

基于小样本故障率的智能电能表 可靠度预估模型

陈祉如^{1,2}, 代燕杰^{1,2}, 杜艳^{1,2}, 董贤光^{1,2}, 张志^{1,2}, 荆臻^{1,2}

(1. 国网山东省电力公司营销服务中心(计量中心), 山东 济南 250000; 2. 国网山东省电力公司电力科学研究院, 山东 济南 250000)

摘要: 基于故障率数据的可靠度评估是智能电能表健康状态管理与维修的重要依据。然而, 异常值及故障率的小样本特性限制了传统智能电能表可靠度预测模型的评估性能。为此, 提出一种基于加权局部离群因子与高斯过程回归的多环境应力下智能电能表故障率预估模型。首先, 建模采用一种加权局部离群因子识别并剔除故障率数据集中的异常值; 然后, 选用不同的核函数匹配典型环境下的多应力输入特征, 选取最优核; 最后, 以高斯过程后验分布预测故障率的置信度 95% 的区间变化, 并据此估计智能电能表可靠度。采用 2 个典型环境地区的智能电能表故障样本进行实例分析, 结果表明所提模型可有效预测智能电能表在多环境应力下故障率变化趋势, 并能准确求解其可靠度。

关键词: 智能电能表; 故障率; 小样本; 高斯过程; 可靠度预估

DOI: 10.19781/j.issn.1673-9140.2023.01.025 中图分类号: TM933.4 文章编号: 1673-9140(2023)01-0218-08

Reliability prediction model based on small sample failure rate of smart meter

CHEN Zhiru^{1,2}, DAI Yanjie^{1,2}, DU Yan^{1,2}, DONG Xianguang^{1,2}, ZHANG Zhi^{1,2}, JING Zhen^{1,2}

(1. State Grid Shandong Electric Power Company Marketing Service Center (Metering Center), Jinan 250000, China;

2. Power Research Institute, State Grid Shandong Electric Power Company, Jinan 250000, China)

Abstract: Reliability evaluation based on the failure rate data is an important basis for the health status management and maintenance of smart meters. However, the small sample characteristics of outliers and failure rates limit the evaluation performance of traditional smart energy meter reliability prediction models. Therefore, a prediction model of smart meter failure rate under multi-environment stress based on weighted local outlier factor and Gaussian process regression is proposed in this paper. Firstly, a weighted local outlier factor is employed with the model to identify and then delete potential outliers in failure rate data sets; then, different kernel functions are selected to match the characteristics of multiple stress inputs in typical environments, and choose the best one. Finally, the interval change of the 95% confidence level of the failure rate is predicted by the posterior distribution of the Gaussian process, and the interval reliability is obtained based on this. Case analysis of fault samples of smart meters in two typical environmental areas shows that the proposed model could effectively predict the trend of failure rate of smart meters under multi-environmental stress, and could accurately solve its reliability.

Key words: smart meter; failure rate; small sample; gaussian process; reliability prediction

收稿日期: 2021-09-18; 修回日期: 2021-11-19

基金项目: 国网山东省电力公司科技项目(520626200021)

通信作者: 代燕杰(1983—), 女, 硕士, 高级工程师, 主要从事电力营销研究; E-mail: daiyanjie@sd.sgcc.com.cn

随着中国智能电网建设的大力推进,智能电能表已广泛应用于工业以及城市电能测量^[1]。近年来,对基于智能电能表故障率数据的可靠性评估的需求日益显现,智能电能表可靠性评估与预测工作对故障诊断及预警方面具有重要指导意义^[2]。智能电能表中包含有大量电子元器件,长期运行在极端环境条件下易导致老化,从而造成设备故障^[3]。智能电能表分布广、数量多的特点导致难以收集大量且有效的智能电能表故障样本信息,同时样本中的噪声干扰也增加了对智能电能表故障率预估的难度。

国内外学者对于智能电能表的故障率预估与可靠性预测问题开展了大量的研究。常用方法可分为两类:直接评估法和数据驱动方法^[4]。直接评估法利用电子设备各个部件的故障率直接分析设备可靠性,故障树分析是目前常用的一种直接评估方法,通过故障树分析可以获得系统的结构信息,进而分析各个子模块的失效概率^[5]。文献[6]提出一种基于故障树的智能电能表可靠性分摊算法,通过选择合适的可靠性指标实现对系统的可靠性分析;文献[7]针对复杂的系统建模问题提出动态故障树,利用动态故障树可以指定系统故障机制的动态行为;文献[8]通过元器件应力法和威布尔分布建立一种智能电表可靠性预计的优化模型。但这些直接评估模型需要电能表组件失效数据,无法应用于系统级失效分析。

数据驱动方法直接利用系统的故障数据来评估系统的可靠性。依据可靠性评估方式的差异,数据驱动方法进一步分为确定性方法和概率性预测方法^[9-10]。常用的确定性方法有人工神经网络、支持向量回归、自回归滑动模型及其组合算法。文献[11]结合主成分分析与支持向量回归对电能计量设备进行状态评价和误差实时评估;文献[12]结合人工神经网络与威布尔分布预测轴承的剩余寿命。然而,这些模型对样本需求量高,而实际的故障率具有小样本的特点,并不满足确定性方法对样本量的需求。

概率方法则不要求大量训练样本,常用的概率方法有维纳过程、高斯过程回归等。维纳过程通过逆高斯分析实现对设备性能退化的寿命分布预测

与可靠性分析,文献[13]利用加速退化试验数据建立的非线性维纳过程退化模型实现对智能电能表运行剩余寿命预测,但维纳过程运算复杂,具体实现成本太高。高斯过程回归(Gaussian process regression, GPR)因其严格的统计学基础以及其自身较强的泛化能力,在解决小样本、高维非线性问题上的独特优势受到人们广泛的关注^[14]。文献[15]利用GPR实现对闭环单端初级电感变换器系统级的故障评估;文献[16]通过构建锂离子电池健康因子,利用GPR模型间接预测其剩余使用寿命。然而,上述研究未充分考虑多个环境应力与设备性能特征的融合,也忽略了带异常值的小样本数据对预测性能的影响。

针对上述情况,本文利用国网新疆维吾尔自治区、黑龙江省的智能电能表故障数据,从数据处理与多源信息融合的角度出发,将加权局部离群因子(weighted local outlier factor, wLOF)应用于原始故障率数据中的异常值检测,对加权后的数据进行高斯过程核匹配,并基于最优核函数对智能电能表的故障率进行区间预测,借助故障率与可靠度机理分析对智能电表可靠度进行建模。最后,结合2个典型地区的真实故障率数据,利用软件库PyMC3验证所提方法的有效性。

1 故障率数据来源与异常值检测

1.1 故障数据来源

为研究不同典型环境对智能电能表故障率的影响,本文选取新疆维吾尔自治区(A)、黑龙江省(B)各7个地区的故障率数据进行分析。其中所有智能电能表样本均来源于同一电表供应商,故障率统计时间跨度为2011年到2017年。

A、B的智能电能表年故障率与时间的变化关系如图1、2所示。图1、2中,两地区的智能电能表在进行采样的前两年其故障率近乎为0。此外,各采样地区的智能电能表的故障率呈现出较强的时间相关性。同时有部分数据明显偏移数据整体的密集区域,如图1中2012年地区2的智能电能表故障率,图2中2013年地区6的智能电能表故障率。

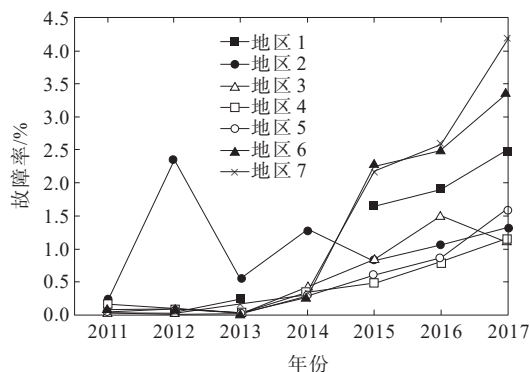


图1 A智能电能表故障率数据

Figure 1 Failure rate data of smart electric energy meters in A

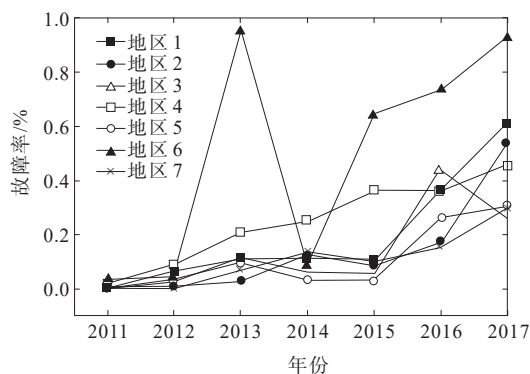


图2 B智能电能表故障率数据

Figure 2 Failure rate data of smart electric energy meters in B

1.2 基于wLOF的故障率数据噪声点检测

由于数据传输故障及操作人员的人为误差等因素,实际收集的故障数据中会引入异常值数据。异常值影响智能电能表故障率的预估结果,也易导致模型过拟合。传统的故障率分析方法难以在小样本与含噪声的故障率数据的变化中达到平衡。为提高数据的有效性,必须对异常值进行检测并处理。LOF算法通过衡量数据点与其邻域数据点之间的密度比值来检测异常数据,适用于在小样本条件下的异常数据检测。

设原始故障率数据 $D_r = \{(t, X_{i,j,t}, y_{i,j,t})\}$, 其中 t 为统计时间, $y_{i,j,t}$ 为省份 j 的第 t 年的第 i 个地区的故障率值。 X 为环境应力, 由温度应力 (T_c) 与湿度应力 (R_H) 组成。

对于数据点 $q(x_q, y_q) \in D_r$ 与其第 k 近邻 $p(x_p, y_p)$ 两者之间的距离表示为

$$D_k(q) = \sqrt{(x_q - x_p)^2 + (y_q - y_p)^2} \quad (1)$$

其中, LOF算法采用欧式距离进行距离度量。

然而欧式距离并未考虑到各维度之间的差异性, 采用欧氏距离难以兼顾时间应力与环境应力同智能电能表故障率之间的相关性。对此, 本文提出了考虑应力与故障率相关性的加权欧式距离, 并以此改写式(1)为

$$D_k^w(q) = \sqrt{[w_d(x_q - x_p)^2] + (y_q - y_p)^2} \quad (2)$$

其中, 权值 w_d 表示不同应力对故障率的影响程度。

利用斯皮尔曼相关性分析来进行求解 w_d 值, 对于样本大小为 n 的故障率数据集 D_r , 斯皮尔曼相关性分析的表达式为

$$w_d = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (3)$$

其中, d_i 为各个应力与故障率之间的秩差。

$$d_i = \text{rank}(t, X_{i,j,t}) - \text{rank}(y_{i,j,t}) \quad (4)$$

由此方法检测异常值, 并基于不含异常值的数据集 $D'_r = \{(t, X_{i,j,t}, y'_{i,j,t})\}$ 提出一种高斯过程优选核方法进行故障评估与可靠性预测。

2 可靠度评估模型构建

2.1 故障率求解可靠度

故障率 $\lambda(t)$ 是定义在 t 时刻未故障的产品在 t 时刻以后单位时间内故障的概率, 即

$$\begin{aligned} \lambda(t) &= \frac{\Delta n(t)}{\Delta t (N - n(t))} = \frac{\Delta n(t)}{N \cdot \Delta t} \cdot \frac{N}{N - n(t)} = \\ &= f(t) \cdot \frac{1}{R(t)} = \frac{dF(t)}{dt} \cdot \frac{1}{R(t)} = \\ &= \frac{d(1 - R(t))}{dt} \cdot \frac{1}{R(t)} = \frac{-R'(t)}{R(t)} \end{aligned} \quad (5)$$

式中, $\Delta n(t)$ 为 Δt 时间内新增故障数; N 为样本总数; $n(t)$ 为已发生故障数; 概率密度 $f(t)$ 为累计故障率分布函数 $F(t)$ 的导数。

对故障率 $\lambda(t)$ 进行积分, 有

$$\int_0^t \lambda(t) dt = \int_0^t \frac{-R'(t)}{R(t)} dt = -\ln R(t) \quad (6)$$

$$R(t) = \exp\left(-\int_0^t \lambda(t) dt\right) \quad (7)$$

因此, 准确可靠度的评估工作是建立在准确的

故障率预估基础之上,采用概率算法估计出区间范围内的故障率往往比点估计方法的预测结果更为可信。

2.2 高斯过程原理

高斯过程(Gaussian process, GP)是由 Williams 和 Rasmussen 提出的一种融合机器学习和统计学的贝叶斯概率建模方法,其本质是进行具有概率意义的核函数学习,通过对历史样本数据的学习得到最优超参数,进而得到预测模型,对新样本进行预测^[17]。GPR 高斯过程回归能给出预测数据样本的后验分布,并提供结果的不确定性,保证了模型的可解释性。与线性回归模型相比,GPR 具备更强的非线性拟合能力;与支持向量机相比,GPR 结果更好解释;相对于神经网络,GPR 不需要额外的优化或正则化方案。但 GPR 涉及协方差矩阵求逆,复杂度为 $O(n^3)$,所以更适合小样本数据的训练。

GP 假定随机变量的任意有限集合,每个变量都有一个联合高斯分布,其性质完全由均值函数和协方差函数决定,即

$$f(\boldsymbol{x}) \sim N(f|m, K) \quad (8)$$

式中, \boldsymbol{x} 为带环境变量的输入向量;均值函数 m 为样本均值;协方差函数 $K = k(x_i, x_j)$ 为样本之间相似性, $x_i, x_j \in \mathbf{R}_d$ 为任意随机变量。

考虑到观测目标值 y 的噪声,可以建立高斯过程回归问题的一般模型,即

$$y = f(\boldsymbol{x}) + \varepsilon \quad (9)$$

式中, ε 为独立的高斯白噪声,满足 $\varepsilon \sim N(0, \sigma_n^2)$ 。

观察值 y 的先验分布为

$$y \sim \text{GP}(m(\boldsymbol{x}), k(x_i, x_j) + \sigma_n^2 \delta_{ij}) \quad (10)$$

式中, δ_{ij} 为 Kronecker delta 函数,当 $i = j$ 时, δ_{ij} 为单位矩阵 I 。

然后得到训练集的输出 y 和测试集的输出 $f(\boldsymbol{x})$ 的联合概率分布为

$$\begin{bmatrix} y \\ f \end{bmatrix} \sim \text{GP} \begin{pmatrix} k(x_i, x_j) + \sigma_n^2 \delta_{ij}, & k(\boldsymbol{x}, \boldsymbol{x}^*) \\ k(\boldsymbol{x}^*, \boldsymbol{x}), & k(\boldsymbol{x}^*, \boldsymbol{x}^*) \end{pmatrix} \quad (11)$$

式中, $k(\boldsymbol{x}, \boldsymbol{x})$ 为训练数据自身协方差函数; $k(\boldsymbol{x}, \boldsymbol{x}^*)$ 为测试数据与训练数据协方差函数。

由此计算预测值后验分布为

$$\begin{cases} f | \boldsymbol{x}, y, \boldsymbol{x}^* \sim N(\bar{f}, \text{Cov}(f)) \\ \bar{f} = k(\boldsymbol{x}, \boldsymbol{x}^*) [k(\boldsymbol{x}, \boldsymbol{x}^*) + \sigma_n^2 \delta_{ij}]^{-1} y \\ \text{Cov}(f) = k(\boldsymbol{x}, \boldsymbol{x}^*) - k(\boldsymbol{x}^*, \boldsymbol{x}) \\ [k(\boldsymbol{x}, \boldsymbol{x}^*) + \sigma_n^2 \delta_{ij}]^{-1} k(\boldsymbol{x}, \boldsymbol{x}^*) \end{cases} \quad (12)$$

式中, \bar{f} 为测试点 \boldsymbol{x}^* 对应的预测值 f 的均值; $\text{Cov}(f)$ 为其方差。

2.3 高斯过程核函数

GP 常见的核函数有指数平方核(squared-exp, SE), 周期核(periodic, PE), 有理二次核(rational quadratic, RQ), 线性核(linear, LIN) 以及马顿核(matérn, MA) 等。每类核函数都代表对目标函数预先的假设,适合不同的数据特征。SE 核适合局部变化特征的数据; PE 核适合周期重复变化结构的数据; RQ 核适合多尺度变化结构的数据; MA 核假定数据具有更细节的变化。对于具有复杂结构的数据,则可由简单核函数进行“加”或者“乘”的组合形式进行表示。例如,“SE+PE”核表示数据具有局部周期变化特征,“LIN×SE”核表示数据加速变化的特征等。

$$k_{\text{LIN}} = \eta_i^2 (x - c)(x' - c) \quad (13)$$

$$k_{\text{SE}} = \eta_i^2 \exp \left[-2 \frac{(x - x')^2}{2t^2} \right] \quad (14)$$

$$k_{\text{MA}} = \eta_i^2 \left(1 + \frac{\sqrt{5(x - x')^2}}{t} + \frac{5(x - x')^2}{3t^2} \right) \cdot \exp \left[-\frac{\sqrt{5(x - x')^2}}{t} \right] \quad (15)$$

$$k_{\text{RQ}} = \eta_i^2 \left(1 + \frac{(x - x')^2}{2\alpha t^2} \right)^{-\alpha} \quad (16)$$

式(13)~(16)中, η_i 为垂直尺度因子; t 为水平尺度因子; α 为 RQ 核函数调节尺度变化的调节因子。

2.4 参数求解

为建立基于故障率预测模型,需求解参数,不同于传统寻优方法,将参数初值设定为固定数值进行网格搜索或遗传搜索等,本文为 Weibull 的参数设置先验分布。再利用马尔科夫链-蒙特卡洛(Markov chain Monte Carlo, MCMC)方法对故障率观测数据进行贝叶斯推断,以推断得到的后验均值

为最优值。

参数先验的选取需满足一定条件,在已知参数先验信息均大于0的条件下,可取其先验分布 f 为具有正支撑性的 Half Cauchy 分布,并简记为 $\theta \sim \text{HalfCauchy}(\nu)$;对于某些参数如水平尺度 ι ,由于长度尺度多变,可取信息量丰富的 Gamma 分布,简记为 $\theta \sim \text{Gamma}(b, a)$ 。

$$f(x, \nu) = \frac{2}{\pi \nu [1 + (x/\nu)^2]} \quad (17)$$

$$f(x, b, a) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx) \quad (18)$$

高斯过程核函数的垂直变化尺度 η_i 为正,水平变化尺度 ι 及 RQ 核的调节因子 α 多变,本实验设定核函数超参数的先验为 $\eta_i \sim \text{HalfCauchy}(5)$; $\iota \sim \text{Gamma}(4, 0.1)$; $\alpha \sim \text{Gamma}(5, 5)$ 。设定好先验分布之后执行贝叶斯推断进行参数求解。

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \quad (19)$$

式中, y 为基于 GP 预测的后验数据样本; θ 为超参数集合。

3 智能电能表实例分析

本文利用 PyMC3 库作为故障率建模工具,以 Python3.8 作为实验分析的仿真环境。

3.1 高斯过程优选核

为确定匹配数据集的最优核函数,必须对不同核函数进行数据特征匹配。核函数选择需在模型复杂度与模型对数据集描述能力之间寻求最佳平衡,赤池信息准则 (Akaike information criterion, AIC) 和贝叶斯信息准则 (Bayesian information criterion, BIC) 是常用的信息准则,但广泛应用信息准则 (W_{AIC}) 更能全面地描述模型信息。鉴于组合核函数较之单核函数超参数数量增加使模型复杂度增加的弊端,先采用 WAIC 准则作为核函数组合与选择的依据,然后选择均方根误差 (R_{MSE}) 与决定系数 (R^2) 来进行预测结果评价。

$$W_{\text{AIC}} = \sum_{j=1}^N \log p(y^*) - \sum_{j=1}^N V(\log p(y^*)|\theta) \quad (20)$$

式中, $p(y^*)$ 为 GP 后验分布; $V(\log p(y^*)|\theta)$ 为似然函数方差。

$$R_{\text{MSE}} = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - y^*)^2} \quad (21)$$

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - y^*)^2}{\sum_{i=1}^m (y_i - \bar{y})^2} \quad (22)$$

式(21)、(22)中, y_i 为第 i 个故障率真值; y^* 为相应预测值。

高斯过程回归多种核模型在 WAIC、RMSE 和 R^2 的比较结果如表 1 所示。首先对几种常见的单核进行 WAIC 信息准则判定, LIN 核函数包含了对数据呈线性变化的假定, WAIC 值为 -78.83 远小于其他单核,说明数据不符合线性变化。SE 核、MA 核、RQ 核的 WAIC 值较为接近,说明故障率数据具有多尺度变化的特征。MA 核是 SE 核的变体,且 WAIC 值较大,故选取 MA 核与 RQ 核进行组合,结果表明,组合核 MA+RQ、MA*RQ 和 LIN+MA*RQ 的 WAIC 值仍在 $[-55, -53]$ 内,其中 MA*RQ 的 WAIC 最大,表明数据具有细微多长度尺度变化特征。然后基于 RMSE 与 R^2 的结果进行再次判定。由表 1 可知,智能电能表故障率数据在 LIN+MA*RQ 核模型中, $R_{\text{MSE}}=0.0239$ 最小,而 R^2 值最大,该核在预测性能上为最优核。

表 1 不同核模型预测性能对比

Table 1 Comparison of prediction performance of different kernel models

核函数	W_{AIC}	R_{MSE}	R^2
LIN	-78.83	0.704 7	-1.101 2
SE	-54.16	0.040 5	0.936 7
MA	-53.46	0.027 6	0.956 8
RQ	-53.90	0.027 1	0.957 6
MA+RQ	-53.44	0.027 5	0.957 0
MA*RQ	-53.00	0.025 3	0.960 5
LIN+MA*RQ	-54.24	0.023 9	0.962 6

3.2 回归模型预测性能对比

进一步验证所提方法的预测性能,对采取 wLOF 异常值处理后数据集进行不同模型的对比实验如表 2 所示。

表 2 不同模型预测性能对比
Table 2 Comparison of prediction performance of different models

模型	R_{MSE}	R^2
神经网络	0.005 8	0.990 9
支持向量机	0.088 4	0.861 9
多核高斯过程	0.042 1	0.935 7
本文	0.023 9	0.962 6

由表 2 可知,神经网络具有最大的 R^2 和最小的 RMSE,支持向量机在预测性能上最差, R^2 仅为 0.861 9 远小于本文模型和神经网络,同时预测误差较大。这从侧面说明了神经网络在非线性预测时的优势,但神经网络内部结构不透明,可解释性较

差。本文基于 GP 的模型能够很好地将数据特征进行阐明,有利于工程应用,并且能给出概率预测结果;预测效果稍次于神经网络的情况下本文模型仍为最佳模型。同时,经过 wLOF 处理的本文模型在预测性能和拟合能力上也比多核高斯过程有所提升。

3.3 故障率预测与可靠度求解

利用高斯过程回归预测每年的智能电能表故障率,基于数据执行 MCMC 对进行参数推断,以 A 智能电能表数据为例,求解结果如表 3 所示。对于变化规律一致或相似的数据集合,高斯过程回归的后验参数仅在垂直尺度因子和水平尺度因子数值有所不同。B 智能电能表故障率数据的后验参数为 $\theta = [0.499, 34.059, 0.994, 38.902, 3.439, 37.306, 0.093]$ 。

表 3 最优核 GP 参数求解结果
Table 3 Calculation results of each cluster group centroid

超参数 θ	η_1	ι_1	α	ι_{rq}	η_{mq}	ι_{ma}	δ
均值	0.594	34.680	0.995	35.694	9.027	32.737	0.645
均方差	0.920	18.753	0.437	19.747	8.220	17.964	0.072
2.5% 置信下限	0.000	6.250	0.277	6.311	0.878	5.337	0.522
97.5% 置信上限	1.828	69.928	1.811	70.719	21.125	65.147	0.789

表 3 中给出了各超参数的贝叶斯推断结果均值及其 2.5% 置信下限和 97.5% 置信上限的数值。 $\eta_1 = 0.594$ 、 $\eta_{mq} = 9.027$ 表明数据线性变化的幅度小于不规则局部变化的幅度。而 LIN 核、MA 核及 RQ 核函数的水平变化尺度分别为 $\iota_1 = 34.68$ 、 $\iota_{ma} = 32.737$ 、 $\iota_{rq} = 35.694$, 这表明在本模型会保持长期线性局部多尺度的不规则变化趋势,执行本模型进行短期故障率预测是准确的。

预测 A、B 智能电能表运行 10 a 的故障率变化情况,并与神经网络、支持向量机的预测结果作对比如图 3 所示。

由图 3 可知,神经网络对历史数据的拟合跟踪能力最强,但在 7~10 a 的预测值上出现了加速增大的现象,且不被本文模型结果 95% 置信区间所涵盖;而支持向量机的预测结果被涵盖在本文模型的 95% 置信区间内。由图 4 可知,从 B 智能电能表的预测数据来看,神经网络和支持向量机的预测结果均涵盖在本文模型的 95% 置信区间内。这种差别在于小样本数据不能使神经网络得到足够的样本训练,其预测易受小样本数据的变化趋势差别影响,这也从侧面说明本文模型的鲁棒性。同时,在算例数据预测前期,会出现高斯过程回归算法概率预测区间有负值的现象,这与执行算法过程中的早期故障率数据接近于 0 有关。高斯过程在函数视角下由均值函数 m 和核函数 K 决定,高斯过程回归预测数据的原理是依靠其与原始数据的距离关系, m 控制全区间高斯过程值的均值,这里将 m 设置为 0。因此,在前期原始数据接近 0 的情况下会出现概率区间预测下限为负值的情况。

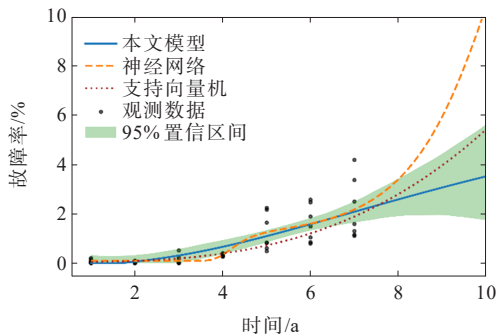


图 3 A 故障率不同模型预测

Figure 3 Failure rate prediction diagram of smart meter in A with different methods

在核函数耦合预测数据点和原始数据点的关系方面,位于历史数据区间内可以完全跟踪拟合数

据变化情况;对于短期预测,可以对数据趋势变化做出预测,但由于预测点与训练数据点距离增加,相关性降低,预测区间下限会出现下降的情况;对于长期预测高斯过程回归将会出现均值为0且带有固定带宽的概率预测结果。因此,在图3、4中,7~10 a属于短期预测,会有概率区间略微下降的现象。基于本文方法由故障率数据得出智能电能表的可靠度预测结果如图5、6所示。

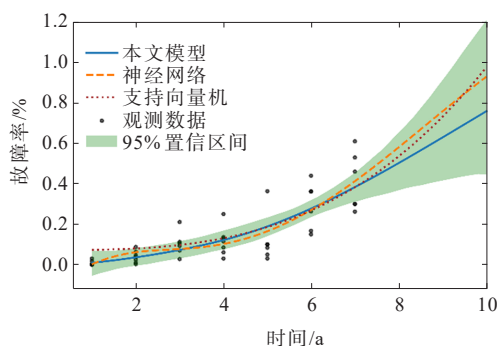


图4 B故障率不同模型预测

Figure 4 Failure rate prediction diagram of smart meter in B using different methods

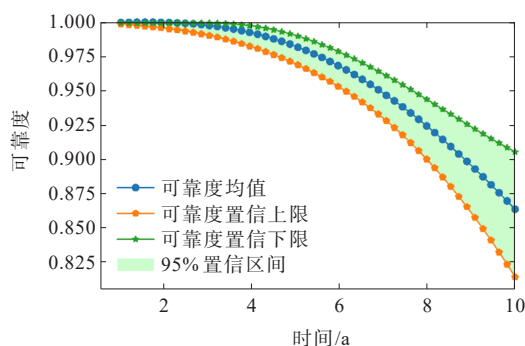


图5 A智能电能表可靠度预测

Figure 5 Reliability prediction diagram of smart meter in A

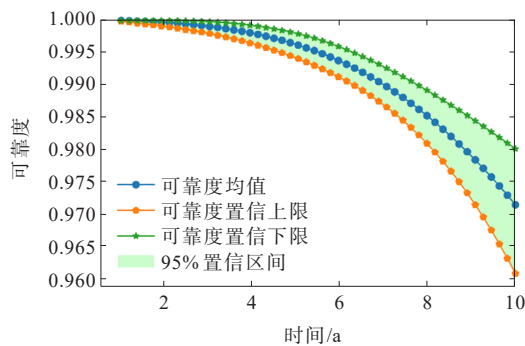


图6 B智能电能表可靠度预测

Figure 6 Reliability prediction diagram of smart meter in B

从故障率原始数据的变化幅度来看,投放在A的智能电能表故障率大于B。可靠度预测结果表明,运行10 a后,A的智能电能表的可靠度下降为0.863;而B省智能电能表的可靠度仍高达0.97。因此,应着力对A的智能电能表的故障原因进行排查,并做出对应的检修策略改进或智能电能表制造厂商的工艺改进。

4 结语

本文提出一种基于智能电能表小样本故障率的可靠度评估模型,利用wLOF对多源故障率数据进行异常值检测,去除异常值后的数据集能明显提高预测方法的预测精度。采用多评价指标WAIC、RMSE等指导高斯过程核函数的选择与建立。贝叶斯推断方法求解超参数更为精确。来自实际统计的智能电能表故障率数据实验结果表明,本文模型能够有效预测智能电能表可靠度的变化趋势,并可为投放在不同地区的智能电能表可靠性评估工作提供参考。

参考文献:

- [1] 杜蜀薇,成达,邱伟,等.基于分层贝叶斯的电能表短期故障评估与预测[J].电子测量与仪器学报,2018,32(12):178-184.
DU Shuwei, CHENG Da, QIU Wei, et al. Short-term fault evaluation and prediction of energy meter based on hierarchical Bayesian[J]. Journal of Electronic Measurement and Instrument, 2018, 32(12): 178-184.
- [2] 宋晓川,郑宽昀,李俊臣,等.基于计量自动化系统的电能计量异常诊断技术研究[J].电网与清洁能源,2022,38(4):110-116.
SONG Xiaochuan, ZHENG Kuanyun, LI Junchen, et al. Research on abnormal diagnosis technology of electric energy metering based on metering automation system[J]. Power System and Clean Energy, 2022, 38(4): 110-116.
- [3] 刘永光,谭赣江,刘型志,等.基于域自适应的二屏电能表显示屏质量检测方法[J].电力系统保护与控制,2022,50(23):180-187.
LIU Yongguang, TAN Ganjiang, LIU Xingzhi, et al. Display quality detection method for a two-screen watt-hour meter based on domain adaptation[J]. Power System Protection and Control, 2022, 50(23): 180-187.

- [4] QIU W, TANG Q, YAO W X, et al. Probability analysis for failure assessment of electric energy metering equipments under multiple extreme stresses[J]. IEEE Transactions on Industrial Informatics, 2021, 17(6):3762-3771.
- [5] 詹坤, 高广德, 李建忠, 等. 基于定频 PWM 稳压的电流互感器取能电源设计方法[J]. 智慧电力, 2021, 49(12):45-51.
ZHAN Kun, GAO Guangde, LI Jianzhong, et al. Design method of current transformer energy extracting power supply based on constant frequency PWM voltage stabilization[J]. Smart Power, 2021, 49(12):45-51.
- [6] 曹宏宇, 刘惠颖, 殷鑫, 等. 典型环境下智能电能表可靠性指标体系及指标量化[J]. 电测与仪表, 2021, 58(3):190-194.
CAO Hongyu, LIU Huiying, YIN Xin, et al. Reliability index system and quantification of smart meter in typical environment[J]. Electrical Measurement & Instrumentation, 2021, 58(3):190-194.
- [7] MO Y C. A multiple-valued decision-diagram-based approach to solve dynamic fault trees[J]. IEEE Transactions on Reliability, 2014, 63(1):81-93.
- [8] 孙谊嫒, 李宁, 董小顺, 等. 智能电表可靠性预计的优化模型[J]. 电力科学与技术学报, 2017, 32(3):15-21.
SUN Yisheng, LI Ning, DONG Xiaoshun, et al. Optimization model for reliability prediction of smart meters[J]. Journal of Electric Power Science and Technology, 2017, 32(3):15-21.
- [9] 彭宇, 刘大同. 数据驱动故障预测和健康管理综述[J]. 仪器仪表学报, 2014, 35(3):481-495.
PENG Yu, LIU Datong. Data-driven prognostics and health management: a review of recent advances[J]. Chinese Journal of Scientific Instrument, 2014, 35(3):481-495.
- [10] 胡昌华, 施权, 司小胜, 等. 数据驱动的寿命预测和健康管理技术研究进展[J]. 信息与控制, 2017, 46(1):72-82.
HU Changhua, SHI Quan, SI Xiaosheng, et al. Data-driven life prediction and health management technology research progress[J]. Information and Control, 2017, 46(1):72-82.
- [11] 张鼎衢, 杨路, 宋强, 等. 一种用于关口电能计量装置远程校验的多路模拟采集器的研制[J]. 电测与仪表, 2022, 59(9):181-187.
ZHANG Dingqu, YANG Lu, SONG Qiang, et al. Development of a multi-channel analog collector for remote calibration of gateway electric energy metering device[J]. Electrical Measurement & Instrumentation, 2022, 59(9):181-187.
- [12] ALI J B, 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(5):150-172.
- [13] 刘影, 刘岩, 燕凯, 等. 基于智能电能表的智慧城市峰值负荷概率估计[J]. 电测与仪表, 2021, 58(9):166-171.
LIU Ying, LIU Yan, YAN Kai, et al. Peak load probability estimation of smart city based on smart meter[J]. Electrical Measurement & Instrumentation, 2021, 58(9):166-171.
- [14] SHENG H M, XIAO J, CHENG Y H, et al. Short-term solar power forecasting based on weighted gaussian process regression[J]. IEEE Transactions on Industrial Electronics, 2018, 65(1):300-308.
- [15] 陈叶, 韩彤, 魏龄, 等. 基于多分类融合模型的智能电能表故障预测[J]. 电测与仪表, 2022, 59(11):162-168.
CHEN Ye, HAN Tong, WEI Ling, et al. Prediction on fault classification of smart meters based on multi-classification integration model[J]. Electrical Measurement & Instrumentation, 2022, 59(11):162-168.
- [16] 黄友朋, 路韬, 陈亮, 等. 智能电能表批次故障预警和寿命预估方法[J]. 中国电力, 2022, 55(7):87-92.
HUANG Youpeng, LU Tao, CHEN Liang, et al. Failure early warning and life estimation method of batch smart meter[J]. Electric Power, 2022, 55(7):87-92.
- [17] RASMUSSEN C E. Gaussian processes for machine learning[M]. Cambridge: MIT Press, 2006, 56-61.