

[状态评估]

数据驱动的油浸式变压器评估与寿命预测

文/王璨¹ 史家云² 张赵阳³ 王巍³ 张亮¹ 杨龙⁴ 李智威³

1. 国网湖北省电力有限公司武汉供电公司; 2. 国网湖北省电力有限公司荆州供电公司;
3. 国网湖北省电力有限公司经济技术研究院; 4. 国网湖北省电力有限公司孝感供电公司

在电力变压器的故障诊断及寿命预测领域，结合双向长短期记忆网络（Bi-LSTM）与基于Weibull分布的统计方法，以提高对油浸式变压器状态评估的准确性，并对其剩余寿命进行预测。以某电网公司2007—2016年间大型油浸式电力变压器的实际数据为例，实现变压器的寿命预测。最终，通过仿真试验验证所提出方法的有效性。

电力变压器作为电能高效转换与传输的核心设备，构成电力系统中至关重要的节点型电气设备。它们的安全运行，对于整个电网的稳定性和可靠性具有至关重要的影响。然而，电力变压器在日常运行中，绝缘老化、线圈形变和套管失效等多种因素均可能潜藏故障风险，成为其稳定运行的隐患。因此，电力变压器的运行状态监测、状态评估以及寿命预测，一直是国内外电力行业和相关研究领域的热点问题^[1-3]。

与此同时，传感技术、人工智能与物联网技术的快速发展，极大地丰富了对设备健康状态的感知手段。这些技术的飞速发展，为设备运行状态监测大数据的获取开辟更多途径。剩余寿命（RUL）预测，作为连接系统运行状态信息感知与基于运行状态实现个性化精准健康管理的关键环节，其主要方法包括失效机理分析方法、数据驱动方法，以及机理模型与数据混合驱动方法。鉴于工程设备固有的复杂性以及任务与运行环境的多样性，其健康状态的演化规律往往难以通过物理机理建模来准确描述，或者获取失效机理模型的成本过高。因此，数据驱动RUL预测技术已经成为可靠性工程与自动化技术领域研究的前沿^[4-7]。

设备故障状态评估与寿命预测方法

随着先进传感技术的迅猛发展，工程设备健康状态的感知手段日益丰富，为设备运行监测大数据的获取提供更多可能性。因此，数据驱动技术在剩余寿命预测方面持续进步，针对大数据环境下随机退化设备的剩余寿命预测问题，引起学者们的广泛关注。数据驱动的设备寿命预测方法主要分为两大类：一类是基于机器学习的方法，另一类是基于统计数据的方法。

1. 基于机器学习的设备故障状态评估

近年来，智能算法与机器学习算法的快速发展，促使相关学者将其应用于电力领域，成功构建多种设备故障智能诊断与识别模型，显著提升设备故障状态评估的精度与可靠性。文献[8]提出利用油中气体分析进行充油电力变压器故障诊断和评估的人工神经网络（ANN）方法。该方法采用两步ANN方法来检测是否存在纤维素故障，这种方法能够获得较高的故障评估准确性。

文献[9]通过分析基于多分类最小二乘支持向量机（LS-SVM），提出一种适用于电力变压器故障评估的方法，利用粒子群算法搜索得到最优化LS-SVM评估参数，通过构造多个LS-SVM2分类器实

现变压器故障的多分类评估。文献[10]提出一种基于贝叶斯网络的变压器故障评估方法,该方法利用贝叶斯网络处理不确定性问题及不完备数据的能力,将变压器的运行历史数据纳入评估数据中,从而使得评估结果更加接近实际且过程具有可解释性。这一方法已在变压器可靠性分析与全寿命周期管理中得到广泛应用,贝叶斯网络不仅用于故障概率估算和维护保养策略制定,还涉及故障原因和频率分析、故障诊断方法等。

尽管基于ANN、SVM和贝叶斯网络的变压器故障评估方法具有一定的有效性,但它们各自存在局限性。例如,ANN在训练阶段容易陷入局部最优解,导致过拟合;SVM在处理大规模数据集时面临挑战;而贝叶斯网络虽然在处理不确定性问题上表现出色,但在表达变压器故障表象与故障模式的复杂关系方面仍有限制。本文使用的Bi-LSTM模型,其前后向网络的输入信息相同但传递方向相反,可以达到对整个输入序列关联的目的。Bi-LSTM模型独特的前后向结构,可以有效提高变压器故障状态评估的精度。

2. 基于统计数据的设备寿命预测

传统设备寿命预测方法基于统计数据,主要对设备失效时间进行分析,常用模型涵盖指数分布、Weibull分布及Gamma分布等。通过构造寿命 T 的分布函数,由此设备在 t 时刻的剩余寿命为 $T-t|T>t, z$,其中 z 是该类设备的事件数据集(主要指失效时间数据),通过分布拟合得到寿命的概率分布,再通过随机变量的关系实现剩余寿命预测^[11]。

随着信息技术和传感器技术的发展,通过设备性能退化变量的监测数据,建立描述设备性能演化过程的随机模型,以概率统计理论为基础,在随机模型框架下建模性能退化变量演变规律,以概率分布的形式预测设备寿命^[12]。常用的随机模型包括Wiener过程、Gamma过程、隐Markov过程和逆高斯过程等。通过监测数据实现模型参数的估计,求解所建立的随机退化过程首次失效阈值时间的概率分布实现RUL预测。Gamma过程和逆高斯过程属于单调随机过程,它们仅适用于描述单调退化现象,如磨损、疲劳裂纹增长等。Wiener过程是由Brownian运动驱动的一类扩散过程,其增量独立且为高斯分布,适合刻画非单调退化过程,在退化测

量信号的建模上更灵活。

基于概率统计理论的统计数据方法,通过随机模型对监测数据进行建模,从而推断剩余寿命,得出剩余寿命的概率分布。这种方法在量化剩余寿命预测的不确定性方面具有天然优势,并且由于随机模型参数与设备退化失效过程紧密相关,使得模型具有较强的可解释性。

基于Bi-LSTM油浸式变压器评估

油中溶解气体能够迅速地检测出变压器的故障类型,并且这一方法无需在停电情况下进行,基本不会受外界影响。利用油中溶解气体和智能算法融合来进行变压器故障分析,能提高故障评估的正确率。

1. 变压器故障特征构造与预处理

变压器绝缘系统的主要绝缘材料是绝缘油。该油会在温度变化和放电影响下发生老化,并通过离子反应或氧化反应产生氢气(H_2)、甲烷(CH_4)、乙烷(C_2H_6)、乙烯(C_2H_4)和乙炔(C_2H_2)等气体。当气体分解过多时,会导致绝缘油的电气性能退化,进而引起变压器运行状态异常和使用寿命缩短。

由于特征气体含量的分散性,相同环境中不同气体的含量可能存在较大差异。相同气体在不同环境下含量差异显著。因此,将气体原始浓度值直接作为输入可能会导致算法模型误差增加并降低稳定性。通过变压器气体在线监测装置报警信息及油中气体色谱分析结果,本文对5种特征气体含量进行了规范化处理,并重新构造了对应的特征参量。令 P_{sum} 为5种气体总含量, PTH 为除 H_2 外4种烃类气体总含量,取 H_2 的相对含量为 $PH_2\%=PH_2/P_{sum}$,其余气体相对含量依次为 $PCH_4\%=PCH_4/PTH$, $PC_2H_6\%=PC_2H_6/PTH$, $PC_2H_4\%=PC_2H_4/PTH$, $PC_2H_2\%=PC_2H_2/PTH$,最终形成特征序列 $[H_2\%\rightarrow CH_4\%\rightarrow C_2H_6\%\rightarrow C_2H_4\%\rightarrow C_2H_2\%]$ 。

在故障评估前需要对监测数据进行预处理。首先,重新构造气体数据以获取每种特征气体的相对含量;接着,采用Z-score标准化方法统一特征数据尺度;最后,根据特征状态转移序列和网络模型确定气体数据的输入方式。

2. 变压器故障评估模型构建

经过上一节重新构造后的特征气体序列具有形

式上的时序结构。循环神经网络能够较好地适应此类数据的学习需求。变压器故障评估在深度学习领域可视为分类问题，输入过多相同作用的特征气体冗余数据会导致分类模型的过拟合，降低模型的泛化能力。因此，引入门控机制的长短期记忆网络（LSTM）作为变压器故障评估的计算框架，其网络单元内部结构如图1所示。

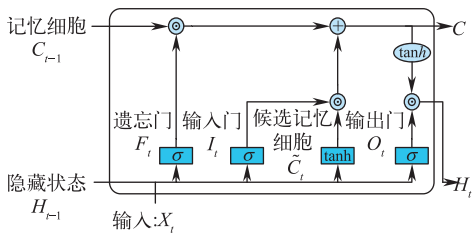


图1 LSTM网络单元内部结构

LSTM网络通过具有学习功能的门来自适应监测数据特征数量的变化。其能够自动调节信息的流动，以应对特征气体的累积作用。同时，还可以灵活地调整长期和短期特征之间的依赖关系，从而有效地挖掘特征气体之间的作用强度。在网络模型完成对特征气体序列的遍历后，会将多个特征的最终作用状态保存在隐藏层中，并最终利用全连接层进行计算，得出网络的预测结果。

变压器故障特征之间的作用并非一定是单向传播的，可能是相互作用和影响的多向互动关系。然而，LSTM作为单向结构的神经网络，在处理变压器故障评估时存在局限性。因此，使用结构更复杂的Bi-LSTM网络是一种更好的选择，可进一步挖掘特征气体与设备故障之间的关联性。

Bi-LSTM在原有LSTM基础上通过增加从后向前传递信息的网络单元来灵活处理特征气体序列，网络中每个特征在当前时刻的输入状态取决于该时刻之前和之后输入的特征子序列。故障评估流程分为两个阶段：首先，依据故障机理初步确定设备运行中的特征作用关系及初始特征状态转移序列；其次，将设备监测数据依据特征序列输入Bi-LSTM故障评估网络，通过LSTM单元双向传播隐藏状态，累积作用信息，连接形成最终状态矩阵，再经全连接网络层融合计算，映射特征作用状态为设备状态类型，得出故障评估结果。

3. 实验结果与算法对比

本文对某油浸式变压器的预防性试验结果中得

到120组油中溶解气体含量数据分析。选取重新构造的特征参量（ H_2 ， CH_4 ， C_2H_6 ， C_2H_4 ， C_2H_2 ）的相对含量作为判断变压器运行状态的主要特征。

算法模型选用交叉熵函数衡量预测类别与实际状态间的误差，并以预测正确率（正确样本数占总样本数的比例）作为评估指标。通过4折交叉验证比较了不同超参数下Bi-LSTM故障评估网络在训练数据集上的表现效果，最终确定的超参数组合，见表1。

表1 Bi-LSTM故障评估网络超参数组合

Bi-LSTM序列模型	故障评估准确率/%
$H_2\% \rightarrow CH_4\% \rightarrow C_2H_6\% \rightarrow C_2H_4\% \rightarrow C_2H_2\%$	94.6
$C_2H_2\% \rightarrow C_2H_4\% \rightarrow C_2H_6\% \rightarrow CH_4\% \rightarrow H_2\%$	91.0
$C_2H_4\% \rightarrow H_2\% \rightarrow C_2H_6\% \rightarrow C_2H_2\% \rightarrow CH_4\%$	89.3

模型经500次迭代训练后，训练集和测试集的准确率分别达到96.3%和94.6%。为了体现出Bi-LSTM算法的优势，实验还对粒子群优化的支持向量机（PSO-SVM）、反向传播神经网络（BPNet）与LSTM评估算法进行了效果对比。4种算法的测试效果对比见表2。可以看出Bi-LSTM对比其他算法具有更高的故障评估准确率。

表2 不同算法故障评估效果对比

故障评估算法	PSO-SVM	BPNet	LSTM	Bi-LSTM
准确率/（%）	82.9	88.7	91.2	94.6

为体现本文所重新构建的特征气体序列对变压器故障评估的积极作用，实验选择另两种特征序列，使用Bi-LSTM评估模型进行故障评估效果对比，对比结果见表3。云南电网创新申请的多尺度多通道CNN-LSTM模型专利显示，结合CNN和LSTM的深度学习模型在变压器绕组变形的检测中具有显著效果。这表明，将深度学习技术应用于电力变压器的故障评估，特别是结合变压器绝缘油产气机理确定的特征状态转移序列，能够有效提高Bi-LSTM模型对故障的评估能力。表3展示了在不同特征序列下，Bi-LSTM故障评估算法与其他算法相比的性能对比。

表3 不同特征序列下的Bi-LSTM故障评估算法效果对比

学习率	批样本数量	随机失活比例	权重衰减系数	模型深度	模型宽度
1×10^{-3}	64	0.2	1×10^{-8}	2	32

油浸式变压器故障预测

在长期的可靠性研究实践中，研究人员发现油浸式变压器的失效率与时间存在一定的关系，遵循所谓的“浴盆曲线”规律，分为三个阶段：早期故障、随机故障和损耗故障，如图2所示。早期故障主要由设计和生产上的不足引起，其特点是在初期故障率较高，但随着时间的推移，故障率迅速下降。随机故障阶段通常持续时间较长，故障率趋于稳定，可以用指数分布来描述。而损耗故障阶段，随着时间的推移，故障率迅速增加，主要原因是设备的磨损和老化。Weibull分布是可靠性分析和寿命检验的理论基础，由于能描述“浴盆曲线”的各个阶段，被广泛应用到油浸式变压器寿命预测中。Weibull分布函数及其对应概率密度函数和失效率函数分别为

$$F(x)=1-\exp[-(\frac{x-\gamma}{\alpha})^{\beta}],x\geq\lambda$$
$$f(x)=\frac{\beta}{\alpha}(\frac{x-\gamma}{\alpha})^{\beta-1}\exp[-(\frac{x-\gamma}{\alpha})^{\beta}],x\geq\gamma$$
$$r(x)=\frac{\beta}{\alpha}(\frac{x-\gamma}{\alpha})^{\beta-1}$$

式中， α 为尺度参数； β 为形状参数； γ 为位置参数。

需要注意的是，Weibull分布还需要满足课题具备随机性条件。

设备故障发生事件是一个随机过程，变压器故障率很难用一个精确函数来模拟，实践中通常采用 $\rho(t)=F_x/T_x$ 来表示，其中 F_x 表示一段时间间隔内变压器故障总数， T_x 表示总运行时间间隔。根据油浸式变压器故障率与相应的运行年限统计数据对Weibull分布曲线进行拟合，可以得到尺度参数 α ，形状参数 β 和位置参数 γ ，其主要流程如图3所示。

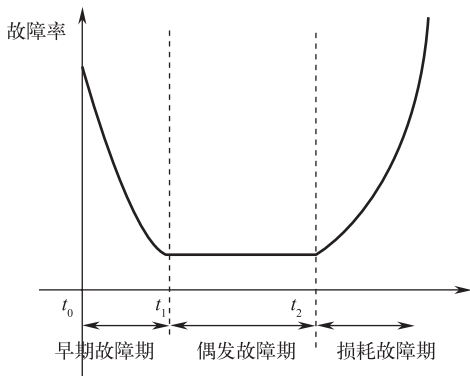


图2 “浴盆曲线”规律

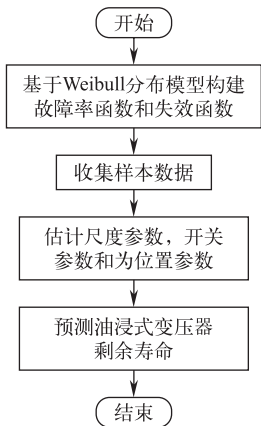


图3 基于Weibull分布的变压器寿命预测步骤

某电网安装的大型油浸式电力变压器的统计数据见表4。对表4进行数据分析得到故障率与时间的拟合曲线如图4所示。该时段内变压器总运行时间为 $T_i=5\ 200$ ，变压器故障总数为， $F_i=5\ 200$ 进而根据最小二乘法得到 k 值，再将 k 值代入故障率函数得到变压器故障率曲线，如图5所示。

表4 某电网公司变压器故障统计

安装年份	安装台数	故障总台数	寿命/年									
			1	2	3	4	5	6	7	8	9	10
2007	321	12	1	4	0	1	1	0	2	2	1	1
2008	336	6	2	1	0	0	1	0	0	1	1	0
2009	372	11	1	2	1	0	1	2	2	0	1	2
2010	408	7	1	0	2	0	2	0	1	1	0	0
2011	472	13	3	3	2	1	1	2	1	0	0	0
2012	486	8	1	1	2	1	1	1	1	0	0	0
2013	493	6	2	0	1	1	2	0	0	0	0	0
2014	507	7	2	1	2	1	0	1	0	0	0	0
2015	529	8	2	3	1	2	0	0	0	0	0	0
2016	543	6	1	2	2	1	0	0	0	0	0	0

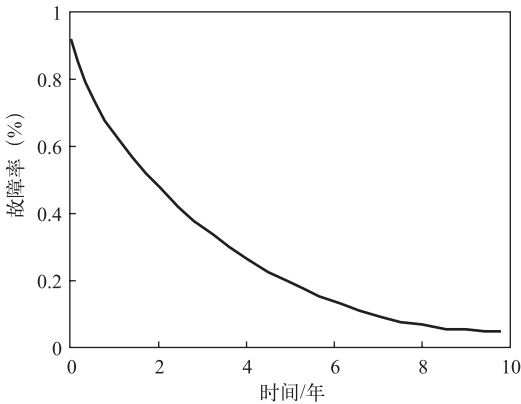


图4 变压器故障率与时间的关系曲线

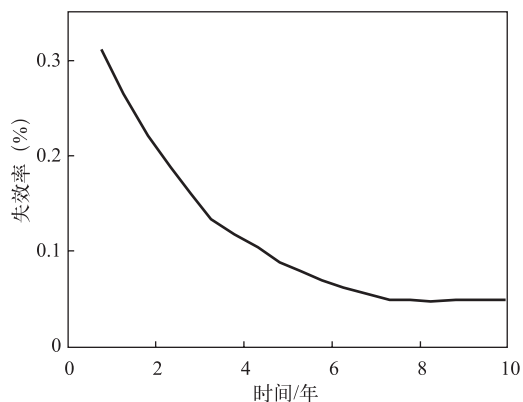


图5 基于Weibull分布的变压器故障率曲线

结束语

本研究综合运用双向长短期记忆网络(BiLSTM)与基于Weibull分布的统计方法,对油浸式变压器的状态进行评估并预测其剩余寿命。首先,本研究系统总结基于数据驱动的设备故障状态评估与剩余寿命预测技术,并分别从机器学习和统计学两个角度进行了详细分类阐述。其次,依据变压器的运行原理,确定特征状态转移序列。国家电网华东分部申请一项关于油浸式变压器状态评估的专利,该专利通过深度学习技术,特别是BiLSTM模型,对变压器的快速反应项目和慢速反应项目进行综合分析,以实现变压器工作状态的精确评估。此外,通过Weibull分布函数拟合油浸式变压器的使用寿命,构建了一个基于二参数Weibull分布的变压器寿命模型,实现对某电力公司大型油浸式变压器的寿命预测。最终,通过仿真试验验证所提出方法的有效性。EA

(收稿日期: 2025-05-09)

参考文献

[1] 谢毓城. 电力变压器手册[M]. 北京: 机械工业出版社, 2003.

- [2] 曾丰赞. 基于状态监测数据的核电厂设备寿命预测算法研究[D]. 北京: 清华大学, 2017.
- [3] 丁玉琴, 张乔根, 高萌, 等. 油浸式配电变压器分布式热路模型[J]. 高电压技术, 2019, 45(3): 968-974.
- [4] 司小胜, 胡昌华, 周东华. 带测量误差的非线性退化过程建模与剩余寿命估计[J]. 自动化学报, 2013, 39(5): 530-541.
- [5] 王加昌, 郑代威, 唐雷, 等. 基于机器学习的剩余使用寿命预测实证研究[J]. 计算机科学, 2022, 49(6): 937-945.
- [6] 文成林, 吕菲亚, 包哲静, 等. 基于数据驱动的微小故障诊断方法综述[J]. 自动化学报, 2016, 42(9): 1285-1299.
- [7] 韩中, 程林, 熊金泉, 等. 大数据结构化与数据驱动的复杂系统维修决策[J]. 自动化学报, 2020, 46(2): 385-396.
- [8] Zhang Y, Ding X, Liu Y, et al. An artificial neural network approach to transformer fault diagnosis[J]. IEEE Transactions on Power Delivery, 1996, 11(4): 1836-1841.
- [9] 郑含博, 王伟, 李晓纲, 等. 基于多分类最小二乘支持向量机和改进粒子群优化算法的电力变压器故障诊断方法[J]. 高电压技术, 2014, 40(11): 3424-3429.
- [10] Carita A J Q, Leite L C, Medeiros A P P, et al. Bayesian networks applied to failure diagnosis in power transformer[J]. IEEE Latin America Transactions, 2013, 11(4): 1075-1082.
- [11] 吴晓欣, 何怡刚, 段嘉璐, 等. 考虑复杂时序关联特性的Bi-LSTM变压器DGA故障诊断方法[J]. 电力自动化设备, 2020, 40(8): 184-193.
- [12] Kapur K C, Pecht M. Reliability Engineering[M]. New Jersey: John Wiley, 2014.