

# 基于 Copula 相似性的航空发动机 RUL 预测

许先鑫<sup>1</sup>, 李娟<sup>1</sup>, 孙秀慧<sup>1</sup>, 戴洪德<sup>2</sup>

(1. 鲁东大学 数学与统计科学学院 统计系, 山东 烟台 264025;

2. 海军航空大学 航空基础学院, 山东 烟台 264001)

**摘 要:** 针对航空发动机性能退化特征众多, 以及特征相互影响等问题, 考虑退化特征间的非线性相关关系, 提出了基于 Copula 相似性的航空发动机 RUL (remaining useful life) 预测方法。通过 K-means 聚类将航空发动机的工作状态分类, 建立退化模型, 选取退化性能趋势最明显的 3 组传感器。基于 Copula 函数对选取的 3 组传感器进行相关性建模分析, 构建发动机传感器之间的 Copula 结构。基于 Copula 相似性实现对航空发动机的剩余寿命预测。结果表明: 基于 Copula 相似性的航空发动机 RUL 预测方法相较传统方法, 在发动机运行周期的 50%、70%、90% 预测误差分别减少 13.053%、31.328%、74.602%, 预测精度得到提高。

**关 键 词:** 预测与健康管理; 特征退化; 剩余使用寿命; Copula 相似性; 非线性

中图分类号: V263.6

文献标志码: A

## RUL prediction for aero-engines based on Copula similarity

XU Xianxin<sup>1</sup>, LI Juan<sup>1</sup>, SUN Xiuhui<sup>1</sup>, DAI Hongde<sup>2</sup>

(1. Department of Statistics, School of Mathematics and Statistics,

Lu Dong University, Yantai Shandong 264025, China;

2. School of Basic Sciences for Aviation,

Naval Aviation University, Yantai Shandong 264001, China)

**Abstract:** In view of many degradation features of aero-engine performance and their mutual influence, the RUL (remaining useful life) prediction method of aero-engine based on Copula similarity was proposed considering the nonlinear correlations of the degradation features. The working state of the aero-engine was classified through K-means clustering, and a degradation model was established to select three sets of sensors with the most obvious degradation performance trend. Based on the Copula function, the correlation modeling and analysis of the selected three sets of sensors were carried out to build the Copula structure between engine sensors. The prediction of the remaining life of aero-engine was realized based on Copula similarity. The results showed that compared with traditional methods, the prediction errors of the aero-engine RUL based on Copula similarity were reduced by 13.053%, 31.328% and 74.602% during the aero-engine operation cycle of 50%, 70%, 90%, respectively, and the prediction accuracy was improved.

**Keywords:** prediction and health management; performance degradation; remaining useful life; Copula similarity; nonlinearity

收稿日期: 2022-08-08

基金项目: 科技项目基金(F062102009)

作者简介: 许先鑫(1999—), 女, 主要从事预测与健康管理研究。

通信作者: 李娟(1981—), 女, 副教授、硕士生导师, 博士, 主要从事预测与健康管理研究。E-mail: daidaiquanquan123@126.com

引用格式: 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): 20220576. XU Xianxin, LI Juan, SUN Xiuhui, et al. RUL prediction for aero-engines based on Copula similarity[J]. Journal of Aerospace Power, 2024, 39(8): 20220576.

1960 年以来,随着社会的不断变革,航空发动机作为工业皇冠上的明珠,是一个国家综合实力的重要体现。目前在应用最多的加力式涡扇发动机中,部件故障问题非常突出。在飞机运行中,如果出现部件故障问题不能及时解决,很可能会发生灾难性事故<sup>[1]</sup>。然而发动机的维护成本非常昂贵,而预测与健康管理的(prognostics health management, PHM)可以降低费用和减少风险。同时发动机中的部件传感器能够感知各种物理参数(如温度、压力、转子速度等)并准确获取飞机发动机控制和管理所需的数据<sup>[2]</sup>。因此利用实时监测到的传感器信息,对航空发动机进行预测与健康管理的是一种有效的方法。

前期研究中,研究者基于经验的模型、数据驱动模型在 RUL(remaining useful life)预测中取得了很多研究成果。其中,基于经验的模型是将专家知识和工程经验与实际相结合来推断 RUL,包括专家系统和模糊逻辑方法,具有一定的主观性。数据驱动模型是指借助以往观察到的数据或历史上相匹配样本来实现剩余寿命预测,包括但不限于统计模型和人工智能模型<sup>[3]</sup>。在统计模型中,如回归模型、基于 Gamma 过程、基于 Wiener 过程、隐马尔可夫模型(hidden Markov model, HMM)已被应用于预测问题。Li 等<sup>[4]</sup>同时考虑退化特征的线性和非线性两种模型,分别引入线性模型和二次回归模型来预测发动机的 RUL。王华伟等<sup>[5]</sup>采用贝叶斯线性模型评估发动机的性能衰退,基于 Gamma 过程实现对航空发动机的寿命预测。Son 等<sup>[6]</sup>将每个操作模式对应的数据应用主成分分析,形成 6 个主成分空间来构建每个模式下每个单元的退化指标,基于 Wiener 随机过程实现对发动机的寿命预测。Giantomassi 等<sup>[7]</sup>通过人工神经网络(artificial neural network, ANN)提取退化特征,基于 HMM 确定航空发动机的剩余寿命。Xiang 等<sup>[8]</sup>利用多细胞长短期记忆人工神经网络和深度神经网络(deep neural networks, DNN)的一个分支提取航空发动机健康指标(health indicators, HI),利用 DNN 的另一部分从人工特征中生成 HI,最后基于串联的 HI 预测 RUL。

然而,统计学方法往往基于较多的假设,而机器学习方法输入和输出之间的关系经常是未知的。这些不确定因素的存在为预测与建模带来了不便。在数据驱动方法中,基于历史样本退化特征相似

性的方法克服了对系统了解有限和一些潜在未知因素(工程差异、环境条件、人为因素等)的影响,面对复杂系统,表现出很好的 RUL 预测能力。该方法的基本思想是利用已知故障时间的多个历史样本数据来建立一个退化模型库。对于一个待测样本,根据退化轨迹定义的距离指标来评估它和库中的每个退化模型之间的相似性,基于历史相似实现待测样本的 RUL 预测<sup>[9]</sup>。Lam 等<sup>[10]</sup>提出了基于皮尔逊相关的相似性线性回归和动态时间规整(dynamic time warping, DTW)方法,用于确定与测试数据最相似的退化模型来实现寿命预测。Khelif 等<sup>[11]</sup>用线性回归模型构建离线状态的 HI 库,通过比较待测组件的 HI 与离线阶段构建库的 HI 的相似性,利用相似性加权估计选择得分最高的实例预测待测组件的剩余寿命。Liu 等<sup>[12]</sup>充分利用原始数据集提取有效特征构造健康指数相似度,基于距离相似度和空间方向相似度实现 RUL 预测方法。Yu 等<sup>[13]</sup>提出一种基于相似度曲线匹配得到的 RUL 估计方法。解决了在匹配训练实例和需要确定规则的测试实例的 HI 曲线时,不同实例的初始健康程度不同的问题。但在以往相似性模型中,对于系统的寿命预测往往从系统中单个退化特征或退化特征的线性组合出发,寻找部件 HI 相似的样本,忽略了退化特征变量之间的非线性相关性。而面对具有多个退化特征的复杂工业系统,加力式涡扇发动机是一个典型复杂系统的例子。温度、压力等皆含有发动机退化的重要信息,并且温度的改变会造成压力的改变。因此为了尽可能准确地评估产品的剩余寿命,在退化特征不独立的情况下,描述退化特征的相关性尤为重要。针对退化特征的非线性相关性建模常用的方法有基于多维退化模型<sup>[14-15]</sup>、基于退化率相关性建模<sup>[16]</sup>和基于 Copula 相关性建模<sup>[17-18]</sup>等。然而多维退化模型需要假设特征变量的分布形式。退化率相关性建模仅考虑组件间退化速率的影响,忽视了部件之间的相依性。而在相关性建模中, Copula 函数是建立随机变量间相关性的一种有效而灵活的方法,不需要事先假定相同的边缘分布,它可以连接任意边缘分布,因此不会造成数据的失真;并能够对尾部、偏态和非正态数据进行相关性分析。近年来在退化特征相关性分析及寿命预测领域得到广泛应用。Sun 等<sup>[17]</sup>考虑退化特征间的非线性相关性,基于贝叶斯动态线性模型来计算退化特征的 RUL,利用 Copula

函数描述 RUL 的相关关系来实现设备的 RUL 预测。Yang 等<sup>[18]</sup> 基于非线性 Wiener 过程和阿基米德 Copula 函数, 建立了一个退化特征相互作用模型来描述两退化变量间的依赖关系。Xi 等<sup>[19]</sup> 利用半 Copula 模型对不同退化程度下的失效时间进行相关性建模, 利用抽样方法对电风扇失效时间和剩余寿命的概率密度函数进行了有效预测。宋仁旺等<sup>[20]</sup> 针对齿轮箱的 RUL 问题, 利用 Copula 函数描述退化变量之间的相关性, 得到 RUL 的联合分布函数, 进而得到齿轮箱 RUL 的联合概率密度函数来实现预测。梳理参考文献发现, 尚缺乏深入考虑退化特征非线性相关结构的相似性来进行寿命预测的研究实例。本文提出了一种基于 Copula 相关结构相似性的航空发动机 RUL 预测方法。

## 1 航空发动机 Copula 相关结构构建

当边缘分布不同的随机变量之间不独立时, 对于联合分布的建模会变得十分困难。Copula 函数不仅可以刻画单个随机变量的非正态非线性性质, 也能描述不同变量之间复杂的相关关系。

由 Sklar 定理<sup>[21]</sup> 可知 Copula 函数是一种将多个随机变量联合分布函数与它们各自的边缘分布函数连接在一起的函数, 因此使用 Copula 函数得到传感器间的联合分布函数。

假设发动机传感器变量的边缘分布函数为矩阵  $F$ :

$$F = \begin{bmatrix} F_1^1[\mathbf{x}_1^1(t)] & F_2^1[\mathbf{x}_2^1(t)] & \cdots & F_R^1[\mathbf{x}_R^1(t)] \\ F_1^2[\mathbf{x}_1^2(t)] & F_2^2[\mathbf{x}_2^2(t)] & \cdots & F_R^2[\mathbf{x}_R^2(t)] \\ \vdots & \vdots & & \vdots \\ F_1^l[\mathbf{x}_1^l(t)] & F_2^l[\mathbf{x}_2^l(t)] & \cdots & F_R^l[\mathbf{x}_R^l(t)] \end{bmatrix} \quad (1)$$

其中  $\mathbf{x}_R^l(t)$  表示发动机  $l$  第  $R$  个传感器变量的值,  $F_R^l[\mathbf{x}_R^l(t)]$  发动机  $l$  的第  $R$  个传感器变量的边缘分布函数。发动机  $l$  传感器组合 (1,2) 的联合分布为  $W[\mathbf{x}_1^l(t), \mathbf{x}_2^l(t)]$ 。则存在函数  $C(\cdot)$  将边缘分布和联合分布“连接”起来得

$$W[\mathbf{x}_1^l(t), \mathbf{x}_2^l(t)] = C\{F_1^l[\mathbf{x}_1^l(t)], F_2^l[\mathbf{x}_2^l(t)]\} \quad (2)$$

根据边缘分布的累积分布函数逆变换  $\mathbf{x}_i^l(t) = F_i^{-1}[\mathbf{u}_i^l(t)]$  ( $i = 1, 2, \dots, R$ ), 则发动机  $l$  传感器组合 (1,2) 的 Copula 函数的表达形式:

$$C(\mathbf{u}_1^l, \mathbf{u}_2^l) = W\left[F_1^{-1}(\mathbf{u}_1^l), F_2^{-1}(\mathbf{u}_2^l)\right] \quad (3)$$

几种常见的二元 Copula 函数形式如表 1 所示。

表 1 二元 Copula 函数简介

Table 1 Introduction to binary Copula functions

函数类型	分布函数表达式	特性
Clayton Copula	$C_{\delta}^{Cl} = (\mu^{-\delta} + \nu^{-\delta} - 1)^{-\frac{1}{\delta}}$	上尾相关
Gumbel Copula	$C_{\theta}^{Gu} = \exp\left\{-\left[\left(-\ln \mu\right)^{\theta} + \left(-\ln \nu\right)^{\theta}\right]^{\frac{1}{\theta}}\right\}$	下尾相关
Frank Copula	$C_{\lambda}^{Fr} = -\frac{1}{\lambda} \ln\left\{1 - \left[\frac{(1 - e^{-\lambda\mu})(1 - e^{-\lambda\nu})}{(1 - e^{-\lambda})}\right]\right\}$	不反映尾部相关性

为保证 Copula 函数中变量的边缘分布  $\mathbf{u}_i^l(t)$  服从  $[0, 1]$  上的均匀分布, 先对发动机的传感器的边缘分布作概率积分变换, 即  $\mathbf{u}_i^l(t) = F[\mathbf{x}_i^l(t)]$ 。采用两阶段极大似然估计<sup>[22]</sup> 对 Copula 函数的未知参数进行估计。以发动机  $l$  传感器组合 (1,2) 为例, 估计过程如下:

$$l(\theta) = \sum_{i=1}^k \ln\left\{c\left\{F_1[\mathbf{x}_1^l(t), \theta_1], F_2[\mathbf{x}_2^l(t), \theta_2]; a\right\}\right\} + \sum_{i=1}^k \sum_{j=1}^2 \ln f_i[\mathbf{x}_i^l(t); \theta_i] \quad (4)$$

### 步骤 1

$$\hat{\theta}_i = \arg \max \sum_{i=1}^k \ln f_i[\mathbf{x}_i^l(t); \theta_i] \quad (5)$$

### 步骤 2

$$\hat{a} = \arg \max \sum_{i=1}^k \ln\left\{c\left\{F_1[\mathbf{x}_1^l(t), \theta_1], F_2[\mathbf{x}_2^l(t), \theta_2]; a\right\}\right\} \quad (6)$$

为选择合适的 Copula 函数来描述已选传感器间的联合分布函数, 选用 Akaike 信息准则 (AIC, 量符号记为  $C_{Ai}$ ) 来量化拟合效果。AIC 值较小的 Copula 函数的拟合效果较好。AIC 表达式为

$$C_{Ai} = 2K - 2\ln L \quad (7)$$

其中  $K$  表示参数的数量,  $L$  为对应的似然函数。由式 (5)、式 (6) 估计求得发动机  $l$  的传感器组合的 Copula 函数的未知参数  $\hat{\theta}_i$ 、 $\hat{a}$ , 由式 (7) 计算对应 Copula 结构的 AIC 值选择能描述已选传感器间相关性的最优 Copula 结构。

## 2 基于 Copula 相关结构相似性的寿命预测方法

基于 Copula 相似性的航空发动机 RUL 预测的框架图如图 1 所示。

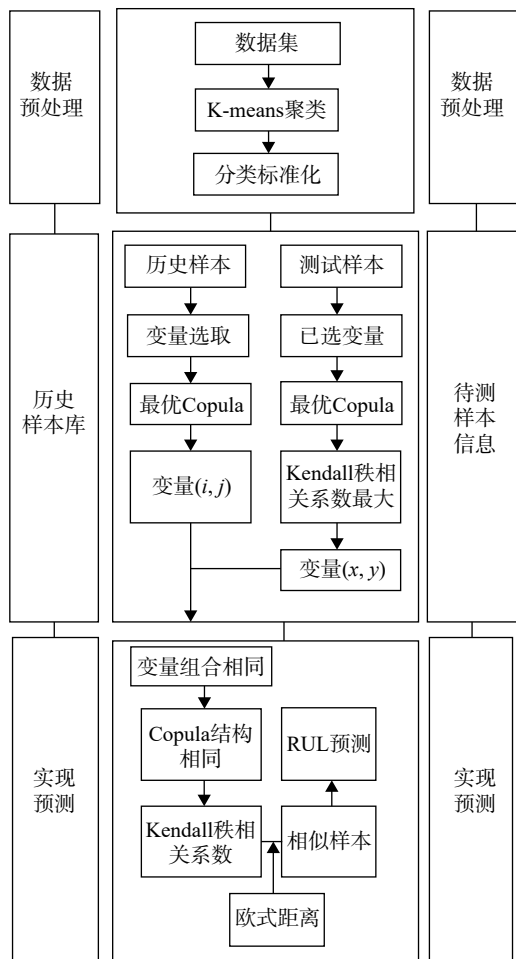


图 1 预测框架图

Fig. 1 Chart of predictive framework

航空发动机在运行周期的运行状态通过多个性能参数表现出来,每个传感器的退化趋势在整个运行周期的敏感性不同,退化趋势无法直接观察到。有明显退化趋势的传感器能够及时反映发动机的运行状态,一些传感器由于噪声高或对退化的灵敏度低,并没有表现出明显的趋势。因此,传感器并不是越多越好。如果选择无意义的传感器在增加计算量的同时预测精度也会降低。所以,首先要对发动机原始传感器参数进行筛选,获得最能代表发动机从健康状态到衰退状态的传感器数据信息。

1) 假设发动机传感器变量的参数值为矩阵  $X$

$$X = \begin{bmatrix} x_1^1(t) & x_2^1(t) & \cdots & x_R^1(t) \\ x_1^2(t) & x_2^2(t) & \cdots & x_R^2(t) \\ \vdots & \vdots & & \vdots \\ x_1^l(t) & x_2^l(t) & \cdots & x_R^l(t) \end{bmatrix} \quad (8)$$

则发动机  $l$  第  $R$  个传感器变量有  $x_{R,1}^l$  到  $x_{R,k}^l$  共  $k$  组监测数据,因此传感器变量的历史观测数据可表示:

$$\mathbf{x}_R^l(t) = [x_{R,1}^l, x_{R,2}^l, \cdots, x_{R,k}^l]^T \quad (9)$$

式中  $x_R^l(t)$  表示发动机  $l$  第  $R$  个传感器变量的值,  $x_{R,k}^l$  表示发动机  $l$  第  $R$  个传感器变量在  $k$  时刻的观测值。

2) 根据先验知识,给定  $\theta(t)$  的分布  $\theta(t) \sim N(\mu_\theta, \sigma_\theta^2)$ , 其中  $\mu_\theta$  为  $\theta(t)$  的平均值,  $\sigma_\theta^2$  为  $\theta(t)$  的方差。 $\varphi$  为截距项,  $\varepsilon(t) \sim N(0, \sigma_\varepsilon^2)$ 。 $\theta(t) \in [0, 1]$ , 该过程仅在检测到每组发动机传感器历史观测数据斜率水平显著时开始被检测。

3) 建立线性退化模型,线性退化模型将退化描述为线性随机过程:

$$S_R(t) = \theta_R(t)t + \varepsilon_R(t) + \varphi_R \quad (10)$$

式中  $S_R(t)$  为第  $R$  个传感器随着发动机从运行到故障全过程观测数据值,  $\theta_R(t)$  为检测到的  $l$  组发动机第  $R$  个传感器数据的总体退化水平。每次检测要重新更新参数的后验分布。

4) 基于退化模型选定退化趋势最明显的 3 个传感器  $x_i(t), i \in [1, 21]$  进行相关结构相似性计算。其中,对于传感器变量的参数值  $\mathbf{x}_1^l(t)$  和  $\mathbf{x}_2^l(t)$ , 由公式有令  $v = F_1[\mathbf{x}_1^l(t)]$ ,  $w = F_2[\mathbf{x}_2^l(t)]$ ,  $v, w \in [0, 1]$ , 由式(2)、式(3)得到相应的 Copula 函数为  $C(v, w)$  时, Kendall 秩相关系数<sup>[23]</sup> 可以表示为

$$\tau = 4 \int_0^1 \int_0^1 C(v, w) dC(v, w) - 1 = 4E[C(v, w)] - 1 \quad (11)$$

式中  $E(\cdot)$  为求期望函数, Kendall 秩相关系数用来衡量变量变化趋势是否相同。即当  $\tau > 0$ , 变量间变化趋势相同; 当  $\tau < 0$ , 变量间变化趋势不相同。

5) 每个发动机在达到故障条件或预设的不良条件阈值时,发动机停止运行。因此第  $l$  个航空发动机的运行时间可表示为  $T_l^l(t = 1, 2, \cdots, k)$ 。则发动机在  $k-2$  时刻的剩余寿命为  $L_{ru,T}^l = T_k^l - T_{k-2}^l$ 。其中  $T_1^l$  为发动机  $l$  开始运行的时刻值,  $T_{k-2}^l$  为运行过程中  $k-2$  的时刻值,  $T_k^l$  为发动机出故障停止运行的时刻值。

6) 假设发动机退化趋势最明显的 3 组传感器间的结构为矩阵  $M$ :

$$M = \begin{bmatrix} M_1^1 & M_2^1 & M_3^1 \\ M_1^2 & M_2^2 & M_3^2 \\ \vdots & \vdots & \vdots \\ M_1^l & M_2^l & M_3^l \end{bmatrix} \quad (12)$$

式中  $M^l = [M_1^l, M_2^l, M_3^l]$ ,  $M^l$  表示发动机  $l$  退化趋势最明显的 3 组传感器 3 种组合。

则  $M$  对应的最优 Copula 结构矩阵  $Q$ :

$$Q = \begin{bmatrix} Q_1^1 & Q_2^1 & Q_3^1 \\ Q_1^2 & Q_2^2 & Q_3^2 \\ \vdots & \vdots & \vdots \\ Q_1^l & Q_2^l & Q_3^l \end{bmatrix} \quad (13)$$

其中  $Q^l = [Q_1^l, Q_2^l, Q_3^l]$ ,  $Q^l$  表示由式(7)筛选出的发动机  $l$  退化趋势最明显的 3 组传感器组合间的最优 Copula 结构。

并计算出最优 Copula 结构的 Kendall 秩相关系数矩阵  $\tau$ :

$$\tau = \begin{bmatrix} \tau_1^1 & \tau_2^1 & \tau_3^1 \\ \tau_1^2 & \tau_2^2 & \tau_3^2 \\ \vdots & \vdots & \vdots \\ \tau_1^l & \tau_2^l & \tau_3^l \end{bmatrix} \quad (14)$$

其中  $\tau^l = [\tau_1^l, \tau_2^l, \tau_3^l]$ ,  $\tau^l$  表示发动机  $l$  传感器组合的最优 Copula 结构的 Kendall 秩相关系数, 选取发动机  $l$  的  $\max\{\tau^l\}$  的传感器组合  $M_p^l$ ,  $p \in [1, 3]$ 。

7) 记待测样本筛选出的传感器组合的 Copula 结构为  $Q_p^*$ , 秩相关系数为  $\tau_p^*$ 。则寻找历史样本库  $M$  中的样本信息  $\{M_p^1, M_p^2, \dots, M_p^l\}$ ,  $\{Q_p^1, Q_p^2, \dots, Q_p^l\}$ ,  $\{\tau_p^1, \tau_p^2, \dots, \tau_p^l\}$ , 并且满足对应的 Copula 结构同为  $Q_p^*$ , 该待测样本与第  $l$  个历史样本的距离为  $d_l = |\tau_p^l - \tau_p^*|^2$ , 则样本距离集合为  $D = \{d_1^2, d_2^2, \dots, d_l^2\}$ 。其中  $d_l$  表示待测样本与第  $l$  个历史样本的距离。距离越小, 与历史样本的相似度越高; 距离越大, 与历史样本的相似度越低。因此对  $D$  集合中的数据值进行从小到大排序, 选取前  $n$  ( $n \leq 50$ ) 个历史样本的平均值作为发动机寿命预测。因此现役发动机的 RUL 预测值可表示:

$$L_{ru}^* = \frac{1}{n} \sum_{l=1}^n L_{ru,l}^l \quad (15)$$

8) 计算与发动机真实剩余寿命的误差,  $E = L_{ru} - L_{ru}^*$ 。其中  $L_{ru}$  为该样本剩余寿命的真实值,  $L_{ru}^*$  为该样本的寿命预测值。如图 2 所示为发动

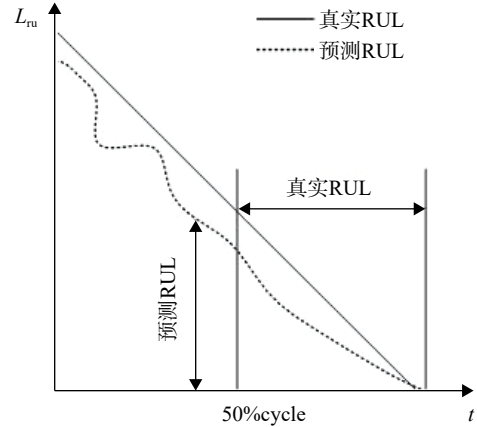


图2 发动机剩余寿命对比图

Fig. 2 Comparison chart of engine remaining life

机在运行过程中某个时刻的预测值与 50% 运行周期的真实寿命图。

为了衡量模型的优劣, 本文选用标准差  $\sigma_s$  (standard deviation) 来进行模型对比验证:

$$E_{\sigma_s} = \sqrt{\frac{1}{N-1} \sum_{l=1}^N (L_{ru}^l - L_{ru}^{ls})^2} \quad (16)$$

式中  $L_{ru}^l$  和  $L_{ru}^{ls}$  分别表示待测发动机的剩余寿命真实值和预测值,  $L_{ru,j}^{ls}$  为  $n$  ( $n \leq 50$ ) 个相似样本的均值,  $N$  为待测样本的总数。

### 3 实验分析

本文数据来源于 NASA 的 C-MAPSS 航空发动机仿真模型数据, 航空发动机主要部件描述图如图 3 所示。该数据集由多维时间序列数据组成, 每个时间序列数据分别为监测到的同一复杂工程系统下不同发动机的运行状况。本文采用 217 组航空发动机在各自飞行周期的 24 个维度性能参数, 包括 3 个操作设置变量和 21 个传感器的参数值, 对 21 个传感器的具体描述如表 2 所示。

在机组的运行中, 故障规模会不断扩大, 直至机组停止运行。如表 3 展示的为一组发动机从开始运行到出故障的全过程数据。数据集中选取 175 组训练集, 42 组验证集。每组训练数据都包含发动机完整的运行周期及退化性能参数。训练集用于建立寿命预测模型库。验证集中包含真实寿命值, 用于模型评价。

记发动机  $l$  的操作设置变量为  $Y$ ,  $Y_j^l(t)$  ( $j = 1, 2, 3$ ) 表示 3 个操作设置的值, 则该组发动机第  $j$  个操作设置变量有  $y_{j,1}^l$  到  $y_{j,k}^l$  共  $k$  组监测数据, 因此操作设置变量的历史观测数据可表示为

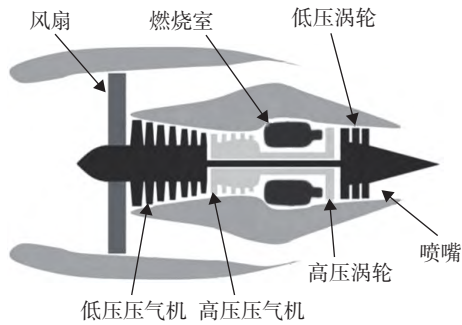


图 3 发动机主要部件简图  
Fig. 3 Sketch of engine main components

表 2 传感器变量介绍<sup>[24]</sup>  
Table 2 Introduction of sensor variables<sup>[24]</sup>

序号	符号	描述
1	T2	风扇入口总温
3	T30	高压压气机出口总温
4	T50	低压涡轮出口总温
5	P2	风扇入口压力
6	P15	外涵总压
7	P30	高压压气机出口总压
8	Nf	风扇物理转速
9	Ne	核心机物理转速
10	epr	发动机压比
11	Ps30	高压压气机出口静压
12	Phi	燃油流量与高压压气机出口总压比值
13	NRf	风扇换算转速
14	Nre	核心机换算转速
15	BPR	涵道比
16	farB	燃烧室燃气比
17	htBleed	引气焓值
18	Nf_dmd	设定风扇转速
19	PCNfR_dmd	设定核心机换算转速
20	W31	高压涡轮冷却引气流量
21	W32	低压涡轮冷却引气流量

注: 发动机压比为高压压气机出口总压和风扇入口压力之比。

$$\mathbf{Y}_j^l(t) = [y_{j,1}^l, y_{j,2}^l, \dots, y_{j,k}^l] \quad (17)$$

式中  $\mathbf{Y}_j^l(t)$  表示发动机  $l$  第  $j$  个操作设置变量的工作状态数据集,  $y_{j,k}^l$  表示发动机  $l$  第  $j$  个操作设置变量在  $k$  时刻的值。

由于工况的影响, 难以看到退化趋势。首先进行聚类, 将训练数据集中  $l$  组发动机的 3 个操作设置变量串联为  $\mathbf{Y}_1(t)$ 、 $\mathbf{Y}_2(t)$ 、 $\mathbf{Y}_3(t)$ , 对其进行 K-means 聚类分析。首先, 操作设置被分为 6 类, 形成 6 个聚类质心点; 然后, 计算  $\mathbf{Y}_1^l(t)$ 、 $\mathbf{Y}_2^l(t)$ 、 $\mathbf{Y}_3^l(t)$  分别距离 6 个质心点的距离, 寻找最近距

表 3 某发动机运行数据  
Table 3 Operating data of an engine

$t$	$y_1(t)$	$\dots$	$x_1^l(t)$	$x_2^l(t)$	$\dots$	$x_{21}^l(t)$
1	10.004 7		489.05	604.13		17.173 5
2	0.001 5		518.67	642.13		23.361 9
3	34.998 6		449.44	555.42		8.855 5
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$
223	34.999 2	$\dots$	449.44	556.6	$\dots$	8.669 5

离进行归类。最终 6 个聚类中心位置情况如图 4 所示, 6 个聚类中心各操作设置参数如表 4 所示。

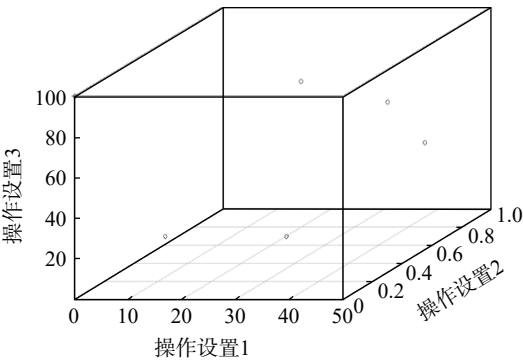


图 4 操作设置聚类结果图  
Fig. 4 Clustering graph of operation settings

表 4 聚类操作设置参数表  
Table 4 Clustering parameters table of operation setting

类别	$\mathbf{Y}_1(t)$	$\mathbf{Y}_2(t)$	$\mathbf{Y}_3(t)$
1	10.003	0.251	20
2	0.001 2	0.001	100
3	25.003	0.621	80
4	42.003	0.841	40
5	20.003	0.701	0
6	35.003	0.841	60

其次按不同的工况进行标准化, 计算各聚类下各传感器的均值  $h_{i,j}$  及标准差  $\sigma_{i,j}$ , 得到数据值矩阵  $\mathbf{H}$

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,21} \\ h_{2,1} & h_{2,2} & \dots & h_{2,21} \\ \vdots & \vdots & \ddots & \vdots \\ h_{6,1} & h_{6,2} & \dots & h_{6,21} \end{bmatrix} \quad (18)$$

其中 6 为发动机工况类别数, 21 为传感器总数。  
标准化方法为:  $z_* = h_* - h_{i,j}/\sigma_{i,j}$ 。其中  $h_*$  表示为任一传感器检测值, 则  $z_*$  为任一传感器标准化后的数据值。



基于退化模型,以 $|\theta_i(t)|$ 为选取标准。对 $|\theta_i(t)|$ , $i \in [1,21]$ 进行排序,退化趋势最强的 3 组传感器,即 Ps30、T50 和 BPR。由图 5 为训练数据集集中编号为 1 的发动机已选传感器退化趋势图,可以看出 3 组传感器都有明显的变化趋势。

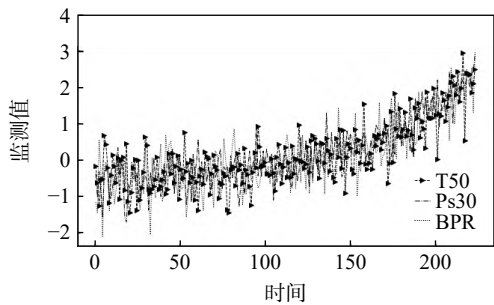


图 5 已选传感器退化趋势图  
Fig. 5 Degradation trend graph of the selected sensors

利用 Copula 函数对 $M^1$ 进行相关性建模分析,得到 $Q^1$ ,并计算出对应的 $\tau^1$ 。表 5 为训练数据集部分发动机的 Copula 结构相关信息。

以编号为 11 的现役发动机为例展示寿命预测过程。首先拟合现役发动机传感器组合 Copula 结构 $Q^{11}$ 并计算出对应的 $\tau^{11}$ 。根据表 6 参数值信息得出 $\max\{\tau^{11}\}$ 对应的传感器变量组合为 $M_p^{11}$ ,并且 $M_p^{11}$ 对应的为 Frank Copula,因此选取历史样本中传感器变量组合 Copula 结构与 $Q_p^{11}$ 一致,同为 Frank Copula,并且以相应的秩相关系数的欧式

表 5 历史样本 Copula 参数值

Table 5 Copula parameter values of historical sample						
$L_{ru}$	$Q_1^1$	$Q_2^1$	$Q_3^1$	$\tau_1^1$	$\tau_2^1$	$\tau_3^1$
223	Fra	Cla	Cla	0.479	0.454	0.495
181	Fra	Joe	Joe	0.531	0.460	0.524
209	Joe	Joe	Joe	0.472	0.332	0.333
⋮	⋮	⋮	⋮	⋮	⋮	⋮
220	Cla	Cla	Cla	0.420	0.390	0.418
255	Fra	Fra	Fra	0.512	0.492	0.536
156	Joe	Fra	Fra	0.513	0.450	0.514

注:表中 Fra 表示 Frank Copula, Cla 表示 Clayton Copula, Joe 表示 Joe Copula。

表 6 现役发动机参数值

Table 6 Parameter values of in-service engine		
传感器组合	Copula	$\tau$
(4,11)	Clayton	0.474
(4,15)	Frank	0.458
(11,15)	Frank	0.519

距离为相似性计算依据。  
如表 7 所示,寻找与现役发动机距离最近的前 $n(n \leq 50)$ 个历史样本。分别为编号 178, 60, 151, ⋯。将运行不同周期的剩余寿命与真实寿命进行对比。可以看出运行周期越长,可测数据越多,预测误差越小。

表 7 现役发动机寿命预测值  
Table 7 Life prediction of in-service engine

相似样本	$L_{ru}$	50%cycle	70%cycle	90%cycle
178	255	127.500	76.500	25.500
60	300	150	90	30
151	317	158.500	95.100	31.700
⋮	⋮	⋮	⋮	⋮
108	230	115	69	23
35	188	94	56.4	18.8
16	172	86	51.600	17.200
预测结果	217.86	108.930	65.358	21.786
真实寿命	210	105	63	21
预测误差	7.860	3.930	2.358	0.786

表 8 是通过退化特征非线性相关结构的相似性来寻找相似样本与通过拟合 HI 来寻找相似性历史样本<sup>[25]</sup>方法的对比误差表。由表 8 可以看出,随着发动机运行时间的增加,均方误差越来越小。但是基于 Copula 相似性方法的误差均小于传统方法。本文方法相较于传统方法而言,在发动机运行周期的 50%、70%、90% 分别低 13.053%、31.328%、74.602%。由此可以看出基于 Copula 相似性的航空发动机 RUL 预测方法预测效果精度更高。且深入考虑退化特征非线性相关结构的相似性,适用于具有多个非独立退化特征的航空发动机寿命预测。

表 8 方法对比误差表  
Table 8 Error comparison table of two methods

方法	估计误差		
	50%cycle	70%cycle	90%cycle
传统方法 <sup>[25]</sup>	26.346	20.014	18.037
基于 Copula 相似性	22.907	13.744	4.581

4 结 论

1) 利用 Copula 函数对每个系统内部退化特征间相关性进行建模,不需要事先对退化特征的分布做出任何假设,数据失真程度低,增加了拟

合结果的准确性。

2) 利用秩相关系数作为相似性度量依据, 将退化特征间的相关结构的相似性考虑到寿命预测中, 提出了一种寻找相似样本的新方法。

3) 选取 NASA 的 C-MAPSS 航空发动机仿真模型数据对基于 Copula 相似性的 RUL 预测方法进行了验证, 通过与传统方法相比, 预测精度得到提高。

## 参考文献:

- [1] XIONG Minglan, WANG Huawei, FU Qiang, et al. Digital twin-driven aero-engine intelligent predictive maintenance[J]. The International Journal of Advanced Manufacturing Technology, 2021, 114(11/12): 3751-3761.
- [2] XU Jiuping, WANG Yusheng, XU Lei. PHM-oriented integrated fusion prognostics for aircraft engines based on sensor data[J]. IEEE Sensors Journal, 2014, 14(4): 1124-1132.
- [3] LIAO Linxia, KÖTTIG F. Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction[J]. IEEE Transactions on Reliability, 2014, 63(1): 191-207.
- [4] LI Yiguang, NILKITSARANONT P. Gas turbine performance prognostic for condition-based maintenance[J]. Applied Energy, 2009, 86(10): 2152-2161.
- [5] 王华伟, 吴海桥. 基于信息融合的航空发动机剩余寿命预测[J]. 航空动力学报, 2012, 27(12): 2749-2755.  
WANG Huawei, WU Haiqiao. Residual useful life prediction for aircraft engine based on information fusion[J]. Journal of Aerospace Power, 2012, 27(12): 2749-2755. (in Chinese)
- [6] SON K L, FOULADIRAD M, BARROS A, et al. Remaining useful life estimation based on stochastic deterioration models: a comparative study[J]. Reliability Engineering & System Safety, 2013, 112: 165-175.
- [7] GIANTOMASSI A, FERRACUTI F, BENINI A, et al. Hidden Markov model for health estimation and prognosis of turbofan engines[C]// International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Washington: ASME, 2011: 681-689.
- [8] XIANG Sheng, QIN Yi, LUO Jun, et al. Multicellular LSTM-based deep learning model for aero-engine remaining useful life prediction[J]. Reliability Engineering & System Safety, 2021, 216: 107927.
- [9] MALINOWSKI S, CHEBEL-MORELLO B, ZERHOUNI N. Remaining useful life estimation based on discriminating shapelet extraction[J]. Reliability Engineering & System Safety, 2015, 142: 279-288.
- [10] LAM J, SANKARARAMAN S, STEWART B. Enhanced trajectory based similarity prediction with uncertainty quantification[C]//Annual Conference of the PHM Society. Texas: IJPHM, 2014: 2513.1-2513.12.
- [11] KHELIF R, MALINOWSKI S, CHEBEL-MORELLO B, et al. RUL prediction based on a new similarity-instance based approach [C]//2014 IEEE 23rd International Symposium on Industrial Electronics. Piscataway, US: IEEE, 2014: 2463-2468.
- [12] LIU Yingchao, HU Xiaofeng, ZHANG Wenjuan. Remaining useful life prediction based on health index similarity[J]. Reliability Engineering & System Safety, 2019, 185: 502-510.
- [13] YU Wennian, KIM I Y, MECHEFSKE C. An improved similarity-based prognostic algorithm for RUL estimation using an RNN autoencoder scheme[J]. Reliability Engineering & System Safety, 2020, 199: 106926.
- [14] PAN Zhengqiang, BALAKRISHNAN N. Reliability modeling of degradation of products with multiple performance characteristics based on gamma processes[J]. Reliability Engineering & System Safety, 2011, 96(8): 949-957.
- [15] ZHENG Jianfei, SI Xiaosheng, HU Changhua, et al. A nonlinear prognostic model for degrading systems with three-source variability[J]. IEEE Transactions on Reliability, 2016, 65(2): 736-750.
- [16] RASMEKOMEN N, PARLIKAD A K. Condition-based maintenance of multi-component systems with degradation state-rate interactions[J]. Reliability Engineering & System Safety, 2016, 148: 1-10.
- [17] SUN Fuqiang, WANG Ning, LI Xiaoyang, et al. Remaining useful life prediction for a machine with multiple dependent features based on Bayesian dynamic linear model and copulas[J]. IEEE Access, 2017, 5: 16277-16287.
- [18] YANG Zhiyuan, ZHAO Jianmin, CHENG Zhonghua, et al. Reliability modeling of two-component system with degradation interaction based on copulas[C]//2018 Prognostics and System Health Management Conference. Piscataway, US: IEEE, 2019: 138-143.
- [19] XI Zhimin, WANG Pingfeng. A Copula based sampling method for residual life prediction of engineering systems under uncertainty[C]//2012 IEEE Conference on Prognostics and Health Management. Piscataway, US: IEEE, 2012: 1-9.
- [20] 宋仁旺, 张岩, 石慧. 基于 Copula 函数的齿轮箱剩余寿命预测方法[J]. 系统工程理论与实践, 2020, 40(9): 2466-2474.  
SONG Renwang, ZHANG Yan, SHI Hui. Prediction method for the remaining useful life of gearbox based on copula function[J]. Systems Engineering-Theory & Practice, 2020, 40(9): 2466-2474. (in Chinese)
- [21] SUKHANOVA E M. A test for independence of two multivariate samples[J]. Mathematical Methods of Statistics, 2008, 17(1): 74-86.
- [22] 蔡菲, 严正, 赵静波, 等. 基于 Copula 理论的风电场风速及输出功率相依结构建模[J]. 电力系统自动化, 2013, 37(17): 9-16.  
CAI Fei, YAN Zheng, ZHAO Jingbo, et al. Dependence structure models for wind speed and wind power among different wind farms based on copula theory[J]. Automation of Electric Power Systems, 2013, 37(17): 9-16. (in Chinese)
- [23] ATIQUÉ F, ATTOH-OKINE N. Using copula method for pipe data analysis[J]. Construction and Building Materials, 2016, 106: 140-148.
- [24] SAXENA A, GOEBEL K, SIMON D, et al. Damage propagation modeling for aircraft engine run-to-failure simulation[C]//2008 International Conference on Prognostics and Health Management. Piscataway, US: IEEE, 2008: 1-9.
- [25] WANG Tianyi, YU Jianbo, SIEGEL D, et al. A similarity-based prognostics approach for Remaining Useful Life estimation of engineered systems[C]//2008 International Conference on Prognostics and Health Management. Piscataway, US: IEEE, 2008: 1-6.

(编辑: 陈 越)