

利用偏差数据估算电力变压器剩余寿命的方法

何 育

国网江苏省电力有限公司设备部 江苏 南京 210000

摘 要: 电力变压器作为电力系统中的关键组件,其性能和可靠性对于保障电力供应的稳定性和效率起着重要作用。鉴于电力变压器的维护计划和资本支出决策依赖于对其剩余寿命的准确估算,因此,开发一种精确的寿命预测方法对于能源公司具有重大的经济和运营意义。现有寿命估算方法常受限于历史数据的不完整性,特别是2015年之前报废设备的记录严重缺失,导致无法提供可靠的估算结果。为解决这一问题,本研究提出了一种创新的基于生存分析的电力变压器寿命估算方法。该方法首先筛选了2015年之后仍在运行的电力变压器数据集,确保了所用数据的完整性和相关性。进一步的,本研究构建了贝叶斯生存模型,该模型采用截断韦布尔分布来描述变压器的失效时间,并考虑到数据中的偏差。最后通过最大化部分似然函数,本研究对模型参数进行了精确估计。在方法验证阶段,本研究对不同电压等级的电力变压器进行了实验分析。实验结果表明,与现有方法相比,本研究所提出的贝叶斯生存模型能够提供更为精确和合理的寿命预测,证明了该方法在电力变压器寿命估算方面的有效性和优越性。

关键词: 数据偏差; 剩余寿命估算; 贝叶斯生存模型; 截断式韦布尔分布; 电力变压器

1 引言

电力系统又称电网或电力基础设施,对社会的运转和发展至关重要。作为电力系统的重要组成部分,电力变压器有助于整个电力系统的电压转换、升压和降压,确保向终端用户提供可靠、高效的电力。

电力变压器通常被认为是成本最高的投资项目,对于电力变压器,高压变电站投资约60%。巨额的投资促使电力公司想要更加准确地估算电力变压器的剩余寿命,即从投产到报废的时间。Hong 等人^[1]基于左截断和右删减的寿命数据,开发了一种基于统计的预测方法,可生成单个电力变压器剩余寿命的预测区间。Li 等人^[2]利用领域适应性进行剩余寿命预测,并提出了一种通过在特征级和语义级对齐分布的新方法。Aizpurua 等人^[3]提出了一种新型变压器状态评估方法,将不确定性建模、数据驱动的预测模型和基于模型的实验模型整合在一起,以提高预测精度并处理不确定性。此外,也有人还采用了一些深度学习方法,包括递归神经网络(RNN)^[4]及其变体,如长短期记忆(LSTM)来对变压器剩余寿命进行预测。

尽管前期研究取得了不错的成果,但数据缺失问题一直影响着研究进程。由于国家电网近十年来刚刚开始数字化转型,设备报废的历史数据并不完整,2015年之

前报废的设备记录严重缺失。因此,大量缺失的数据不足以支撑传统数据的预测结果,将传统的寿命预测方法直接应用于这些数据,势必会高估设备寿命,从而影响设备更换决策。

受到这一问题的推动,本研究提供了一种电力变压器寿命估算方法。本研究所提出的方法仅考虑2015年之后仍在运行的电力变压器寿命数据。尽管这些数据在完整性上有所保证,但依然存在偏差,原因是许多短寿命变压器的数据未被记录,这高估了设备的平均寿命。为了纠正这一偏差,本研究采用了贝叶斯生存模型,该模型利用截断韦布尔分布对寿命数据进行拟合,以更准确地反映电力变压器是否报废。在参数估计方面,本研究通过最大化部分似然函数来估计参数。这种方法允许模型在给定的先验信息和数据条件下,对参数进行更新和估计,从而提高了预测的准确性和可靠性。为了验证所提出方法的有效性,本研究对四个不同电压等级的变压器进行了实验。实验结果表明,与现有方法相比,本研究中的方法能够更合理地预测电力变压器的寿命,为电力系统的维护和资产更换决策提供了更为准确的依据。

2 模型说明

生存分析是统计学的一个分支,通常用于分析相关事件发生前的时间,这里指的是电力变压器的报废。其优点之一是能够考虑到在实际的研究场景中,并非所有电力变压器在观察期限内都会经历报废这一终点事件^[5]。生存分析的优势在于其能够适应性地处理截尾数据(Censored Data),即那些在观察期结束时尚未报废的电力变压器

作者电子邮件: yvav@foxmail.com

鸣谢

本研究得到了国家电网科技项目(批准号:1400-202356331A-1-1-ZN)的支持。

数据。通过这种方法,研究者可以对整个样本群体进行分析,而不仅仅是那些在观察期内已经报废的电力变压器,从而避免了由于数据不足导致的估计偏差。生存分析通常采用生存函数(Survival Function)、危险函数(Hazard Function)等概念来描述个体在特定时间点存活或报废的概率。这些函数为评估和预测电力变压器的寿命提供了数学基础。

2.1 标准生存模型

让 T 表示变压器的寿命,由于电力变压器的寿命较长,在有限的观测期内可能无法观测到完整的电力变压器的整个寿命周期。所以,本研究利用 $S(t)$ 表示生存函数,代表电力变压器在时间 t 之后的生存概率,除此之外,通常通过指定一个危险函数 $h(t)$ 来模拟这种情况,该危险函数是仍在运行的电力变压器在时间 t 时的瞬时死亡率。

2.2 截断韦布尔分布

应用韦布尔分布来模拟特定事件发生的时间已有很长的历史,在电力相关的其他领域也有广泛的运用^[6]。本研究采用双参数韦布尔分布,该分布由参数集 $\Theta = k, \lambda$ 表示。危险函数由 $h(t)$ 表示,生存函数用 $S(t)$ 表示,密度函数即为生存函数与危险函数的乘积,用 $f(t)$ 来进行表示。

请注意,标准韦布尔分布是在无限域上定义的,然而,电力变压器的寿命是有限的(通常不会超过其技术寿命 L),并以固定的周期(每年)来衡量。因此,我们首先截断韦布尔分布,使存活函数为

$$S_r(t) = \frac{S(t) - S(L)}{1 - S(L)} = \frac{\left[\exp\left[-\left(\frac{t}{\lambda}\right)^k\right] - \exp\left[-\left(\frac{L}{\lambda}\right)^k\right] \right]}{\left[1 - \exp\left[-\left(\frac{L}{\lambda}\right)^k\right] \right]} \quad (1)$$

$S(0) = 1$ 和 $S(L) = 0$ 。然后,我们将密度函数离散化:

$$f_r(t) = S_r(t-1) - S_r(t) = \frac{\left[\exp\left[-\left(\frac{t-1}{\lambda}\right)^k\right] - \exp\left[-\left(\frac{t}{\lambda}\right)^k\right] \right]}{\left[1 - \exp\left[-\left(\frac{L}{\lambda}\right)^k\right] \right]} \quad (2)$$

因此,电力变压器的寿命遵循离散的韦布尔分布。如有需要,可通过将速率参数指定为:

$$\log(\lambda) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (3)$$

2.3 贝叶斯扩展

公式(2)和(3)存在数据缺失问题,即只有2015年后报废的电力变压器的报废时间是准确的,会导致一定的数据偏差。例如,1980年有9台电力变压器投入

运行,其中8台已报废,但报废时间不详。数据中只记录了1台电力变压器,且仍在运行。直接使用这些数据会导致错误的结论,即这类电力变压器可以存活很长时间,从而使寿命估算出现偏差。

为了解决这个问题,数据中只对2015年后有精确记录的电力变压器进行采样。因此,应采用贝叶斯方法修改寿命分布。在不失一般性的前提下,让 O 表示记录完整的日历年,即我们数据集中的2015年。则生存函数变为

$$S_r(t) = \Pr(T > t | T \geq O - C) = \frac{S_r(t)}{S_r(O)} = \frac{\exp\left[-\left(\frac{t}{\lambda}\right)^k\right] - \exp\left[-\left(\frac{L}{\lambda}\right)^k\right]}{\exp\left[-\left(\frac{O}{\lambda}\right)^k\right] - \exp\left[-\left(\frac{L}{\lambda}\right)^k\right]} \quad (4)$$

密度函数改变为: $f_r(t) = S_r(t-1) - S_r(t)$ 。

3 参数估计

3.1 似然函数

假设研究中有 N 台电力变压器,每台电力变压器都被观察到日历年年底 E 。则第 i 个电力变压器用 t_i, δ_i 来表示,其中 $\delta_i = 1$ 表示电力变压器被观测到报废, $\delta_i = 0$ 表示电力变压器被观测到正常运行。如果 $\delta_i = 1$,则 $t_i = T_i$,否则 $t_i = O_i - C_i$ 。

因此,电力变压器停止工作的可能性为 $f_r(t_i)$,即 $S_r(t-1) - S_r(t)$,而在研究期间存活的变压器的可能性为 $S_r(t_i)$,即公式(4)。

将这两种情况和数据中的所有个体结合起来,得出的可能性表达式被规定为 $l(k, \lambda) = \prod_{i=1}^N f_r(t_i)^{\delta_i} S_r(t_i)^{1-\delta_i}$,接着将参数 $l(k, \lambda)$ 通过极大似然估计进行参数估计。

3.2 相关的量

对于所有电力变压器,预期寿命利用 $E[T]$ 进行表示,即求出 T 的期望来表示,对于运行了 Z 年的电力变压器,条件预期寿命利用 $E[T | T > Z]$ 表示,即在运行 Z 年的条件下 T 期望,那么预期剩余寿命就可以利用条件预期寿命减去已经运行的时间 Z 进行计算,即 $E[T | T > Z] - Z$ 。

为了更好地规划维护和调整资本支出,以及未来运营地计划,企业还希望计算出未来某年可能报废的电力变压器数量,将该电力变压器数量设为 Y 。预期报废电力变压器的数量用 $EN(Y)$ 表示, $EN(Y)$ 就是预期剩余寿命在 $Y-E-1$ 和 $Y-E$ 之间的电力变压器数量。

$$EN = \sum_{i=1}^N g(Y - E - 1 < ERL_i \leq Y - E) \quad (5)$$

其中 $g(TRUE) = 1, g(FALSE) = 0$ 。

4 案例研究

4.1 数据说明

我们对三个不同电压等级中台数最多的交流电力变压器应用了所提出的方法。数据来自国家电网华东分部，观测期截至 2022 年。2021 年之前所获得的数据作为训练样本。相关的统计量如表1所示。

表1 数据

电压	观察次数	观测报废次数	最长寿命	最低寿命	技术寿命
500 千伏	358	46	47	8	40
220 千伏	1086	167	43	2	36
110 千伏	3237	391	42	3	35

4.2 结果和讨论

为了验证该方法，我们采用了截断韦布尔分布的传统生存模型进行比较。参数估计也是通过极大似然估计进行的，在 $l(k, \lambda)$ 中， $f_r(t_i)$ 由 $f(t_i)$ 代替， $S_r(t_i)$ 由 $S(t_i)$ 代替。结果见表2。

表2 参数估计结果

电压	截断法			贝叶斯法		
	ν	λ	预期寿命	ν	λ	预期寿命
500 千伏	13.71	45.71	37.65	3.86	47.35	31.30
220 千伏	7.85	46.52	32.29	4.76	38.15	29.15
110 千伏	6.91	34.32	29.85	4.57	34.25	27.61

从表 2 中可以发现，传统方法明显高估了电力变压器的寿命，有着较大的偏差。这两组参数分别用于估算电力变压器的条件剩余寿命，并进一步用于预测 2022 年的预期报废数量。图 1 比较了预测值和实际值，其中我们提出的方法叫做Bayesian，用黄色表示；而与其对比的是传统的Truncated方法，用橙色表示；真实值Real Value用绿色表示。

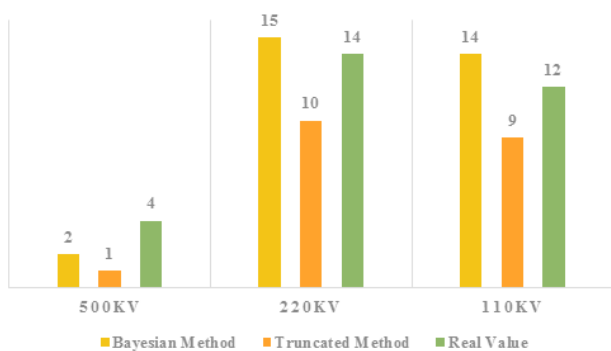


图1 2022年三种电压等级下报废电力变压器的预测数量与实际数量的比较

而根据图 1 显示，本研究提出的方法能更加准确地

预测电力变压器的预期报废量。

5 结论

在电力系统维护领域，准确预测电力变压器的寿命对于资产管理和成本效益分析至关重要。然而，历史数据的不完整性往往导致传统寿命预测模型存在固有的偏差。为了解决这一问题，本研究提出了一种新的方法和模型，该方法通过截断韦布尔分布对生存模型进行调整，以适应数据缺失的现实情况。

具体而言，本研究所提出的方法首先识别并剔除了因数据缺失导致偏差的样本，然后采用截断韦布尔分布对模型进行校准。在此基础上，进一步引入贝叶斯生存模型，对模型参数进行估计和更新。为了实验的严谨性，本研究在三个不同电压等级的交流电力变压器上进行了实验。实验结果表明，与传统方法相比，本研究的方法能够提供更为合理的寿命估计，并且能够更准确地预测预期电力变压器报废数量。此外，本研究的方法还具有扩展性，可以通过引入额外的预测变量来丰富模型，进一步提高预测精度。

参考文献

- [1] Hong, Y., Meeker, W.Q., McCalley, J.D. (2009) Prediction of remaining life of power transformers based on left truncated and right censored lifetime data. The Annals of Applied Statistics, 3(2): 857-879.
- [2] Li, X., Li, J., Zuo, L., Zhu, L., Shen, H. T. (2022) Domain adaptive remaining useful life prediction with transformer. IEEE Transactions on Instrumentation and Measurement, 71: 1-13.
- [3] Aizpurua, J. I., McArthur, S. D., Stewart, B. G., Lambert, B., Cross, J. G., Catterson, V. M. (2018) Adaptive power transformer lifetime predictions through machine learning and uncertainty modeling in nuclear power plants. IEEE Transactions on Industrial Electronics, 66(6): 4726-4737.
- [4] 邓斌, 张楠, 王江, 葛磊蛟. (2022). 基于LTC-RNN模型的中长期电力负荷预测方法. 天津大学学报(自然科学与工程技术版), 55(10), 1026-1033.
- [5] Allison, P.D. (2010) Survival Analysis Using SAS: A Practical Guide. Sas Institute.
- [6] 王锐, 杨帆, 袁静, 徐新平. (2019). 基于威布尔分布和极大似然法的智能电能表寿命预测方法研究. 计量学报, 40(6A), 125-129.