

小样本情况下的 PHM 系统预测能力评估方法研究

解江, 王唯良

西北政法大学经济学院, 陕西 西安 710063

DOI: 10.61369/TACS.2025010027

摘 要 : 针对产品在使用初期阶段缺少数据情况下难以对 PHM 系统预测能力进行有效评估的问题, 论文对评估过程的数据收集、数据样本分析及转化等问题进行了系统阐述, 在对产品研制和样机试制阶段的验前信息进行相容性检验和可信用度分析的基础上, 采用基于 Bayes 方法的多源验前信息融合方法对各阶段数据进行综合应用, 并通过具体的案例对论文提出的方法进行了验证。

关 键 词 : PHM; 验前信息; 融合

Research on the Evaluation Method of PHM System Prediction Ability under Small Sample Conditions

Jie Jiang, Wang Weiliang

Northwest University of Political Science and Law, School of Economics, Xi'an, Shaanxi 710063

Abstract : In response to the problem of difficulty in effectively evaluating the predictive ability of PHM systems in the early stages of product use when there is a lack of data, this paper systematically elaborates on the issues of data collection, data sample analysis, and transformation in the evaluation process. Based on the compatibility test and credibility analysis of the pre-test information in the product development and prototype trial production stages, a multi-source pre-test information fusion method based on Bayes method is adopted to comprehensively apply the data in each stage, and the method proposed in the paper is verified through specific cases.

Keywords : PHM; pre inspection information; fusion

引言

PHM 系统是近年来提出的集故障诊断、故障预测和健康管理能力于一体的新型综合系统, 它借助于信息化、智能化手段对关键单元的故障进行实时检测与隔离, 从而在故障发生之前触发主动控制活动。故障预测能力是 PHM 系统的核心技术, 也是区别于以往系统诊断技术的重要特点。预测技术相比诊断技术而言, 远没有成熟, 在其投入使用前, 需要综合验证和评价。因此, 对 PHM 系统故障预测能力做好评估, 并明确相关数据收集和处理方法是非常必要的。

一、故障预测能力数据收集及处理方法

PHM 系统评估所需信息存在于研制和使用的各个环节, 信息获取的渠道主要有以下几方面: (1) 从使用阶段初期使用中获取信息。这是在真实使用环境下设备预测能力的真实反映; (2) 从历史资料中获取信息。主要来源包含了单元及分系统试验信息、国内外相似系统的信息、单元或系统以往的各种试验信息; (3) 通过理论分析或仿真试验获取信息。仿真试验信息是获取信息的重要途径; (4) 专家信息。在数据较少的情况下, 可依靠专家丰富的工程经验。

二、故障预测能力评估总体思路

在使用阶段初期样本数量较小的情况下可采用 Bayes 方法, 综合运用使用阶段、研制和样机测试阶段的历史数据和专家信息, 对 PHM 预测指标进行验证评估。论文对数据进行相容性检验和可信用度分析后, 采用多源信息融合方法以获得先验信息, 并结合似然函数, 最终得到预测指标的后验分布。Bayes 方法应用于 PHM 系统故障预测能力指标评估过程如下:

- (1) 列出需要用 Bayes 方法进行分析的诊断能力评估指标;
- (2) 收集相关的验前信息进行相容性、可信性等检验和分析;
- (3) 根据获取的验前信息, 确定相应的验前分布;
- (4) 获取设备使用阶段初期统计数据, 形成故障预测能力评估指标的似然函数;
- (5) 运用 Bayes 公式, 将先验分布与似然函数进行融合, 得到后验分布;

(6) 根据后验分布, 进行故障预测能力评估指标参数统计推断。

对 PHM 系统故障预测能力指标的评估过程如图 1 所示:

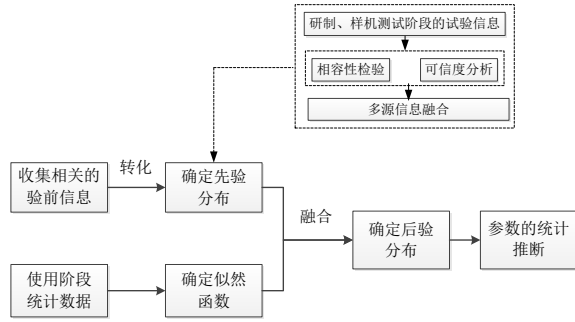


图 1 基于 Bayes 方法的故障预测能力评估过程

三、基于 Bayes 方法的多源验前信息的概率分布

(一) Bayes 方法概述

Bayes 方法是小样本数据分析的一种主要方法, Bayes 方法的优势在于可以利用先验信息增加信息量, 从而获得比传统方法更可靠、更准确的结论。Bayes 方法通过先验信息的收集、挖掘和加工, 使它参加到统计推断中, 解决由于样本信息不足导致统计推断的可信度低的问题, 从而弥补 PHM 系统诊断能力评估时统计数据不足或统计周期过长的问题。Bayes 方法应用的基础是 Bayes 公式, 其连续形式为:

$$\pi(\theta | x) = \frac{L(x | \theta)\pi(\theta)}{\int_{\Theta} L(x | \theta)\pi(\theta) d\theta} \quad (1)$$

式中 θ 表示要推断的参数或指标, Θ 表示 θ 的取值范围, x 为样本观测值, $L(x | \theta)$ 为似然函数, $\pi(\theta)$ 为 θ 的先验分布密度函数。

(二) 验前信息的相容性检验

Bayes 小子样理论需要对验前信息做相容性检验或一致性检验, 即验前信息和验后信息应近似服从同一总体, 在方法上可采用显著性检验。在有验前信息的时, θ 的验前分布为 $\pi(\theta)$, 在加入装备的使用数据后可得验后分布 $\pi(x | \theta)$ 。若假定显著性水平为 α , 则显著性检验的置信区间 (θ_1, θ_2) 可由式 (2) 确定:

$$\int_{\theta < \theta_1} \pi(\theta | x) d\theta = \alpha / 2; \quad \int_{\theta > \theta_2} \pi(\theta | x) d\theta = \alpha / 2 \quad (2)$$

如果 θ 落在该区间内, 可认为验前、验后数据是相容的。若将其与实际结合并做拓展, 即假设同一类型的验前信息有 L 组, 这 L 组的数据与获得使用数据后的样本相容, 于是, L 组验前信息对应得到的验前分布密度为: $\pi^{(l)}(\theta), l=1, \dots, L$ [15-9]。

(三) 验前信息的可信性分析

在验前信息通过相容性检验后, 从理论上可认为验前信息的可信性较高, 可以采信。在实际工作中, 常常需对多组不同的验前子样做分析, 假定 L 组验前信息的可信度分别为 $b_l, l=1, \dots, L$, 则关于指标 θ 的验前分布密度表示为 [15-9]:

$$\pi(\theta) = \sum_{l=1}^L w_l \pi^{(l)}(\theta), \quad w_l = b_l / \sum_{l=1}^L b_l, l=1, \dots, L \quad (3)$$

(四) 多源验前信息的融合验后分布

通过相容性检验和可信性分析后, 在获得初期使用阶段子样 X 之后, 利用 Bayes 公式得到指标 θ 的验后密度 $\pi(\theta | X)$, 表示为:

$$\pi(\theta | X) = \frac{f(X | \theta)\pi(\theta)}{\int_{\Theta} f(X | \theta)\pi(\theta) d\theta} \quad (4)$$

(4) 式可改写为:

$$\pi(\theta | X) = \frac{1}{m(\theta | \pi)} \sum_{l=1}^L w_l \pi^{(l)}(\theta) f(X | \theta) \quad (5)$$

式中, $m(\theta | \pi) = \int_{\Theta} f(X | \theta)\pi(\theta) d\theta$, 表示 X 的边缘密度。

当 $L=2$ 时, 表示当前有两种验前信息, 将其融合后得到验前密度 [10-15]:

$$\pi(\theta) = w_1 \pi^{(1)}(\theta) + w_2 \pi^{(2)}(\theta)$$

式中: $w_1 + w_2 = 1$, 且有: $w_1 = \frac{b_1}{b_1 + b_2}, w_2 = 1 - w_1 = \frac{b_2}{b_1 + b_2}$

式中, b_1, b_2 分别表示两种验前信息的可信度, 即:

$$\pi(\theta | X) = \frac{[w_1 \pi^{(1)}(\theta) + (1 - w_1) \pi^{(2)}(\theta)] f(X | \theta)}{m(X | \pi)}$$

又因为:

$$\pi^{(i)}(\theta | x) = \frac{\pi^{(i)}(\theta) f(X | \theta)}{m(X | \pi^{(i)})}, i=1, 2$$

综上可得:

$$\pi(\theta | X) = \frac{[w_1 m(X | \pi^{(1)}) \pi^{(1)}(\theta | x) + (1 - w_1) m(X | \pi^{(2)}) \pi^{(2)}(\theta | x)]}{m(X | \pi)}$$

$$\text{令: } \lambda_1(X) = \frac{w_1 m(X | \pi^{(1)})}{m(X | \pi)},$$

$$\text{则: } 1 - \lambda_1(X) = \frac{1 - w_1}{m(X | \pi)} m(X | \pi^{(2)})$$

$$\text{又因为: } m(X | \pi) = \int_{\Theta} f(X | \theta) \pi(\theta) d\theta$$

$$= \int_{\Theta} f(X | \theta) [w_1 \pi^{(1)}(\theta) + (1 - w_1) \pi^{(2)}(\theta)] d\theta$$

$$= w_1 m(X | \pi^{(1)}) + (1 - w_1) m(X | \pi^{(2)})$$

可得:

$$(1 - w_1) m(X | \pi^{(2)}) = m(X | \pi) - w_1 m(X | \pi^{(1)})$$

$$\pi(\theta | X) = \lambda_1(X) \pi^{(1)}(\theta | X) + [1 - \lambda_1(X)] \pi^{(2)}(\theta | X) \quad (6)$$

上面的推导给出了在两种验前信息的情况下, 指标 θ 的验后分布是由两个验后分布融合而成, 一个是在验前密度 $\pi^{(1)}(\theta)$ 下初期使用阶段子样的验后分布 $\pi^{(1)}(\theta | X)$; 另一个是在验前密度 $\pi^{(2)}(\theta)$ 下

初期使用阶段子样的 θ 验后分布 $\pi^{(i)}(\theta|X)$ 。对二者取加权值后得到融合验后分布。若将其做拓展, 则可得到 L 种验前信息下融合验后分布:

$$\pi(\theta|X) = \sum_{l=1}^L \lambda_l(X) \pi^{(l)}(\theta|X) \quad (7)$$

$$\text{式中, } \lambda_l(X) = \frac{w_l m(X|\pi^{(l)}(\theta))}{m(X|\pi^{(l)}(\theta))}; l=1, \dots, L。$$

四、算例分析

(一) PHM 评估指标分析和数据采集

PHM 系统预测能力度量指标体系的基准诊断度量包含预测成功率 (PHS), 虚预测率 (FPR), 平均偏差 (AB), 平均绝对偏差 (AAB) 等等。论文以预测成功率 (PHS) 为例, 对本文方法进行说明。

预测成功率指的是在用户要求的提前预警时间内, 正确预测故障数目 (如预测所料实际故障时间落入 $E[t_f] \pm t_p$ 范围内的故障数目) 与所有预测故障数目的比率。要考核预测成功率指标, 关键在于确定两个参数, 一是要确定所有预测故障的数目, 这个数据可通过 PHM 系统自身的记录系统获取; 二是要确定实际故障时间落入 $E[t_f] \pm t_p$ 范围内的故障数目, 对于 $E[t_f] \pm t_p$ 范围可通过 PHM 系统自身的记录系统获取, 而对于实际故障时间, 可从使用阶段初期实际统计获得。对于预测成功率 (PHS) 的点估计, 采用下列公式进行计算:

$$PHS = \frac{N_{t_f \in [t_L, t_U]}}{N_{TF \leq \Delta t_a}} \quad (8)$$

式中:

$N_{t_f \in [t_L, t_U]}$ —— 如预测所料实际故障时间落入 $E[t_f] \pm t_p$ 范围内的故障数目;

$N_{TF \leq \Delta t_a}$ —— 用户要求的提前预警时间内, 所有预测故障数目。

预测成功率指标服从二项分布, 其单侧置信下限依据下式计算:

$$\sum_{i=0}^F \binom{n}{i} P_L^{n-i} (1-P_L)^i = 1 - C_1 \quad (9)$$

式中: P_L —— 预测成功率的单侧置信下限; C_1 —— 单侧置信度; n —— 样本量; F —— 预测失败次数。规定时间内, 单侧置信度置信水平 C_1 条件下, 由 PHM 系统预测失败的故障总数, 数值等于 $n - N_{t_f \in [t_L, t_U]}$

按照文中给出的数据采集和使用方法, 对在研制阶段和使用阶段初期使用阶段初期的统计数据进行分析处理, 得到各阶段的 PHM 故障预测信息统计表如下所示:

表1 小样本情况下的故障预测能力数据统计表

项目 数值	$N_{t_f \in [t_L, t_U]}$	$N_{TF \leq \Delta t_a}$	n (个)	t_p (秒)	$E[t_f]$ (秒)
研制阶段	39	40	40	60	784
样机测试阶段	38	40	40	60	815
使用阶段初期	13	15	15	60	886

(二) 计算步骤

下面对小样本情况下的后验分布确定过程进行说明:

步骤1: 通过对研制和样机测试阶段的试验数据进行分析 and 处理后, 两个阶段的数据均通过一致性检验。综合考虑各阶段试验数据的特点, 分别计算使用阶段初期子样 X 之前 H_0 成立的先验概率 $P(H_0)$ [6-7]。假定通过计算可得到研制阶段的 $P(H_0)_Y = 0.75$, 样机测试阶段的 $P(H_0)_S = 0.85$ 。

步骤2: 考虑存在有信息可以利用的情况, 在研制和样机测试阶段, 采用“风险上平等对待”的原则 [8] 来确定 α 和 β , 即可取 $\alpha = \beta$ 。假定 $\alpha = 5\%$, 计算可得研制阶段的可信度 $P(H_0|A)_Y$ 和样机测试阶段的可信度 $P(H_0|A)_S$ 分别为:

$$P(H_0|A)_Y = \frac{(1-\alpha)P(H_0)_Y}{(1-2\alpha)P(H_0)_Y + \alpha} = 0.9828,$$

$$P(H_0|A)_S = \frac{(1-\alpha)P(H_0)_S}{(1-2\alpha)P(H_0)_S + \alpha} = 0.9908$$

由公式 $w_l = b_l / \sum_{l=1}^L b_l, l=1, \dots, L$ (b_l 为各阶段试验的可信度), 得到研制和样机测试阶段的可信度权重分别为: 0.4980, 0.5020。

步骤3: 由各阶段数据分别获得验前分布密度 $\pi^{(l)}(\theta), l=1, \dots, L$ (由于只针对研制和样机测试数据, 所以 $l=2$), 并在步骤2的基础上得到其融合验前密度 $\pi(\theta)$ 。

在 Bayes 方法中, 最常用的先验分布为共轭分布, 即先验分布与后验分布具有相同的函数形式。先验分布中所含的未知参数称为超参数。PHS 服从二项分布, 其共轭先验分布是贝塔分布。因此, 可选取 Beta 分布 $\beta(\alpha_0, \beta_0)$ 作为 PHS 的先验分布, 先验分布密度函数为:

$$\pi(q; \alpha_0, \beta_0) = \frac{\Gamma(\alpha_0 + \beta_0)}{\Gamma(\alpha_0)\Gamma(\beta_0)} q^{\alpha_0-1} (1-q)^{\beta_0-1} \quad (10)$$

其中 α_0, β_0 即为待确定的超参数, q 为待评估的参数。根据先验信息获得 q 的先验均值 $\hat{q}_0$ 和先验方差 $\hat{\sigma}_0^2$, 根据矩等效法, 令这两个参数分别等于 Beta 分布的期望和方差, 即:

$$\frac{\alpha_0}{\alpha_0 + \beta_0} = \hat{q}_0; \quad \frac{\alpha_0 \beta_0}{(\alpha_0 + \beta_0)^2 (\alpha_0 + \beta_0 + 1)} = \hat{\sigma}_0^2 \quad (11)$$

求解上式方程组, 可得先验分布的超参数 α_0, β_0 。

通过计算可得: 研制阶段的 $\alpha_0 = 0.074, \beta_0 = 0.007$; 样机测试阶段的 $\alpha_0 = 0.078, \beta_0 = 0.006$ 。

由此, 就可以分别确定研制和样机测试阶段的验前分布密度 $\pi^{(l)}(\theta), l=1, \dots, L$

当考虑研制和样机测试两个阶段的试验数据的情况下, 即取 $L=2$ 时, 此时有两种验前信息, 其融合验前密度 $\pi(\theta)$ 为:

$$\pi(\theta) = w_1 \pi^{(1)}(\theta) + w_2 \pi^{(2)}(\theta) \\ = 0.498 \times \beta(0.074, 0.007) + 0.502 \times \beta(0.078, 0.006)$$

步骤4：由使用阶段初期数据获得似然函数 $f(X|\theta)$ 。根据使用阶段初期统计数据，采用最大似然估计，得到似然函数 $f(X|\theta)$ 。在二项分布条件下，假定 q 为评估参数， n 为使用阶段初期数据样本量， i 为预测失败次数，则关于 PHS 的似然函数为： $f(X|q) = \binom{n}{i} q^{n-i} (1-q)^i$ ，其中 n 为使用阶段初期数据样本量， $n=15$ ， i 为预测失败次数， $i=2$ 。

步骤5：后验分布 $\pi(\theta|X)$ 的确定。由 Bayes 公式可得：

$$\pi(\theta|X) = \frac{f(X|\theta)\pi(\theta)}{\int_{\Theta} f(X|\theta)\pi(\theta)d\theta}$$

将步骤3得到的融合验前密度 $\pi(\theta)$ 及步骤4得到的似然函数 $f(X|\theta)$ 代入上式即可求得后验分布 $\pi(\theta|X)$ 。后验分布 $\pi(\theta|X)$ 可写为如下形式，以方便计算。

$$\pi(\theta|X) = \frac{1}{m(\theta|\pi)} \sum_{i=1}^L w_i \pi^{(i)}(\theta) f(X|\theta)$$

式中， $m(\theta|\pi) = \int_{\Theta} f(X|\theta)\pi(\theta)d\theta$ ，它为 X 的边缘密度。

在给单侧置信度 $C = 0.9$ 下，通过计算得到故障预测成功率的置信区间为：[0.8357, 1]。

五、结论

论文针对产品使用初期阶段缺少数据的情况，对 PHM 系统预测能力指标评估过程中的故障预测能力数据收集、数据样本分析及转化等问题进行了阐述，采用基于 Bayes 方法的多源验前信息融合方法，并通过具体的案例对论文提出的方法进行了验证。这对于缺少使用数据情况下的 PHM 系统能力指标评估问题具有一定的意义。

参考文献

- [1] 孙博，康锐，谢劲松. 故障预测与健康管理研究现状综述 [J]. 系统工程与电子技术，2007，29(10): 1762–1767.
- [2] 邢国强，闫雪，侯代文. 基于验前信息可信度的鱼雷实航工作可靠性 Bayes 融合评估方法 [J]. 鱼雷技术，2012.20 (2): 139–142.
- [3] 姚志军，王国平，王广伟. 验前信息可信度及其在 Bayes 评估中的应用 [J]. 火力与指挥控制. 2007.32(7): 51–53.
- [4] 李鹏波，谢红卫，张金槐. 考虑验前信息可信度时的 Bayes 估计 [J]. 国防科技大学学报. 2003.25(4): 107–109.
- [5] Hess A, Frith P, Calvello G. Challenges, Issues, and Lessons Learned Chasing the Big ‘P’: Real Prognostics Part 1[C]. 2005 IEEE Conference, March 2005.
- [6] Calvello G, Dabney T, Hess A. PHM a key enabler for the JSF autonomic logistics support concept[C]. 2004 IEEE Conference, March 2004.
- [7] Smeulders M, Zeelen R, Bos A. PROMIS—a generic PHM methodology applied to aircraft subsystems[C]. 2002 Aerospace Conference Proceedings, 2002, 6: 3153–3159.
- [8] Wu Mingqiang, Fang Hongzheng, Wen Bowu. Research on Aircraft Prognostics and PHM Integrated Engineering Environment [J]. Computer Measurement & Control, 2011, 19(1): 98–101.
- [9] 李松，武文华，姚伟岸. 基于实测数据的半潜式平台贝叶斯模型修正方法研究 [J]. 船舶力学，2024，28(8): 179–190.
- [10] 周单. 基于空变数据的相控贝叶斯—序贯高斯模拟方法 [J]. 石油物探，2024.63 (3): 610–622.
- [11] 刘昊邦，陈童，胡涛，李明贵，杜凯. 基于多批次增长实验的舰空导弹命中概率贝叶斯估计 [J]. 系统工程与电子技术. 2024. <https://link.cnki.net/urlid/11.2422.TN.20240709.1004.008>.
- [12] 刘昊邦，史宪铭，屈金标，荣丽卿. 基于多项分布的导弹命中精度贝叶斯验收方案设计 [J]. 北京航空航天大学学报，2024，50(6): 1991–2000.
- [13] 于小岚，熊伟. 基于 DBN 的天基信息支援装备体系作战效能评估方法 [J]. 指挥控与仿真. 2023.45(4): 14–23.
- [14] 赵师，刘洪坤，孙文纪. 贝叶斯在装备技术实验结果信息解释决策中的应用 [J]. 火力与指挥控制，2011，36(8): 56–59.
- [15] 郭玉鹏，巢蕾. 通讯装备诊断贝叶斯网络结构构建方法研究 [J]. 现代电子技术. 2010.23: 99–101.