引用格式:方梦泓,邹阳,陈啸轩,等.基于Stacking异质集成的油纸绝缘受潮程度量化方法[J].电力科学与技术学报,2025,40(4):294-304.
Citation:FANG Menghong, ZOU Yang, CHEN Xiaoxuan, et al. A quantification method of moisture degree of oil-paper insulation based on Stacking heterogeneous integration[J]. Journal of Electric Power Science and Technology,2025,40(4):294-304.

基于Stacking异质集成的油纸绝缘受潮程度量化方法

方梦泓[1],邹　阳[1],陈啸轩[2],黄　煜[1],金　涛[1]
(1.福州大学电气与自动化学院,福建 福州 350108;2.国网泉州供电公司,福建 泉州 362000)

摘　要:针对当前变压器油纸绝缘系统受潮评估模型遴选受限,且在老化—水分协同作用下,特征指标对绝缘受潮情况的指向性较低等问题,提出基于Stacking异质集成的油纸绝缘受潮程度量化方法。首先,对老化—水分协同作用下的宏观绝缘状态变化趋势进行微观介质层面的机理分析,提取可剔除老化作用干扰的多源异构受潮特征指标,构建与水分强相关的特征集合;其次,基于Stacking的融合学习思想集成4类异质算法,建立模型的初级评估体系,并通过Optuna超参数调优框架降低输出噪声;再次,利用权重赋值后的堆栈数据训练次级评估体系,构建基于加权改进的Stacking模型,量化油纸绝缘系统的受潮程度;最后,以实测数据为例,验证所提模型在受潮评估中的有效性。该模型可为变压器油纸绝缘系统受潮评估提供借鉴。

关　键　词:油纸绝缘;Stacking模型;异质集成;受潮定量评估;频域介电谱
DOI:10.19781/j.issn.1673-9140.2025.04.026　　**中图分类号**:TM863　　**文章编号**:1673-9140(2025)04-0294-11

A quantification method of moisture degree of oil-paper insulation based on Stacking heterogeneous integration

FANG Menghong[1], ZOU Yang[1], CHEN Xiaoxuan[2], HUANG Yu[1], JIN Tao[1]
(1. School of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108, China;
2. State Grid Quanzhou Power Supply Company, Quanzhou 362000, China)

Abstract: To address the issues of the limitations of the current transformer oil-paper insulation system moisture assessment model selection and the low directivity of the feature indexes for insulation moisture assessment under the synergistic effect of aging and moisture, a quantification method of the moisture degree of oil-paper insulation based on Stacking heterogeneous integration is proposed. Firstly, the mechanism analysis is conducted about the change trend of the macro insulation state under the synergistic effect of aging and moisture synergism at the micro-medium level, and multi-source heterogeneous moisture feature indexes that can eliminate the interference of aging are extracted, forming a highly correlated feature set with moisture. Secondly, based on four types of heterogeneous algorithms that are integrated based on the fusion learning idea of Stacking, the primary assessment system of the overall model is constructed, and output noise is reduced through the Optuna hyper-parameter tuning framework. Finally, the stack data after weight assignment is utilized to train the sub-assessment system, and then a weighted and improved Stacking model is built to quantify the moisture degree of the oil-paper insulation system. Lastly, the effectiveness of the proposed model is validated in moisture assessment by taking the measured data as an example. The model can serve as a reference for assessing moisture in transformer oil-paper insulation systems.

Key words: oil-paper insulation; Stacking model; heterogeneous integration; quantitative moisture assessment; dielectric spectrum of frequency domain

收稿日期:2024-08-13;**修回日期**:2024-12-27
基金项目:国家自然科学基金(92266110,52377088)
通信作者:邹　阳(1980—),男,博士,副教授,主要从事电气系统智能化故障诊断等方面的研究;E-mail: 24001744@qq.com

为服务于“十四五”时期现代能源体系建设的主要目标，未来电力系统将着重提升其安全保障能力。油浸式变压器作为新型电力系统中传输与分配的关键枢纽之一，其油纸绝缘系统是在设备层面上的实现电网稳定运行重要基础之一[1]。然而，大规模的电力电子拓扑结构使其需要长期面对机械、过热、电磁等复杂应力环境，这将导致其绝缘性能大幅降低，进而缩短变油浸式压器的整体服役寿命。其中，水分含量的超标不仅加快了其绝缘系统劣化的速度，其作为介质损耗的主导因素还将直接影响油浸式变压器工作的可靠性[2]。因此，为给安全稳定的新型电力系统提供必要的技术支持，量化油纸绝缘受潮程度，并依据该程度辅助制定科学的检修策略具有重要的工程意义。

目前，用于评估油纸绝缘受潮状态最为直接的方法包括露点法、卡尔费休法等传统检测方法。但这些方法均存在取样困难、易损坏绝缘等不足，导致这些方法的在实际中的可行性下降[3-4]。与传统绝缘检测方法相比，基于电介质极化理论的频域介电谱法（frequency domain spectroscopy，FDS）作为一种新型无损诊断技术，因其频谱数据具有高精度、携带绝缘信息丰富等优势，逐渐成为研究热点[5]。有学者通过直接建立FDS谱线与宏观绝缘受潮状态间的映射关系来挖掘谱线中蕴含的绝缘受潮信息。文献[6]提取了复介电常数实部在$f=10^{-4}$ Hz时的变化率绝对值，将其用于油纸水分含量的定量评估。文献[7]通过引入介电模量的概念，将虚部积分定义为新的受潮特征因子，提出了有效的环氧树脂浸渍皱纹纸（epoxy resin impregnated crepe paper，ERICP）绝缘受潮状态评估方法。但这些研究均仅利用单个特征指标进行受潮评估，存在特征参量所含物理意义不明与绝缘信息单一等缺点。对此，部分学者提出利用介电弛豫模型来解析FDS测试曲线的方法，通过建模来提取多个相关电路参量，并用其来表征绝缘系统的受潮程度。文献[8]采用混合极化模型，提取了可用于表征偶极子转向极化与界面极化过程的特征指标，以探究受潮程度与绝缘介质电导和极化特性的关联性。文献[9]引入Dissado-Hill（D-H）模型，对FDS曲线进行重构，建立了多个模型特征参数与水分含量间的映射关系。文献[10]通过采用Havriliak-Negami（H-N）模型的辨识参数拟合绝缘系统中的水分含量，实现了受潮状态的定量评估。这些研究提取的特征参量数量有所增加，在进行数据分析时自由度也随之增大，对试验结果的分析也更全面和准确。然而，这些研究大多忽视了多种不同的单特征指标分别获取的受潮程度评估结果存在差异且难以统一的问题。

随着研究的不断深入，部分学者基于多元回归分析法、改进层次分析法、模糊K近邻等算法，将高纬度且多结构的特征指标进行融合，以此构建油纸绝缘微水含量综合评估体系[11-16]。然而，仅通过单一算法难以取得较高的识别精度，且模型的泛化性能也不尽如人意。

有鉴于此，本文提出基于Stacking异质集成模型的油纸绝缘受潮程度量化方法。首先，基于FDS实测数据，遴选出可剔除老化作用干扰的受潮评估特征集合；其次，引入Stacking模型的融合思想，通过考虑多类异质算法相互的匹配程度，集成最优基模型组合，突破了其他研究仅针对单一评估模型进行优化改进的瓶颈，并基于Optuna超参数调优框架，采用贝叶斯超参数优化方法，减少了模型的输出噪声；再次，在遵循K折交叉验证法以规避过拟合风险的基础上，对初级评估体系纵向堆叠后的评估结果采取赋权处理，降低传统模型存在数据特征遗漏的可能性；最后，采用实测数据验证本文所提模型的有效性。

1　受潮评估指标的选取

1.1　油浸纸试样制备

为遴选用于搭建异质集成模型的受潮程度量化指标，本研究制备不同老化状态及不同微水含量下的油浸纸试样。其中，纤维素绝缘纸选用厚度为1 mm的魏德曼高压绝缘硫酸盐木浆纸板，绝缘油为25号环烷基矿物油。其样本制备步骤可分为4步。

1）为保证试样的初始状态一致，将绝缘纸置于105 ℃/50 Pa的真空烘箱中干燥72 h；同时，依据《电工流体变压器和代替开关用的未使用过的矿物绝缘油》（GB 2536—2011）对绝缘油进行脱气、除水及过滤杂质等预处理。

2）将预处理后所得的35组样本均分为5组，并将其置于130 ℃/50 Pa的恒温真空老化箱中，分别进行为期0、10、20、30、40 d的加速热老化，以获得不同老化程度的油浸纸试样。

3）将各组不同老化程度的样本置于空气中自然受潮，采用精密电子天平，实时称取并记录目标

试样质量，以制备不同水分梯度（0.76%、1.31%、3.11%、4.28%）的油浸纸样品。

4）为确保绝缘纸板内部的空腔被绝缘油完全填充，在30 ℃/50 Pa的环境条件下，依照20∶1的油纸质量比，向盛放受潮绝缘纸试样的磨口玻璃瓶中注入预处理后的绝缘油，使油浸纸样品在真空条件下充分浸渍24 h。

1.2 FDS测试

为探究该模拟油纸绝缘系统在老化-受潮协同作用过程背后蕴含的弛豫响应机理，本文采用奥地利OMICRON公司研发的DIRANA频域介电测量仪，搭建FDS测试平台，对所制油浸纸试样进行FDS测试，该测试平台如图1所示。该试验在30 ℃的恒温条件下进行，设置仪器的输出电压峰值为200 V、测量带宽为10^{-3}～10^{3} Hz。

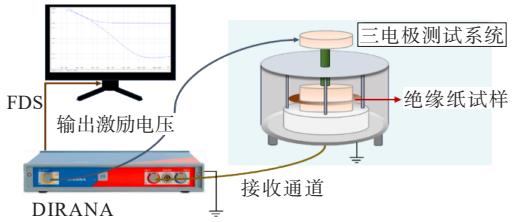


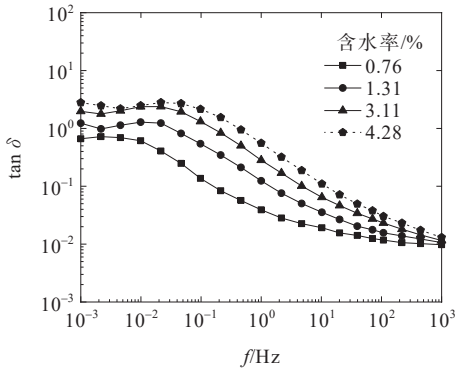
图1 FDS测试平台

Figure 1 FDS test platform

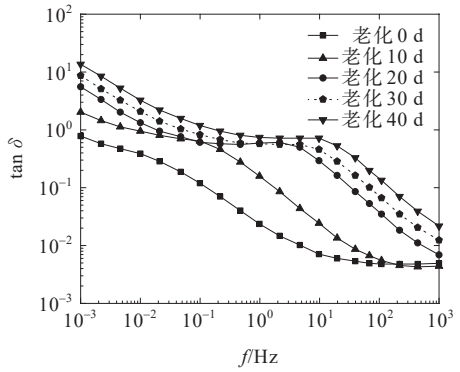
1.3 频谱分析及受潮评估指标选取

因篇幅限制，本文仅展示未老化试样于不同受潮程度下的FDS谱线与含水率为1.2%的试样在不同老化天数下的FDS谱线，不同绝缘状态下的tan δ频谱如图2所示。在图2中，f为测试频率，tan δ为介质损耗因数。

从图2中可以看出，在中、低频段范围内，同一微水含量油浸纸的tan δ随着测试频率的升高而下降，呈现出低频弥散（low frequency dispersion，LDF）效应的典型特征。该效应的微观介电机理在于：被视作载流子团簇的充油腔体内部溶解态水会被作为杂质溶剂，促进载流子的形成。当外施电场的频率较低时，载流子将通过由水分子与CHOH基团构成的渗流网络进行远距离跳跃[17]，进而产生弥散现象。从图2中还可以看出，在同一频率下，随着微水含量的增加，tan δ频谱曲线不断向上移动。油浸纸受潮程度的加深不仅造成更多的水分子参与极化过程，还会驱使绝缘系统内部的杂质离子发生解离，最终导致参与极化过程的自由电荷数目增多，大幅提升了绝缘介质的总极化强度。这一变化在宏观上表现为tan δ随微水含量的增加而变大的现象。



（a）老化时间为0 d的绝缘纸tan δ频谱



（b）含水率约为1.2%的绝缘纸的tan δ频谱

图2 不同绝缘状态下的tan δ频谱

Figure 2 tan δ frequency spectra for different insulation states

目前，已有分析不同受潮状况下tan δ频谱曲线间存在的差异、微水含量与频谱积分值协同起伏现象的研究。文献[18-19]提出了利用介损与频率轴间的积分值来表征油纸绝缘系统的受潮程度的方法。微水含量对tan δ的影响在全测量频率范围内均存在。因此，本文分别对10^{-3}～10^{-1} Hz、10^{-1}～10^{1} Hz与10^{1}～10^{3} Hz这3个频段内的tan δ频谱曲线进行分段积分，并依次定义这3个分段积分结果为特征指标X_1、X_2、X_3，即

$$X_1=\int_{10^{-3}}^{10^{-1}}\tan\delta\mathrm{d}f \tag{1}$$

$$X_2=\int_{10^{-1}}^{10^{1}}\tan\delta\mathrm{d}f \tag{2}$$

$$X_3=\int_{10^{1}}^{10^{3}}\tan\delta\mathrm{d}f \tag{3}$$

从图2中还可以看出，油纸绝缘系统的老化程度与受潮程度对tan δ的影响大致相当。在不同变量条件下，这些曲线形态的演变特性大致相仿。故本文从图2中抽选出4条曲线进行对比。老化及受

潮影响下的tan δ频谱对比如图3所示。分析图3的阴影区域可知，在老化与受潮协同作用下，中、高频段内的tan δ变化趋势将相互掩盖，难以区分每个影响因素对FDS特征参量的独立影响。

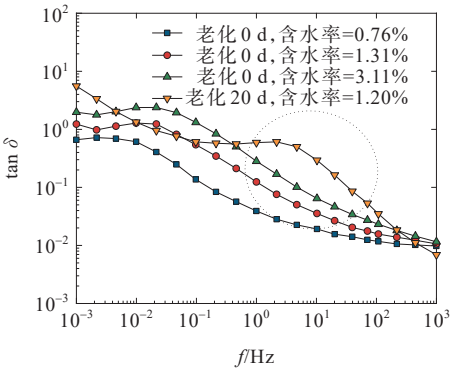


图3　老化及受潮影响下的tan δ频谱对比

Figure 3　Comparison of tan δ frequency spectra under influence of aging and moisture

同时，亦有研究表明，电导率σ_{oil}对受潮这个单一作用的敏感程度较高。当绝缘系统微水含量由1%增加至4%时，σ_{oil}数值呈几何级数的增长趋势；而当绝缘老化程度由良好状态逐渐演变至重度老化状态时，σ_{oil}数值仅呈个位数的增长趋势[20]。因此，为实现在不同老化状态下对油纸绝缘受潮程度的精准评估，本文引入电导率σ_{oil}来修正老化状态在受潮定量评估中产生的误差。由于频域介电测量仪的外施电场$\dot{E}(t)$是连续变化的，故可采用杜阿美尔(Duhamel)积分来分析总吸收电流$\dot{I}_a$随时间的变化[20]，即

$$\begin{cases}\dot{I}_a=\dfrac{\omega^2\tau^2 g}{1+\omega^2\tau^2}S\dot{E}+\mathrm{j}\dfrac{\omega\tau g}{1+\omega^2\tau^2}S\dot{E}\\ g=\dfrac{\varepsilon_0(\varepsilon_s-\varepsilon_\infty)}{\tau}\end{cases}\tag{4}$$

式中，g为等效导电率；ω为交变电压的角频率；τ为松弛极化的松弛时间；ε_0、ε_s和ε_∞分别为真空介电常数、静态相对介电常数及光频相对介电常数；S为假设极板的截面积；$\dot{E}$为外施电场强度。

介质中与交变电场同相变化的电导电流$\dot{I}_R$与瞬时充电电流$\dot{I}_\infty$的计算式分别为

$$\dot{I}_\infty=\mathrm{j}\omega\varepsilon_0\varepsilon_\infty S\dot{E}\mathrm{e}^{\mathrm{i}\omega t}\tag{5}$$

$$\dot{I}_R=\dot{I}-\dot{I}_a+\dot{I}_\infty\tag{6}$$

式(5)、(6)中，t为时间；$\dot{I}$为介质总电流。

结合式(5)、(6)，可得介质的总电流复指数形式$\dot{I}$，即

$$\dot{I}=\left(\frac{\omega^2\tau^2 g}{1+\omega^2\tau^2}\right)S\dot{E}+\mathrm{j}\omega\varepsilon_0\left(\varepsilon_0+\frac{\varepsilon_s-\varepsilon_\infty}{1+\omega^2\tau^2}\right)S\dot{E}+\dot{I}_R\tag{7}$$

引进复介电常数后，若不考虑该电导电流分量，则总电流$\dot{I}$计算式为[20]。

$$\dot{I}=S\varepsilon_0\omega(\mathrm{j}\varepsilon_p'+\varepsilon_p'')\dot{E}\tag{8}$$

式中，ε_p'和ε_p''分别为弛豫复介电常数的实部与弛豫复介电常数虚部。

结合式(7)，可将式(8)拆解为

$$\varepsilon_p'=\varepsilon_\infty+\frac{\varepsilon_s-\varepsilon_\infty}{1+\omega^2\tau^2}\tag{9}$$

$$\varepsilon_p''=\frac{(\varepsilon_s-\varepsilon_\infty)\omega\tau}{1+\omega^2\tau^2}\tag{10}$$

然而，复合油纸绝缘系统中的各个弛豫过程相互耦合，彼此间存在重叠与掩盖现象。故令$\omega=\mathrm{e}^x$，进行坐标变换，进而求解复介电常数实部的一阶微分，以凸显复合油纸绝缘系统中的弛豫特征，即

$$-\frac{\partial\varepsilon_p'(\omega)}{2\partial\ln(\omega)}=\frac{\varepsilon_p''(\omega)^2}{(\varepsilon_s-\varepsilon_\infty)}\tag{11}$$

为解耦电导率σ_{oil}，本文基于普适介电响应理论，剔除实测$\varepsilon''(\omega)$中经由式(11)获取的弛豫复介电常数虚部分量$\varepsilon_p''(\omega)$，余下部分即为电导分量$\varepsilon_g''(\omega)$，即

$$\varepsilon_g''(\omega)=\varepsilon''(\omega)-\varepsilon_p''(\omega)\tag{12}$$

则鉴于电导的频域特性计算式为

$$\varepsilon_g^*(\omega)=\frac{\sigma_{oil}}{\mathrm{j}\omega\varepsilon_0}\tag{13}$$

该$\varepsilon_g^*(\omega)$可在双对数坐标下利用最小二乘法进行拟合。令其截距为B，建立电导率σ_{oil}与截距B间的函数。故记电导率σ_{oil}为辅助特征指标X_4，其计算式为

$$X_4=\sigma_{oil}=\mathrm{e}^B\varepsilon_0\tag{14}$$

2　Stacking融合模型构建

2.1　Stacking模型原理

Stacking融合模型最早由David H. Wolpert提出，其建模方式的本质是基于集成多类异构算法对原始样本并行学习[21-23]。该模型基本框架由2部分构成：其初级评估体系一般选取多个学习能力强、复杂度高的基模型来挖掘数据信息，以期获得与原始样本相关性高的新特征集；其次级评估体系通常仅包含较为简单且泛化性强的单一模型，简称其为元模型。Stacking基本框架如图4所示。

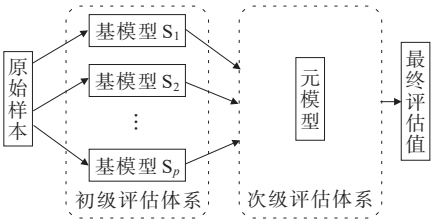


图 4 Stacking 基本框架

Figure 4 Basic framework of Stacking

2.2 基模型遴选

不同算法的区别在于观测数据空间与数据结构时角度的不同[24-27]。基模型从不同观测角度对原始杂乱无章的原始样本进行特征抽选，将输出结果作为初始特征的补充来训练元模型。因此，初级评估体系中的基模型须满足以下2方面要求：

1）各类基模型之间应具备足够的差异性及较强的学习能力，以保证新特征集的多样性和有效性。

2）各类基模型的算法原理不同，故其对不同特性数据的挖掘能力也有所不同。因此，基模型的选择还应考虑到FDS测试数据可能出现样本量偏少、样本分布不平衡与数据可能存在缺失值等特殊情况。

针对这些要求，本文最终集成支持向量机回归（support vector regression，SVR）、随机森林（random forest，RF）、K最邻近法（K-nearest neighbor，KNN）、XGBoost这4种异质算法来构建Stacking模型的初级评估体系，各算法优势及缺点见表1。

KNN是一种基于距离衡量的监督学习算法。KNN通过直接利用样本间的关系，最大程度地减少了分类过程中的误差项，且其对数据的输入形式无特殊要求[28-29]，但当数据样本存在近邻噪声或样本分布不均匀时，其计算精度显著降低。RF是一种基于Bagging的决策树集成算法。RF随机采样的数据采样方式使其具有较强的抗噪性与处理样本不平衡性的能力[30]，但其对小样本数据的预测效果较差。SVR能较好地处理非线性数据和异常值。

表 1 各算法优势及缺点

Table 1 Advantages and disadvantages of each algorithm

算法	优势	缺点
KNN	模型简单	无抗噪性
RF	有效处理样本不平衡的问题	不适用小样本数据
SVR	样本数量要求低	数据缺失情况敏感
XGBoost	拓展性强，可并行学习	效率低，内存损耗大

SVR利用核技巧和软间隔优化的手段，增强了算法的鲁棒性，且其对样本数量的要求不高，能有效解决FDS数据样本量较少的问题[31]，但需要调整的参数较多，这导致SVR算法对数值缺损情况的处理效果不佳。作为树集成学习模型的代表，XGBoost的可拓展性强，XGBoost可根据需求对参数进行调整与优化；同时，其将缺失数据泛化为稀疏矩阵，有效避免了节点分裂过程中缺失值的干扰[32]。但XGBoost的并行运算与即时缓存的工作方式使得其整体运行效率较低，造成了大量冗余的内存损耗。

综上所述，本文遴选的4种基模型均有严谨的理论支撑，算法之间相互异质，能甄别并保留FDS实测数据的有效特征，为分析油纸绝缘系统受潮程度与特征指标间的函数关系提供了必要的基础模型。

2.3 元模型的选择

元模型的本质是将基模型所抽取出的新特征集进行泛化与再学习。因此，元模型更注重模型对数据特征的总结归纳与防止过拟合现象发生的能力。核岭回归（kernel ridge regression，KRR）是一种将岭回归与核技术相结合的回归算法[33-34]，该算法通过添加L_2范数的正则化项与核函数来保证模型的泛化性，并以这种方式解决了最小二乘法中存在的多重共线性问题，为处理庞杂的数据提供了有效途径。故本文拟采用KRR算法作为Stacking评估模型的元模型。

2.4 超参数调优

与模型内部配置的参数自行根据数据进行驱动调整不同，超参数旨在模型运行前，通过手动设置来对模型的收敛性、泛化性、运行速度等关键性能进行调优。Stacking融合模型的迭代特性将极大地放大超参数对模型精度的影响。为解决该问题，本文基于Optuna超参数调优框架，采用贝叶斯超参数优化方法，使构建的模型在效果和效率上均取达到较好的效果。

3 油纸绝缘受潮评估Stacking模型

3.1 基于传统Stacking的油纸绝缘受潮评估模型

为全面挖掘FDS的弛豫响应信息，精准量化变压器油纸绝缘系统的受潮状态，本文将SVR、RF、KNN与XGBoost这4种异质算法作为基模型，将

KRR作为元模型，提出一种基于Stacking异质集成的油纸绝缘受潮状况定量评估模型。以介质损耗因数 $\tan\delta$ 在指定频段内对不同频率的积分值的4个特征指标作为输入变量，以微水含量作为输出变量，对模型进行训练。Stacking模型训练流程如图5所示。其步骤可分为6步。

1）从数据库中取若干样本数据，对其进行归一化处理，归一化后的样本数据范围为[0,1]，即

$$a'_{ij}=\frac{a_{ij}-\min\limits_{i}\left\{a_{ij}\right\}}{\max\limits_{i}\left\{a_{ij}\right\}-\min\limits_{i}\left\{a_{ij}\right\}} \tag{15}$$

式中，a_{ij} 为第 i 份油浸纸试样第 j 个特征量；a_{ij}' 为第 i 份油浸纸试样第 j 个特征量规范值。

2）利用归一化处理后的样本数据，构建原始数据集 $D=\{(\boldsymbol{x}_t,y_t)\}(t=1,2,\cdots,T)$。其中，$T$ 为该原始数据集的样本总数；$\boldsymbol{x}_t$ 为原始数据集的特征向量，$\boldsymbol{x}_t=(X_1,X_2,X_3,X_4)^{\mathrm{T}}$；$X_1$、$X_2$、$X_3$、$X_4$ 为选取的特征指标；y_t 为对应指标的油纸绝缘微水含量。将该原始数据集按照一定比例分为初始训练集 $D_{\mathrm{train}}=\left\{\left(\boldsymbol{x}_{\mathrm{train}},y_{\mathrm{train}}\right)\right\}$ 和初始测试集 $D_{\mathrm{test}}=\left\{\left(\boldsymbol{x}_{\mathrm{test}},y_{\mathrm{test}}\right)\right\}$。定义元模型 S_0 与各类基模型 S_h，$h=1,2,3,4$。

3）将初始训练集 D_{train} 随机均分为 k 份子集，记其为 $d_1,d_2,\cdots,d_k$。依次选取其中一个子集作为测试集，其余 $k-1$ 个子集作为训练集，输入基模型进行模型训练，直至 k 份测试子集全部训练完，且这些测试子集的并集可以完全覆盖初始训练集 D_{test}。因此，每一个基模型都会得到 k 组输出结果，设 R_h^k 为第 h 个基模型的以子集 d_k 为测试集的输出值；R_h，$h=1,2,3,4$ 为第 h 个基模型的所有 k 个测试集所训练得到的输出值的集合。

4）遵循K折交叉验证法，将初始测试集 D_{test} 依次输入各个基模型中进行训练，得到 k 个子模型，对得到的这 k 组评估值进行均值化处理，记处理后的结果为 r_h，$h=1,2,3,4$。

5）令 $x'_{\mathrm{train}}=R_h$，合并基模型的输出值为新训练集 $D'_{\mathrm{train}}=\left\{\left(R_1,y_{\mathrm{train}}\right),\left(R_2,y_{\mathrm{train}}\right),\cdots,\left(R_h,y_{\mathrm{train}}\right)\right\}$；类似地，令 $x'_{\mathrm{test}}=r_h$，构建新测试集 $D'_{\mathrm{test}}=\left\{\left(r_1,y_{\mathrm{test}}\right),\left(r_2,y_{\mathrm{test}}\right),\cdots,\left(r_h,y_{\mathrm{test}}\right)\right\}$，完成模型的初级训练。

6）将步骤1)～5)所获取的新训练集 D_{train} 和新测试集 D_{test}，输入次级评估体系中元模型进行训练与测试，其所得到的输出即为最终评估结果。

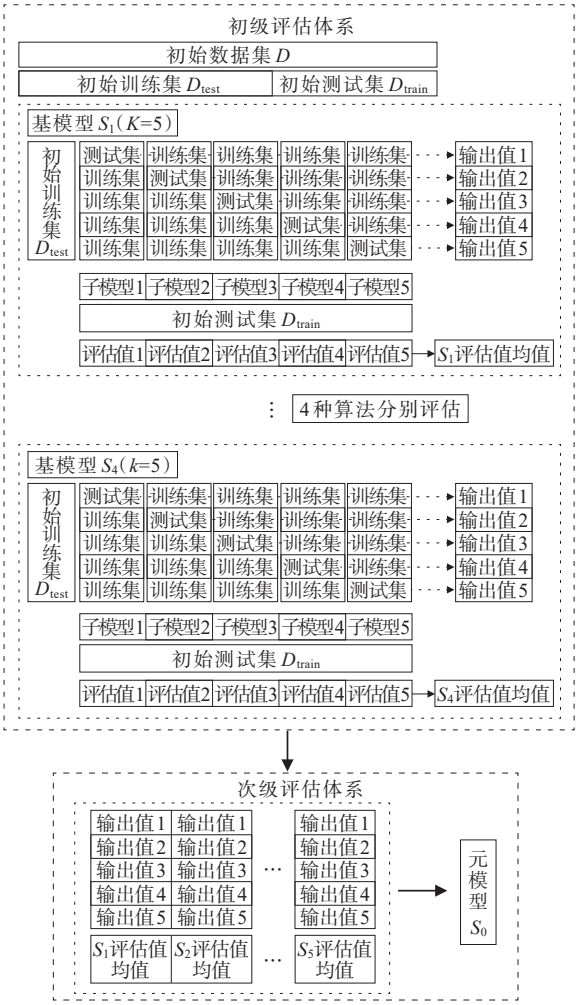


图5　Stacking模型训练流程

Figure 5　Stacking model training flowchart

3.2　考虑基模型赋权的Stacking改进

Stacking模型的内核是通过集成多个算法的评估结果以获取泛化的结果，并利用各个基模型之间相互独立的优势，采取K折交叉验证法，使其能够避免过拟合情况的发生。为处理在K折交叉验证法中造成的维度灾难问题，传统的Stacking模型一般采取将多个基模型的评估结果的均值作为元模型输入的方式。但该方式并未综合考虑各个子模型拟合精度不一致的实际情况，其本质是给所有模型的评估结果赋予了相等的权值。这可能掩盖了评估效果较好的模型的优越性，放大了评估效果较差的模型的缺陷，进而影响最终Stacking模型的评估精度。

为解决该问题，本文分别对同一基模型不同次K折交叉验证训练的预测结果进行加权处理。基于各个子模型K折交叉验证训练的输出值与真实值之间的误差，对该子模型预测的评估值赋予不同的

权重。$e_h^k = \left| y_{\text{test}} - R_h^k \right|$为该模型预测输出结果$y_{\text{test}}$与真实值$R_h^k$间的误差值。先利用$e_h^k$在误差总和$e_h$中所占比率分别确定权值$\rho_h^k$, $h = 1, 2, 3, 4$;然后,赋予k组评估结果相应的权重,并实施纵向堆叠处理;最后,将处理过后的新特征集输入元模型并进行二次评估。考虑权重赋值的改进Stacking评估模型流程如图6所示。

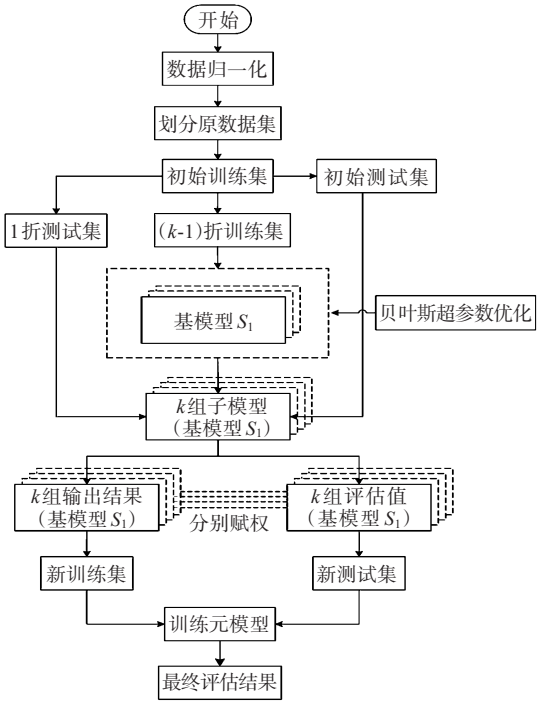


图6 考虑赋权的改进Stacking评估模型流程

Figure 6 Flowchart of improved Stacking evaluation model considering empowerment

4 算例分析

4.1 衡量指标

为检验该受潮定量评估模型的有效性,本文采用平均相对误差(mean absolute percentage error, MAPE)e_{MPAE}、均方误差(root mean square error, RMSE)e_{RMSE}和最大误差e_{MAX}来验证模型的评估精度。这3种误差越小,模型的精度越高。这3种误差的计算公式分别为

$$e_{\text{MPAE}} = \frac{1}{n}\sum_{i=1}^{n}\left(\left|\frac{Y(i) - y(i)}{Y(i)}\right|\right) \times 100\% \quad (16)$$

$$e_{\text{RMSE}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[Y(i) - y(i)\right]^2} \times 100\% \quad (17)$$

$$e_{\text{MAX}} = \max_{1 \leqslant i \leqslant n}\left\{\left|Y(i) - y(i)\right|\right\} \times 100\% \quad (18)$$

式(16)~(18)中,n为样本总数;$Y(i)$为水分含量实际值;$y(i)$为水分含量预测值。

4.2 基模型超参数调优

基模型超参数将在相当程度上影响融合模型的最终效果。为避免调参失误造成的失败,本文调用频域介电数据库中200组FDS测试数据,基于Optuna超参数调优框架,采用贝叶斯超参数优化方法以实现这4类基模型超参数的智能寻优。寻优结果如下:SVR模型核函数为RBF,惩罚系数为2;RF模型树数量为75,树深度取4,样本权重等于0.8;KNN模型邻居数为7;XGBoost模型树数量为120,树深度为3,随机采样比例等于0.73,子节点最小样本权重和取4,子节点分裂最小损失取0.1,学习率为0.15;KRR模型核函数为RBF。

4.3 异质集成模型对比验证

为检验该Stacking模型在油纸绝缘受潮量化评估中的准确性,本文另取300组不同绝缘状态下油浸纸试样的FDS测试数据进行异质集成模型对比试验。由于不同批次制备的样本聚合度存在差异,故参考《电力设备预防性试验规程》(DL/T 596—2005)中的相关规定,将油纸绝缘老化状态分为4类等级,其绝缘状态分级标准见表2。不同老化状态等级的样本数比例为1∶1∶1∶1,水分含量范围为[0.31%,4.65%]。按照5∶1的比例将数据集分为训练集与测试集,并采用文4.2提出的寻优后的基模型超参数,将这4种单一算法与随机抽选2类算法集成的Stacking模型的评估效果进行比对,各算法评估误差相关性指标如图7所示。

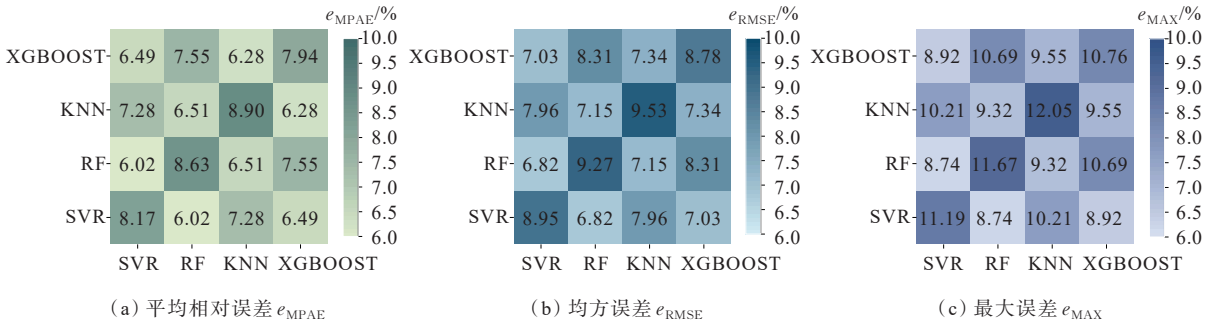
从图7中可以看出,Stacking融合模型的评估效果优于各单一模型的,这验证了该模型的有效性。空间受限可能导致单一模型无法搜寻出当前目标的真实值,但Stacking融合模型所具备迭代特性不仅降低了单一模型极易陷入局部极小值的风险,同时由于其极大地扩大了相应假设空间,故还显著提升了该模型逼近至真实函数的概率。此外,虽然RF与XGBoost算法均具有较强的决策能力,但这二者本质存在相似的树状结构机理,导致训练后的同质集成评估模型不存在显著优势。

为进一步验证Stacking异质集成模型可实现油纸绝缘受潮程度的精准量化评估,先基于提取的数

表2 绝缘状态分级标准

Table 2 Insulation state grading standards

聚合度	[900,+∞)	[600,900)	[300,600)	(−∞,300)
绝缘等级	优秀	良好	一般	较差



（a）平均相对误差 e_{MPAE}　（b）均方误差 e_{RMSE}　（c）最大误差 e_{MAX}

图 7　各算法评估误差相关性指标

Figure 7　Evaluation error correlation indexes for each algorithm

据集，并在元模型恒定的条件下，通过嵌套多类异质算法组成不同的模型框架以构建Stacking初级评估体系，不同基模型组合命名见表3。再运用消融试验与控制变量法对基于赋权改进策略及基模型遴选策略的所有模型进行训练并评估。

不同组合异质集成模型性能对比如图8所示。从图8中可以看出，多模型融合后的改进Stacking评估模型具备最佳基模型的组合方式，其观测数据的角度方位多，充分发挥了各基模型的优势。此外，该模型对不同K折的预测结果进行赋权处理，这进一步提高了其评估精度。

表 3　不同基模型组合命名

Table 3　Names of different base model combinations

第一层模型	定义名称
SVR＋RF＋ KNN	M_1
RF＋ KNN ＋ XGBoost	M_2
SVR ＋ KNN ＋ XGBoost	M_3
SVR＋RF＋ XGBoost	M_4
Stacking评估模型	M_5
改进Stacking评估模型	M_6
Stacking评估模型＋决策树	M_7
Stacking评估模型＋AdaBoost	M_8

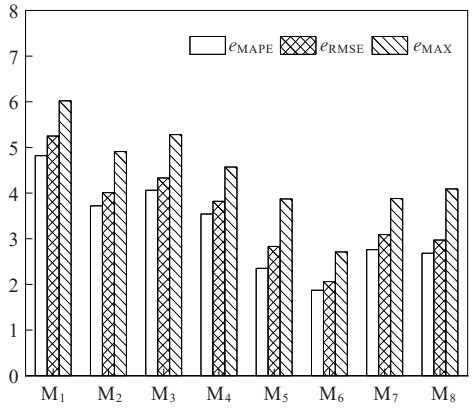


图 8　不同组合异质集成模型性能对比

Figure 8　Comparison of heterogeneous integration model ability for different combinations

从图8中还可以看出，对比M_6与M_7、M_8，基模型数量的增加并未改善最终模型的评估效果。这说明基模型的堆叠程度与评估精度并非呈正比。造成该现象的主要原因有2个：一方面是决策树算法其自身泛化性与稳定性比其他模型相比并无显著优势；另一方面是AdaBoost与XGBoost算法在数据特征结构与空间上存在较高的吻合度。因此，本文所提出Stacking异质集成模型利用赋权处理有效提升了模型对数据特征的敏感性；所提模型的异质集成特点不仅突破了其他大多数研究中的仅针对单一评估模型进行优化改进的瓶颈，降低了单一算法评估带来的偏差，还提高了融合模型的整体评估精度，这种特点也为其带来了相较于其他模型更高的准确性。

4.2　可行性验证

为检验本文所提模型在实际工况下的可行性，现对福建省莆田、宁德市等地多台不同运行年份与不同受潮程度的变压器进行FDS测试。该测试借助DIRAN评估软件获取测试对象的水分含量值。测试变压器基本信息与含水率值见表4。

由表4可知，根据IEEE Std 62—1995中的规定，这4台测试变压器的绝缘状态均为干燥。其中，南

表 4　测试变压器基本信息与含水率值

Table 4　Basic information and moisture content value of tested transformers

变压器型号	位置	运行时长/a	水分含量/%
SFPS－180000/220	南平市沙溪口水电厂	13	0.910
SZ10－50000/110	莆田市城南变电站	14	0.870
SFSZ10－180000/220	莆田市盖尾变电站	18	1.230
SF08－31500/110	宁德市古田溪电厂	备用	1.300

平市沙溪口水电厂的型号为SFPS-180000/220的变压器与莆田市城南变电站的型号为SZ10-50000/110的变电器均存放在干燥的环境中，外部水分入侵可能性较小，其水分含量均较低；莆田市盖尾变电站的型号为SFSZ10-180000/220变电器长期运行在温度较高的环境中，其绝缘纸不易吸收水分，其水分含量的参考值比宁德市古田溪电厂的型号为SF08-31500/110的备用变压器的略低。综上所述，DIRANA软件的评估结果符合所测试变压器的实际运行情况。

对这4种变压器进行频域介电谱试验，基于FDS测试数据，并利用式(1)～(14)，计算表征测试样本受潮程度的特征指标X_1～X_4，测试变压器的微水含量评估值见表5。

表5 测试变压器的微水含量评估值

Table 5 Assessment value of micro-water content of tested transformers

变压器型号	X_1	X_2	X_3	X_4
SFPS-180000/220	0.001 2	0.216 0	0.485 0	1.100 0
SZ10-50000/110	0.001 3	0.372 0	0.476 0	0.920 0
SFSZ10-180000/220	0.005 7	0.681 0	1.823 0	2.100 0
SF08-31500/110	0.002 4	0.563 0	1.267 0	23.000 0

按照图5所示流程，将该4组特征指标值代入改进Stacking评估模型中，对其油纸绝缘受潮情况进行定量评估，评估结果对比见表6。

表6 评估结果对比

Table 6 Comparison of assessment results

变压器型号	实际水分含量/%	评估结果/%	相对误差δ/%
SFPS-180000/220	0.910	0.931	2.31
SZ10-50000/110	0.870	0.849	2.41
SFSZ10-180000/220	1.230	1.201	2.36
SF08-31500/110	1.300	1.327	2.08

由表6可知，本文所提模型的4组评估结果均符合实际情况，其相对误差δ范围为2.0%～2.5%。这说明该模型可有效剔除变压器因运行周期过久导致的绝缘老化对受潮定量评估所造成的误差，充分验证了改进Stacking模型在油纸绝缘受潮程度评估中的可行性及实用价值。

5 结语

本文计及老化作用对受潮评估所造成的误差，提出了一种基于Stacking异质集成模型的油纸绝缘受潮程度量化方法，得出如下主要结论：

1）在老化-受潮协同影响下，FDS曲线的变化趋势出现交叉掩盖现象。本文所提取的特征集合可有效辨识受潮单一作用对油纸绝缘系统频谱曲线的独立贡献，为精准量化受潮程度提供必要的模型基础。

2）将SVR、RF、KNN及XGBoost各算法动态融合后的异质集成模型可充分发挥各算法优势，减少噪声输出并过滤冗余特征，有效提升受潮评估模型整体的泛化能力。

3）经赋权改进后的Stacking模型可丰富数据的观测角度并甄别非线性数据中的特征信息。经实例对比验证，该模型具备更高的评估精度，可有效实现对油纸绝缘受潮程度的精准量化。

参考文献：

[1] 刘庆珍，黄昌硕．基于FCBF特征选择和XGBoost原则的油纸绝缘介电响应特征量优选研究[J]．电力系统保护与控制，2022，50(15)：50-59.
LIU Qingzhen, HUANG Changshuo. Optimization of dielectric response characteristics of oil paper insulation based on FCBF feature selection and the XGBoost principle[J]. Power System Protection and Control, 2022, 50(15):50-59.

[2] 曾庆美，蔡金锭．基于电流极化谱特征量的油纸绝缘变压器微水含量评估[J]．电力科学与技术学报，2018，33(3)：75-80.
ZENG Qingmei, CAI Jinding. Evaluation of moisture content in the oil-paper insulation transformer based on the characteristic of current polarization spectrum[J]. Journal of Electric Power Science and Technology, 2018, 33(3):75-80.

[3] 邓昭宏，赵春明，冷俊，等．基于IWOA-BP算法的变压器绝缘老化评估方法[J]．电力科学与技术学报，2023，38(5)：253-261.
DENG Zhaohong, ZHAO Chunming, LENG Jun, et al. Evaluation method of transformer insulation aging state based on IWOA-BP algorithm[J]. Journal of Electric Power Science and Technology, 2023, 38(5):253-261.

[4] 刘庆珍，陈俊鸿．基于等级云模型的油纸绝缘老化状态评估[J]．高压电器，2023，59(1)：176-184.
LIU Qingzhen, CHEN Junhong. Assessment of ageing state of oil-paper insulation based on grade cloud model [J].High Voltage Apparatus, 2023, 59(1):176-184.

[5] 邹阳，陈啸轩，张云霄，等．基于界面电荷极化特性的变压器油纸绝缘寿命预测方法研究[J]．电力系统保护与控制，2024，52(7)：138-148.
ZOU Yang, CHEN Xiaoxuan, ZHANG Yunxiao, et al. A

life prediction method of transformer oil-paper insulation based on interfacial charge polarization characteristics[J]. Power System Protection and Control, 2024,52(7):138-148.

[6] 张涛,李林多,王尧,等.基于复介电常数实部变化率的油纸绝缘受潮状态研究[J].电测与仪表,2019,56(16):45-49+60.
ZHANG Tao, LI Linduo, WANG Yao, et al. Study on wetting state of oil paper insulation based on the real part change rate of complex permittivity[J]. Electrical Measurement & Instrumentation,2019,56(16):45-49+60.

[7] 江阳,赵贺,宋一凡,等.基于频域介电谱的环氧树脂浸渍皱纹纸绝缘受潮状态评估[J].绝缘材料,2023,56(7):60-66.
JIANG Yang, ZHAO He, SONG Yifan, et al. Moisture condition evaluation of epoxy resin impregnated crepe paper based on frequency domain spectroscopy[J]. Insulating Materials,2023,56(7):60-66.

[8] 邹阳,翁祖辰,金涛.考虑界面极化因素的油纸绝缘样品中含水量评估方法的研究[J].电机与控制学报,2022,26(7):88-97.
ZOU Yang, WENG Zuchen, JIN Tao. Water content evaluation method of oil-paper insulating sample considering interfacial polarization factor[J]. Electric Machines and Control,2022,26(7):88-97.

[9] 郭蕾,张传辉,廖维,等.基于Dissado-Hill模型的油纸绝缘受潮参数特征与评估方法[J].电工技术学报,2021,36(23):5058-5068.
GUO Lei, ZHANG Chuanhui, LIAO Wei, et al. Oil-paper insulation moisture parameter characteristics and evaluation method based on dissado-hill model[J]. Transactions of China Electrotechnical Society, 2021, 36(23):5058-5068.

[10] XU Q C,WANG S K,LIN F C,et al.Extracting frequency spectroscopy of oil-immersed paper based on havriliak－negami model without known insulation structure of transformer[J]. IEEE Transactions on Instrumentation and Measurement,2022,71:6003312.

[11] 石宜金,谭贵生,赵波,等.基于模糊综合评估模型与信息融合的电力变压器状态评估方法[J].电力系统保护与控制,2022,50(21):167-176.
SHI Yijin, TAN Guisheng, ZHAO Bo, et al. Condition assessment method for power transformers based on fuzzy comprehensive evaluation and information fusion [J]. Power System Protection and Control, 2022, 50(21):167-176.

[12] 田旭,张逸伦,张桂红,等.器件混合型直流变压器控制建模及扰动[J].电网与清洁能源,2024,40(10):13-23.
TIAN Xu,, ZHANG Yilun, ZHANG Guihong, et al. The control modeling and disturbance of the hybrid DC transformer[J]. Power System and Clean Energy, 2024,40(10):13-23.

[13] 范贤浩,刘捷丰,张镱议,等.融合频域介电谱及支持向量机的变压器油浸纸绝缘老化状态评估[J].电工技术学报,2021,36(10):2161-2168.
FAN Xianhao, LIU Jiefeng, ZHANG Yiyi, et al. Aging evaluation of transformer oil-immersed insulation combining frequency domain spectroscopy and support vector machine[J]. Transactions of China Electrotechnical Society,2021,36(10):2161-2168.

[14] 孙世明,岑红星,白建民,等.基于集成SAO优化互相关熵极限学习机模型的变压器故障诊断方法[J].电测与仪表,2024,61(9):56-64.
SUN Shiming, CEN Hongxing, BAI Jianmin, et al. Transformer fault diagnosis method based on integrated correntropy extreme learning machine model optimized by SAO[J]. Electrical Measurement & Instrumentation, 2024,61(9):56-64.

[15] 邹阳,俞豪奕,金涛.融合模糊K近邻及证据理论的变压器油纸绝缘状态评估方法[J].电力系统保护与控制,2023,51(14):55-63.
ZOU Yang, YU Haoyi, JIN Tao. Evaluation method of the oil-paper insulation condition of a transformer based on fuzzy K nearest neighbor and evidence theory[J]. Power System Protection and Control,2023,51(14):55-63.

[16] BANDARA K,EKANAYAKE C,SAHA T K.Analysis of frequency domain dielectric response of pressboard insulation impregnated with different insulating liquids [J]. IEEE Transactions on Dielectrics and Electrical Insulation,2016,23(4):2042-2050.

[17] 杨丽君,彭攀,高竣,等.基于频域介电响应特征指纹的油纸绝缘受潮及老化状态区间识别[J].电工技术学报,2018,33(9):2105-2114.
YANG Lijun, PENG Pan, GAO Jun, et al. The range recognition of moisture and aging status of oil-paper insulation based on frequency domain dielectric response characteristic fingerprint[J]. Transactions of China Electrotechnical Society,2018,33(9):2105-2114.

[18] 刘捷丰.基于介电特征量分析的变压器油纸绝缘老化状态评估研究[D].重庆:重庆大学,2015.
LIU Jiefeng. Condition assessment of transformer oil-paper insulation based on analysis of dielectric characteristic parameters[D]. Chongqing: Chongqing University,2015.

[19] 张德文,张健,曲利民,等.基于PSO-ELM的变压器油纸绝缘状态无损评估方法[J].电力工程技术,2024,43(3):201-208.
ZHANG Dewen, ZHANG Jian, QU Limin, et al. The assessment method of transformer oil-paper insulation state based on PSO-ELM[J]. Electric Power Engineering Technology,2024,43(3):201-208.

[20] 殷之文.电介质物理学[M].2版.北京:科学出版社,2003.

YIN Zhiwen. Dielectric physics[M]. 2nd ed. Beijing: Science Press,2003.

[21] WOLPERT D H. Stacked generalization[J]. Neural Networks,1992,5(2):241-259.

[22] 朱文广,李映雪,杨为群,等.基于K-折交叉验证和Stacking融合的短期负荷预测[J].电力科学与技术学报,2021,36(1):87-95.
ZHU Wenguang,LI Yingxue,YANG Weiqun,et al.Short-term load forecasting based on the K-fold cross-validation and stacking ensemble[J]. Journal of Electric Power Science and Technology,2021,36(1):87-95.

[23] 寿绍安,罗海荣,王晓康,等.基于VMD-ISODATA-DBN的配电台区短期负荷预测方法[J].智慧电力,2023,51(11):53-60.
SHOU Shaoan, LUO Hairong, WANG Xiaokang, et al. Forecasting method of short-term load in distribution station area based on VMD-ISODATA-DBN[J]. Smart Power,2023,51(11):53-60.

[24] 史佳琪,张建华.基于多模型融合Stacking集成学习方式的负荷预测方法[J].中国电机工程学报,2019,39(14):4032-4042.
SHI Jiaqi, ZHANG Jianhua. Load forecasting based on multi-model by stacking ensemble learning[J]. Proceedings of the CSEE,2019,39(14):4032-4042.

[25] 李志军,徐博,张家安,等.基于TD3可变长度时间窗口最优加权的短期负荷预测策略[J].电力建设,2024,45(6):140-148.
LI Zhijun,XU Bo,ZHANG Jia'an ,et al.Short-term load optimal weighted forecasting strategy based on TD3 variable length time window[J]. Electric Power Construction,2024,45(6):140-148.

[26] 魏书荣,张鑫,符杨,等.基于GRA-LSTM-Stacking模型的海上双馈风力发电机早期故障预警与诊断[J].中国电机工程学报,2021,41(7):2373-2383.
WEI Shurong, ZHANG Xin, FU Yang, et al. Early fault warning and diagnosis of offshore wind DFIG based on GRA-LSTM-stacking model[J]. Proceedings of the CSEE,2021,41(7):2373-2383.

[27] 张宏杰,陈贵凤,闫宏伟,等.基于SMOTE与Bayes优化的LSTM网络变压器故障诊断[J].中国电力,2023,56(10):164-170.
ZHANG Hongjie, CHEN Guifeng, YAN Hongwei, et al. Fault diagnosis of LSTM network tansformer based on SMOTE and Bayes optimization[J].Electric Power,2023,56(10):164-170.

[28] 唐登平,蔡文嘉,周翔宇,等.基于VMD和样本熵的电磁式电流互感器故障诊断[J].电力科学与技术学报,2021,36(6):144-150.
TANG Dengping,CAI Wenjia,ZHOU Xiangyu,et al.Fault diagnosis of current transformer based on VMD and sample entropy[J].Journal of Electric Power Science and Technology,2021,36(6):144-150.

[29] 马向南,杨涛,王云龙,等.基于轻量化神经网络及模型压缩的变压器油色谱故障诊断方法[J].供用电,2023,40(5):84-91.
MA Xiangnan, YANG Tao, WANG Yunlong, et al. Transformer dissolved gas fault diagnosis method based on lightweight neural network and model quantization[J]. Distribution & Utilization,2023,40(5):84-91.

[30] 肖湘奇,肖宇,黄瑞,等.基于随机森林算法的低压系统漏电检测技术研究[J].电力科学与技术学报,2024,39(3):38-47+115.
XIAO Xiangqi,XIAO Yu,HUANG Rui,et al.Research on leakage detection technology of low-voltage power system based on random forest algorithm[J]. Journal of Electric Power Science and Technology, 2024, 39(3): 38-47+115.

[31] 江永鑫,陈丽安.基于IIST与SVM的串联故障电弧诊断方法研究[J].电力科学与技术学报,2023,38(5):159-168.
JIANG Yongxin,CHEN Li'an.Research on arc diagnosis method of series faults based on IIST and SVM[J]. Journal of Electric Power Science and Technology,2023,38(5):159-168.

[32] 罗昊,杨琪,王炜宇,等.基于XGBoost的输电线路火灾风险等级评估方法[J].电力科学与技术学报,2023,38(6):132-141.
LUO Hao, YANG Qi, WANG Weiyu, et al. XGBoost-based assessment method for fire risk levels of transmission lines[J]. Journal of Electric Power Science and Technology,2023,38(6):132-141.

[33] 王泽,张玉敏,吉兴全,等.基于深度学习与内核岭回归的电力系统鲁棒状态估计[J].高电压技术,2022,48(4):1332-1342.
WANG Ze, ZHANG Yumin, JI Xingquan, et al. Robust state estimation of power system based on deep learning and kernel ridge regression[J].High Voltage Engineering, 2022,48(4):1332-1342.

[34] 叶健强,孙敦虎.碳交易条件下基于鲁棒优化的电源规划研究[J].发电技术,2024,45(3):566-574.
YE Jianqiang, SUN Dunhu. Research on power planning based on robust optimization under carbon trading condition[J]. Power Generation Technology, 2024, 45(3): 566-574.