其MSE为0.2050, 相比VAR模型、GRU网络、PSO-LSTM和GWO-LSTM网络集成方法分别降低了87.3%、79.6%、47.7%和24.7%。ISCSO-LSTM网络集成方法的MAE和MAPE分别为0.3605和0.0072, 表明其在误差控制方面表现最佳。相比之下, PSO-LSTM和GWO-LSTM网络集成方法虽优于VAR模型和GRU网络, 但仍不及ISCSO-LSTM网络集成方法; GRU表现稍显逊色, VAR模型表现最差, 其MAE和MAPE高达0.9870和0.0192。此外, ISCSO-LSTM网络集成方法的 R^2 值也达到了0.9632, 接近理想值1, 显著高于另外4种方法, 这一结果充分证明了本文所提方法在预测精度方面具有明显的优势。

考虑到温度过高可能导致动量轮轴承过度磨损和润滑失效, 同时参考同类动量轮轴承温度管理的工程研究, 当轴承温度持续超过70℃~80℃时, 其润滑性能

表 2 对比实验预测方法性能对比

Table 2 Performance comparison of prediction method in comparative experiment

预测方法	MSE	MAE	MAPE	R^2
VAR	1.6199	0.9870	0.0192	0.7092
GRU	1.0064	0.8192	0.0161	0.8193
PSO-LSTM	0.3925	0.4804	0.0101	0.9295
GWO-LSTM	0.2723	0.4290	0.0086	0.9511
ISCSO-LSTM	0.2050	0.3605	0.0072	0.9632

显著下降, 标志着性能退化进入加速阶段^[35]. 因此, 本文取该范围中间值75℃作为功能失效的判定阈值, 将第1201天作为外推起点, 采用滚动预测机制对动量轮轴温进行外推, 当预测温度达到失效阈值时判定为动量轮失效. 如表3所示, 各方法在RUL预测方面表现出显著差异. ISCSO-LSTM网络集成方法准确捕捉了动量轮轴温关键区间的非线性升温过程, 预测达到75℃失效阈值时的时间与真实寿命仅相差37天, 相比GRU网络的127天误差降低了70.9%. 而VAR模型作为线性模型存在固有局限性, 无法准确捕捉温度非线性演化规律, 最终因未能触及失效阈值而无法输出RUL预测值.

4.5 消融试验

为深入分析ISCSO-LSTM网络集成中各模块的作用, 本文设计消融实验验证单LSTM网络、SCSO-LSTM网络集成方法和ISCSO-LSTM网络集成方法的预测性能差异. 首先, 为验证ISCSO算法改进本身的有效性, 图6对比了权重优化阶段ISCSO算法与原始SCSO算法的适应度收敛曲线. 可以观察到, ISCSO算法的收敛曲线始终位于下方且下降更快, 表明其凭借Tent混沌映射改进的初始化策略, 能以更高的效率找到更优的权重解, 从优化过程上解释了最终集成性能提升的原因. 图7展示了消融实验各方法在测试集上的预测效果对比. ISCSO-LSTM网络集成方法表现最优, 其预测曲线与真实值高度吻合, 即使在温度快速变化的复杂阶段也能保持准确的预测效果. SCSO-LSTM网络集成方法虽在某些局部波动处仍存在轻微的预测偏差, 但在整个测试集上都能较好地跟踪真实温度变化趋势. 相比之下, 单LSTM网络虽在初始阶段的预测表现尚可, 但后期预测值逐渐偏离真实值, 显示

表 3 对比实验动量轮RUL预测误差

Table 3 Prediction error of momentum flywheel RUL in comparative experiment

预测方法	真实值 (天)	预测值 (天)	预测误差 (天)
VAR	2495	—	—
GRU	2495	2368	127
PSO-LSTM	2495	2417	78
GWO-LSTM	2495	2443	52
ISCSO-LSTM	2495	2532	37

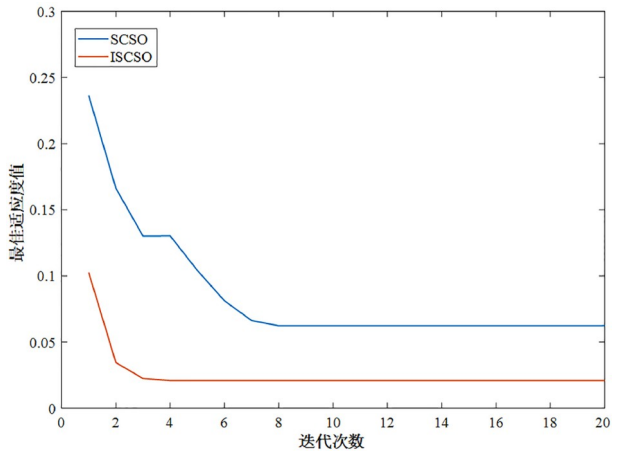


图 6 (网络版彩图) ISCSO与SCSO算法适应度曲线对比
Figure 6 (Color online) Comparison of fitness curves between ISCSO and SCSO algorithms.

出其预测能力的不足.

表4给出了消融实验的四种评价指标结果. 从结果可以看出, 随着模型复杂度的增加, 各项性能指标呈现出明显的递进改善趋势. 单一LSTM网络作为基线方法, 其MSE为0.8204, MAE为0.7186, MAPE为0.0140, R^2 的值达到了0.8527, 反映了其在动量轮轴温预测中的基本性能水平. 当引入标准的SCSO算法与集成策略后, SCSO-LSTM网络集成方法的MSE降低至0.2591, 降幅达68.4%, R^2 也从0.8527提升至0.9535, 这表明SCSO算法与集成策略相结合在提升预测精度方面具有显著效果. 然而, SCSO-LSTM网络集成方法的MAPE为0.0082, 说明其在某些局部时间段仍存在较大的相对误差. 相比之下, ISCSO-LSTM网络集成方法在所有指标上均达到最优, 其MSE进一步降至0.2050, 相比SCSO-LSTM网络集成方法又改善了20.9%. 对于MAE、MAPE和 R^2 三种指标, SCSO-LSTM网络集成

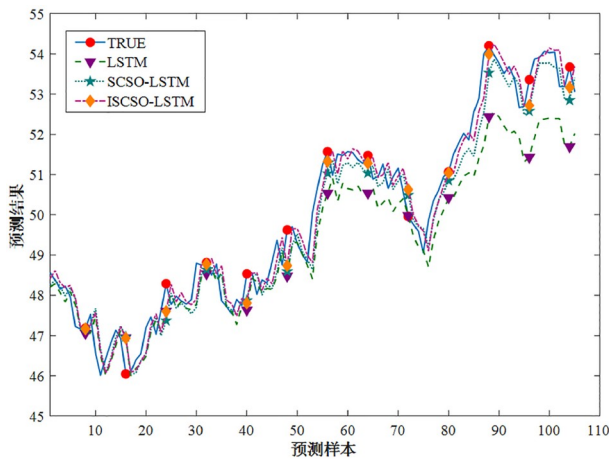


图 7 (网络版彩图)消融实验预测值与真实值对比

Figure 7 (Color online) Comparison between the predicted and true values in ablation experiment.

表 4 消融实验预测方法性能对比

Table 4 Performance comparison of prediction method in ablation experiment

预测方法	MSE	MAE	MAPE	R^2
LSTM	0.8204	0.7186	0.0140	0.8527
SCSO-LSTM	0.2591	0.4103	0.0082	0.9535
ISCSO-LSTM	0.2050	0.3605	0.0072	0.9632

方法的结果也均有提升,表明ISCSO算法能够找到更优的权重组合,进一步提升LSTM网络集成的预测精度。

表5给出了消融实验中各方法从第1201天开始进行轴温外推的RUL预测值及其误差。单一LSTM网络通过更复杂的门控结构和细胞状态机制,其预测误差为95天。SCSO-LSTM网络集成方法则通过集成策略有效缓解了单一LSTM网络在外推预测中的随机误差累积,将预测误差从95天降至66天。而ISCSO-LSTM网络集成方法仅产生37天的预测误差,将预测误差降低了43.9%,这一结果证明了ISCSO算法与集成策略的结合,能够通过抑制误差传播,在长时序外推预测中实现更稳定、更精确的RUL预测。

表 5 消融实验动量轮RUL预测值及其误差

Table 5 RUL prediction values and errors of momentum flywheel in ablation experiment

预测方法	真实值 (天)	预测值 (天)	预测误差 (天)
LSTM	2495	2590	95
SCSO-LSTM	2495	2429	66
ISCSO-LSTM	2495	2532	37

5 结论与展望

本文提出了一种基于ISCSO-LSTM网络集成的动量轮RUL预测方法。该方法通过构建多个结构相同但初始权重不同的LSTM网络作为基预测器,并引入ISCSO算法实现基预测器的最优权重分配,形成了具有强泛化能力的集成预测框架。在数据处理阶段,首先从原始轴温退化序列中提取了5种候选特征,构建了多维特征数据集。随后,构建了一种基于单调性与趋势性多指标评估的特征优选方法,通过构建综合评分函数筛选出对退化过程最敏感的候选特征。在RUL预测阶段,先将特征优选后的数据作为ISCSO-LSTM网络集成方法的输入,预测未来时刻的轴温。随后,设定失效阈值,通过滚动预测机制完成动量轮RUL的预测。为验证所提方法的有效性,本文采用动量轮轴温仿真数据进行了系统验证。实验结果表明,ISCSO-LSTM网络集成方法在预测精度和稳定性方面均表现出显著优势。在对比实验中,该方法在MSE、MAE、MAPE和 R^2 四个评价指标上均达到最优值,RUL预测结果显著优于对比方法。消融实验结果表明,改进的优化算法与集成策略均能提升预测性能,其中ISCSO算法通过优化基预测器的权重分配,进一步增强了所提方法的预测精度与稳定性。

尽管本文提出的ISCSO-LSTM网络集成方法在卫星动量轮RUL预测中表现出良好的性能,但构建能够融合振动、电流等多源监测参数的预测框架仍具有重要研究价值。未来可在多维特征协同表征、异构数据对齐及信息互补机制等方面深化研究,以进一步提升退化表征的完整性。

参考文献

- Yang J, Li B, Fan K, et al. Analysis of laser intersatellite links and topology design for mega-constellation networks. *IEEE Internet Things J*,

- 2024, 11: 34554–34566
- 2 Li J, Su J. Activating the principle of ‘taking into account the special needs of the developing countries’ in the distribution of low Earth orbit frequency/orbit resources. *Telecommun Policy*, 2026, 50: 103082
 - 3 Hassanien A E, Darwish A, Abdelghafar S. Machine learning in telemetry data mining of space mission: Basics, challenging and future directions. *Artif Intell Rev*, 2020, 53: 3201–3230
 - 4 Zheng G K, Zhou Z, Yan R Q. Method for predicting the remaining useful life of mechanical equipment based on deep reinforcement learning (in Chinese). *Sci Sin Tech*, 2023, 53: 1175–1188 [郑国康, 周峥, 严如强. 基于深度强化学习的机械装备剩余使用寿命预测方法. 中国科学: 技术科学, 2023, 53: 1175–1188]
 - 5 Zio E. Prognostics and health management (PHM): Where are we and where do we (need to) go in theory and practice. *Reliab Eng Syst Saf*, 2022, 218: 108119
 - 6 Ferreira C, Gonçalves G. Remaining useful life prediction and challenges: A literature review on the use of machine learning methods. *J Manuf Syst*, 2022, 63: 550–562
 - 7 Xu Y, Zhou J, Zhou R Z, et al. Data mining and analysis in satellite on-orbit faults of meteorological and remote sensing satellites (in Chinese). *Qual Reliab*, 2022, 5: 18–22+27 [许莹, 周捷, 周汝志, 等. 遥感卫星在轨故障挖掘分析. 质量与可靠性, 2022, 5: 18–22+27]
 - 8 Li W, Zhang L C, Wu C H, et al. A data-driven approach to RUL prediction of tools. *Adv Manuf*, 2024, 12: 6–18
 - 9 Halouani A, Khatir Z, Lallemand R, et al. New model of crack propagation of aluminium wire bonds in IGBT power modules under low temperature variations. *Microelectron Reliab*, 2023, 147: 115066
 - 10 Pei X, Li X, Gao L. A novel machinery RUL prediction method based on exponential model and cross-domain health indicator considering first-to-end prediction time. *Mech Syst Signal Process*, 2024, 209: 111122
 - 11 Wang T, Liu Z, Mrad N. A probabilistic framework for remaining useful life prediction of bearings. *IEEE Trans Instrum Meas*, 2021, 70: 1–12
 - 12 Zhang Y, Fang L, Qi Z, et al. A review of remaining useful life prediction approaches for mechanical equipment. *IEEE Sens J*, 2023, 23: 29991–30006
 - 13 Ansari S, Ammirul Atiqi Mohd Zainuri M, Ayob A, et al. Expert deep learning techniques for remaining useful life prediction of diverse energy storage systems: Recent advances, execution features, issues and future outlooks. *Expert Syst Appl*, 2024, 258: 125163
 - 14 Chae S G, Bae S J, Oh K Y. State-of-health estimation and remaining useful life prediction of lithium-ion batteries using DnCNN-CNN. *J Energy Storage*, 2025, 106: 114826
 - 15 Boujamza A, Elhaq S L. Optimizing remaining useful life predictions for aircraft engines: A dilated recurrent neural network approach. *IFAC-PapersOnLine*, 2024, 58: 811–816
 - 16 Yu Y, Si X, Hu C, et al. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Computation*, 2019, 31: 1235–1270
 - 17 Cahuantzi R, Chen X Y, Güttel S. A comparison of LSTM and GRU networks for learning symbolic sequences. In: *Proceedings of the Science and Information Conference*. London, 2023. 771–785
 - 18 Gu J, Wang Y, Su M Q, et al. Multi unmanned vehicle cooperative encirclement control based on bidirectional long short-term memory and mixed reward functions (in Chinese). *Sci Sin Tech*, 2024, 54: 1665–1675 [顾健, 王寅, 苏牧青, 等. 基于双向长短时记忆与混合奖励函数的多无人车协同围捕控制. 中国科学: 技术科学, 2024, 54: 1665–1675]
 - 19 Baruah R D, Organero M M. Explicit context integrated recurrent neural network for applications in smart environments. *Expert Syst Appl*, 2024, 255: 124752
 - 20 Yunita A, Pratama M I, Almuzakki M Z, et al. Performance analysis of neural network architectures for time series forecasting: A comparative study of RNN, LSTM, GRU, and hybrid models. *MethodsX*, 2025, 15: 103462
 - 21 Wu J, Hu K, Cheng Y, et al. Data-driven remaining useful life prediction via multiple sensor signals and deep long short-term memory neural network. *ISA Trans*, 2020, 97: 241–250
 - 22 Qiu X H, Zhang L, Ren Y, et al. Ensemble deep learning for regression and time series forecasting. In: *Proceedings of 2014 IEEE Symposium on Computational Intelligence in Ensemble Learning (CIEL)*. Orlando, Florida, 2014. 1–6
 - 23 Dietterich T G. Ensemble methods in machine learning. In: Kittler J, Roli F (eds). *Multiple Classifier Systems*. Berlin: Springer, 2000. 1–15
 - 24 Chu W, Li J, Wang D, et al. Ensemble deep bidirectional GRU network with grey wolf optimizer for spacecraft fault diagnosis. *IEEE Trans Ind Electron*, 2026, 73: 3140–3149
 - 25 Cheng R, Yang D, Liu D, et al. A reconstruction-based secondary decomposition-ensemble framework for wind power forecasting. *Energy*, 2024,

308: 132895

- 26 Deng L, Zhao C, Yan X, et al. A novel approach for bearing fault diagnosis in complex environments using PSO-CWT and SA-FPN. *Measurement*, 2025, 249: 117027
- 27 Zhang W C, Ni H, Shu P, et al. Fault mode recognition for variable shape vehicles based on ACO-LSTM (in Chinese). *Flight Control & Detection*, 2023, 6: 72–77 [张万超, 倪昊, 舒鹏, 等. 基于蚁群优化的长短时神经网络变外形飞行器故障模式识别. *飞控与探测*, 2023, 6: 72–77]
- 28 Seyyedabbasi A, Kiani F. Sand Cat swarm optimization: A nature-inspired algorithm to solve global optimization problems. *Eng Comput*, 2023, 39: 2627–2651
- 29 He Y, Zhao C, Shen W. Cross-domain compound fault diagnosis of machine-level motors via time-frequency self-contrastive learning. *IEEE Trans Ind Inf*, 2024, 20: 9692–9701
- 30 Miao Y H, Li C H, Shi H F, et al. Remaining useful life prediction of the aircraft engine and its uncertainty quantification based on ConvJANET (in Chinese). *Sci Sin Tech*, 2023, 53: 1189–1201 [苗永浩, 李晨辉, 石惠芳, 等. 基于ConvJANET的航空发动机剩余寿命预测及其不确定性量化. *中国科学: 技术科学*, 2023, 53: 1189–1201]
- 31 Zhu J, Ma J, Wu J, et al. DTFFNet: A dual-branch time–frequency feature fusion network for remaining useful life prediction of mechanical equipment. *Mech Syst Signal Process*, 2025, 236: 113006
- 32 Li B, Li J, Wu X, et al. Gated recurrent unit and temporal convolutional network with soft thresholding and attention mechanism for tool wear prediction. *Measurement*, 2025, 240: 115546
- 33 Baptista M L, Goebel K, Henriques E M P. Relation between prognostics predictor evaluation metrics and local interpretability SHAP values. *Artif Intell*, 2022, 306: 103667
- 34 Zhou Y, Wang S, Xie Y, et al. Remaining useful life prediction and state of health diagnosis for lithium-ion batteries based on improved grey wolf optimization algorithm-deep extreme learning machine algorithm. *Energy*, 2023, 285: 128761
- 35 Li H T, Jin G, Zhou J L, et al. Momentum wheel Wiener process degradation modeling and life prediction (in Chinese). *J Aerospace Power*, 2011, 26: 622–627 [厉海涛, 金光, 周经伦, 等. 动量轮维纳过程退化建模与寿命预测. *航空动力学报*, 2011, 26: 622–627]
- 36 Zhang W, Han G, Wang J, et al. A BP neural network prediction model based on dynamic cuckoo search optimization algorithm for industrial equipment fault prediction. *IEEE Access*, 2019, 7: 11736–11746

Remaining useful life prediction of satellite momentum flywheel based on ISCSO-LSTM network ensemble

YU MuYe¹, XU WenQing¹, QIAN Zhen¹, FU FangZhou¹ & WANG DaYi²

¹ School of Aeronautics and Astronautics, Sun Yat-sen University, Shenzhen 518107, China

² Beijing Institute of Spacecraft System Engineering, Beijing 100094, China

As the core actuator of the attitude control system, the momentum flywheel's performance degradation directly impacts the satellite's lifespan. Therefore, it is essential to research predicting the remaining useful life (RUL) of the momentum flywheel. To address the issues of insufficient prediction accuracy and unstable prediction performance in momentum flywheel RUL, this paper proposes an integrated prediction method based on an improved sand cat swarm optimization (ISCSO) algorithm and long short-term memory (LSTM) network. First, multi-dimensional candidate features are extracted from the original bearing temperature sequence of the momentum flywheel, and a comprehensive evaluation method based on monotonicity and tendency is employed to select sensitive features. Next, multiple LSTM networks with different initialization parameters are constructed as base predictors to enhance the model's generalization ability. Finally, the ISCSO algorithm, incorporating a Tent chaotic map, is introduced to optimize the output weights of each base predictor. Experimental results show that the proposed method outperforms comparison methods across multiple evaluation metrics. Ablation experiments further verify the effectiveness of the improved optimization algorithm and integration strategy.

remaining useful life prediction, momentum flywheel, long short-term memory network, ensemble strategy, improved sand cat swarm optimization

doi: [10.1360/SST-2025-0386](https://doi.org/10.1360/SST-2025-0386)