

分类号\_\_\_\_\_

学校代码 10487

学号 M201971512

密级\_\_\_\_\_

# 华中科技大学

# 硕士学位论文

(学术型☒ 专业型☐)

## 定时截尾数据下的智能电表可靠性模 型优选与可靠性评估

学位申请人：李若茜

学 科 专 业：电气工程

指 导 教 师：肖霞 副教授

答 辩 日 期：2022 年 5 月 19 日

**A Dissertation Submitted in Partial Fulfillment of the Requirements  
for the Master Degree in Engineering**

**Reliability Model Selection and Reliability Estimation of  
Smart Meters with Type-I Censoring Samples**

**Candidate : LI Ruoqian**  
**Major : Electrical Engineering**  
**Supervisor : Assoc. Prof. XIAO Xia**

**Huazhong University of Science and Technology**  
**Wuhan 430074, P. R. China**  
**May, 2022**

## 摘要

智能电表作为高级计量体系和智能电网的基础设施已伴随着智能电网的发展得到广泛应用，其运行可靠性极大地影响着计量公平和电网的安全稳定运行，合理评估智能电表可靠性对其周期检定、轮换工作具有指导意义。但目前针对智能电表可靠性模型优选的研究不多，现场数据未能在可靠性评估中得到充分利用，且融合多源数据的可靠评估方法未考虑到截尾样本的情形。针对上述问题，在获得可处理为定时截尾数据的现场数据和定时截尾加速寿命试验数据的基础上，本文研究了智能电表的可靠性模型优选与可靠性评估，具体内容包括：

（1）结合智能电表的特点分析了备选可靠性模型，根据智能电表可靠性数据来源及其数据特点给出了相应的数据预处理方法，最后分析了两种基于数据融合的智能电表可靠性评估方法的原理及其适用情形。

（2）结合信息论与差分进化算法对智能电表的可靠性模型进行优选：针对截尾样本下可靠性模型参数的极大似然估计难以求解的问题，采用差分进化算法进行求解后，计算 Akaike 信息准则优选了对数正态分布作为该批次智能电表的可靠性模型，结果表明相较于秩回归与拟合优度结合的方法，本文方法所得结果不受累积失效概率估计式的影响。

（3）融合现场数据和加速寿命试验数据提高智能电表可靠性评估的准确度：在结合最大熵与 Bootstrap 方法分析现场数据获得先验分布的基础上，基于贝叶斯方法融合加速寿命试验数据对智能电表的可靠性进行评估，结果表明本文方法所得结果能在缺少现场数据的时段利用加速寿命试验数据对可靠度进行修正的同时，仍对现场数据保有较好的拟合效果，且评估的准确度相较于单一数据源的结果得到提高。

本文的研究对于在截尾数据下优选智能电表可靠性模型、融合不同数据源获得更加准确的智能电表可靠性评估结果有参考价值，可用于指导智能电表的轮换与检修工作。

**关键词：** 可靠性评估；模型优选；差分进化算法；最大熵法； 贝叶斯

## Abstract

As infrastructure of advanced metering system and smart grid, smart meters have been widely used with the development of smart grid. The operational reliability of smart meters greatly affects the fairness of metering and the safe and stable operation of the power grid. Therefore, reasonable evaluation of the reliability of smart meters is of significance for its periodic verification and rotation. However, at present there are not many studies on the optimal selection of the reliability model for smart meters. The field data has not been fully utilized in the reliability evaluation, and the censored samples are not taken into account in the reliability evaluation methods integrating multi-source data. In view of the above problems, on the basis of obtaining field data that can be processed as time-censored data and time-censored accelerated life test data, this paper studies the reliability model selection and reliability evaluation of smart meters. The main researches and contributions of this paper are as follows:

(1) Alternative reliability models are analysed combined with the features of smart meters. According to the reliability data sources and data characteristics of smart meters, the corresponding data preprocessing methods are presented. Finally, the principles and application situations of two smart meter reliability assessment methods based on data fusion are analyzed.

(2) Optimal reliability model is selected by combining information theory and differential evolution algorithm. After using the differential evolution algorithm to solve the maximum likelihood estimates of the reliability model parameters under censored samples, the Akaike information criterion is calculated and the log-normal distribution is selected as the reliability model of this batch of smart meters. It is shown that the results obtained by the method used in this paper are not affected by the cumulative failure probability formula compared with the rank regression method combined with the goodness-of-fit method.

(3) The accuracy of smart meter reliability assessment is improved by fusing field data and accelerated life test data. After obtaining the prior distribution through combining the maximum entropy method and Bootstrap method to analyze the field data, accelerated

life test data is fused based on Bayesian method to evaluate the reliability of smart meters. The findings show that the results obtained by the method used in this paper can use the accelerated life test data to correct the reliability in the period lacking field data, while still maintaining a good fitting effect on the field data and improving the accuracy of the assessment compared to the results based on single data source.

The research in this paper has reference value for optimal selection of reliability model and fusing data from different sources to obtain more accurate reliability evaluation result of smart meters under censored samples, and can be used to guide the rotation and maintenance of smart meters.

**Key words:** Reliability estimation, model selection, differential evolution algorithm, maximum entropy method, Bayes

# 目 录

1 绪论.....	1
1.1 研究背景与意义.....	1
1.2 国内外研究现状.....	2
1.3 存在的问题.....	6
1.4 本文主要内容.....	7
2 智能电表可靠性评估理论基础 .....	9
2.1 常用可靠性模型.....	9
2.2 智能电表可靠性数据处理 .....	11
2.3 基于数据融合的可靠性评估方法 .....	14
2.4 本章小结.....	17
3 基于差分进化算法的智能电表可靠性模型优选 .....	18
3.1 可靠性模型优选准则.....	18
3.2 可靠性模型参数估计方法 .....	19
3.3 基于差分进化算法的智能电表可靠性模型优选 .....	25
3.4 智能电表可靠性模型优选结果分析 .....	32
3.5 本章小结.....	33
4 基于贝叶斯方法的智能电表可靠性评估 .....	34
4.1 贝叶斯可靠性评估相关方法 .....	34
4.2 基于贝叶斯方法的智能电表可靠性评估 .....	37
4.3 智能电表可靠性评估结果分析 .....	45
4.4 本章小结.....	47
5 总结与展望.....	48
5.1 本文主要内容及结论.....	48
5.2 本文的主要创新点.....	49
5.3 展望.....	49
参考文献.....	50

## 1 绪论

### 1.1 研究背景与意义

可持续发展一直是我国关注的焦点，智能电网也是基于该背景所提出的，通过数字化信息网络系统连接发输配用各环节用能设备，对每个用户和节点进行实时监控以构建安全可靠、可实时互动以及高能源效率的电网，通信技术、传感检测技术、控制技术和可再生与分布式能源接入技术等构成了智能电网的关键技术。基于传感检测技术的高级量测体系（Advanced Metering Infrastructure, AMI）能为智能电网运行调控、用户用电行为分析等提供所需的数据，而 AMI 中用于采集用户用电信息的智能电表也随着智能电网的发展得到了广泛的应用，根据统计，截至 2018 年国家电网已累计安装智能电表超过 4.96 亿只、南方电网 8815.72 万只，截至 2019 年国家电网内在运智能电表总数达 4.7 亿只，基本实现全覆盖。

智能电表与主站双向通信和数据传输的功能对计量公平有着极大影响，海量智能电表的投入使用也给电网的安全稳定运行带来了更大的挑战，因此智能电表的可靠性对电网的计量公平、安全运行有重要意义。此外，我国自 2009 年以来安装的智能电表已逐步运行达 8 年需根据计量检定规程的要求开展检定工作<sup>[1]</sup>，但实际检定数据表明超期服役的智能电表其总体合格率仍可达 99%<sup>[2]</sup>，8 年到期即轮换的方式必将造成大量资源的浪费，合理的轮换周期应当以可靠性评估结果为依据。

现阶段在计量自动化系统中存储有尚未被利用的、现场返回的故障电表数据，但因智能电表技术规范要求其寿命不小于 16 年且 10 年的累计故障率不超过 4.25%<sup>[3,4]</sup>，使得在该时间内各批次电表实际收集到的故障表数量有限，现场数据是深截尾样本。部分电表生产厂家通过对其产品开展可靠性试验获得了一定的可靠性试验数据，尽管受成本的影响该数据也为截尾数据，但其截尾程度远小于现场数据，涵盖更长的时间范围，唯一的不足在于试验条件与现场存在差异。如能充分利用智能电表的现场故障数据、寿命试验数据对智能电表开展可靠性评估工作，势必能够为其可靠性保障、周期检定和轮换工作提供参考与指导。

本文基于前述实际情况，将智能优化算法、最大熵法、Bootstrap 方法和贝叶斯方法结合应用于智能电表的可靠性评估工作，旨在解决截尾数据存在于实际所得的可靠性数据中给智能电表可靠性模型优选以及可靠性评估带来不利影响的问题，从数据出发选取最优寿命分布模型并通过融合不同来源的数据充分利用可靠性信息，为智能电表提供更为合理的可靠性评估结果。

## 1.2 国内外研究现状

### 1.2.1 智能电表可靠性评估研究现状

目前智能电表可靠性评估的相关研究按数据源大体可分为基于单一数据源的方法和基于多源数据的方法。

IEEE 给出了 4 种基于单一数据源的电子系统与设备可靠性评估方法，分别为基于手册、基于试验数据、基于现场数据和基于失效物理模型的可靠性评估方法<sup>[5]</sup>，其中除基于失效物理模型的方法受限于失效机理的复杂性而暂时未在智能电表中得到应用外，其余方法均有相关的研究。

基于手册的预计方法主要有元器件计数法和元器件应力法，前者因为缺少对元器件种类和应力条件的考量，所以准确率不高，主要用于比较产品可靠性的差异以初步确定方案，后者建立在元器件可靠度服从指数分布的假设之上，结合元器件种类、应力条件等因素预计可靠性，应用于设计阶段，结果准确性较高。国外对基于手册的可靠性预计方法研究较早，在美国军方颁布 MIL-HDBK-217 标准后，其他一些国家和机构又陆续颁布了 Telcordia SR-332、SN29500、IEC62059-41 和 IEC62380 等预计手册，英国燃气和电气行业监管机构 OFGEM 目前采用的预计标准为 SN29500<sup>[6]</sup>，N Li 对较为复杂的 IEC62380 应用于智能电表可靠性评估进行了讨论<sup>[7]</sup>。我国将国外的成果转化得到常用的 GJB/Z299C 标准，杨洪旗基于该标准对智能电表常规元件进行可靠性预计<sup>[8]</sup>，李向峰则认为军标类可靠性预计手册较为严格，应用于智能电表时所得的可靠性预计结果与实际情况存在较大差异，而商用类可靠性预计手册中的 SN29500 涵盖多个领域的基础数据且更新及时，因此更适用于电能表<sup>[9]</sup>；袁金灿将可从厂家处获取的信息同智能电能表与各预计手册的特点综合考虑，选择



基于 Telcordia SR-332 对智能电表进行可靠性预计并提出可靠性提升措施<sup>[10]</sup>。

基于可靠性试验的方法由于能在较短的时间内获得更多的可靠性信息而在可靠性评估中得到广泛的研究和应用。Z. Yang 基于故障模式分析的结果设计了以温度、湿度和电流作为加速应力的多应力恒加速退化试验 (Accelerated Degradation Test, ADT), 采用线性模型拟合计量误差退化数据的轨迹并进一步得到加速寿命模型参数和可靠度函数<sup>[11]</sup>; K. Lin 对文献[11]的数据进一步分析发现退化数据存在异质性, 在对伪寿命数据分析时以混合 Weibull 分布作为可靠性模型<sup>[12]</sup>; D. Xu 针对智能电表故障数据存在异质性而无法采用单一失效分布描述的问题提出以 Birnbaum-Saunders 模型作为通用模型对基本误差退化数据建模<sup>[13]</sup>; K. Lin 针对智能电表退化增量相互不独立且服从混合分布的问题, 采用自适应核密度估计和 copula 函数拟合退化轨迹, 结果表明该方法得到的退化轨迹优于基于 Wiener 过程的结果<sup>[14]</sup>; H. Zheng 采用四参数模型对 ADT 得到的退化数据建模以更好地描述退化轨迹, 利用粒子滤波器进行模型参数识别以实现单个电表的实时剩余寿命预测<sup>[15]</sup>; 李亦非将高温高湿恒应力加速寿命试验 (Accelerated Life Test, ALT) 应用电能表寿命评估<sup>[16]</sup>, 罗冉冉基于历史数据分析、环境影响分析和强化试验设计了恒定温度和过电流的双应力定时截尾 ADT, 对获得的日计时误差、基本误差超差退化数据拟合线性退化轨迹后获得伪寿命数据, 根据伪寿命数据得到各应力组合下的可靠性模型分布参数, 通过广义 Eyring 模型求得外推至正常条件下的可靠性模型分布参数及可靠度函数<sup>[17]</sup>; 查国清则开展了以温度、湿度、电流为加速应力的多应力 ADT, 以 Weibull 分布作为可靠性模型, 基于形状参数与应力无关、尺度参数与应力满足对数线性模型的假定获取正常条件下的智能电表可靠度函数<sup>[18]</sup>; 张景元在查国清的研究基础上认为形状参数与应力间的关系也满足对数线性模型, 通过对数据的分析提出了新的多应力退化模型, 最终获得能反映电应力对寿命影响的可靠度函数<sup>[19]</sup>; 辛承霖针对恒定温湿度应力下智能电表计量误差会出现可逆性超差的问题, 通过优化的湿热循环应力可靠性强化试验收集计量性能退化数据, 采用双随机效应 Wiener 过程建模并由马尔科夫链蒙特卡洛 (Markov Chain Monte-Carlo, MCMC) 方法求得模型参数, 但该方法还需以恒定应力 ALT 作为参考求解等效的加速因子<sup>[20]</sup>。

现场数据相对试验数据的收集周期更长，但能反映实际运行环境下的可靠性水平。Y. Li 基于指数分布和现场累积失效率评估智能电表的寿命<sup>[21]</sup>，J. Zhang 针对每种故障模式首先根据现场故障数据和 Weibull 分布确定基准失效率，再结合气候数据得到温湿度对失效率的影响从而修正基准失效率，最终实现考虑应力影响的智能电表短期寿命预计<sup>[22]</sup>。W. Qiu 和杜蜀薇针对目前仪器的可靠性分析未考虑不同区域条件影响的问题，基于分层贝叶斯方法分析区域故障率数据，以半柯西分布作为形状参数的先验分布，对数线性模型描述地区和时间对尺度参数的影响并以正态分布作为模型中回归系数的先验分布，实现短期故障分析<sup>[23,24]</sup>。对于智能电表中现场失效率高且多为厂家定制而无法通过预计手册得到可靠性预计结果的磁保持继电器，李翰斌通过收集现场数据进行可靠性评估，结果表明电流、电压等外界环境冲击造成了现场失效率远高于预计手册预计的失效率<sup>[25]</sup>。刘旭明评估智能电表的运行可靠性时，考虑现场故障数据样本量小、零值多的特点采用零膨胀负二项分布对数据建模，基于层次贝叶斯得到模型参数实现智能电表可靠性预估<sup>[26]</sup>。

产品可靠性的不断提升使得单一数据源提供的可靠性信息有限，多源数据融合的可靠性评估方法因能够通过扩大信息量提高预测精度而得到发展。D. Xu 以线性 Wiener 过程退化模型对 ADT 所得的基本误差数据建模，融合现场状态数据得到智能电表运行状态下的评估结果<sup>[27]</sup>。X. Liu 和田正其通过开展 ADT 得到了性能退化数据和整表失效数据，以退化数据求得的正常条件下伪寿命数据作为先验分布信息源，采用贝叶斯方法结合外推至正常条件的整表失效数据得到后验分布，从而得到了更符合工程实际的智能电表可靠性评估结果<sup>[28,29]</sup>。李贺龙采用非线性 Wiener 过程描述 ADT 退化数据，在假设参数先验分布为正态分布的基础上依据贝叶斯方法得到后验分布，再以其为先验信息对现场所得的退化数据采用粒子滤波进行融合以更新模型参数，可对智能电表的剩余寿命进行实时预测<sup>[30]</sup>。张家琦以异常数据中各异常类型次数为协变量并结合检修数据中的寿命数据，基于生存分析和神经网络构建 CoxPH 模型评估智能电表的可靠性<sup>[31]</sup>。

## 1.2.2 可靠性模型优选研究现状

合适的可靠性模型能准确描述可靠性数据的特点从而能提高可靠性评估的准确

性,可靠性模型优选即是从多个备择模型中选取能对产品可靠性有最佳描述模型,常用的方法可分为假设检验方法和基于信息论的方法<sup>[32]</sup>,其中假设检验方法又可分为基于似然比检验的方法和基于拟合优度检验的方法。

基于似然比检验的方法在获得备选模型的参数极大似然估计的基础上根据原假设和备择假设的似然函数比值选择合适的模型,可以比较两个模型在给定数据下的表现,通常用于两个备选模型中的一个模型为另一模型特例的情形,如 W. Wang 基于该方法从 Arrhenius 模型和 Eyring 模型中优选了智能电表的加速模型<sup>[33]</sup>,董贤光根据现场某批次电能表的故障数据求得 Weibull 分布和指数分布的参数极大似然估计后,采用似然比检验确定该批次电能表的最优寿命分布为 Weibull 分布<sup>[2]</sup>。该方法也可拓展至两种以上无关联关系模型的选择中,如 A. K. Dey 从 Weibull 分布、对数正态分布和广义指数分布中以似然函数最大为依据进行模型优选,并研究了完全样本和定时截尾样本下该方法正确选择模型的概率<sup>[34]</sup>。但基于似然比检验选择模型的缺点在于该方法仅能比较各模型在给定数据下的表现,不能提供该模型正确与否的信息,且该方法在分析正确选择模型的概率时需要借助大量仿真获取统计量的渐进分布。

基于拟合优度检验的方法有 Kolmogorov-Smirnov 检验、Cramer-von Mises 检验和 Anderson-Darling 检验等,其原理为在求得拟合分布函数后结合经验分布函数构造相应的统计量并根据该统计量按一定原则从通过检验的分布选取最优模型,拟合分布函数的参数需通过样本求得。在可靠性模型优选中,拟合分布模型的参数可采用秩回归方法求得,该方法将获得的故障数据按升序排列后由中位秩、平均秩等估计累积失效概率,再基于最小二乘回归求得模型参数,如肖勇基于秩回归方法求得模型参数后以 K-S 检验统计量最小为依据选取智能电表通信模块的最优寿命模型<sup>[35]</sup>;文献[19]则根据归一化相关系数选取,申桂香等人基于熵值-模糊综合评判法以拟合分布和经验分布的检验统计量最小、平均累积误差最小、关联度最大、正负残差最大值之差最小、概率密度函数均方误差最小为依据并考虑模型特点,综合多种主、客观因素确定随机截尾试验数据的最佳分布<sup>[36]</sup>;X. Yang 则综合 K-S 检验和 A-D 检验确定了油气分离器中三种元器件的维修时间分布<sup>[37]</sup>。

相较于前两种方法，基于信息论的方法在进行模型选择时不仅考虑了数据和理论模型的差异度，同时考虑了模型的复杂度，常用的准则有 Akaike 信息准则（Akaike Information Criterion, AIC）、Bayes 信息准则（Bayesian Information Criterion, BIC）和 DIC 准则，其中 AIC 和 BIC 准则是以模型极大似然估计为基础的，BIC 准则对模型参数增加的惩罚系数更大会选择参数更少的模型，X. Liu 和王智明基于 AIC 和 BIC 准则分别选择了对数正态分布描述集中器的伪寿命数据、两参数 Weibull 分布描述加工中心的可靠性<sup>[28]</sup>；T. Emura 在分析电力变压器的左截尾右删失寿命数据时根据 AIC 准则选择了对数正态模型<sup>[39]</sup>。DIC 准则多用于复杂层次模型和贝叶斯参数估计下的模型选择，如 E. A. Nasir 基于贝叶斯框架优化设计加速寿命试验后对试验数据采用 DIC 准则选择最合适加速模型<sup>[40]</sup>。也有相关研究针对以上准则不足之处进行改进，如 E. Amalnerkar 就 AIC 准则应用于小样本表现不佳的问题将 AIC 偏差项中的模型参数个数替换为由 Bootstrap 方法得到的统计量并将该方法用于可靠性分析<sup>[41]</sup>。

## 1.3 存在的问题

通过研究现状可知，目前智能电表可靠性评估的相关研究存在以下三个问题：

（1）智能电表可靠性评估以基于单一数据源的方法为主，且大多数研究围绕可靠性试验中 ADT 退化数据的建模展开，对智能电表现场数据应用于可靠性评估的研究有限，缺乏对更能反映实际可靠性水平的现场数据的充分利用。

（2）尽管基于多源数据的可靠性评估有较多研究，但将该方法应用于智能电表可靠性评估的文献却十分有限，且大多研究采用 ADT 数据外推的伪寿命数据作为信息源之一，ADT 在试验过程中对所有样品均进行退化量监测，因此外推得到的伪寿命数据构成了完全样本，并未考虑到实际情况高可靠性产品现场返回的故障数据、厂家开展 ALT 得到的寿命数据均包含截尾数据而非完全样本的情况。

（3）可靠性模型选择已有相当一部分的研究，但围绕智能电表的研究并不多，目前从多个备选模型中优选智能电表的可靠性模型采用基于拟合优度检验结合秩回归的方法，但该方法构造的统计量包含了有多种估计方式的累积失效概率，加之求

解模型参数时由于对因变量进行了相应变换实现线性化，难以保证变换前的残差平方和最小，影响了参数估计的精度，故该方法并不适合模型优选。

本文针对以上问题，首先将现场数据作为智能电表可靠性评估的数据来源之一，充分挖掘其包含的信息实现对现场数据的充分利用；其次，在基于贝叶斯理论融合现场数据和 ALT 数据的过程中，针对两种数据源均包含截尾数据的问题，修正了似然函数的表达式并采用参数 Bootstrap 方法挖掘先验信息；最后，结合秩回归方法和智能优化算法求解模型参数，采用基于信息论的方法综合考虑模型与数据的差异和模型复杂度确定智能电表的可靠性模型。

## 1.4 本文主要内容

本文共分为 5 章，章节内容之间的关联如图 1-1 所示，主要内容如下：

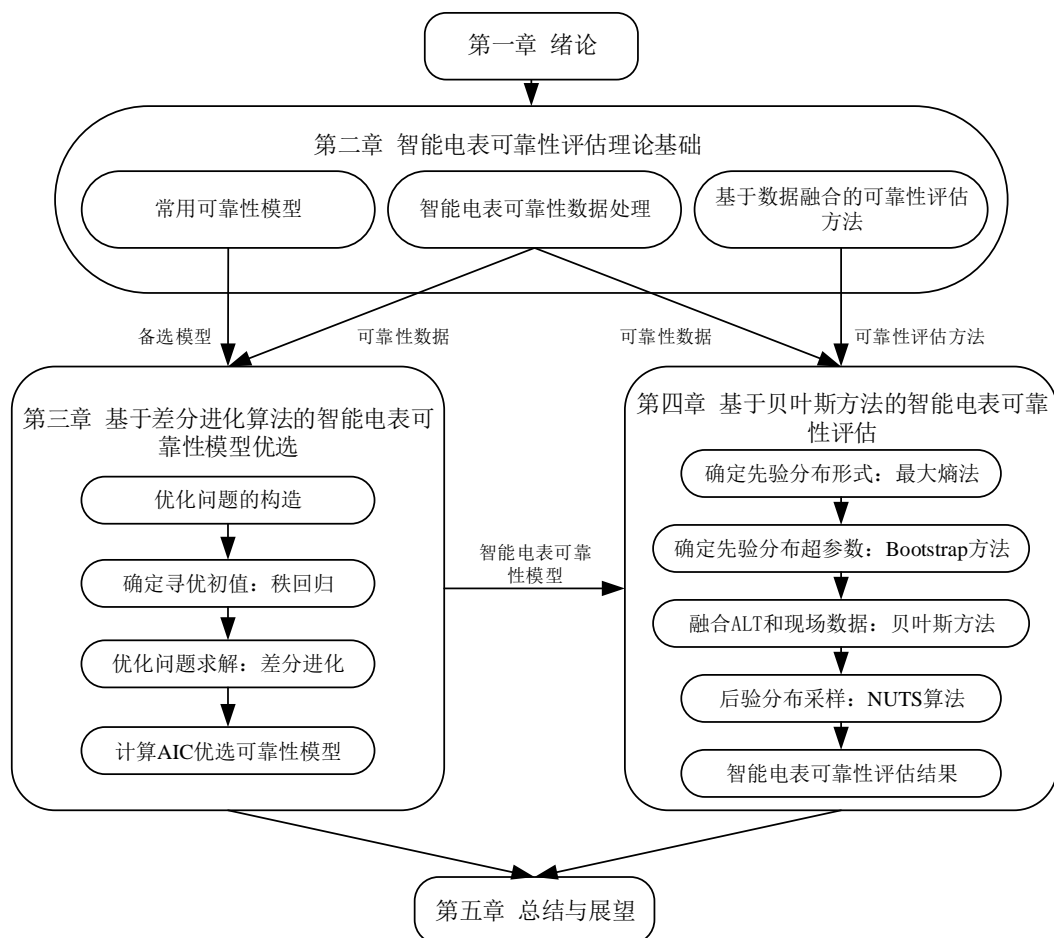


图 1-1 组织结构图

# 华中科技大学硕士学位论文

---

第一章 绪论：介绍本文的研究背景、国内外研究现状、存在的问题和主要研究内容。

第二章 智能电表可靠性评估理论基础：介绍智能电表可靠性的相关理论，包括常用可靠性模型、智能电表可靠性数据来源及其数据处理方法，以及基于数据融合的可靠性评估方法。

第三章 基于差分进化算法的智能电表可靠性模型优选：针对目前智能电表可靠性模型优选方法的不足，在选择合适的模型优选准则的基础上，将秩回归法和差分进化算法结合用于分析可靠性数据，获得智能电表的最优可靠性模型。

第四章 基于贝叶斯方法的智能电表可靠性评估：在第三章确定可靠性模型的基础上，研究融合的可靠性数据含有截尾数据时的贝叶斯可靠性评估方法，并重点研究了该数据下先验分布的获取方法。

第五章 总结与展望：总结本文的主要工作和结论，并对未来工作进行展望。

## 2 智能电表可靠性评估理论基础

智能电表可靠性评估是通过对收集到的智能电表可靠性数据选择合理的可靠性模型，并采用合适的可靠性评估方法对数据进行分析获得模型参数实现的，因此可靠性模型、可靠性数据的处理和可靠性评估方法对可靠性评估结果产生直接的影响。本章围绕以上三个要素对智能电表可靠性评估涉及的理论进行介绍，作为后续章节的理论基础。

### 2.1 常用可靠性模型

可靠性是能反映产品质量的特性之一，通常可以采用可靠度来进行衡量，智能电表的可靠度是指智能电表在给定条件下运行给定时间而不失去其功能的概率，描述可靠度随时间变化的模型称为可靠性模型，常用的概率分布均可用于描述智能电表的可靠度，通过可靠度函数 $R(t)$ 可求得不可靠度函数 $F(t)$ 、概率密度函数 $f(t)$ 和失效率函数 $\lambda(t)$ 等指标。

指数分布模型在可靠性建模中的研究应用最早，是基于可靠性预计手册评估智能电表可靠性的基本假设之一<sup>[42]</sup>，其可靠度、不可靠度、概率密度和失效率函数的表达式如式（2-1）至式（2-4）所示，该模型的失效率不随时间变化，可用于经过老化筛选后处于偶然失效期的产品<sup>[43]</sup>。

$$R(t) = e^{-\lambda t} \quad (2-1)$$

$$F(t) = 1 - e^{-\lambda t} \quad (2-2)$$

$$f(t) = \lambda e^{-\lambda t} \quad (2-3)$$

$$\lambda(t) = \lambda \quad (2-4)$$

Weibull 分布模型是基于最薄弱环节提出的，适用于一个子系统或单元失效则引发系统失效的情形，该模型因很多产品在可靠性评估时以串联模型为基础而得到了广泛的应用，是评估智能电表可靠性最常用的模型，其可靠度、不可靠度、概率密度和失效率函数的表达式如式（2-5）至式（2-8）所示，式中 $\beta$ 为形状参数、 $\eta$ 为尺度

参数也称为特征寿命。该模型的优势在于其失效率根据  $\beta$  取值的不同可以为上升、恒定或递减函数，同时也包含或可近似部分常用分布，如对比指数分布可知，当  $\beta$  取值为 1 时 Weibull 分布即为指数分布，因此在后文中不再单独讨论指数分布。

$$R(t) = e^{-\left(\frac{t}{\eta}\right)^\beta} \quad (2-5)$$

$$F(t) = 1 - e^{-\left(\frac{t}{\eta}\right)^\beta} \quad (2-6)$$

$$f(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^\beta} \quad (2-7)$$

$$\lambda(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1} \quad (2-8)$$

智能电表包含了大量电子元器件，而正态分布常被用于分析电子元器件因损耗和工作时间延长引起的失效，该模型的可靠度、不可靠度、概率密度和失效率函数的表达式如式（2-9）至式（2-12）所示。

$$R(t) = 1 - \int_{-\infty}^t \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\tau-\mu}{\sigma}\right)^2} d\tau \quad (2-9)$$

$$F(t) = \int_{-\infty}^t \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\tau-\mu}{\sigma}\right)^2} d\tau \quad (2-10)$$

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} \quad (2-11)$$

$$\lambda(t) = \frac{1}{1 - \int_{-\infty}^t \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\tau-\mu}{\sigma}\right)^2} d\tau} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} \quad (2-12)$$

对数正态分布被广泛应用于由一般自然规律导致的失效和加速寿命试验数据的统计分析中，而加速寿命试验是智能电表可靠性数据的重要来源之一，因此该分布也可作为智能电表的可靠性模型，该模型的优势在于限制变量取值为正数，其可靠度、不可靠度、概率密度和失效率函数的表达式如式（2-13）至式（2-16）所示。

$$R(t) = 1 - \int_{-\infty}^t \frac{1}{\sigma\tau\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln \tau - \mu}{\sigma}\right)^2} d\tau \quad (2-13)$$



$$F(t) = \int_{-\infty}^t \frac{1}{\sigma\tau\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln\tau-\mu}{\sigma}\right)^2} d\tau \quad (2-14)$$

$$f(t) = \frac{1}{\sigma t\sqrt{2\pi}} e^{-\frac{(\ln t - \mu)^2}{2\sigma^2}} \quad (2-15)$$

$$\lambda(t) = \frac{1}{1 - \int_{-\infty}^t \frac{1}{\sigma\tau\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln\tau-\mu}{\sigma}\right)^2} d\tau} \frac{1}{\sigma t\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln t-\mu}{\sigma}\right)^2} \quad (2-16)$$

考虑到 Weibull 分布包含了指数的分布的情形, 本文在后续章节中仅以 Weibull 分布、正态分布和对数正态分布中作为备选可靠性模型, 并从中优选最佳模型用于智能电表的可靠性评估。

## 2.2 智能电表可靠性数据处理

可靠性数据作为可靠性评估工作的基础可根据数据来源分为现场数据和加速试验数据, 本节对两种数据及其相应的预处理进行介绍。

### 2.2.1 现场数据

随着电网公司计量自动化系统的建成, 对投运智能电表的批次信息、投运日期、故障日期均有详细记录, 为智能电表的可靠性评估提供了可信的现场数据。对于智能电表这类可靠性水平要求较高的产品, 当可靠度下降至限值时会停止同一批次产品的使用, 这使得收集现场运行记录获得的样本同时包含了失效样本和仍可继续正常使用的未失效样本, 相应的现场可靠性数据也包括两部分: 失效样本的寿命数据和未失效样本的已运行时间数据, 后者也称为截尾数据, 这类包含截尾数据的样本称为截尾样本。

对于同一批次的智能电表, 由于投运日期不同, 样本在时间轴上的分布示意图如图 2-1 (a) 所示, 其中  $x_1$  和  $x_2$  为截尾数据,  $x_3$  和  $x_4$  为寿命数据。对于这类复杂结构的数据, 需对所有数据的投运日期做归一化处理, 即以最早投运设备的投运日期为基准, 将其他样本的投运时间与之对齐, 可等效为如图 2-1 (b) 所示的随机截尾样本。

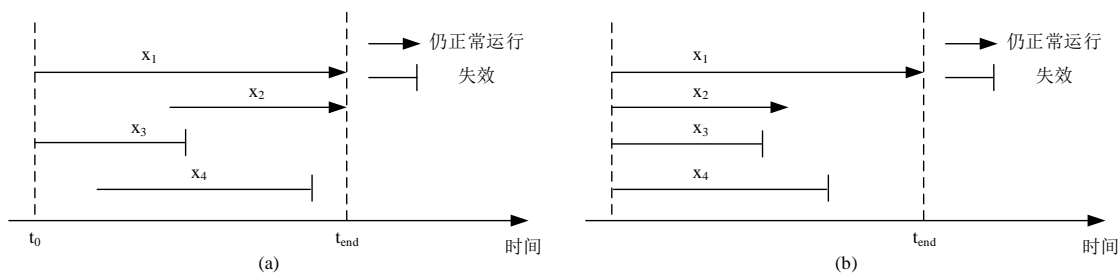


图 2-1 现场数据及其预处理后的等效数据

图中 $t_0$ 为最早投运的设备的投运日期， $t_{\text{end}}$ 为停止数据收集的日期。

### 2.2.2 加速试验数据

加速试验数据是对产品开展加速试验时收集所得的数据，相较于现场数据，该数据能覆盖更长生命周期时段，是可靠性数据的另一重要来源。按照试验过程中监测量的不同又可分为记录样本故障时间的加速寿命试验 ALT 和记录样本性能退化量的加速退化试验 ADT。

#### (1) 加速寿命试验及其数据预处理

ALT 的原理为在保证不引入新失效模式的基础上通过给试验样品施加远大于实际运行环境的应力加速失效的发生从而得到可靠性数据，目前研究较多的试验方案有恒定应力试验、步进应力试验和序进应力试验，其中恒定应力试验因理论成熟得到了广泛的应用。

ALT 在试验过程中根据样本故障与否记录样本的故障前运行时间，即可以直接获得试验样本的寿命数据，适用于突发性失效和退化型失效，但由于试验中一般采用固定间隔时间对样本故障情况进行一次记录的方式，因此得到的原始数据并非样本实际寿命数据且出现多个样本故障记录时间相同的情况，需对原始数据进行相应的离散化处理：对于在一次观测间隔时间内发生故障的样本，假定其故障时间均匀散布于该间隔时间内<sup>[44]</sup>，则故障样本的寿命数据可由式（2-17）计算得到：

$$t_{mi} = (n-1)T + \frac{i}{p+1}T \quad (i=1,2,\dots,p) \quad (2-17)$$

式中 $t_{mi}$ 为第  $n$  个间隔时间内第  $i$  个样本的寿命数据， $p$  为第  $n$  个间隔时间内故障样本的总数， $T$  为观测间隔时间。

## (2) 加速退化试验及其数据预处理

ADT 是针对部分高可靠性产品在有限时间的 ALT 中故障样本数量少或无故障发生的问题而提出的, 与 ALT 监测记录样本故障情况不同, ADT 在试验过程中监测记录样本性能退化量, 因此仅该方法只能用于退化型失效<sup>[6]</sup>, 其寿命数据是通过拟合退化量的退化轨迹并计算其到达失效阈值的时间得到的, 也称为伪寿命数据<sup>[45]</sup>, 其计算式如 (2-18) 所示:

$$t_i = \inf \{t | \varphi_i(t) \geq D_{th}\} \quad (2-18)$$

式中  $t_i$  为第  $i$  个样本的伪寿命数据,  $\varphi_i(t)$  为第  $i$  个样本的退化轨迹,  $D_{th}$  为退化量的失效阈值。

## (3) 加速试验数据的外推处理

经由式 (2-17) 或 (2-18) 得到的寿命数据仍为加速应力条件下的数据, 为获得正常使用环境下的数据还需通过相应的加速模型进行外推处理。在加速试验中应用较多的加速模型有描述温度与寿命关系的 Arrhenius 模型、描述电应力与寿命关系的逆幂律模型, 以及描述温湿度应力与寿命关系的 Peck 模型, 相应的加速因子 (Accelerated Factor, AF) 如式 (2-19) 至 (2-21) 所示:

$$AF_{Arr} = e^{\frac{E_a}{k} \left( \frac{1}{T_u} - \frac{1}{T_a} \right)} \quad (2-19)$$

$$AF_{IPL} = \left( \frac{S_u}{S_a} \right)^{-c} \quad (2-20)$$

$$AF_{Peck} = \left( \frac{RH_u}{RH_a} \right)^{-n} e^{\frac{E_a}{k} \left( \frac{1}{T_u} - \frac{1}{T_a} \right)} \quad (2-21)$$

式中  $c$  和  $n$  为常数;  $E_a$  为活化能常数一般取为  $0.9\text{eV}$ ;  $k$  为玻尔兹曼常数取值为  $8.617 \times 10^{-5} \text{eV}$ ;  $T_u$ 、 $S_u$  和  $RH_u$  分别为正常应力条件下温度、电应力和相对湿度,  $T_a$ 、 $S_a$  和  $RH_a$  分别为加速应力条件下温度、电应力和相对湿度。

选择适用于智能电表可靠性评估的加速试验方法以获取试验数据时, 需要考虑智能电表的故障模式。本文对某省故障电表的情况进行调研收集到数据 64.3 万条, 通过分析发现智能电表除日计时超差等退化型失效外, 还存在 485 通信异常等突发性失效, 因此本文采用恒定应力 ALT 获取加速试验数据。

## 2.3 基于数据融合的可靠性评估方法

融合多源数据对智能电表的可靠性进行评估能够提高评估结果的准确性，在主要的融合方法中，相较于需要一定数量训练数据的神经网络和模糊积分方法，贝叶斯方法和 D-S 证据理论目前在可靠性评估中有相对深入的研究和应用，本节对这两种方法应用于可靠性评估的原理进行介绍，根据各自特点以明确其用于融合智能电表可靠性数据时适用的情形。

### 2.3.1 基于贝叶斯理论的可靠性评估方法

基于贝叶斯理论的可靠性评估方法在获得参数的贝叶斯估计结果的基础上对可靠性进行评估，贝叶斯参数估计方法则以贝叶斯定理为依据，连续分布形式下的贝叶斯定理如式（2-22）所示：

$$g(\theta|X) = \frac{\pi(\theta)p(X|\theta)}{\int \pi(\theta)p(X|\theta)d\theta} \quad (2-22)$$

式中 $\pi(\theta)$ 称为 $\theta$ 的先验分布； $p(X|\theta)$ 为条件 $\theta$ 已知时事件 $X$ 发生的概率； $g(\theta|X)$ 为根据事件发生情况更新后的 $\theta$ 的概率分布，也称为后验分布；分母部分为关于 $X$ 的常数。

通过式（2-22）可知，贝叶斯方法的特点在于：该方法不将未知参数视为确定值而是将其视为随机变量，在充分利用样本信息和先验信息得到用于对参数进行统计推断的后验分布后，再将依据后验分布导出的相关统计量作为参数估计值，统计量可以是后验众数估计 $\theta_{\text{MAP}}$ 或后验期望估计 $\theta_E$ ，相应的计算式分别为式（2-23）和（2-24）。

$$\theta_{\text{MAP}} = \arg \max_{\theta} g(\theta|X) \quad (2-23)$$

$$\theta_E = \int \theta g(\theta|X)d\theta \quad (2-24)$$

该方法应用于可靠性模型参数估计时， $\theta$ 为可靠性模型的参数，事件 $X$ 的条件概率 $p(X|\theta)$ 则为根据样本和选定的可靠性模型构建的似然函数，由于后验分布是根据样本信息更新得到的，故该方法多用于多源数据的形式相同的情况（如均为寿命数据）。

基于贝叶斯方法融合两种来源的数据进行可靠性评估时存在两种方式，如图 2-2（a）和（b）所示。

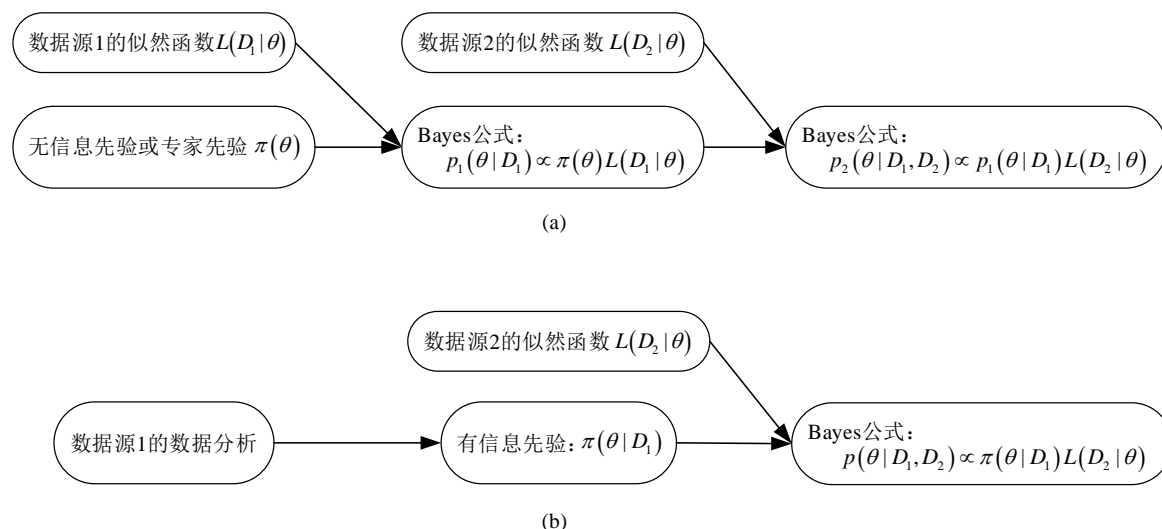


图 2-2 贝叶斯可靠性评估方法融合多源数据的两种方式

两种方法在先验分布的确定上有所区别：如图 2-2（a）所示的方式，先验分布为专家根据经验或由贝叶斯假设事先确定，再根据贝叶斯定理利用其中一个数据源样本构造的似然函数对其更新，得到的后验分布作为另一数据源的先验分布，对另一数据源重复前述更新过程即可得到融合两种数据源后的参数估计结果；图 2-2（b）的方式则将两种数据源中的一个作为先验信息源并通过对其样本的分析获取先验分布，另一数据源作为观测样本构建似然函数，再基于贝叶斯定理得到融合两种数据源后的参数估计结果。

后验分布受到先验分布的影响，在小样本情形下由可信的先验信息得到的参数估计结果相较于经典统计理论方法得到的结果更加准确，因此先验分布的确定对贝叶斯可靠性评估结果至关重要，也是基于贝叶斯理论的可靠性评估方法饱受争议的原因。

## 2.3.2 基于 D-S 证据理论的可靠性评估方法

D-S 证据理论根据基本概率分配函数和证据合成规则进行多源数据的融合，实现对辨识框架中各假设的推理。

对于一个包含  $n$  个元素的辨识框架  $\Theta$ ，每个子集都对应一个命题  $S_i$ ，全部的  $2^n$

个命题构成了该辨识框架对应的幂集，证据的基本概率分配函数 $m(S_i)$ 能反映其对命题 $S_i$ 的支持力度，且 $m(S_i)$ 能满足式（2-25）：

$$\begin{cases} m(\phi) = 0 \\ \sum_{i=1}^{2^n} m(S_i) = 1 \end{cases} \quad (2-25)$$

不同的证据可以对各命题给出不同的基本概率分配函数，在已知  $k$  个证据的基本概率分配函数时可由式（2-26）的 Dempster 证据合成公式得到融合后用于决策的  $m(S_i)$ 。

$$m(C) = \begin{cases} 0 & X \cap Y = \phi \\ \frac{\sum_{X \cap Y = C} m_i(X) m_j(Y)}{1 - \sum_{X \cap Y = \phi} m_i(X) m_j(Y)} & X \cap Y \neq \phi \end{cases} \quad (2-26)$$

D-S 证据理论具有表达不确定性的能力，能对不确定信息建模和推理，因此该方法多用于专家经验这类不确定信息的融合，而该方法要求辨识框架中的元素互斥并能够穷举的特点，使得其应用于可靠性评估时更适合多源可靠性数据特征层或决策层的融合，基于 D-S 证据理论的可靠性评估方法有以下两种方式：

（1）通过 D-S 证据理论直接融合多源数据：将可靠度或已知的寿命取值范围离散化为  $n$  个区间得到辨识框架 $\Theta = \{(a_0, a_1), (a_1, a_2), \dots, (a_{n-1}, a_n)\}$ ，各数据源作为证据，以分析数据源所得的可靠度或寿命区间估计为依据构造基本概率分配函数，证据合成后的 $m(S_i)$ 表示可靠度或寿命各区间的取值概率，可由 $m(S_i)$ 直接得到最终的可靠度或寿命评估结果。

（2）通过 D-S 证据理论间接融合多源数据：对于难以预先确定取值范围的可靠性模型参数 $\theta$ ，以各数据源求得的参数值 $\theta_i$ 为辨识框架 $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ ，以数据源的样本量、可信度等信息为证据构造基本概率分配函数，证据合成后的 $m(S_i)$ 表示各数据源的权重，可由 $m(S_i)$ 对 $\theta_i$ 加权融合得到最终的可靠性模型参数估计结果。

相较于贝叶斯理论，D-S 证据理论的优势在于不需要先验信息，适用于先验信息难以获得的情形，但该方法难以客观地构造会对融合结果产生较大影响的基本概率分配函数、证据合成方法缺乏数学依据，且要求证据之间相互独立，尤其是证据

间相互冲突时会得到完全不合理的结论。

## 2.4 本章小结

本章主要围绕智能电表可靠性评估的相关理论展开。本章首先介绍了四种常用可靠性模型，并以 Weibull 分布、正态分布和对数分布作为后续可靠性模型优选的备选模型；同时对现场数据、加速试验数据这两种来源的可靠性数据及其预处理方法进行了介绍，综合考虑加速试验和智能电表故障的特点选择采用恒定应力 ALT 获取智能电表的可靠性试验数据；最后对两种常用的基于数据融合的可靠性评估方法进行了介绍。

## 3 基于差分进化算法的智能电表可靠性模型优选

智能电表可靠性评估是在可靠性模型的基础上进行相应的推断，因此可靠性模型的选择对可靠性评估结果有着至关重要的影响，应当在后续推断之前结合智能电表的可靠性数据对可靠性模型进行优选。针对智能电表可靠性数据含有截尾数据使得极大似然估计求解困难的问题，本章将参数的极大似然估计视为优化问题并采用差分进化算法求解，随后计算用于模型优选的 AIC 值，通过对数据的分析得到最合理的可靠性模型。

### 3.1 可靠性模型优选准则

如 1.2.2 节中所述，相较于基于假设检验方法的模型优选，基于信息论方法优选模型时兼顾了数据与拟合模型的差异情况以及模型复杂度两个维度，该方法已广泛用于模型选择，因此本章采用基于信息论的方法作为智能电表可靠性模型优选的依据。

AIC 准则是以 Kullback-Leibler 距离和极大似然估计为基础的，Kullback-Leibler 距离能反映用拟合模型替代真实模型时造成的信息损失，Akaike 研究了 Kullback-Leibler 距离与对数似然函数的关系给出了 AIC 准则的表达式如式 (3-1) 所示，在模型优选时选择 AIC 最小的模型：

$$AIC = -2\ln L(\theta) + 2k \quad (3-1)$$

其中  $L(\theta)$  为拟合模型在给定数据下的极大似然函数值， $\theta$  为参数的极大似然估计， $k$  为模型参数个数。

应用于小样本情形时，对式 (3-1) 修正后的  $AIC_c$  准则如式 (3-2) 所示，式中的参数  $n$  为样本总数：

$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1} \quad (3-2)$$

类似 AIC 准则，BIC 准则同样基于极大似然估计但两者在表示模型复杂度的修正项上有所区别，BIC 准则的表达式如式 (3-3) 所示，式中各参数含义同上：



$$BIC = -2 \ln L(\theta) + k \ln n \quad (3-3)$$

不同于 AIC 和 BIC 准则，DIC 准则以参数的贝叶斯估计为基础，采用整个参数后验分布上的对数似然函数平均值与参数后验点估计的对数似然函数值之差描述模型复杂度，DIC 准则的表达式如式（3-4）所示：

$$DIC = 2E[-2 \ln L(\theta)] - (-2 \ln L(\hat{\theta})) \quad (3-4)$$

其中  $E[-2 \ln L(\theta)]$  为对数似然函数在参数  $\theta$  的后验分布上的平均值； $\ln L(\hat{\theta})$  是参数后验点估计对应的对数似然函数值； $\hat{\theta}$  是参数后验点估计，一般为后验期望估计。

以上三种常用的模型优选准则各有特点：从实际应用场景来看，AIC 准则选择模型时兼顾了对已有数据和未来数据的较好拟合，通常用于预测问题的模型选择，BIC 准则不受未来数据的影响，通常用于选择能更好解释现有数据的模型，且由于其隐含假设为模型参数的先验分布是以极大似然估计值为中心的多元正态分布，对于该假设不合理的情形，BIC 准则用于模型选择的表现不佳，DIC 准则由于依托贝叶斯理论而常用于复杂层次模型的优选；从选择结果来看，AIC 准则倾向于选择对数据拟合最好的模型，BIC 准则倾向于选择真实模型，DIC 准则的选择结果则受所选先验分布的影响较大<sup>[40,46]</sup>。

考虑到本文的备选可靠性模型不涉及复杂层次模型，可靠性模型优选的目的在于对智能电表做可靠性评估，且不能确定真实的可靠性模型是否包含在本文的候选模型中，故采用更为通用的 AIC 准则进行可靠性模型优选。同时如本章后续章节所述，通过 ALT 获取的智能电表寿命数据属于小样本（ $n=30$ ），采用 AIC<sub>c</sub> 准则选择最优模型，而现场返回的故障数据则可采用 AIC 准则选取最优模型。

## 3.2 可靠性模型参数估计方法

可靠性模型优选的前提是得到可靠性模型参数的估计值，再计算用于模型优选的指标。目前智能电表可靠性模型的参数估计多采用基于秩回归的方法，该方法在求得近似累积失效概率的基础上对线性变换后的分布函数做最小二乘回归求得模型参数，由于累积失效概率的近似估计有多种形式，加上线性变换后拟合求解难以保证变换前的残差平方和最小，故该方法计算所得的参数估计结果精度不理想难以据

此计算用于模型优选的指标<sup>[38,47]</sup>。针对这一问题，本节采用差分进化算法求解截尾数据下的可靠性模型参数极大似然估计。

## 3.2.1 截尾样本下的模型参数极大似然估计问题

在高可靠性产品的可靠性试验中，受限于试验成本，通常不会采用完全试验的方式将试验进行至所有样本失效，而是采用试验结束时仍有部分样本可正常运行的截尾试验方式。因此对于智能电表而言，其试验数据同现场数据一样，均为截尾样本。

寿命数据和截尾数据均提供了产品可靠性的相关信息，但两者略有差异：寿命数据给出了样本的准确寿命时间，截尾数据则给出了样本寿命的取值范围。在构建似然函数时，前者的概率由概率密度函数 $f(t)$ 描述、后者的概率由可靠度函数 $R(t)$ 描述，可得根据包含截尾数据的可靠性数据构建的似然函数表达式如式（3-5）所示<sup>[48,49]</sup>：

$$L(\theta) = \prod_{i=1}^r f(t_i, \theta) \prod_{j=1}^{n-r} R(T, \theta) \quad (3-5)$$

其中  $\theta$  为可靠性模型参数， $r$  为失效样本个数， $n$  为样本总数， $t_i$  为第  $i$  个失效样本的寿命数据， $T$  为截尾时间。

参数的极大似然估计是使似然函数取最大值时参数的取值，一般的求解流程为：对式（3-5）所示的似然函数取对数后再对参数求导，根据导数为 0 列出方程组求解。

选择 Weibull 分布作为可靠性模型时，将其可靠度函数式（2-5）和概率密度函数式（2-7）代入式（3-5）可得对数似然函数和对应的方程组如式（3-6）和（3-7）所示：

$$\ln L(\beta, \eta) = r \ln \beta - r \beta \ln \eta + (\beta - 1) \sum_{i=1}^r \ln t_i - \sum_{i=1}^r \left( \frac{t_i}{\eta} \right)^\beta - (n-r) \left( \frac{T}{\eta} \right)^\beta \quad (3-6)$$

$$\begin{cases} \frac{\partial \ln L}{\partial \beta} = \frac{r}{\beta} - r \ln \eta + \sum_{i=1}^r \frac{t_i}{\eta^\beta} - \frac{1}{\eta^\beta} \left[ \sum_{i=1}^r t_i^\beta (\ln t_i - \ln \eta) + (n-r) T^\beta (\ln T - \ln \eta) \right] = 0 \\ \frac{\partial \ln L}{\partial \eta} = -\frac{r\beta}{\eta} + \frac{\beta}{\eta^{\beta+1}} \left[ \sum_{i=1}^r t_i^\beta + (n-r) T^\beta \right] = 0 \end{cases} \quad (3-7)$$

选择正态分布作为可靠性模型时，将其可靠度函数式（2-9）和概率密度函数式

(2-11) 代入式 (3-5) 可得对数似然函数和对应的方程组如式 (3-8) 和 (3-9) 所示:

$$\ln L(\mu, \sigma) = -r \ln \sigma - \frac{r}{2} \ln 2\pi - \frac{1}{2} \sum_{i=1}^r \left( \frac{t_i - \mu}{\sigma} \right)^2 + \sum_{j=1}^{n-r} \ln \left[ 1 - \int_{-\infty}^T \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\tau - \mu}{\sigma} \right)^2} d\tau \right] \quad (3-8)$$

$$\begin{cases} \frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^r \frac{t_i - \mu}{\sigma^2} + \frac{(n-r) \int_{-\infty}^T \frac{\mu - \tau}{\sigma^3 \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\tau - \mu}{\sigma} \right)^2} d\tau}{1 - \int_{-\infty}^T \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\tau - \mu}{\sigma} \right)^2} d\tau} = 0 \\ \frac{\partial \ln L}{\partial \sigma} = -\frac{r}{\sigma} + \sum_{i=1}^r \frac{(t_i - \mu)^2}{\sigma^3} + \frac{(n-r) \int_{-\infty}^T \frac{1}{\sigma^2 \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\tau - \mu}{\sigma} \right)^2} \left[ 1 - \frac{(\tau - \mu)^2}{\sigma^2} \right] d\tau}{1 - \int_{-\infty}^T \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\tau - \mu}{\sigma} \right)^2} d\tau} = 0 \end{cases} \quad (3-9)$$

选择对数正态分布作为可靠性模型时, 将其可靠度函数式 (2-13) 和概率密度函数式 (2-15) 代入式 (3-5) 可得对数似然函数和对应的方程组如式 (3-10) 和 (3-11) 所示:

$$\ln L(\mu, \sigma) = -r \ln \sigma - \frac{r}{2} \ln 2\pi - \sum_{i=1}^r \left[ \ln t_i + \frac{(\ln t_i - \mu)^2}{2\sigma^2} \right] + \sum_{j=1}^{n-r} \ln \left[ 1 - \int_{-\infty}^T \frac{e^{-\frac{(\ln \tau - \mu)^2}{2\sigma^2}}}{\sigma \tau \sqrt{2\pi}} d\tau \right] \quad (3-10)$$

$$\begin{cases} \frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^r \frac{\ln t_i - \mu}{\sigma^2} + \frac{(n-r) \int_{-\infty}^T \frac{\ln \tau - \mu}{\tau \sigma^3 \sqrt{2\pi}} e^{-\frac{(\ln \tau - \mu)^2}{2\sigma^2}} d\tau}{1 - \int_{-\infty}^T \frac{1}{\sigma \tau \sqrt{2\pi}} e^{-\frac{(\ln \tau - \mu)^2}{2\sigma^2}} d\tau} = 0 \\ \frac{\partial \ln L}{\partial \sigma} = -\frac{r}{\sigma} + \sum_{i=1}^r \frac{(\ln t_i - \mu)}{\sigma^3} + \frac{(n-r) \int_{-\infty}^T \frac{1}{\tau \sigma^2 \sqrt{2\pi}} e^{-\frac{(\ln \tau - \mu)^2}{2\sigma^2}} \left[ 1 - \frac{(\ln \tau - \mu)^2}{\sigma^2} \right] d\tau}{1 - \int_{-\infty}^T \frac{1}{\sigma \tau \sqrt{2\pi}} e^{-\frac{(\ln \tau - \mu)^2}{2\sigma^2}} d\tau} = 0 \end{cases} \quad (3-11)$$

从式 (3-6) 至式 (3-11) 可知, 不同于完全样本的情形, 截尾数据的存在使得似然函数及其相应的极大似然方程组复杂化, 尤其是正态分布和对数正态分布不再可以根据样本数据直接计算均值  $\mu$ 、方差  $\sigma^2$  的极大似然估计值。对于参数无封闭表达式的情形一般采用 New-Raphson 法求解。

Newton-Raphson 法求解非线性方程组时需计算方程组对于待求参数的偏导以获

取 Jacob 矩阵, 应用于上述三组形式复杂的方程组时存在较大的计算难度, 且应用于智能电表的可靠性数据时, 由于 ALT 所得试验数据属于小样本、现场数据的截尾程度较深, Newton-Raphson 法可能存在不收敛的情况<sup>[39]</sup>, 故本文 3.2.2 节将采用智能优化算法求解可靠性模型参数的极大似然估计值。

### 3.2.2 基于差分进化算法的截尾样本模型参数求解方法

极大似然估计求解参数估计值的核心在于使似然函数在参数可取值的范围内达到最大值, 而优化问题则是使目标函数在一定约束条件下达到最优。对比可知求解参数的极大似然估计可视为优化问题, 因此可采用智能优化算法求解, 由 3.2.1 节的对数似然函数和样本导出的模型参数约束条件, 可得本节处理的三个优化问题如式 (3-12) 至 (3-14) 所示:

$$\begin{cases} \max & r \ln \beta - r \beta \ln \eta + (\beta - 1) \sum_{i=1}^r \ln t_i - \sum_{i=1}^r \left( \frac{t_i}{\eta} \right)^{\beta} - (n-r) \left( \frac{T}{\eta} \right)^{\beta} \\ \text{s.t.} & \beta > 0, \Gamma\left(1 + \frac{1}{\eta}\right) \beta > \bar{X} \end{cases} \quad (3-12)$$

$$\begin{cases} \max & -r \ln \sigma - \frac{r}{2} \ln 2\pi - \frac{1}{2} \sum_{i=1}^r \left( \frac{t_i - \mu}{\sigma} \right)^2 + \sum_{j=1}^{n-r} \ln \left[ 1 - \int_{-\infty}^T \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{\tau - \mu}{\sigma} \right)^2} d\tau \right] \\ \text{s.t.} & \sigma^2 > S^2, \mu > \bar{X} \end{cases} \quad (3-13)$$

$$\begin{cases} \max & -r \ln \sigma - \frac{r}{2} \ln 2\pi - \sum_{i=1}^r \left[ \ln t_i + \frac{(\ln t_i - \mu)^2}{2\sigma^2} \right] + \sum_{j=1}^{n-r} \ln \left[ 1 - \int_{-\infty}^T \frac{e^{-\frac{(\ln \tau - \mu)^2}{2\sigma^2}}}{\sigma \tau \sqrt{2\pi}} d\tau \right]^{\beta} \\ \text{s.t.} & \sigma^2 > S^2, \mu > \bar{Y} \end{cases} \quad (3-14)$$

其中  $\bar{X}$  为样本均值,  $\bar{Y}$  为样本取对数后的均值,  $S^2$  为样本方差,  $\Gamma(\cdot)$  为伽马函数。

相较于传统优化算法, 智能优化算法的优势在于计算过程不基于目标函数的导数而采用启发式方法搜寻最优解, 因此可用于求解较为复杂的寻优问题。在现有的智能优化算法中, 遗传算法、差分进化算法、蚁群算法、粒子群算法和模拟退火算法目前的研究和应用较多, 其各自的优缺点如表 3-1 所示<sup>[50]</sup>。

# 华中科技大学硕士学位论文

表 3-1 常用智能优化算法及其优缺点

优化算法	优点	缺点
遗传算法	搜索具有并行性、鲁棒性；易与其他优化算法结合；可求解只有代码概念的问题	局部搜索能力差、易早熟；高维问题收敛慢
差分进化算法	实数编码；搜索具有并行性且鲁棒性强；同样精度要求下收敛速度更快、求解效率更高	可能出现早熟
蚁群算法	搜索具有并行性；易与其他优化算法结合；不依赖初始值	收敛慢、搜索时间长
粒子群算法	对种群大小不敏感；收敛快	易陷入局部最优且解的质量与参数的选择有较大关系
模拟退火算法	能以一定概率接受目标变差的值、全局搜索能力强；不依赖初始值；对目标函数要求少、适用范围广	并行性差；需较高初始温度、较慢降温和足够抽样次数，优化时间长

大部分智能优化算法在提供一个较好的寻优初值时最终能收敛到问题的最优解，一种可行的方法是将基于秩回归方法求得的参数估计结果作为寻优初值<sup>[51]</sup>。如前所述，尽管该方法所得可靠性模型参数的估计结果由于累积失效概率估计式有多种形式、无法保证因变量变换前的残差平方和最小而存在精度不够理想的问题，但其值仍在最优解附近，因此采用智能优化算法求解可靠性模型参数的极大似然估计时配合搜索能力强、收敛速度快的算法可以较快获得比较理想的解。

结合表 3-1 中各算法的特点，本文结合搜索效率高、解的精度较高且鲁棒性好的差分进化算法与基于秩回归的方法求解截尾数据下可靠性模型参数极大似然估计问题，原因在于该问题难以采用常规迭代法求解但可视为以对数似然函数为目标函数、可靠性模型参数为优化变量的优化问题。为了避免算法早熟问题，采用如式 (3-15) 所示的自适应变异算子，式中  $G$  为当前迭代次数、 $G_{\max}$  为最大迭代次数，由式可知：在算法前期较大的变异算子增加种群多样性，在后期较小的变异算子提高算法精度。

$$\begin{cases} \lambda = e^{\frac{1-G_{\max}}{1+G_{\max}-G}} \\ F = F_0 \cdot 2^\lambda \end{cases} \quad (3-15)$$

采用前述方法求解问题的流程如下：首先随机生成初始种群，再将初始种群中的一个向量替换为由秩回归法得到的模型参数估计值，达到将基于秩回归法求得的参数估计结果作为寻优初值的目的；随后随机选取 3 个包含在种群中且互不相同的向量，将其中一个向量与经变异系数放大后的另外 2 个向量的差值相加，得到变异向量，完成变异操作；再根据设定的规则将变异向量与目标向量混合获得试验向量，完成交叉操作；对种群中每个向量均进行变异交叉后，将超出边界的向量替换为边界内随机生成的向量，完成边界条件的处理，得到临时种群；以对数似然函数为目标函数计算两个种群中每个个体的适应度，保留目标向量及其对应的试验向量中适应度大的个体，得到新一代种群；重复以上种群生成过程直到满足终止条件，得到最终的可靠性模型参数估计值。

相应的流程图如图 3-1 所示。

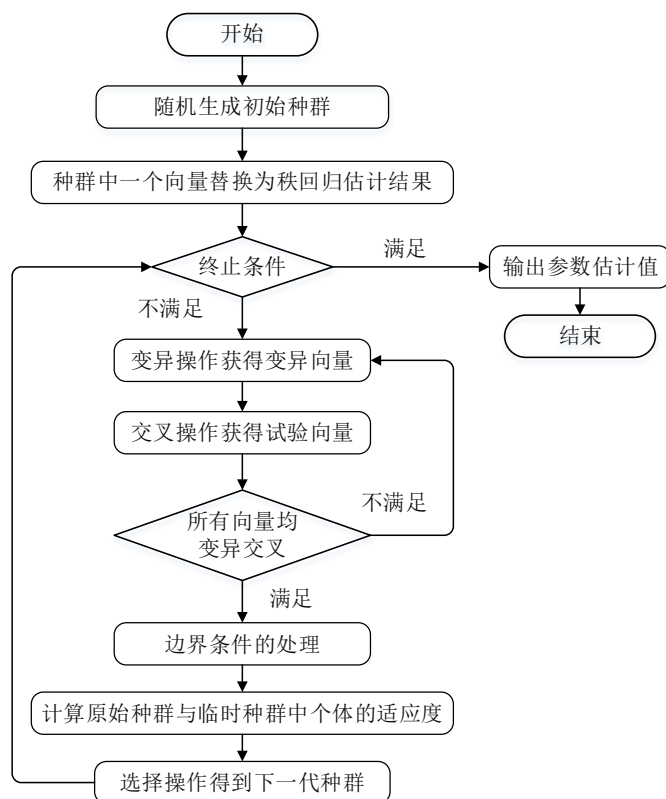


图 3-1 基于差分进化算法的截尾数据可靠性模型参数求解流程

### 3.3 基于差分进化算法的智能电表可靠性模型优选

#### 3.3.1 基于秩回归的寻优初始值

本文在温度为 85℃、相对湿度为 75%的条件下对 30 只智能电表开展为期 22 天的无替换、定时截尾的恒应力 ALT，试验中采用每间隔 24h 对智能电表的故障情况进行一次观测记录的方式，得到试验的原始记录数据如表 3-2 第 3 列所示。

由于试验采用每间隔 24h 对样本故障情况进行一次记录的方式，使得原始数据存在记录时间相同的情况，可根据 2.2.2 节中加速寿命试验数据预处理的式（2-17）将原始数据离散化，得到故障时间数据见表 3-2 第 4 列。

试验采用的加速应力为温度和湿度，相应的寿命应力关系由 Peck 模型描述，根据 Peck 模型对应的加速因子式（2-21）外推至正常条件。考虑到现场数据实际的使用环境情况，本文取正常使用条件为温度 20℃、相对湿度 75%，得到对应 AF 取值为 647，外推后的故障时间数据如表 3-2 第 5 列所示。

表 3-2 智能电表 ALT 数据及其处理结果

序号	故障现象	故障记录时间 (h)	加速应力下故障时 间 (h)	正常条件下故障时 间 (h)
1	背光故障	72	56	36232
2	背光故障	72	64	41408
3	背光故障	96	78	50466
4	背光故障	96	84	54348
5	背光故障	96	90	58230
6	背光故障	120	102	65994
7	背光故障	120	108	69876
8	载波通信异常	120	114	73758
9	载波通信异常	144	128	82816
10	载波通信异常	144	136	87992
11	载波通信异常	168	152	98344
12	载波通信异常	168	160	103520
13	载波通信异常	192	174	112578
14	载波通信异常	192	180	116460

续表 3-2

序号	故障现象	故障记录时间 (h)	加速应力下故障时 间 (h)	正常条件下故障时 间 (h)
15	载波通信异常	192	186	120342
16	液晶故障	336	320	207040
17	液晶故障	336	328	212216
18	液晶故障	360	348	225156
19	液晶故障	384	368	238096
20	液晶故障	384	376	243272
21	液晶故障	408	392	253624
22	液晶故障	408	400	258800
23	红外不通信	456	444	287268
24	日计时误差超差	504	488	315736
25	电池功耗异常	504	496	320912
26	上电不停复位	528	516	333852

秩回归法求解可靠性模型参数的原理为：首先由秩估计式得到各失效样本的累积失效概率 $F(t_i)$ ，再将可靠性模型转化为 $y=Ax+B$ 的线性形式后，由各样本 $(t_i, F(t_i))$ 采用最小二乘回归根据故障时间和累积失效概率求得回归系数。

累积失效概率 $F(t)$ 的秩估计式有多种形式，常用的 Bernard 中位秩估计、Blom 中位秩估计和平均秩估计的计算式分别如式（3-16）至（3-18）所示：

$$F(t_i) = \frac{i-0.3}{n+0.4} \quad (i=1,2,\dots,r) \quad (3-16)$$

$$F(t_i) = \frac{i-0.375}{n+0.25} \quad (i=1,2,\dots,r) \quad (3-17)$$

$$F(t_i) = \frac{i}{n+1} \quad (i=1,2,\dots,r) \quad (3-18)$$

其中  $i$  为故障样本序号， $t_i$  是序号为  $i$  的样本的故障时间； $n$  为样本总数； $r$  为失效样本数。

当待求解模型为 Weibull 分布时，该模型的不可靠度式（2-6）经过两次对数变换后可得到线性化的表达式如式（3-19）所示。



$$\ln\left(\ln\frac{1}{1-F(t)}\right) = \beta \ln t - \beta \ln \eta \quad (3-19)$$

此时，令  $x = \ln t$  和  $y = \ln\left(\ln\frac{1}{1-F(t)}\right)$ ，可得模型参数与线性函数系数的关系为  $A = \beta$ 、 $B = -\beta \ln \eta$ 。

当待求解模型为正态分布和对数正态分布时，变换后的线性化表达式如式（3-20）和式（3-21）所示。

$$\Phi^{-1}(F(t)) = \frac{1}{\sigma}t - \frac{\mu}{\sigma} \quad (3-20)$$

$$\Phi^{-1}(F(t)) = \frac{1}{\sigma}\ln t - \frac{\mu}{\sigma} \quad (3-21)$$

对于正态分布令  $x = t$  和  $y = \Phi^{-1}(F(t))$ ，对于对数正态分布则令  $y = \Phi^{-1}(F(t))$  和  $x = \ln t$ ，则得到模型参数和线性函数系数满足  $A = \frac{1}{\sigma}$ 、 $B = -\frac{\mu}{\sigma}$  的关系，其中  $\Phi^{-1}(\cdot)$  表示正态分布的逆函数。

由各失效样本  $(t_i, F(t_i))$  计算定义的  $x_i$  和  $y_i$ ，代入式（3-22）和（3-23）中可计算得到各备选模型对应的  $A$  和  $B$ ，再根据各备选模型参数与  $A$  和  $B$  的关系即可得到基于秩回归法的可靠性模型参数估计值。

$$A = \frac{\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2} \quad (3-22)$$

$$B = \frac{1}{n} \sum_{i=1}^n y_i - \frac{A}{n} \sum_{i=1}^n x_i \quad (3-23)$$

本节采取前述的三种秩估计计算累积失效概率，将表 3-2 中给出的 ALT 故障时间数据和根据式（3-16）到（3-18）得到的累积失效概率整理如表 3-3 所示。

表 3-3 智能电表 ALT 数据及累积失效概率

序号	故障时间 (h)	Benard 中位秩估计	Blom 中位秩估计	平均秩估计
1	36232	0.02303	0.02066	0.03226
2	41408	0.05592	0.05372	0.06452
3	50466	0.08882	0.08678	0.09677

续表 3-3

序号	故障时间 (h)	Benard 中位秩估计	Blom 中位秩估计	平均秩估计
4	54348	0.12171	0.11983	0.12903
5	58230	0.15461	0.15289	0.16129
6	65994	0.18750	0.18595	0.19355
7	69876	0.22039	0.21901	0.22581
8	73758	0.25329	0.25207	0.25806
9	82816	0.28618	0.28512	0.29032
10	87992	0.31908	0.31818	0.32258
11	98344	0.35197	0.35124	0.35484
12	103520	0.38487	0.38430	0.38710
13	112578	0.41776	0.41736	0.41935
14	116460	0.45066	0.45041	0.45161
15	120342	0.48355	0.48347	0.48387
16	207040	0.51645	0.51653	0.51613
17	212216	0.54934	0.54959	0.54839
18	225156	0.58224	0.58264	0.58065
19	238096	0.61513	0.61570	0.61290
20	243272	0.64803	0.64876	0.64516
21	253624	0.68092	0.68182	0.67742
22	258800	0.71382	0.71488	0.70968
23	287268	0.74671	0.74793	0.74194
24	315736	0.77961	0.78099	0.77419
25	320912	0.81250	0.81405	0.80645
26	333852	0.84539	0.84711	0.83871

除 ALT 数据外, 本文还从南方电网辖区下收集了与试验样本同型号的一批智能电表现场数据, 该批次智能电表共 2312 只、其中故障表 53 只, 由于该批次智能电表的投运日期较为集中, 故对投运日期做归一化后可将其处理为定时截尾样本。对其故障样本同样采用式 (3-16) 到 (3-18) 计算累积失效概率近似值, 并同故障时间整理如表 3-4 所示。

# 华中科技大学硕士学位论文

表 3-4 智能电表现场数据及累积失效概率

序号	故障时间 (h)	Benard 中位秩估计	Blom 中位秩估计	平均秩估计
1	54528	0.00030	0.00027	0.00043
2	57696	0.00074	0.00070	0.00086
3	57840	0.00117	0.00114	0.00130
4	59064	0.00160	0.00157	0.00173
5	60072	0.00203	0.00200	0.00216
6	60312	0.00246	0.00243	0.00259
7	64128	0.00290	0.00287	0.00303
8	65040	0.00333	0.00330	0.00346
9	65256	0.00376	0.00373	0.00389
10	65544	0.00419	0.00416	0.00432
11	66918	0.00463	0.00460	0.00476
12	66924	0.00506	0.00503	0.00519
13	66930	0.00549	0.00546	0.00562
14	67080	0.00592	0.00589	0.00605
15	67152	0.00636	0.00633	0.00649
16	67944	0.00679	0.00676	0.00692
17	68160	0.00722	0.00719	0.00735
18	70952	0.00765	0.00762	0.00778
19	70960	0.00809	0.00805	0.00821
20	71832	0.00852	0.00849	0.00865
21	72944	0.00895	0.00892	0.00908
22	72952	0.00938	0.00935	0.00951
23	73632	0.00982	0.00978	0.00994
24	73656	0.01025	0.01022	0.01038
25	74208	0.01068	0.01065	0.01081
26	74472	0.01111	0.01108	0.01124
27	74544	0.01155	0.01151	0.01167
28	74664	0.01198	0.01195	0.01211
29	74880	0.01241	0.01238	0.01254

续表 3-4

序号	故障时间 (h)	Benard 中位秩估计	Blom 中位秩估计	平均秩估计
30	75360	0.01284	0.01281	0.01297
31	75888	0.01328	0.01324	0.01340
32	76344	0.01371	0.01368	0.01383
33	76488	0.01414	0.01411	0.01427
34	76512	0.01457	0.01454	0.01470
35	76728	0.01501	0.01497	0.01513
36	78648	0.01544	0.01541	0.01556
37	78840	0.01587	0.01584	0.01600
38	78864	0.01630	0.01627	0.01643
39	79824	0.01674	0.01670	0.01686
40	82032	0.01717	0.01714	0.01729
41	82200	0.01760	0.01757	0.01773
42	82368	0.01803	0.01800	0.01816
43	82392	0.01847	0.01843	0.01859
44	82896	0.01890	0.01887	0.01902
45	83784	0.01933	0.01930	0.01946
46	84096	0.01976	0.01973	0.01989
47	84606	0.02020	0.02016	0.02032
48	84612	0.02063	0.02060	0.02075
49	84618	0.02106	0.02103	0.02118
50	87552	0.02149	0.02146	0.02162
51	88056	0.02193	0.02189	0.02205
52	89760	0.02236	0.02233	0.02248
53	89784	0.02279	0.02276	0.02291

将表 3-3 和 3-4 的数据结合式 (3-19) 至 (3-23), 可得到 ALT 数据、现场数据在三种备选可靠性模型下模型参数基于秩回归的的估计结果, 分别整理如表 3-5 和 3-6 所示。

表 3-5 ALT 数据可靠性模型参数秩回归估计结果

秩估计方法	Weibull 分布		正态分布		对数正态分布	
	$\beta$	$\eta$	$\mu$	$\sigma$	$\mu$	$\sigma$
Benard 中位秩	1.478	214111	188229	135361	11.952	0.913
Blom 中位秩	1.499	213658	188386	134108	11.953	0.903
平均秩	1.405	215859	187714	140095	11.949	0.948

表 3-6 现场数据可靠性模型参数秩回归估计结果

秩估计方法	Weibull 分布		正态分布		对数正态分布	
	$\beta$	$\eta$	$\mu$	$\sigma$	$\mu$	$\sigma$
Benard 中位秩	7.300	140700	141774	28686	12.128	0.391
Blom 中位秩	7.383	139808	141245	28434	12.120	0.387
平均秩	7.012	143941	143682	29604	12.155	0.404

### 3.3.2 加速寿命试验数据和现场数据的可靠性模型

本文通过 Python 的第三方库 scikit-opt 实现基于差分进化算法求解定时截尾数据下可靠性模型参数的极大似然估计。

当求解由 ALT 数据基于极大似然原理构建的优化问题时，样本总数  $n$  为 30，失效样本数  $r$  为 26，截尾样本数为 4，截尾时间  $T$  为 333852h，并将 3.3.1 节中根据 Bernard 中位秩求得的估计结果加入算法的初始种群作为寻优初值，设置种群个体数量  $N_p$  为 50，采用前述自适应变异算子，最大迭代次数  $G_{\max}$  为 800，得到问题的最优解和相应的  $AIC_c$  值如表 3-7 所示：

表 3-7 ALT 数据备选模型最优解及  $AIC_c$

	Weibull 分布	正态分布	对数正态分布
模型参数寻优结果	$\beta = 1.475$ $\eta = 215726$	$\mu = 188184$ $\sigma = 123153$	$\mu = 11.927$ $\sigma = 0.815$
$AIC_c$	688.318	696.882	686.501

可以看出，对数正态分布的参数最优解对应的  $AIC_c$  是三种备选模型中最小的，由此可知：采用对数正态分布模型描述该型号智能电表加速寿命试验数据最为合适。

同理，差分进化算法求解由现场数据基于极大似然原理构建的优化问题，此时样本总数  $n$  为 2312，失效样本数  $r$  为 53，截尾样本数为 2259，截尾时间  $T$  为

89784h, 其余参数设置不变, 得到的最优解和 AIC 值如表 3-8 所示:

表 3-8 现场数据备选模型最优解及 AIC

	Weibull 分布	正态分布	对数正态分布
模型参数寻优结果	$\beta = 5.023$ $\eta = 189898$	$\mu = 168790$ $\sigma = 39649$	$\mu = 12.418$ $\sigma = 0.508$
AIC	1632.171	1634.997	1628.229

同样, 由对数正态分布的 AIC 最小可以得到: 采用对数正态分布描述该型号智能电表的现场数据最为合适。

## 3.4 智能电表可靠性模型优选结果分析

为进一步说明本文方法的优越性, 分别将表 3-5 和 3-6 中 Blom 中位秩估计、平均秩估计对应的模型参数估计结果作为寻优初值再次运行差分进化算法求解相应的优化问题, 得到的备选模型参数寻优结果与 3.3.2 节相同。此外, 表 3-9 和 3-10 分别给出了 ALT 数据和现场数据采用三种秩回归方法求得模型参数时各备选模型的 AIC<sub>C</sub> 和 AIC 值。

表 3-9 ALT 数据三种秩回归方法的 AIC<sub>C</sub>

	Weibull 分布	正态分布	对数正态分布
Benard 中位秩	688.322	697.291	687.099
Blom 中位秩	688.336	697.216	686.997
平均秩	688.406	697.629	687.523

表 3-10 现场数据三种秩回归方法的 AIC

	Weibull 分布	正态分布	对数正态分布
Benard 中位秩	1655.138	1650.915	1638.311
Blom 中位秩	1656.313	1651.572	1638.810
平均秩	1651.414	1649.155	1636.677

结合 3.3.2 节的结果, 可分析得到如下结论:

- (1) 从 3.3.2 节优化问题的寻优结果来看, 在描述该型号智能电表的 ALT 数据

和现场数据时，三种备选模型中均为对数正态分布表现最佳，故采用对数正态分布作为该型号智能电表的可靠性模型。

(2) 不同参数估计方法的优劣也可以通过对比 AIC 进行判断。由表 3-9、3-10 和表 3-7、3-8 中的  $AIC_c$  和 AIC 可知：对于秩回归方法，不同的累积失效概率秩估计式得到的结果不同，当样本量足够大时平均秩优于 Bernard 中位秩和 Blom 中位秩的结果，而对于小样本当可靠性模型为正态分布或对数正态分布时 Blom 秩的表现最优、可靠性模型为 Weibull 分布时 Bernard 秩表现最优；本文采用的方法所得  $AIC_c$  和 AIC 均略小于基于秩回归方法的结果，说明本文方法表现更优且基于秩回归方法的结果确实接近最优解。

(3) 在采用差分进化算法求解问题最优解时，分别将三种秩回归得到的参数估计值加入寻优初值均得到相同的模型参数寻优结果，说明相较于仅采用秩回归的方法，本文采用秩回归与差分进化算法相结合的方法所得模型参数估计结果不会受累积失效概率估计式的影响，原因在于：本文采用秩回归法仅为优化算法提供较好的寻优初值，在构造待求解优化问题时并不需要事先确定累积失效概率。

### 3.5 本章小结

本章研究了智能电表的可靠性模型优选。本章以综合考虑了拟合效果和模型复杂度的 AIC 准则作为模型优选准则，针对 Newton-Raphson 法求解由非完全样本构建的极大似然方程组时计算复杂、小样本与深截尾情形可能不收敛的问题，本章将可靠性模型参数的极大似然估计求解问题转化为优化问题进行处理，以基于秩回归的参数估计值作为寻优初始值，采用搜索效率高、解的精度较高且鲁棒性好的差分进化算法求得问题的最优解，结果表明本文所用方法得到的可靠性模型参数估计结果不受累积失效概率估计式影响，最终优选了对数正态分布作为该型号智能电表的可靠性模型。

## 4 基于贝叶斯方法的智能电表可靠性评估

在确定寿命服从的可靠性模型后，通过对可靠性数据的分析求得模型参数即可对产品的可靠性进行评估。本章在第三章的基础上，针对由单一数据源得到的可靠性评估结果准确度不高的问题，基于贝叶斯方法对智能电表的可靠性进行评估，由 ALT 数据修正采用最大熵法和 Bootstrap 方法挖掘现场数据包含的可靠性信息获取的可信先验分布，实现对现场数据的充分利用、ALT 数据与现场数据的融合，提高可靠性评估结果的准确性。

### 4.1 贝叶斯可靠性评估相关方法

考虑到本文通过加速寿命试验和现场收集的数据均为寿命数据的形式，采用更适合处理多源数据具有相同数据形式的基于贝叶斯理论的可靠性评估方法。本节介绍贝叶斯可靠性评估流程中的先验分布确定、后验分布采样相关方法。

#### 4.1.1 先验分布的确定方法

先验分布的获取可分为主观法和客观法两类方法，主观法确定先验分布时由专家根据经验确定分布的形式及参数，但该方法所得结果的可信程度完全受限于专家的主观判断；客观法按是否掌握先验信息又可分为无信息先验和有信息先验，前者主要为贝叶斯假设，后者根据一定的原则确定先验分布的形式后，再从已有的资料和数据入手确定先验分布的超参数，主要方法有共轭先验分布、最大熵先验分布。

##### (1) 贝叶斯假设

贝叶斯假设的思想是同等无知原则：在参数可能的取值范围内不对任一取值有所偏好而是假定概率相等，即采用均匀分布构建先验分布，如式（4-1）所示，式中  $c$  为常数。

$$\pi(\theta) = c \quad (4-1)$$

这一方法仅适用于参数取值范围已知的情形，但应用于高可靠性产品时会有不利影响<sup>[52]</sup>。



## (2) 共轭先验分布

共轭先验分布的基本思想是使得先验分布和后验分布具有相同的形式，总体分布中的每一个参数均有各自的共轭分布，如完全样本下正态分布的均值和方差的共轭先验分布分别为正态分布和逆伽马分布。

该方法因便于后验分布的计算而在贝叶斯方法中得到广泛的应用，但一方面该方法仅考虑数学形式上的方便却缺少了对实际情况合理性的考量，另一方面似然函数形式复杂时难以确定相应的共轭先验分布。

## (3) 最大熵法

最大熵法以最大熵准则作为先验分布形式的确定原则，相较于前述的几种方法，该方法采用熵对先验分布的客观性进行评价，最大熵准则采取在满足先验信息所得约束条件的同时最大化 Shannon-Jaynes 熵的策略构建先验分布，如式 (4-2) 所示：

$$\begin{cases} \max & -\int \pi(\theta) \ln \pi(\theta) d\theta \\ \text{s.t.} & S_i(\pi(\theta)) = c_i \end{cases} \quad (4-2)$$

其中  $-\int \pi(\theta) \ln \pi(\theta) d\theta$  为先验分布的熵， $S_i(\pi(\theta)) = c_i$  表示由先验信息构造的约束条件。

该方法能够兼顾对先验信息的充分利用和尽可能少的未知假设的引入，熵概念的应用从一定程度上解决了先验分布的确定缺少客观准则的问题。

### 4.1.2 后验分布的采样方法

后验分布是经正则化的先验分布与似然函数之积，除共轭先验分布情形下可直接写出后验分布表达式外，其余情形下的后验分布形式均较为复杂。为获取相应的后验分布及其相关统计量，可采用 MCMC 方法生成后验分布的模拟样本。

根据遍历马尔科夫链在任意初始分布下经足够多的状态转移次数后最终收敛至平稳分布的特性，MCMC 方法构造一个马尔科夫链使其平稳分布为目标分布，则经过足够多次的状态转换后所得的序列  $\{x_m, x_{m+1}, \dots, x_n\}$  即可视为来自目标分布的样本。马尔可夫链收敛的充分条件由细致平衡方程 (4-3) 给出：

$$\pi(i) p(i, j) = \pi(j) p(j, i) \quad (4-3)$$

式中  $\pi(i)$  为马尔科夫链的平稳分布在状态  $i$  的概率； $p(i, j)$  为状态转移矩阵  $\mathbf{P}$  的元素，

表示状态  $i$  向状态  $j$  转移的概率。

构造满足式 (4-3) 的状态转移矩阵  $\mathbf{P}$  是 MCMC 方法的核心, 不同的状态转移矩阵构造方法对应了不同的 MCMC 方法。

## (1) Metropolis-Hasting 算法

Metropolis-Hasting 算法在采样时任意给定一个提议分布得到状态转移矩阵  $\mathbf{Q}$ , 通过向  $\mathbf{Q}$  引入接受率  $\alpha$  获得能满足式 (4-3) 的状态转移矩阵  $\mathbf{P}$ , 即以一定的概率  $\alpha$  接受按照矩阵  $\mathbf{Q}$  进行的状态转移可以使马尔科夫链收敛:

$$\pi(i)q(i,j)\alpha_{ij} = \pi(j)q(j,i)\alpha_{ji} \quad (4-4)$$

由式 (4-4) 可知, 一种可行的方案为取  $\alpha_{ij} = \pi(j)q(j,i)$ 、 $\alpha_{ji} = \pi(i)q(i,j)$ , 此时新的状态转移矩阵  $\mathbf{P}$  的元素为  $p(j,i) = q(j,i)\alpha_{ji}$ 。考虑到  $\alpha$  为两个概率之积取值可能较小, Metropolis-Hasting 算法对原接受率做放大改进后得到最终的接受率表达式为  $\alpha_{ij} = \min\left(1, \frac{\pi(j)q(j,i)}{\pi(i)q(i,j)}\right)$ 。

Metropolis-Hasting 算法的优点在于操作简单, 每次采样仅需要根据提议分布生成下一个候选样本后再根据接受率判断是否接受该样本, 但由于存在接受率小于 1 的情况, 需经过较长时间获得平稳分布的样本, 高维情况下采样效率更低。

## (2) Gibbs 算法

Gibbs 算法在构造状态转移矩阵  $\mathbf{Q}$  时不再随机给定, 而是从高维分布的条件分布出发构造能满足式 (4-3) 的状态转移矩阵  $\mathbf{Q}$ 。对于两个仅有一个维度取值不等的  $n$  维样本  $\mathbf{A}(x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)})$  和  $\mathbf{B}(x_1^{(2)}, x_2^{(1)}, \dots, x_n^{(1)})$ , 由条件概率可得到式 (4-5) 成立:

$$\pi(\mathbf{A})\pi(x_1^{(2)} | x_2^{(1)}, x_3^{(1)}, \dots, x_n^{(1)}) = \pi(\mathbf{B})\pi(x_1^{(1)} | x_2^{(1)}, x_3^{(1)}, \dots, x_n^{(1)}) \quad (4-5)$$

对比式 (4-4) 和 (4-5) 可知, 当状态转移矩阵  $\mathbf{Q}$  中的元素为目标分布的条件分布时马尔科夫链可收敛至平稳分布。

Gibbs 算法相当于  $\alpha$  为 1 的 Metropolis-Hasting 算法, 可更高效地对高维分布抽样, 但也存在计算量大的缺点, 原因是在  $n$  维空间中从样本  $(x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)})$  出发生成下一个样本  $(x_1^{(2)}, x_2^{(2)}, \dots, x_n^{(2)})$  需经过  $n$  次状态转移, 此外该算法还需已知目标分布的条件分布, 对于条件分布难以获取的情形, 该算法难以适用。

## (3) NUTS 算法

NUTS 算法将哈密顿动力学思想应用于状态转移矩阵的构造。对于给定的目标分布 $\pi(\theta)$ ，定义相应的势能函数 $U(\theta)=-\ln\pi(\theta)$ ，同时引入辅助分析的变量  $p$ ，其概率密度函数为 $f(p)$ ，定义相应的动能函数 $K(p)=-\ln f(p)$ ，根据能量守恒及能量的正则分布有式 (4-6) 成立：

$$\frac{1}{Z}e^{-U(\theta_1)-K(p_1)} = \frac{1}{Z}e^{-U(\theta_2)-K(p_2)} \quad (4-6)$$

进一步整理得到式 (4-7)：

$$\pi(\theta_1)f(p_1) = \pi(\theta_2)f(p_2) \quad (4-7)$$

当辅助变量从状态  $p_1$  向状态  $p_2$  转移时，按哈密顿动力学方程求解  $\theta_1$  的状态转移结果  $\theta_2$ ，考虑到离散化求解哈密顿动力学方程引入了计算误差，采用接受率  $\alpha=\min\left(1, \frac{\pi(\theta_2)f(p_2)}{\pi(\theta_1)f(p_1)}\right)$  进行修正，则式 (4-7) 成立，相应的马尔科夫链收敛。

NUTS 算法以哈密顿动力学作为状态转移的依据，避免了随机状态转移造成的接受率不高的问题，适合处理高维数据且收敛速度更快。

## 4.2 基于贝叶斯方法的智能电表可靠性评估

### 4.2.1 基于贝叶斯方法的智能电表可靠性评估流程

如前所述，本文采用贝叶斯可靠性评估方法对智能电表的可靠性进行评估。

为在尽可能不引入主观信息的前提下最大限度利用已有客观信息，本文采用最大熵法确定先验分布形式，并将能反映现场实际可靠性水平的现场数据作为先验信息源，通过 Bootstrap 方法做进一步挖掘以确定先验分布的超参数，从而获取相对客观可信的先验分布，提高贝叶斯可靠性评估结果的准确性。

基于贝叶斯定理融合 ALT 数据和现场数据后，由于定时截尾数据下似然函数形式复杂，可靠性模型参数的后验分布也相对复杂，采用性能更加优良的 NUTS 算法对后验分布进行采样以获取相应的参数估计结果。

上述流程对应的流程图如图 4-1 所示：具体分析过程和相应结果在后续小节中详述。

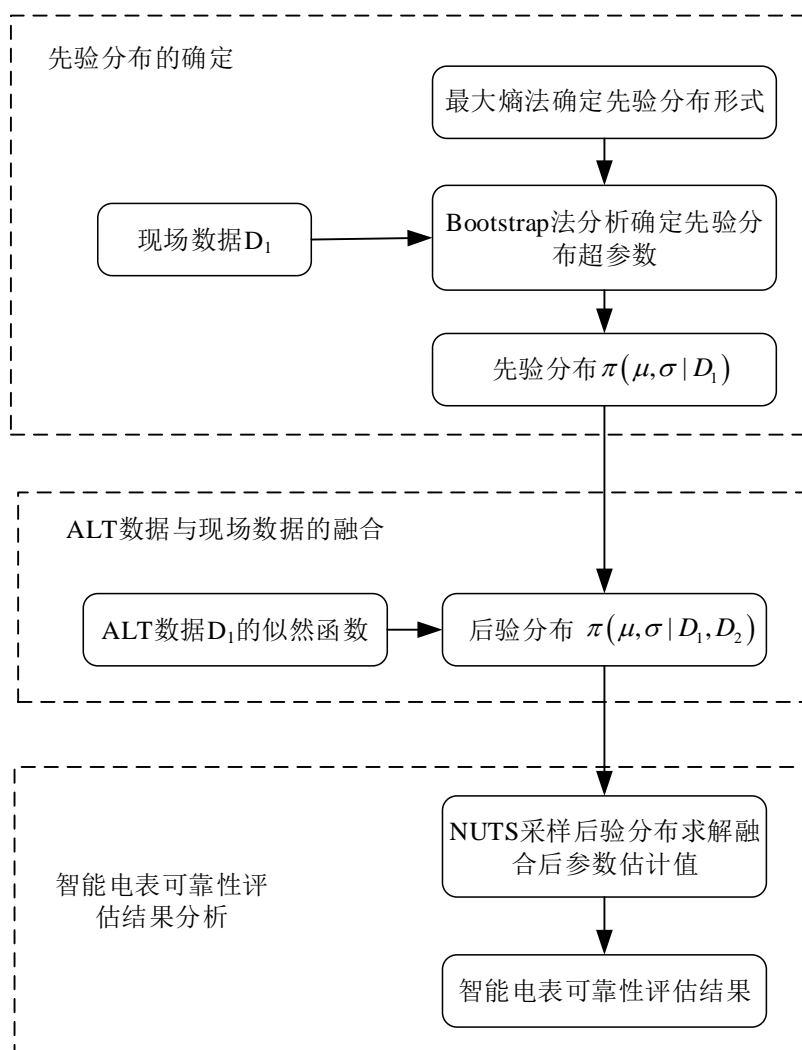


图 4-1 基于贝叶斯方法的智能电表可靠性评估流程

#### 4.2.2 最大熵法确定先验分布形式

由最大熵准则构造  $\mu$ 、 $\sigma$  联合先验分布的问题如式 (4-8) 所示，式中的三个约束条件分别为概率正则化条件和模型参数的矩约束条件：

$$\begin{cases} \max & -\iint \pi(\mu, \sigma) \ln \pi(\mu, \sigma) d\mu d\sigma \\ \text{s.t.} & \\ & \iint \pi(\mu, \sigma) d\mu d\sigma = 1 \\ & \iint \mu^i \pi(\mu, \sigma) d\mu d\sigma = \mu_i \quad (i=1, 2, \dots, k) \\ & \iint \sigma^j \pi(\mu, \sigma) d\mu d\sigma = \sigma_j \quad (j=1, 2, \dots, m) \end{cases} \quad (4-8)$$

其中  $\mu_i$  是参数  $\mu$  的  $i$  阶原点矩， $\sigma_j$  是参数  $\sigma$  的  $j$  阶原点矩。

以对数正态分布的参数  $\mu$ 、 $\sigma$  的一阶和二阶原点矩作为约束条件<sup>[53]</sup>，根据原点矩和中心矩的关系将二阶原点矩约束变为等价的二阶中心矩约束后，引入拉格朗日乘子将有约束优化问题转化为等价的无约束优化问题，相应的目标函数如式（4-9）所示：

$$\begin{aligned} L(\pi(\mu, \sigma)) = & -\iint \pi(\mu, \sigma) \ln \pi(\mu, \sigma) d\mu d\sigma + \lambda_0 \left[ \iint \pi(\mu, \sigma) d\mu d\sigma - 1 \right] \\ & + \lambda_1 \left[ \iint \pi(\mu, \sigma) \mu d\mu d\sigma - \mu_1 \right] + \lambda_2 \left[ \iint \pi(\mu, \sigma) (\mu - \mu_1)^2 d\mu d\sigma - \mu_2 + \mu_1^2 \right] \\ & + \alpha_1 \left[ \iint \pi(\mu, \sigma) \sigma d\mu d\sigma - \sigma_1 \right] + \alpha_2 \left[ \iint \pi(\mu, \sigma) (\sigma - \sigma_1)^2 d\mu d\sigma - \sigma_2 + \sigma_1^2 \right] \end{aligned} \quad (4-9)$$

式中  $\lambda_i$  和  $\alpha_j$  为拉格朗日乘子。将式（4-9）中的常数项合并，记被积函数为  $F(\pi(\mu, \sigma))$ ，整理得到式（4-10）：

$$\begin{cases} L(\pi(\mu, \sigma)) = \iint F(\pi(\mu, \sigma)) d\mu d\sigma + C \\ F(\pi(\mu, \sigma)) = \left( -\ln \pi(\mu, \sigma) + \lambda_0 + \lambda_1 \mu + \lambda_2 (\mu - \mu_1)^2 + \alpha_1 \sigma + \alpha_2 (\sigma - \sigma_1)^2 \right) \pi(\mu, \sigma) \\ C = -\lambda_0 - \lambda_1 \mu_1 - \lambda_2 (\mu_2 - \mu_1^2) - \alpha_1 \sigma_1 - \alpha_2 (\sigma_2 - \sigma_1^2) \end{cases} \quad (4-10)$$

该目标函数的自变量是先验分布  $\pi(\mu, \sigma)$ ，故目标函数为泛函，求解目标函数最大值即求解泛函极值问题，可由 Euler-Lagrange 方程结合  $F(\pi(\mu, \sigma))$  中不包含自变量的导数项得到式（4-11）：

$$-\ln \pi(\mu, \sigma) + \lambda_0 + \lambda_1 \mu + \lambda_2 (\mu - \mu_1)^2 + \alpha_1 \sigma + \alpha_2 (\sigma - \sigma_1)^2 - 1 = 0 \quad (4-11)$$

则由拉格朗日乘子表示的先验分布为：

$$\pi(\mu, \sigma) = e^{\lambda_0 - 1 + \lambda_1 \mu + \lambda_2 (\mu - \mu_1)^2 + \alpha_1 \sigma + \alpha_2 (\sigma - \sigma_1)^2} \quad (4-12)$$

进一步整理：

$$\pi(\mu, \sigma) = e^{\lambda_0 - 1 - \frac{\lambda_1^2}{4\lambda_2} + \lambda_1 \mu_1 - \frac{\alpha_1^2}{4\alpha_2} + \alpha_1 \sigma_1 + \lambda_2 \left( \mu - \mu_1 + \frac{\lambda_1}{2\lambda_2} \right)^2 + \alpha_2 \left( \sigma - \sigma_1 + \frac{\alpha_1}{2\alpha_2} \right)^2} \quad (4-13)$$

由式（4-13）可知  $\mu$  关于  $\mu_1 - \frac{\lambda_1}{2\lambda_2}$  对称， $\sigma$  关于  $\sigma_1 - \frac{\alpha_1}{2\alpha_2}$  对称，而可由一阶矩约束条件得到  $\mu$  关于  $\mu_1$  对称， $\sigma$  关于  $\sigma_1$  对称，根据两条对称性的结论可得  $\lambda_1$  和  $\alpha_1$  均为 0，在此基础上将（4-12）带入正则化条件和二阶中心矩约束条件，得到待求方程组如式（4-14）所示：

$$\begin{cases} \frac{\pi e^{\lambda_0^{-1}}}{\sqrt{\lambda_2} \alpha_2} = 1 \\ \frac{\pi e^{\lambda_0^{-1}}}{2(-\lambda_2)^{\frac{3}{2}}(-\alpha_2)^{\frac{1}{2}}} = \mu_2 - \mu_1^2 \\ \frac{-\pi e^{\lambda_0^{-1}}}{2(-\alpha_2)^{\frac{3}{2}}(-\lambda_2)^{\frac{1}{2}}} = \sigma_2 - \sigma_1^2 \end{cases} \quad (4-14)$$

求解得到:

$$\begin{cases} e^{\lambda_0^{-1}} = \frac{1}{2\pi\sqrt{(\mu_2 - \mu_1^2)(\sigma_2 - \sigma_1^2)}} \\ \lambda_2 = \frac{1}{2(\mu_2 - \mu_1^2)} \\ \alpha_2 = \frac{1}{2(\sigma_2 - \sigma_1^2)} \end{cases} \quad (4-15)$$

根据各参数的结果整理得到参数  $\mu$  和  $\sigma$  的联合先验分布为二元正态分布:

$$\pi(\mu, \sigma) = \frac{1}{2\pi\sigma_\mu\sigma_\sigma} e^{-\frac{1}{2\sigma_\mu^2}(\mu - \bar{\mu})^2 - \frac{1}{2\sigma_\sigma^2}(\sigma - \bar{\sigma})^2} \quad (4-16)$$

式中  $\bar{\mu} = \mu_1$ 、 $\bar{\sigma} = \sigma_1$  分别为参数  $\mu$  和  $\sigma$  的均值;  $\sigma_\mu^2 = \mu_2 - \mu_1^2$ 、 $\sigma_\sigma^2 = \sigma_2 - \sigma_1^2$  分别为参数  $\mu$  和  $\sigma$  的方差。

### 4.2.3 Bootstrap 法确定先验分布超参数

确定了先验分布形式为二元正态分布后, 分布中的超参数  $\bar{\mu}$ 、 $\bar{\sigma}$ 、 $\sigma_\mu^2$  和  $\sigma_\sigma^2$  还需对现场数据进一步挖掘后确定。Bootstrap 方法是由 Efron 于 1979 年提出的一种增广样本统计方法<sup>[54]</sup>, 能对已有样本的信息进行充分挖掘, 其核心思想是通过样本及其 Bootstrap 子样研究未知分布的统计特性, 其流程如下:

(1) 由总体  $F(\theta)$  的样本  $X = \{x_1, x_2, \dots, x_n\}$  计算总体分布的参数估计值  $\hat{\theta}$ , 构造统计量  $T_n = \hat{\theta} - \theta$ ;

(2) 根据样本  $X = \{x_1, x_2, \dots, x_n\}$  进行  $m$  次重抽样, 获得  $m$  个 Bootstrap 子样  $X^i = \{x_1^i, x_2^i, \dots, x_n^i\}$ , 对各子样采用与步骤 (1) 同样的参数估计方法得到对应的  $m$  个估计值  $\theta_i (i=1, 2, \dots, m)$ ;

(3) 由统计量 $R_n = \theta_i - \hat{\theta}$ 可获得统计量 $T_n$ 的分布及其统计特性。

上述流程也可用于获取总体参数 $\theta$ 的分布，即用 $\theta_i (i=1,2,\dots,m)$ 的分布作为总体参数 $\theta$ 的分布。

根据 Bootstrap 子样获取方式的不同，该方法可分为非参数 Bootstrap 方法和参数 Bootstrap 方法，前者直接从样本中进行重复有放回抽样获取 Bootstrap 子样，后者从根据样本得到的分布模型中抽取 Bootstrap 子样，应用于可靠性数据分析时前者适用于完全样本、后者适用于截尾样本。

本文获取的智能电表现场数据为截尾样本，因此本文采用参数 Bootstrap 方法对现场数据进行处理从而获取先验分布超参数，相应流程如图 4-2 所示。

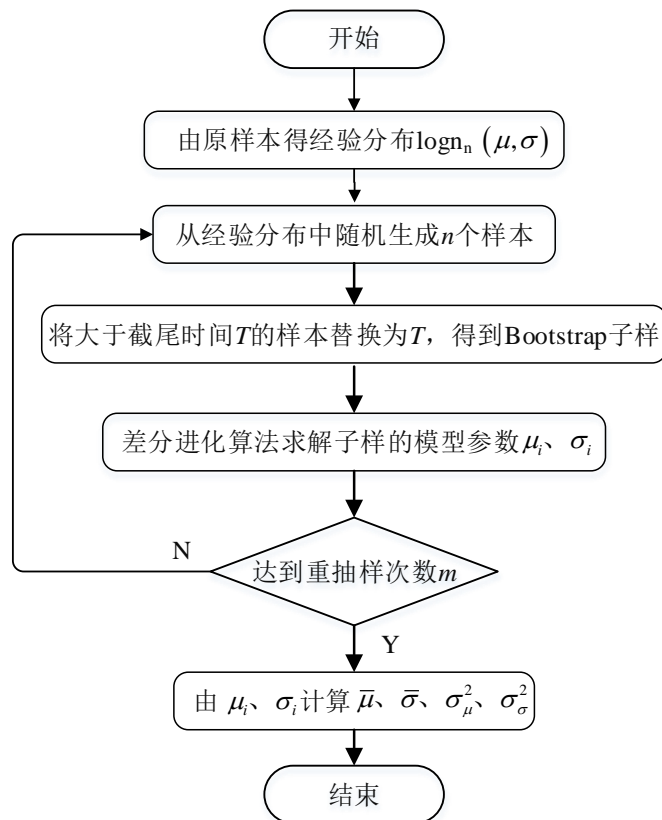


图 4-2 参数 Bootstrap 方法确定先验分布超参数的流程

图 4-2 的流程中：首先确定现场数据对应的总体为对数正态分布，并通过 3.2.2 节的方法求得参数估计值，记相应的经验分布为 $\text{logn}_n(\mu, \sigma)$ ；再依据 $\text{logn}_n(\mu, \sigma)$ 进行重抽样，由于现场数据为定时截尾样本，为了不引入现场数据不包含的故障信息，

在重抽样时从经验分布中随机生成总数与现场数据相同的  $n$  个样本，将不大于定时截尾时间  $T$  的样本予以保留、其余样本用  $T$  代替以构造定时截尾样本<sup>[55]</sup>，重复  $m$  次获取  $m$  个 Bootstrap 子样；对各子样也采用 3.2.2 节的方法求解参数  $\mu_i$  和  $\sigma_i$  ( $i=1,2,\dots,m$ )；最后由  $\mu_i$  和  $\sigma_i$  ( $i=1,2,\dots,m$ ) 计算相应的统计量  $\bar{\mu}$ 、 $\bar{\sigma}$ 、 $\sigma_{\mu}^2$  和  $\sigma_{\sigma}^2$  得到先验分布的超参数。

由 3.3.2 节的结果，经验分布参数  $\mu$  和  $\sigma$  分别为 12.418、0.508，样本总数  $n$  为 2312，截尾时间  $T$  为 89784h，重抽样次数设为 200，按上述流程运行程序得到超参数的结果如表 4-1 所示。

表 4-1 先验分布的超参数

统计量	$\mu$	$\sigma$
均值	12.409	0.504
方差	0.019	0.004

## 4.2.4 现场数据与 ALT 数据的融合

根据 3.3.2 节的结果可知，该批次智能电表的 ALT 数据也服从对数正态分布，将对数正态分布的概率密度函数表达式 (2-15)、可靠度表达式 (2-13) 带入截尾样本的似然函数表达式 (3-5) 可得到由 ALT 数据构建的似然函数，再依据贝叶斯定理将该似然函数与式 (4-16) 所示的先验分布结合得到后验分布如式 (4-17) 所示：

$$p(\mu, \sigma | D_1, D_2) \propto \frac{1}{2\pi\sigma_{\mu}\sigma_{\sigma}} e^{-\frac{(\mu-\bar{\mu})^2}{2\sigma_{\mu}^2} - \frac{(\sigma-\bar{\sigma})^2}{2\sigma_{\sigma}^2}} \prod_{i=1}^r \frac{1}{\sigma t_i \sqrt{2\pi}} e^{-\frac{(\ln t_i - \mu)^2}{2\sigma^2}} \prod_{j=1}^{n-r} \left( 1 - \Phi\left(\frac{\ln T - \mu}{\sigma}\right) \right) \quad (4-17)$$

其中  $n$  等于 30， $r$  等于 26， $T$  等于 333852h， $\Phi(\cdot)$  表示正态分布函数。

通过 ALT 数据构造的似然函数更新包含现场数据信息的先验分布得到该批次智能电表可靠性模型参数的后验分布，实现现场数据与 ALT 数据的融合。

## 4.2.5 后验分布采样结果

对于如式 (4-17) 所示的形式复杂的后验分布，本文依托 Python 用于贝叶斯分析的 Pymc3 软件库，采用 NUTS 方法对后验分布进行采样以获得后验分布相关信息，重复采样 4 次，每次采样前预采样 2000 个点后，再正式开始采样 1000 个点，得到



融合两种数据后的可靠性模型参数 $\mu$ 和 $\sigma$ 的后验分布概率密度采样结果分别如图 4-3、4-4 所示:

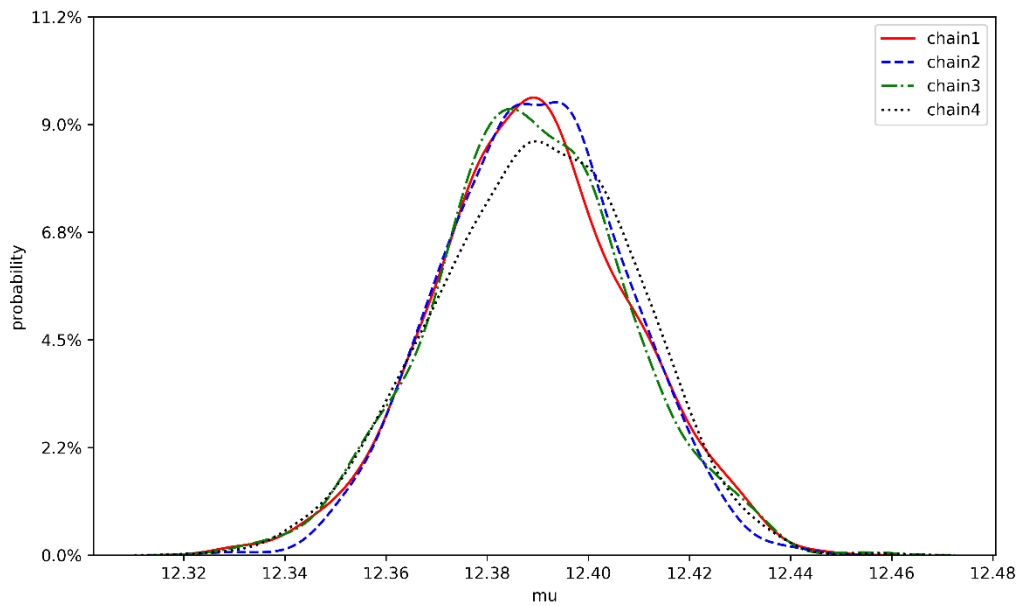


图 4-3 参数  $\mu$  后验分布概率密度的采样结果

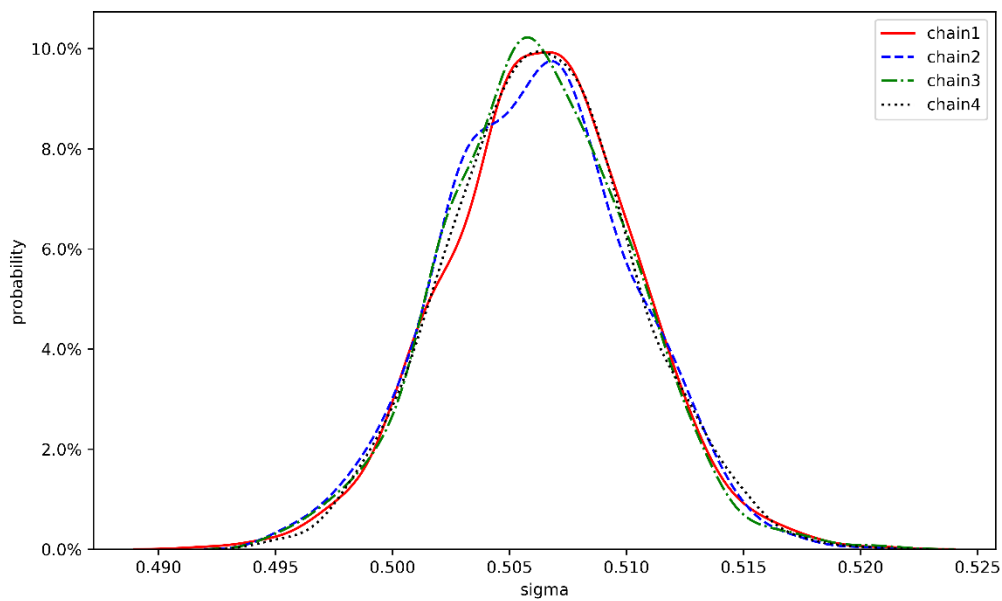


图 4-4 参数  $\sigma$  后验分布概率密度的采样结果

从图 4-3 和 4-4 可以看出, 4 次采样得到后验分布样本相对集中。为进一步说明

采样算法已收敛，图 4-5、4-6 给出了参数  $\mu$  和  $\sigma$  采样过程中的采样轨迹图。

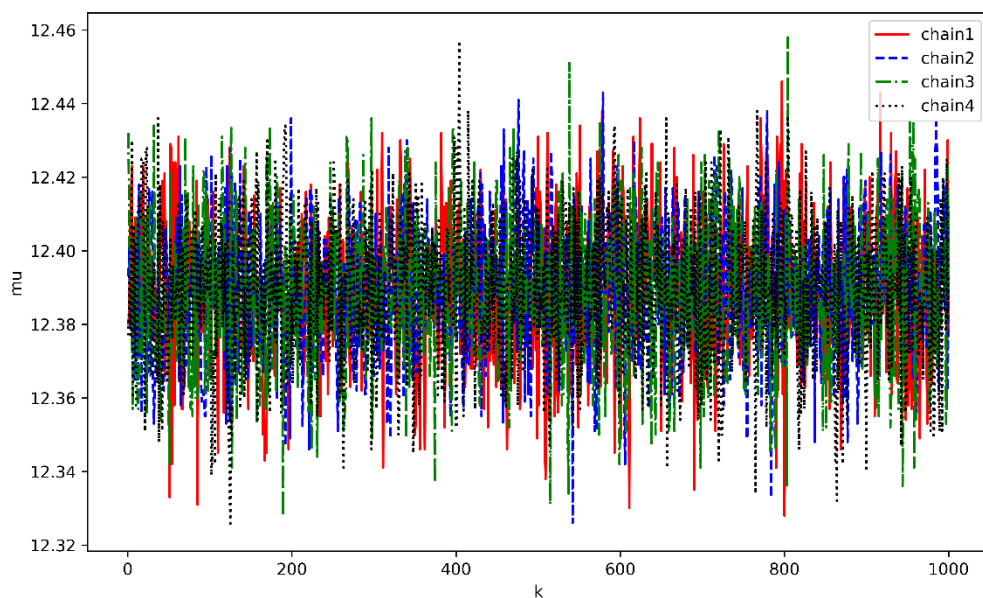


图 4-5 参数  $\mu$  后验分布采样轨迹图

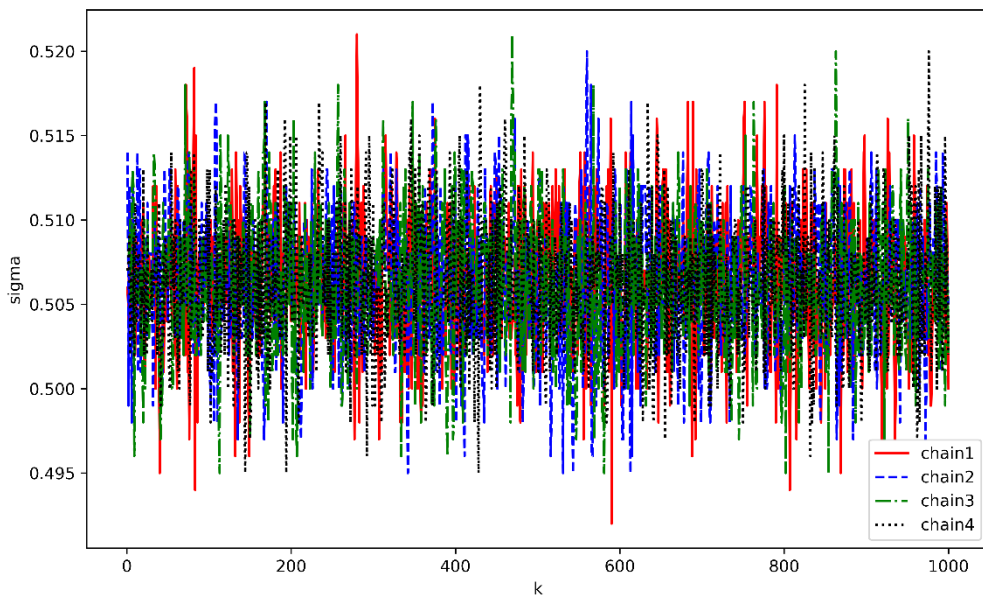


图 4-6 参数  $\sigma$  后验分布采样轨迹图

从采样轨迹图中可以看出每次采样得到的轨迹均在一定的范围内，且 4 次采样

的轨迹在轨迹图中交叉混合, 根据 Gelman-Rubin 收敛性诊断方法可知 NUTS 算法已经收敛, 生成的后验样本可信<sup>[56]</sup>。

对后验分布的采样结果进行统计分析, 相关统计量结果整理如表 4-2 所示。

表 4-2 参数后验分布统计量

参数	统计特征		后验置信区间	
	均值	方差	2.5%	97.5%
$\mu$	12.388	0.019	12.348	12.424
$\sigma$	0.506	0.004	0.498	0.513

### 4.3 智能电表可靠性评估结果分析

以后验均值为参数的点估计值得到融合两种数据源后的可靠性模型参数估计结果, 将数据融合前后的可靠度随时间变化的曲线绘制于一张图中如图 4-7 所示:

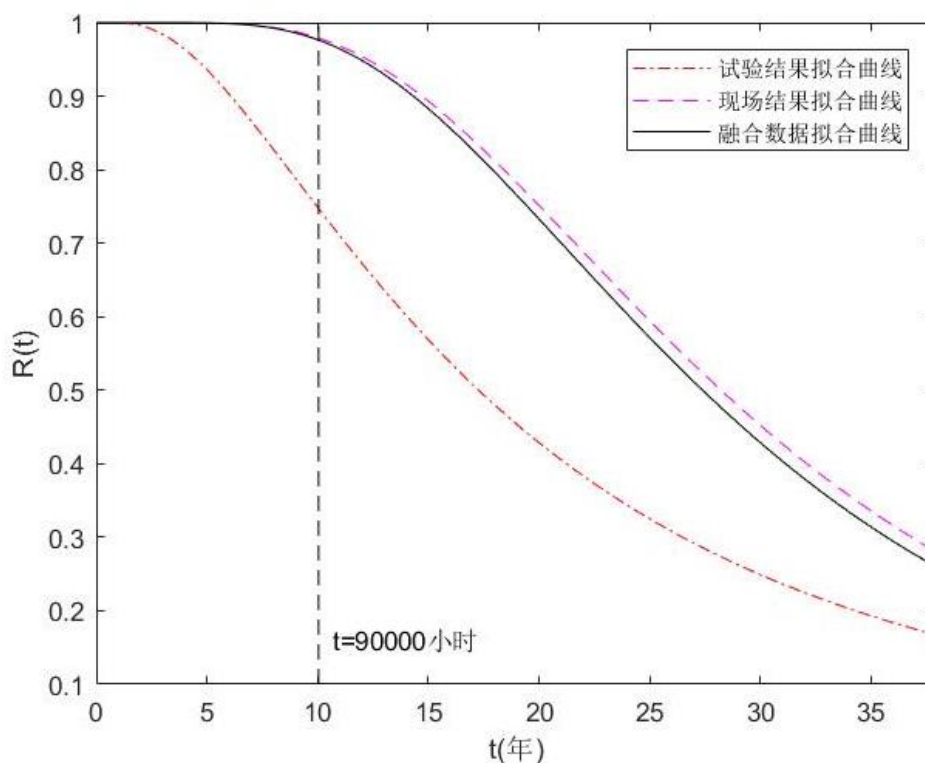


图 4-7 数据融合前后的可靠度曲线

为进一步分析融合两种数据源所得评估结果相较于单一数据源评估结果的优越性, 表 4-3、4-4 给出数据融合前后参数  $\mu$  和  $\sigma$  的 95% 置信区间估计结果及相应的区

间宽度:

表 4-3 数据融合前后参数  $\mu$  的区间估计

数据源	置信区间		区间宽度
	2.5%	97.5%	
单一现场数据源	12.174	12.722	0.548
单一 ALT 数据源	11.604	12.228	0.624
融合两种数据源	12.348	12.424	0.076

表 4-4 数据融合前后参数  $\sigma$  的区间估计

数据源	置信区间		区间宽度
	2.5%	97.5%	
单一现场数据源	0.387	0.653	0.266
单一 ALT 数据源	0.578	1.081	0.503
融合两种数据源	0.498	0.513	0.015

结合图 4-7 和表 4-3、4-4 可得到如下分析结论:

(1) 从整体来看, 依据 ALT 数据得到的智能电表可靠度远低于依据现场数据得到的可靠度, 而融合两种数据源所得的可靠度则介于两者之间且略低于依据现场数据得到的可靠度。

(2) 在不超过 10 年的时间范围内, 融合两种数据得到的可靠度结果与依据单一现场数据得到的可靠度结果几乎完全重合, 能反映实际的现场可靠性水平, 原因在于该时间段涵盖了所有的现场失效数据, 但仅涵盖了少量 ALT 失效数据, 因此该时间段内的可靠性评估结果主要由现场失效数据决定。

(3) 在 10 至 38 年这一时间范围内, 数据融合后的可靠度与依据单一现场数据得到的可靠度的差距逐渐明显, 原因在于该时间段内无现场失效数据, ALT 数据对可靠性评估结果的修正作用逐渐显著。

(4) 对比参数估计的 95% 置信区间宽度发现, 在同样的可靠性模型下, 融合两种数据源所得参数估计的区间宽度比单一数据源明显减小, 说明本文所用方法通过融合现场数据和 ALT 数据充分利用各可用信息使可靠性评估结果的准确度得到了

提高。

## 4.4 本章小结

本章研究基于贝叶斯方法的智能电表可靠性评估。在确定可靠性模型的基础上，本章基于最大熵法确定先验分布的形式，再以现场数据作为先验信息源，通过参数 Bootstrap 方法分析现场数据确定先验分布的超参数，最后基于贝叶斯方法由加速寿命试验数据修正先验分布，得到融合两种数据源的智能电表可靠性评估结果，与依据单一数据源所得的智能电表可靠性评估结果进行对比可知，融合两种数据源的智能电表可靠性评估结果能充分利用现场数据和 ALT 数据包含的可靠性信息，在反映实际现场可靠性水平的同时提高了评估结果的准确度。

## 5 总结与展望

### 5.1 本文主要内容及结论

随着智能电表的广泛应用，其可靠性逐步受到重视，相关研究也已逐步开展。然而现有针对智能电表可靠性的研究存在如下问题：少有的围绕智能电表可靠性模型优选的研究以拟合优度和秩回归方法为基础，但该方法求得的模型参数估计值、模型优选指标均受到累积失效概率多种秩估计式的影响，难以用于优选可靠性模型；智能电表的可靠性评估未能充分利用现场数据，且融合多源数据的方法以完全样本为基础，难以应用于实际非完全样本的现场数据和 ALT 数据。

本文主要针对以上问题围绕智能电表的可靠性模型优选与可靠性评估开展研究，主要内容及相应的结论如下：

首先，本文给出了三种备选可靠性模型、现场数据和加速试验数据及其预处理方法，以及基于数据融合的可靠性评估方法，作为后续章节的基础。

其次，本文结合秩回归方法与差分进化算法求解定时截尾数据下的可靠性模型参数极大似然估计，在此基础上以 AIC 准则为依据进行可靠性模型的优选。通过对比三种备选可靠性模型 AIC 值，优选了对数正态分布作为该型号智能电表的可靠性模型；通过对比秩回归法和本文方法所得结果的 AIC 值验证了秩回归法求解模型参数精度不高的问题；不同秩回归结果作为寻优初值均得到相同的结果，说明了本文所用方法不受累积失效概率秩估计式形式的影响。

最后，本文在结合最大熵法和参数 Bootstrap 方法分析现场数据获取先验分布的基础上，基于贝叶斯方法融合现场数据与 ALT 数据对智能电表的可靠性进行评估。对比单一数据源与本文所用方法的可靠度随时间变化的曲线可知，融合两种数据源的可靠性评估结果能在缺少现场数据的时段利用 ALT 数据对可靠度进行修正的同时，仍对现场数据保有较好的拟合效果；而通过对比数据融合前后参数估计结果 95% 置信区间的宽度验证了融合两种数据源能提高智能电表可靠性评估结果的准确性。

## 5.2 本文的主要创新点

本文的主要创新点如下：

（1）采用差分进化算法求解可靠性模型参数的极大似然估计，在解决定时截尾数据下可靠性模型参数极大似然估计问题难以用常规方法求解的同时，使得可靠性模型优选不再受累积失效概率估计式的影响。

（2）在基于贝叶斯方法融合现场数据和 ALT 数据时，结合最大熵法和参数 Bootstrap 方法充分挖掘现场数据包含的可靠性信息构造先验分布，使得可靠性评估结果更为合理。

## 5.3 展望

本文研究了定时截尾数据下智能电表的可靠性模型优选、融合现场数据和 ALT 数据的可靠性评估，但仍存在不足之处可做进一步的改进与拓展：

（1）本文收集的现场数据可近似处理为定时截尾数据，而受数据记录的影响，部分智能电表存在安装时间、故障时间记录不全的问题，加之投运时间存在较大差异，使得实际现场数据可能呈现左删失、右截尾的随机截尾特性，后续研究可围绕现场数据的这一特点开展。

（2）本文中的备选模型均为简单模型，但智能电表的故障是多种故障模式共同作用引起的，后续研究可以考虑将用于描述多故障模式下可靠性的混合分布加入备选模型以优选更合适的可靠性模型。

（3）本文开展 ALT 并未发现计量误差相关故障，而计量作为智能电表的关键功能应当在可靠性评估中予以考虑，目前多采用 ADT 方法单独研究计量误差引起的可靠性问题，后续研究中可考虑通过 ADT 获取计量相关故障信息，再与 ALT 数据和现场数据进行多源数据融合得到更为准确的可靠性评估结果。

## 参考文献

- [1] 中华人民共和国国家质量监督检验检疫总局. JJG596-2012 电子式交流电能表检定规程. 北京: 中国标准出版社, 2012
- [2] 董贤光, 代燕杰, 王婷婷, 崔越, 刘凯辛. 基于现场删失数据的电能表可靠性评估方法及轮换策略研究. 电测与仪表, 2021, 58(12): 175-183
- [3] 国家电网有限公司. 单相智能电能表(2020)通用技术规范. 2020
- [4] 国家电网有限公司. 三相智能电能表(2020)通用技术规范. 2020
- [5] IEEE Standards Coordinating Committee 37. IEEE Std 1413.1-2002 IEEE Guide for Selecting and Using Reliability Predictions Based on IEEE 1413. New York: The Institute of Electrical and Electronics Engineers Inc, 2003
- [6] 罗冉冉, 左嘉, 田成明, 江小强, 邹跃. 电子式电能表的可靠性评估方法研究. 电测与仪表, 2013, 50(S1): 1-6
- [7] N. Li, G. Tong, J. Yang, G. Sun, D. Han, G. Wang. Reliability Prediction Approaches For Domestic Intelligent Electric Energy Meter Based on IEC62380. In: 3<sup>rd</sup> International Conference on Environmental Science and Material Application(ESMA), Chongqing, China, 25-26 Nov. 2018, IOP, 2018: 052030
- [8] 杨洪旗, 刘少卿, 黄进永. 智能电能表的可靠性预计方法研究. 电子产品可靠性与环境试验, 2016, 34(03): 65-71
- [9] 李向锋, 宗建华, Joe Imfeld. IEC62059 标准在智能电能表可靠性预计与考核验证方法上的应用. 电测与仪表, 2010, 47(01): 75-80
- [10] 袁金灿, 马进, 王思彤, 周丽霞, 周晖. 智能电能表可靠性预计技术. 电力自动化设备, 2013, 33(07): 161-166
- [11] Z. Yang, Y. Chen, Y. Li, E. Zio, R. Kang. Smart Electricity Meter Reliability Prediction based on Accelerated Degradation Testing and Modeling. International Journal of Electrical Power & Energy Systems, 2014, 56(3):209-219
- [12] K. Lin, Y. Chen, D. Xu. Reliability assessment model considering heterogeneous population in a multiple stresses accelerated test. Reliability Engineering and



- System Safety, 2017, 165(sep.): 134-143
- [13] D. Xu, J. He, Z. Yang. Reliability prediction based on Birnbaum–Saunders model and its application to smart meter. *Annals of Operations Research*, 2019(2): 1-14
- [14] K. Lin, Y. Wang, Y. Chen. A Kernel-Density based Semi-Parametric stochastic degradation model with dependent increments. *Mechanical Systems and Signal Processing*, 2021,161(dec.): 107978
- [15] H. Zheng, Z. Yao, W. Li, X. Huang, T. Qin. Remaining Useful Life Prediction of Smart Electricity Meters Based on Particle Filter and a Four-Parameter Degradation Model. In: 2019 14th IEEE Conference on Industrial Electronics and Applications(ICEI EA). Xian, China, 19-21 Jun. 2019, IEEE, 2019: 1151-1155
- [16] 李亦非, 邹战明, 祁邦彦, 胡薇薇, 孙宇锋, 邓晓波. 基于寿命应力模型的电能表加速寿命试验研究. *现代电子技术*, 2011, 34(20): 167-170
- [17] 罗冉冉. 单相智能电能表可靠性加速寿命试验研究[硕士学位论文]. 北京: 华北电力大学, 2012
- [18] 查国清, 黄小凯, 康锐. 基于多应力加速试验方法的智能电表寿命评估. *北京航空航天大学学报*, 2015, 41(12): 2217-2224
- [19] 张景元, 何玉珠, 崔唯佳. 基于多应力退化模型的智能电表可靠寿命预估. *北京航空航天大学学报*, 2017, 43(08): 1662-1669
- [20] 辛承霖. 基于 AET 的湿热地区智能电表可靠性评估方法研究[硕士学位论文]. 杭州: 中国计量大学, 2019
- [21] Y. Li, C. Xiao, K. Luan, J. Fan, Q. Sui, F. Zhang, et al. Service life estimation of smart electricity meters using operation data. In: 2017 Chinese Automation Congress (CAC). Jinan, China, 20-22 Oct. 2017, IEEE, 2017: 4755-4759
- [22] J. Zhang, L. Yao, J. Lv, L. Ni, J. Yuan, Y. Hu. Life Prediction and Verification of Energy Meter Based on Multi-stress Influence Weibull Distribution Model. In: 2020 21st International Conference on Electronic Packaging Technology (ICEPT). Guangzhou, China, 12-15 Aug. 2020, IEEE, 2020: 1-6
- [23] W. Qiu, Q. Tang, W. Yao, Y. Qin, J. Ma. Probability Analysis for Failure Assessment of Electric Energy Metering Equipment Under Multiple Extreme Stresses. *IEEE*

- Transactions on Industrial Informatics. 2021, 17(6): 3762-3771
- [24] 杜蜀薇, 成达, 邱伟, 袁翔宇. 基于分层贝叶斯的电能表短期故障评估与预测. 电子测量与仪器学报, 2018, 32(12): 178-184
- [25] 李翰斌. 基于失效率等级的智能电能表可靠性预计应用方法研究[硕士学位论文]. 哈尔滨: 哈尔滨理工大学, 2018
- [26] 刘旭明, 唐求, 邱伟, 成达, 李宁. 基于 ZINB 层次贝叶斯智能电能表的可靠性预估. 电子测量与仪器学报, 2019, 33(07): 28-36
- [27] D. Xu, H. He, Y. Li. Operating state evaluation of smart electricity meter based on data fusion method. In: 2019 IEEE International Conference on Prognostics and Health Management(ICPHM). San Francisco, CA, 17-20 Jun. 2019, IEEE, 2019: 1-8
- [28] X. Liu, Y. Liu, X. Liu. Fusion method of metering equipment reliability assessment based on degradation information and failure data. In: 1st International Conference on Reliability Systems Engineering(IERSE). Beijing, China, 21-23 Oct. 2015, IEEE, 2015: RP0288
- [29] 田正其, 祝宇楠, 刘建, 徐晴. 基于贝叶斯的智能电能表可靠性评估方法. 计算机测量与控制, 2016, 24(10): 274-277
- [30] 李贺龙, 于海波, 何娇兰. 基于数据融合方法的智能电能表运行剩余寿命预测. 电测与仪表, 2019, 56(18): 126-133
- [31] 张家琦, 郭帅, 李国昌, 陈颖, 宋玮琼, 关慧哲. 基于多元大数据融合的智能电能表可靠性评估模型. 电测与仪表: 1-7. <http://kns.cnki.net/kcms/detail/23.1202.TH.20210303.0843.002.html>
- [32] F. Lewis, A. Butler, L. Gilbert. A unified approach to model selection using the likelihood ratio test. Methods in Ecology & Evolution, 2011, 2(2): 155-162
- [33] W. Wang, X. Ma, S. Chang, R. Kang. A Comparison of Two Accelerated Degradation Models with Temperature and Humidity as Accelerating Stresses. Applied Mechanics & Materials, 2013, 300-301: 1162-1170
- [34] A. K. Dey, D. Kundu. Discriminating Among the Log-Normal, Weibull, and Generalized Exponential Distributions. IEEE Transactions on Reliability, 2009, 58(3): 416-424

- [35] 肖勇, 胡珊珊, 许卓, 王保帅, 罗奕. 智能电表通信模块可靠性评估. 南方电网技术, 2020, 14(08): 52-57
- [36] 申桂香, 陈炳锬, 张英芝, 薛玉霞. 基于熵值-模糊综合评判的可靠性模型优选. 吉林大学学报(工学版), 2008, 38(S2): 117-121
- [37] X. Yang, M. S. Mannan. An uncertainty and sensitivity analysis of dynamic operational risk assessment model: A case study. Journal of Loss Prevention in the Process Industries, 2010, 23(2): 300-307
- [38] 王智明. 数控机床的可靠性评估与不完全预防维修及其应用[博士学位论文]. 上海: 上海交通大学, 2011
- [39] T. Emura, S. K. Shiu. Estimation and model selection for left-truncated and right-censored lifetime data with application to electric power transformers analysis. Communications in Statistics-Simulation and Computation, 2016, 45(9): 3171-3189
- [40] E. A. Nasir, R. Pan. Simulation-based Bayesian optimal ALT designs for model discrimination. Reliability Engineering & System Safety, 2015, 134(feb.): 1-9
- [41] E. Amalnerkar, T. H. Lee, W. Lim. Reliability analysis using bootstrap information criterion for small sample size response functions. Structural and Multidisciplinary Optimization, 2020, 62(6): 2901-2913
- [42] L. Zhou, R. Cao, C. Qi, R. Shi. Reliability prediction for smart meter based on Bellcore standards. In: 2012 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering(ICQR2MSE). Chengdu, China, 15-18 Jun. 2012, IEEE, 2012: 631-634
- [43] E. A. Elsayed. 可靠性工程 (第 2 版). 北京: 电子工业出版社, 2013
- [44] 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会. GB/T 17215.931-2017 电测量设备 可信性 第 311 部分: 温度和湿度加速可靠性试验. 北京: 中国质检出版社, 2017.
- [45] L. Wang, R. Pan, X. Li, T. Jiang. A Bayesian reliability evaluation method with integrated accelerated degradation testing and field information. Reliability Engineering & System Safety, 2013, 112(APR.): 38-47
- [46] E. J. Ward. A review and comparison of four commonly used Bayesian and

- maximum likelihood model selection tools. *Ecological Modelling*, 2008, 211(1-2):1-10
- [47] 荣雪琴, 刘勇, 刘昊, 卜树坡. 基于电能表剩余寿命预测的 Weibull 分布优化研究. *电测与仪表*, 2019, 56(17): 147-152
- [48] 王文焕, 杨国生, 周泽昕, 詹荣荣, 张烈, 郭鹏, 等. 基于随机截尾数据及极大似然估计的继电保护可靠性分布. *电力系统保护与控制*, 2019, 47(12): 125-131
- [49] J. He, J. Chen, X. Guan. Lifetime distribution selection for complete and censored multi-level testing data and its influence on probability of failure estimates. *Structural and Multidisciplinary Optimization*, 2020, 62(1): 1-17
- [50] 包子阳, 余继周, 杨杉. 智能优化算法及其 MATLAB 实例 (第 2 版). 电子工业出版社, 2018
- [51] 刘方亮. 核电站小样本数据贝叶斯处理方法研究[硕士学位论文]. 北京: 清华大学, 2010
- [52] 方良海. 产品可靠性评估中的多源信息融合技术研究[博士学位论文]. 合肥: 合肥工业大学, 2006
- [53] 凌光, 戴怡, 李曦. 基于二阶矩约束的最大熵先验信息解及其在数控系统可靠性评估中的应用. *中国机械工程*, 2010, 21(12): 1466-1468
- [54] B. Efron. Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, 1979, 7(1): 1-26
- [55] 杨兆军, 李洪洲, 陈传海, 李国发, 王彦鹄. 考虑定时截尾数据的数控机床可靠性 Bootstrap 区间估计方法. *机械工程学报*, 2017, 53(07): 129-137.
- [56] 郭骏宇. 多源信息下数控机床关键子系统可靠性建模与评估研究[博士学位论文]. 成都: 电子科技大学, 2019