文章编号:1674-2974(2025)06-0178-09 DOI:10.16339/j.cnki.hdxbzkb.2025187

# 基于改进XGBoost算法硅橡胶击穿场强预测与分析

毕茂强[1†],张世宇[1],张文轩[1],江天炎[1],陈曦[1],郭元吉[2]
(1. 重庆理工大学 电气与电子工程学院,重庆 400054;
2. 山东齐林电力设备股份有限公司 技术研发部,山东 淄博 255000)

**摘 要**:硅橡胶材料因其出色的绝缘性能,常用作高压条件下的绝缘材料,其击穿场强是重要的电气性能指标,与材料配方之间存在复杂的非线性关系.基于此,提出了一种基于遗传算法(genetic algorithm,GA)优化的极端梯度提升(extreme gradient boosting,XGBoost)算法的高效评价模型.该模型将GA和XGBoost相结合,以温度、色母相对含量、$Al(OH)_3$微粉直径、$Al(OH)_3$相对含量和厚度作为输入,建立了优化后的XGBoost模型,对击穿场强进行预测,GA算法在XGBoost模型训练过程中自动选择最优参数.采用皮尔逊相关系数对其影响因素进行分析可知,厚度和温度是影响击穿场强的关键因素,而色母相对含量、$Al(OH)_3$微粉直径和相对含量的影响相对较小.将常用的回归模型与所提出模型的评价指标进行对比分析,该模型决定系数可达0.953,均方根误差和平均绝对误差仅为0.361 kV/mm和0.168.结果表明:GA-XGBoost模型能够较准确地预测该材料的击穿场强,可为研究硅橡胶材料性能和优化材料配方提供参考依据.

**关键词**:硅橡胶;工频击穿试验;相关因素;预测分析
**中图分类号**:TM21;TM215.92 **文献标志码**:A

# Prediction and Analysis of Silicone Rubber Breakdown Field Strength Based on Improved XGBoost Algorithm

BI Maoqiang[1†],ZHANG Shiyu[1],ZHANG Wenxuan[1],JIANG Tianyan[1],CHEN Xi[1],GUO Yuanji[2]
(1. School of Electrical and Electronic Engineering, Chongqing University of Technology, Chongqing 400054, China;
2. Technical Research and Development Department, Shandong Qilin Power Equipment Co., Ltd., Zibo 255000, China)

**Abstract**: Silicone rubber materials are commonly used as insulation materials under high-pressure conditions due to their excellent insulation properties. The breakdown field strength is an important electrical performance index, and there is a complex nonlinear relationship between the breakdown field strength and the material formula. Based on this, an efficient evaluation model based on genetic algorithm (GA) optimized extreme gradient boosting (XGBoost) algorithm is proposed. The model combines GA and XGBoost, and uses temperature, relative content of masterbatch, diameter of $Al(OH)_3$ micropowder, relative content of $Al(OH)_3$ and thickness as inputs to establish an improved XGBoost model to predict the breakdown field strength. The GA algorithm automatically selects the optimal parameters during the training process of the XGBoost model. The Pearson correlation coefficient is used to analyze

* **收稿日期**:2024-06-04
**基金项目**:重庆市自然科学基金资助项目(CSTB2022NSCQ-MSX0267), Natural Science Foundation of Chongqing (CSTB2022NSCQ-MSX0267)
**作者简介**:毕茂强(1984—),男,四川内江人,重庆理工大学副教授
† 通信联系人,E-mail: bimaoqiang@cqut.edu.cn

the influencing factors. It can be seen that the thickness and temperature are the key factors affecting the breakdown field strength, while the influence of the masterbatch content, the diameter and relative content of $Al(OH)_3$ micropowder is relatively small. The evaluation index of the commonly used regression model is compared with the proposed model. The coefficient of determination of the model can reach 0.953, and the root-mean-square error and mean absolute error are only 0.361 kV/mm and 0.168, respectively. The results show that the GA-XGBoost model can accurately predict the breakdown fieid strength of the material, which can provide a reference for studying the properties of silicone rubber materials and optimizing the material formulation.

**Key words**: silicone rubber; power frequency breakdown test; correlators; predictive analytics

硅橡胶因其出色的耐热性、耐寒性、耐候性、电气特性及生理惰性而备受关注. 其绝缘性能相较于普通材料更为卓越,通常在电弧放电的情况下不易发生焦烧,因此在高压环境下使用时显得尤为可靠. 相关研究表明,其绝缘性能受到温度、成分和厚度等多种因素的共同影响[1-4].

目前,有大量学者对硅橡胶的绝缘性能进行了研究. 代冲[5]研究了温度对硅橡胶性能的影响,揭示了低温下硅橡胶的沿面闪络电压和击穿电压的变化规律. 同时,色母的加入可能会改变硅橡胶的交联结构,从而影响其绝缘性能. 此外,选择$Al(OH)_3$作为填料时,其粒径和添加量对硅橡胶的击穿场强和其他性能有着显著影响. 胡建林等[6]通过添加$Al(OH)_3$粉末,提升了硅橡胶的电气性能,尽管这可能会牺牲部分加工性能和力学性能. 周远翔等[7]研究了不同厚度硅橡胶试样的直流击穿电压,发现击穿场强随厚度增加而降低. 尽管众多研究文献已经探讨了影响硅橡胶电气性能的关键因素,但由于硅橡胶配方的复杂性,单一因素的研究难以确定最佳配比及其对绝缘强度的影响.

基于上述问题,有人提出利用算法研究材料的配比或绝缘性能,并取得了一定成果. Nandi[8]利用多目标优化算法设计了硅橡胶复合模具材料,优化了颗粒增强硅胶软模工艺. Shaymaa等[9]运用鲸鱼优化算法和海鸥优化算法计算了理想填料浓度,以提高绝缘子的介电强度. 也有人分析了交联聚乙烯绝缘电缆的晶体形貌、厚度和位置对其短期击穿特性的影响,得到了相应的寿命预测模型.Chen等[10]通过对电极面积较小、层数较少的薄膜介质进行平均击穿场强试验,建立了不同电极面积和薄膜层数下薄膜介质平均击穿场强预测模型,并验证了该预测模型的准确性. 许旭日等[11]基于不同厚度的聚乙烯纳米复合材料进行直流击穿场强试验,并结合BCT模型预测了该材料的直流击穿场强,分析了厚度对预测结果的影响,结果表明该预测模型效果较好. 尽管算法在相关领域已应用广泛,且已用于对硅橡胶材料的部分特性进行理论研究,但是鲜有研究采用算法对不同配比下硅橡胶材料的工频击穿特性进行预测及分析.

因此,本研究依托工频击穿试验,对不同配方和厚度的硅橡胶样品进行了击穿场强测试,并运用皮尔逊相关系数分析了各种影响因素之间的相关性,还利用改进XGBoost算法模型预测了不同影响因素下的击穿场强. 最后比较了5种常见算法的评估指标,为硅橡胶材料的绝缘性能研究提供了新的思路和参考.

## 1　不同配比硅橡胶工频击穿试验

本次试验使用型号为ZJC-50 kV的工频击穿电压测试系统,对硅橡胶样品在不同温度、色母相对含量、$Al(OH)_3$微粉直径、$Al(OH)_3$相对含量以及厚度等条件下进行了工频击穿试验,采用式(1)计算样品的平均击穿场强$E$,单位为kV/mm.

$$E = \frac{U}{d} \tag{1}$$

式中:$d$为对应击穿点的厚度,单位为mm;$U$为击穿点电压,单位为kV.

### 1.1　样品制备

硅橡胶测试样品采用热压法制备. 根据表1所列的不同配方,将原料混合均匀,并倒入不同厚度的模具中,在硫化过程中采用平板硫化机施加15 MPa的压力. 将硫化、冷却后的硅橡胶试片,制成规格为10 cm×10 cm的方形试片.

在研究色母粒的相对含量、$Al(OH)_3$的相对含量

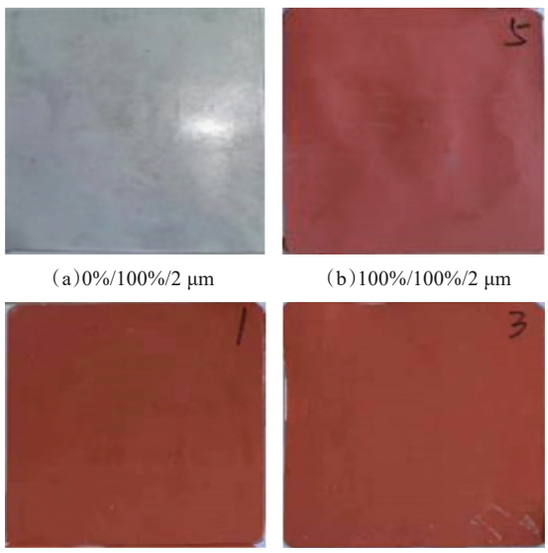
及其微粉的直径时,对于同一配方,分别制备了5组不同含量的样品(每组包括厚度为1 mm和2 mm的样品各5组).在制备不同厚度的样品时,遵循了上述步骤,并使用了不同厚度的热压模具.

**表1 试验变量选取**

**Tab.1 Selection of test variables**

| 取值类型 | 温度/℃ | 厚度/mm | 色母相对含量/% | $Al(OH)_3$微粉直径/μm | $Al(OH)_3$相对含量/% |
|---|---|---|---|---|---|
| 取值范围 | −30~20 | 1~5 | 0~100 | 2~5 | 90~110 |
| 取值梯度 | 10 | 1 | 50 | 2.5 | 10 |

部分样品如图1所示,分图名从左至右依次为色母相对含量、$Al(OH)_3$相对含量、$Al(OH)_3$微粉直径.




(a)0%/100%/2 μm (b)100%/100%/2 μm

(c)100%/100%/5 μm (d)100%/110%/5 μm

图1 不同配比的部分样品试片

Fig.1 Part test pieces with different proportions

### 1.2 样品预处理

硅橡胶材料表面易吸附杂质,如灰尘等,这些杂质可能污染二甲硅油介质,从而导致试验结果产生偏差.为确保试验的准确性,在试验前,必须对试样表面进行严格的清洁预处理.

清洁过程包括使用无水乙醇和去离子水仔细清除表面污秽,确保试样表面彻底清洁.随后,将清洁后的试样放入干燥箱中,直至试样完全干燥.干燥完成后,为避免热冲击和保持试样稳定,应将其放在清洁、无尘的环境中自然冷却至室温.同时,与试样接触的电极也需要进行相应的清洁预处理,以确保试验的可靠性[12-13].

### 1.3 试验平台搭建

根据GB/T 1408.1—2016关于绝缘材料电气强度试验方法中的电极设计标准,装置电极系统为直径25 mm且表面光滑的平板形电极[14];采用型号为SDJ710FB的高低温试验箱调节试验温度,温控范围−70~100 ℃,温度偏差±2 ℃;温度控制器型号为KX-709,精度0.1 ℃,测温范围−70~200 ℃,将其伸入硅油中,试验平台搭建如图2所示.

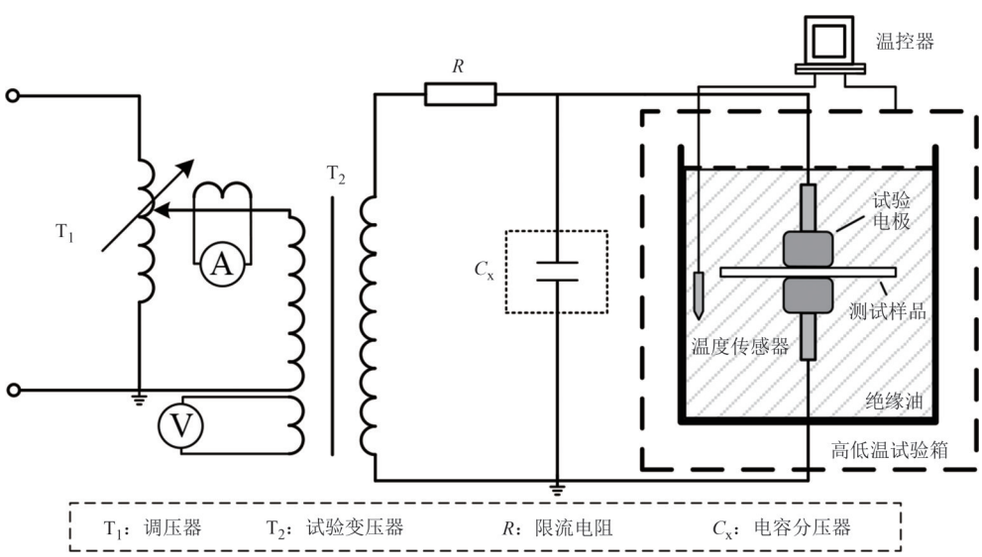



图2 工频击穿测试系统

Fig.2 Power frequency breakdown test system

将整个试验电极平台放置于二甲硅油中,该硅油绝缘性能好,导热效率高,凝固点为−100 ℃,在高低温时依然能保持良好的流动性,一方面可以防止电极之间沿空气的闪络,另一方面可减弱加压时发

生在针尖空气气隙的局部放电[15]. 硅橡胶材料放入绝缘油中静置 5 min,保证硅橡胶温度与外部一致后,再进行击穿试验. 由于加热时间较短,可以认为此时硅橡胶材料还未发生热老化[16].

### 1.4　试验步骤

将清洁的试样置于高低压等直径电极之间,保证试样与电极之间的良好接触,两电极与试片刚好贴合却不挤压[17].将整个试验电极平台置于高低温试验箱中,待温度传感器显示的硅油温度与试验箱温度吻合时,根据GB/T 1695—2005硫化橡胶工频击穿电压强度的测定方法,采用快速均匀升压法,从零开始均匀升高电压,加压速度为1.0 kV/s,直到达到击穿的最大值,并对此时的数据进行记录[18]. 每次更换试片后需要等待温度传感器的温度与试验箱温度接近一致后再施加电压[19]. 为减少单次试验结果的偶然性,提高测试结果的准确性和可靠性,同一配方下的每个样品在不同位置至少进行4次击穿,累计20次. 通过调节试验箱的温度,记录每个温度点和每种配方下硅橡胶材料的击穿电压值.

## 2　影响因素分析

### 2.1　异常值处理

在进行硅橡胶材料的工频击穿试验时,硅橡胶材料中可能存在微观缺陷或不均匀性,如气泡、杂质等,或者由于硅橡胶材料本身的固有变异性,从而产生低于正常范围的异常值. 基于此,采用箱线图对原始数据集进行异常值分析.

箱线图(box plot)是一种用于展示数据分布的统计图表,它能够直观地显示数据的四分位数、异常值和数据的整体分布情况. 由于原始数据较为分散,而相关文献表明击穿场强与厚度有较强关联性[20],故在进行数据处理时,对原始数据按照不同厚度分为5组. IQR(四分位数差)用于描述数据的分散程度和识别异常值,一般大于上四分位数1.5倍IQR的值,或者小于下四分位数1.5倍IQR的值,被认为是可能的异常值,用触须表示,分析结果如图3所示.

由图3可知,原始数据中存在少量异常值,且随着样品厚度增加,其击穿场强整体呈现下降的趋势. 此外,1 mm和2 mm硅橡胶样品击穿场强分布范围较广,相同厚度下不同配比样品的击穿场强差异较大,造成数据分散性较大. 对异常值进行剔除后,创建有效数据集进行皮尔逊相关系数分析.

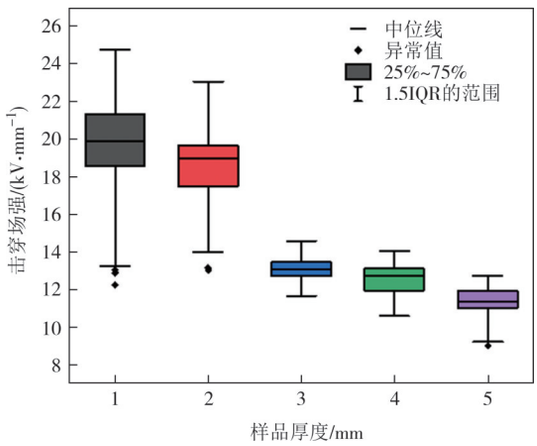



图3　不同厚度击穿场强箱线图

Fig.3　Breakdown strength box plot of different thickness

### 2.2　皮尔逊相关系数分析

皮尔逊相关系数(Pearson correlation coefficient)是用于度量两个变量$x$和$y$之间的线性相关程度的一个统计量,这种简洁的表达方式使得相关性非常直观和易于理解,且被应用于多个领域. 其计算原理如式(2)所示:

$$r_{xy}=\frac{n(\sum xy)-(\sum x)(\sum y)}{\sqrt{[n\sum x^2-(\sum x)^2][n\sum y^2-(\sum y)^2]}} \tag{2}$$

式中:$n$是数据点的数量;$\sum xy$是所有数据点的$x$值与其对应的$y$值的乘积之和;$\sum x$和$\sum y$分别是所有$x$值和$y$值的总和;$\sum x^2$和$\sum y^2$分别是所有$x$值的平方和与所有$y$值的平方和.

现有的大多数研究仅基于单一变量探究其对击穿场强的影响,具有一定的局限性. 因此,采用上述方法对硅橡胶材料不同配比、厚度以及温度之间的相关性进行综合分析具有重要意义,各影响因素分析结果如图4所示.

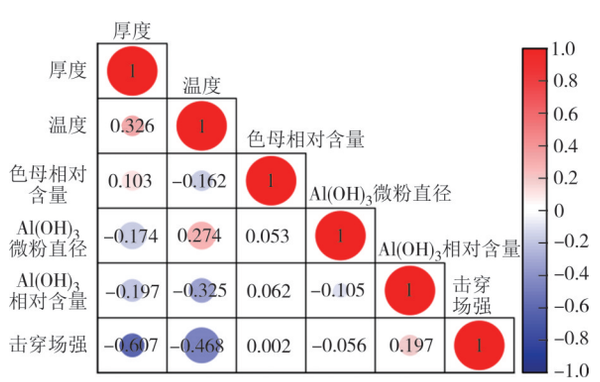



图4　相关性热力图

Fig.4　Correlation heat map

通过皮尔逊相关系数分析,发现材料特征量之间大多无强相关,表明它们之间没有显著的线性关系.具体来看,厚度与击穿场强呈现-0.607的负相关性,这意味着硅橡胶材料中较厚材料可能存在更多缺陷或杂质,从而导致击穿场强降低.但是厚度与温度的正相关性为0.326,表明当材料厚度增加时,会伴随温度的升高,由于更厚的材料在试验温度升高后其内部温升效应明显,从而有效地吸收或保留试验过程中产生的热量,导致测得的温度较高.

温度与击穿场强的负相关性为-0.468,说明温度升高导致击穿场强降低,这可能是由于温度升高增加了材料的导电性和离子移动性,导致击穿场强降低.但是温度与$Al(OH)_3$微粉直径有中等正相关性(0.274),这表明当试验温度升高时,硅橡胶基体可能变得更为柔软或流动性增强,导致$Al(OH)_3$微粉在硅橡胶中的分布或排列方式发生变化,使微粉颗粒间的距离减小或排列变得更为紧密.温度与$Al(OH)_3$含量有中等负相关性(-0.325),表明在较高的温度下,$Al(OH)_3$的添加量可能会减少,这可能是因为温度的升高改变了填料的分散性、反应活性或材料的整体热稳定性,从而影响了$Al(OH)_3$的用量.

色母相对含量与其他参数的线性相关性较弱,$Al(OH)_3$微粉直径除了与温度有中等相关性外,与其他特征的相关性都很低,但二者仍然可能与其他因素共同作用对击穿场强产生影响.

$Al(OH)_3$相对含量与击穿场强呈现0.197的弱正相关性,与其他特征的相关性较低,表明$Al(OH)_3$作为阻燃填料,可能会通过改善材料的热稳定性和电气绝缘性能,对击穿场强产生正面影响.

整体来看,厚度和温度对击穿场强有显著影响,色母相对含量影响较小,$Al(OH)_3$微粉直径和相对含量的影响也符合预期.通过上述分析,创建基于硅橡胶击穿场强与不同影响因素的数据集.

## 3 模型构建及结果分析

### 3.1 数据预处理

为了提高模型的性能,建立具有良好预测精度和泛化能力的XGBoost预测模型,通过异常值处理以及$Z$分数标准化操作,可以使得模型更好地学习数据中的规律,减少过拟合,增加模型的可解释性,确保输入数据的质量[21].

基于上述影响因素分析,针对有相关性的特征,对数据进行$Z$分数标准化处理,使不同特征的数值范围相近,以利于模型训练,将数据转换为具有标准正态分布(均值为0,标准差为1)的数值.该方法可以消除不同特征之间的量级差异,使得每个特征在模型训练时对结果的影响更加公平[22].数据整体分布情况如表2所示,$Z$分数标准化的公式如下:

$$Z = \frac{X - \mu}{\sigma} \tag{3}$$

式中:$X$是原始数据点;$\mu$是原始数据集的均值;$\sigma$是原始数据集的标准差.

以厚度、温度、色母相对含量、$Al(OH)_3$微粉直径、$Al(OH)_3$相对含量等影响硅橡胶击穿场强的影响因素作为输入值,击穿场强作为标签值,将预处理后的数据构建数据集.

**表2 数据分布情况**

**Tab.2 Data distribution**

| 指标 | 温度/℃ | 厚度/mm | 色母相对含量/% | $Al(OH)_3$微粉直径/μm | $Al(OH)_3$相对含量/% |
|---|---|---|---|---|---|
| 最大值 | 20 | 5.01 | 100 | 5 | 110 |
| 最小值 | -30 | 0.97 | 0 | 2 | 90 |
| 平均值 | -5 | 2.99 | — | — | — |

### 3.2 XGBoost算法

极端梯度提升(XGBoost)是一种实现梯度提升框架的优化分布式梯度增强库,旨在提供一种可扩展、可移植、分布式的梯度提升机(gradient boosting machine,GBM)算法,以解决各种数据科学问题.XGBoost是基于决策树(通常是CART)的集成学习方法.其通过不断添加新的树来拟合数据中的残差,每棵树都是基于前一棵树的残差来构建的.每棵树学习的是前一棵树的不完美之处,逐渐提高整体的预测精度[23-24].其目标函数可以表示为:

$$O(b) = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(b)}) + \sum_{k=1}^{K} \Omega(f_k) \tag{4}$$

式中:$O(b)$是第$b$次迭代的目标函数值;$l(y_i, \hat{y}_i^{(b)})$是第$i$个样本的损失函数,用来衡量预测值$\hat{y}_i^{(b)}$和真实值$y_i$之间的差异;$K$代表模型中树的数量上限;$\Omega(f_k)$是树的复杂度惩罚项,用于控制模型的复杂度,其具体计算公式如下:

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \| \omega \|^2 \tag{5}$$

式中:$\gamma$为叶节点惩罚系数;$T$为树中叶节点的个数;$\lambda$为正则化惩罚系数;$\omega$为叶节点权重.

XGBoost使用梯度提升的方法来训练模型.在每一轮迭代中,计算损失函数的负梯度,这被视为残差$g_i$,然后用一棵新树来拟合这些残差.可以表

示为:

$$g_i = -\left[\frac{\partial l(y_i, \hat{y}_i)}{\partial \hat{y}_i}\right]_{\hat{y}_i = \hat{y}_i^{(t-1)}} \tag{6}$$

式中:$\hat{y}_i^{(t-1)}$是前$t$−1次迭代后的预测值.

在训练每一棵树时,XGBoost会寻找最优的分割点,以最小化目标函数.这涉及计算每个分割点的增益,增益可以用如下公式表示:

$$\text{Gain} = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right] - \gamma \tag{7}$$

式中:$G_L$和$G_R$分别是左子节点和右子节点的梯度之和;$H_L$和$H_R$分别是左子节点和右子节点的二阶导数之和;$\lambda$是L2正则化项的权重;$\gamma$是树的复杂度惩罚项的权重.

### 3.3　基于GA优化的XGBoost预测模型

遗传算法(GA)是一种启发式搜索算法,其受到生物进化的启发,通过模拟自然选择和遗传机制来优化问题解[25].将遗传算法应用于XGBoost模型的优化,可以帮助在参数空间中高效搜索最优或近似最优的参数组合,减少陷入局部最优的风险[26].

硅橡胶材料的击穿场强受多种因素(如配比、厚度、温度等)影响,这些因素之间可能存在复杂的非线性关系.GA和XGBoost的结合能够处理这种复杂性,从大量数据中提取出关键信息,建立准确的预测模型.此外,GA可以通过优化XGBoost的模型参数来减少过拟合的风险,同时提高模型的训练效率和预测性能,这对于处理硅橡胶材料击穿场强这类需要高精度预测的问题尤为重要.

基于此,将经过GA优化后的XGBoost算法应用于硅橡胶击穿场强的预测,通过遗传算法,为XGBoost模型的参数优化提供了一种高效的全局搜索方法,有利于对不同配比和厚度下硅橡胶的击穿场强进行精确的预测.其具体流程如图5所示.

个体编码:个体通常通过二进制编码表示为染色体,个体每个基因位编码从中按照等概率随机抽取.采用均方误差作为适应度函数,如式(8)所示,用于指导算法的搜索方向.

$$F(\theta) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{8}$$

式中:$y_i$是实际击穿场强;$\hat{y}_i$是模型预测的击穿场强;$n$是样本数量;$\theta$作为函数的输入(个体参数).

初始化种群:随机生成初始种群,每个个体代表一组XGBoost模型的参数,包括学习率、树的数量、树的深度等.此处将种群大小设置为50.

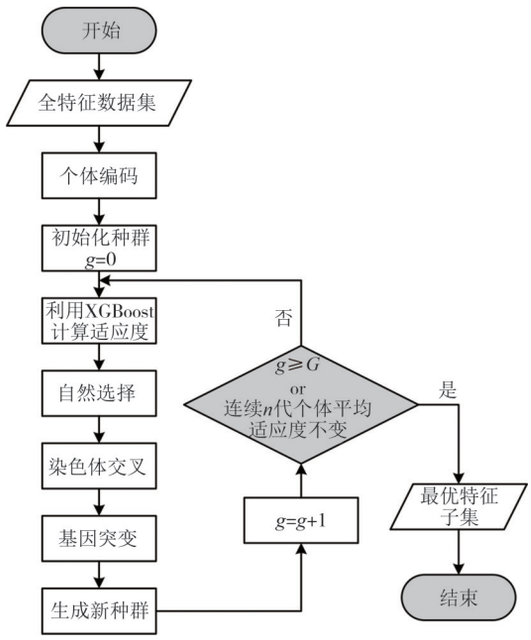



图5　算法流程图

Fig.5　Algorithm flow chart

选择和交叉:采用单点交叉法随机选择一个交叉点,交换交叉点之后的基因段,基于适应度函数的值来选择个体进行繁殖.其中个体$i$被选中的概率为$P(i)$,如式(9)所示:

$$P(i) = \frac{f(i)}{\sum_{j=1}^{N} f(j)} \tag{9}$$

式中:$P(i)$是个体$i$被选中的概率;$f(i)$是个体$i$的适应度;$N$是种群大小.

变异:通过交换两个个体的部分基因来产生新的个体,对新产生的个体执行变异操作,以增加种群的多样性.变异可以表示为$g'$,如式(10)所示:

$$g' = g + \delta \tag{10}$$

式中:$g$是原始基因值;$g'$是变异后的基因值;$\delta$是一个小的随机扰动.

迭代:用新生成的子代个体替换当前种群中的个体.反复执行以上操作,直到达到设置的最大迭代次数(300次).在最终的种群筛选中,挑选出适应度最为卓越的个体,从而锁定并获取针对特定问题的最佳特征子集.

### 3.4　超参数优化

通过遗传算法在对各超参数设定的合适范围内进行最优值搜索,得到如表3所示的XGBoost模型最优超参数.

表 3 超参数设置

Tab.3 Super parameter settings

| 超参数类型 | 取值范围 | 最优参数 | 参数含义 |
|---|---|---|---|
| learning_rate | 0.01~0.3 | 0.28 | 学习率 |
| max_depth | 3~10 | 9.48 | 树的最大深度 |
| min_child_weight | 1~10 | 3.20 | 子节点中所需的最小权重之和 |
| gamma | 0~0.5 | 0.12 | 分裂节点所需最小损失减少阈值 |
| subsample | 0.5~1 | 0.50 | 训练每棵树时采样的行比例 |
| colsample_bytree | 0.5~1 | 0.80 | 训练每棵树时采样的列比例 |
| n_estimators | 50~200 | 80 | 要训练的树的数量 |
| lambda | 0~10 | 3.96 | L2 正则化项的权重 |
| alpha | 0~10 | 1.11 | L1 正则化项的权重 |

### 3.5 模型评估指标

本文采用3种常见的统计学指标分析预测模型性能,分别是均方根误差(root-mean-square error, RMSE),用于度量预测值与实际值之间的差异;平均绝对误差(mean absolute error, MAE)表示预测值与实际值之差的绝对值的平均数,MAE提供了预测误差的一个直观理解;$R^2$称为决定系数,其值的范围为0~1,值越接近1表示模型拟合效果越好[27].

$$\mathrm{RMSE}=\sqrt{\mathrm{MSE}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \tag{11}$$

$$\mathrm{MAE}=\frac{1}{n}\left|y_i-\hat{y}_i\right| \tag{12}$$

$$R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{13}$$

式中:$y_i$是第$i$个观测的实际值;$\hat{y}_i$第$i$个观测的预测值;$n$是观测的数量;$\bar{y}$是实际值的平均值.

### 3.6 模型预测结果与分析

根据前期试验结果,经数据预处理后得到数据372组,设置随机种子数为100,以7∶3的比例划分为训练集与测试集.硅橡胶材料击穿场强试验值和预测值关系如图6所示,其中,测试集的$R^2$=0.953、MAE=0.168、RMSE=0.361 kV/mm.

上述结果表明,针对位于16~22 kV/mm的常见值预测比较准确,在该范围内模型训练后具有较高的预测精度,尽管模型对个别工况点的预测结果稍差,导致预测结果的线性拟合直线有所偏离,但从整体来看,模型对各参数的预测值和试验值的相关性较好且预测精度较高.在预测极端值(过大或过小)时存在一定偏差,这是因为训练数据中这些极端值出现的频率较低,使得模型在训练过程中给予的权重也较小,因此模型更倾向于预测更常见的数值范围.

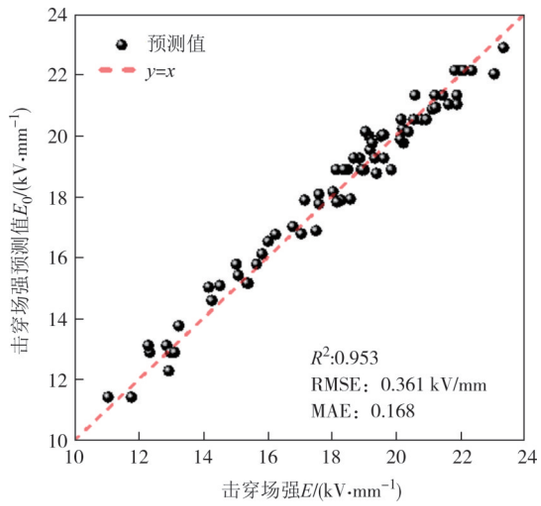



图6 模型预测结果

Fig.6 Model prediction results

为验证GA-XGBoost模型对硅橡胶击穿场强的预测效果,分别采用决策树、BP神经网络、线性回归、SVR和未优化的XGBoost模型与之进行对比,进一步验证优化后模型的有效性.为保证所有预测模型结果的可比性,训练集和测试集数据需保持一致.各算法性能对比结果如表4所示.

表 4 不同算法评估指标

Tab.4 Evaluation indicators of different algorithms

| 算法 | MAE | RMSE/(kV·mm$^{-1}$) | $R^2$ |
|---|---|---|---|
| 决策树 | 0.307 | 0.891 | 0.908 |
| BP神经网络 | 0.635 | 1.207 | 0.857 |
| 线性回归 | 0.436 | 1.024 | 0.903 |
| SVR | 0.289 | 0.648 | 0.912 |
| XGBoost | 0.211 | 0.483 | 0.920 |
| GA-XGBoost | 0.168 | 0.361 | 0.953 |

由表4可得,本研究提出的GA-XGBoost方法可以有效提高大部分数据集的预测精度,且与另外5种算法相比,3种评价指标均有明显提升.其中,GA-XGBoost的决定系数$R^2$与其他算法相比提升了3.59%~11.20%,表明该优化模型预测精度较高;均方根误差与其他算法相比降低了25.26%~70.09%,说明预测值与真实值之间的偏差相较于其他方法更低;平均绝对误差与其他算法相比降低了20.38%~73.54%,同样说明了预测值与真实值之间的平均差

异程度降低.

结果表明,经过遗传算法优化的XGBoost模型提高了对数据集的预测精度.此外,BP神经网络相较于其他算法误差较大,原因可能为神经网络需要通过大量数据进行模型训练,导致算法的准确度和解释性相差较大,鉴于本数据集的样本数量相对较少,更适合那些对数据集需求不高的算法.

## 4 结 论

本文以高温硫化硅橡胶材料为研究对象,基于试验击穿场强数据,利用遗传算法良好的寻优能力,确定了XGBoost算法的超参数,获得了基于硅橡胶材料的具有普适性的击穿场强预测模型,具体结论如下:

1)通过对硅橡胶制备过程中的关键参数进行异常值处理和皮尔逊相关系数分析,可知:厚度和温度对击穿场强有显著影响,色母相对含量影响较小,$Al(OH)_3$微粉直径和相对含量的影响也符合预期.厚度与击穿场强呈现出显著的负相关性,更厚的材料由于缺陷或杂质的存在可能导致击穿场强降低;温度与击穿场强之间也表现出负相关性,温度升高可能会导致导电性和离子移动性的增加,从而降低击穿场强;$Al(OH)_3$相对含量与击穿场强呈现正相关性,$Al(OH)_3$作为阻燃填料可能通过改善材料的热稳定性和电气绝缘性能对击穿场强产生正面影响.

2)采用改进XGBoost算法模型进行硅橡胶击穿场强预测,测试集的RMSE=0.361 kV/mm、MAE=0.168、$R^2$=0.953,预测精度优于线性回归、决策树、SVR、BP神经网络以及未优化的XGBoost算法.该方法有效提升了对数据集的预测能力,依托该模型,可较为准确地预测一定配比及厚度下硅橡胶的击穿场强,对硅橡胶材料制备及绝缘性能预测具有重要意义.

3)本文所采用的GA-XGBoost算法虽然能够在一定程度上表明不同硅橡胶配方与其击穿场强的相关性,但还需要进一步进行试验验证和考虑其他因素(老化程度、材料内部缺陷或杂质等)对其击穿场强的影响,后续将继续开展改进工作,以更好地提升其预测性能.

## 参考文献


[1] 王小龙, 刘亨阳, 顾哲屹, 等. 输电线路用的硅橡胶材料及老化问题研究[J]. 化学与粘合, 2024, 46(1):49-52.
WANG X L, LIU H Y, GU Z Y, et al. Study on the silicone rubber materials and aging problems for transmission lines [J]. Chemistry and Adhesion, 2024, 46(1): 49-52. (in Chinese)

[2] 刘兴杰,沈瑶,贾方瑞, 等. 纳米$SiO_2$/硅橡胶复合材料热裂解进程的分子模拟[J]. 高电压技术, 2024, 50(12):5397-5405.
LIU X J, SHEN Y, JIA F R, et al. Molecular simulation of thermal cracking process of nano-$SiO_2$/silicone rubber composites [J]. High Voltage Engineering, 2024, 50(12): 5397-5405. (in Chinese)

[3] 毕茂强, 帅垭灵, 邓润坤, 等. 雾霾参数对电晕老化硅橡胶外绝缘特性影响研究[J]. 重庆理工大学学报(自然科学), 2023, 37(2): 251-259.
BI M Q, SHUAI Y L, DENG R K, et al. Study on the influence of haze parameters on the external insulation characteristics of corona-aged silicone rubber [J]. Journal of Chongqing University of Technology (Natural Science), 2023, 37(2): 251-259. (in Chinese)

[4] 毕茂强,杨忠平,江天炎,等.雾霾条件下不同配比硅橡胶电晕老化特性[J].湖南大学学报(自然科学版),2022,49(12):167-175.
BI M Q, YANG Z P, JIANG T Y, et al. Corona aging characteristics of silicone rubber with different proportions under haze conditions [J]. Journal of Hunan University (Natural Science), 2022, 49(12):167-175. (in Chinese)

[5] 代冲. 低温下硅橡胶绝缘材料电气性能研究[D]. 北京:华北电力大学,2017.
DAI C. Study on electrical properties of silicone rubber insulation material at low temperature [D]. Beijing: North China Electric Power University, 2017. (in Chinese)

[6] 胡建林,谭陈,钟睿,等. 室温硫化硅橡胶涂层厚度对电晕老化特性的影响[J]. 高电压技术, 2020, 46(4): 1328-1336.
HU J L, TAN C, ZHONG R, et al. Influence of the thickness of room temperature vulcanized silicone rubber coatings on the aging characteristics of corona [J]. High Voltage Engineering, 2020, 46(4):1328-1336. (in Chinese)

[7] 周远翔,侯非,聂琼,等. 温度对硅橡胶电树枝老化特性的影响[J]. 高电压技术,2012,38(10):2640-2646.
ZHOU Y X, HOU F, NIE Q, et al. Temperature effects on electrical tree aging characteristics of silicone rubber [J]. High Voltage Engineering, 2012, 38(10):2640-2646. (in Chinese)

[8] NANDI A K. Genetic algorithm-based design and development of particle-reinforced silicone rubber for soft tooling process [J]. Materials and Manufacturing Processes, 2013, 28(7):753-760.

[9] SHAYMAA A Q, SAUD A A, KAREEM A. Grain size effects on the behavior of silicone rubber high voltage power cables using seagull optimization algorithm [J]. The Indonesian Journal of Electrical Engineering and Computer Science, 2023, 32(3): 1246-1256.

[10] CHEN Z Q, JIA W, CHENG L, et al. Prediction of breakdown field strength for large-area and multilayer film dielectrics [J]. IEEE Transactions on Dielectrics and Electrical Insulation,

2022, 29: 470-477.

[11] 许旭日, 高宇, 李敬,等. 基于等温表面电位衰减法预测不同厚度的聚乙烯纳米复合材料直流击穿场强[J]. 高电压技术, 2022,48(7): 2864-2872.
XU X R, GAO Y, LI J, et al. Prediction of thickness-dependent DC breakdown strength of polyethylene nanocomposite estimated with surface potential decay measurement [J]. High Voltage Engineering,2022,48(7):2864-2872.(in Chinese)

[12] LI Z L, DONG Y M, WU Y, et al. Breakdown performance evaluation and lifetime prediction of XLPE insulation in HVAC cables[J]. Energies, 2024, 17(6): 1337.

[13] 王宏旭,李源,林海泉, 等. 低温下硅橡胶直流击穿特性变化规律研究[J]. 高压电器,2019, 55(3):103-108.
WANG H X, LI Y, LIN H Q, et al. Study on change regularity of DC breakdown characteristics of silicone rubber at low temperature [J]. High Voltage Apparatus, 2019,55(3):103-108.(in Chinese)

[14] 国家质量监督检验检疫总局, 中国国家标准化管理委员会. 绝缘材料电气强度试验方法 第1部分: 工频下试验:GB/T 1408.1—2016[S]. 北京:中国标准出版社, 2016.
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Insulating materials for electrical strength-test methods-Part 1: tests at power frequencies: GB/T 1408.1—2016 [S]. Beijing: Standards Press of China,2016.(in Chinese)

[15] BI M Q, LI K, ZHANG S Y, et al. Electric field simulation analysis of ICOG corona ring and study on flashover characteristics of bird droppings [J]. Electric Power Systems Research,2024,233:110459.

[16] 王若丞,贺云逸,康洪玮,等. 电缆接头绝缘用硅橡胶热老化及超声特性[J]. 高电压技术,2021,47(9): 3181-3188.
WANG R C, HE Y Y, KANG H W, et al. Thermal aging and ultrasonic characteristics of silicone rubber for cable joint insulation[J]. High Voltage Engineering, 2021, 47(9): 3181-3188.(in Chinese)

[17] 杨涛,郝建,陈鑫, 等. 三元混合式绝缘油吸潮、工频击穿及产气特性研究[J]. 高电压技术,2020,46(11):4014-4022.
YANG T, HAO J, CHEN X, et al. Study on moisture absorption, power frequency breakdown and gas production characteristics of three-element mixed insulation oil[J]. High Voltage Engineering, 2020, 46(11): 4014-4022.(in Chinese)

[18] 国家质量监督检验检疫总局,中国国家标准化管理委员会. 硫化橡胶 工频击穿电压强度和耐电压的测定方法:GB/T 1695—2005[S]. 北京: 中国标准出版社,2006.
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Rubber, vulcanized—determination of electrical breakdown strength and voltage resistant at commercial power frequency: GB/T 1695—2005[S]. Beijing:Standards Press of China,2006.(in Chinese)

[19] 高建,严智民,李建英,等.温度对硅橡胶电树枝自恢复特性的影响机制[J].西安交通大学学报, 2022, 56(9):176-184.
GAO J, YAN Z M, LI J Y, et al. The influence mechanism of temperature on the self-recovery characteristics of silicone rubber electrical tree[J]. Journal of Xi'an Jiaotong University, 2022, 56(9): 176-184.(in Chinese)

[20] 王霞,王华楠,陈飞鹏,等. 电晕放电加速硅脂劣化对硅橡胶绝缘表面性能的影响[J]. 高电压技术,2020, 46(11):3977-3985.
WANG X, WANG H N, CHEN F P, et al. Effect of corona-discharge-accelerated deterioration of silicone grease on the surface properties of silicone rubber insulation[J]. High Voltage Engineering,2020,46(11):3977-3985.(in Chinese)

[21] DAS M K, CHAUDHARY A, BRYAN A, et al. Rapid screening evaluation of SARS-CoV-2 IgG assays using Z-scores to standardize results [J]. Emerging Infectious Diseases, 2020, 26(10): 2501-2503.

[22] ZHANG L, WANG F L, SUN T, et al. A constrained optimization method based on BP neural network[J]. Neural Computing and Applications,2018,29(2): 413-421.

[23] OMER S G, LIOR R K, Approximating XGBoost with an interpretable decision tree [J]. Information Sciences, 2021,572: 522-542.

[24] 王晨光,秦浩然, 杨超越, 等. 基于XGBoost算法的冻土强度预测与影响因素分析[J/OL]. 金属矿山, 2024:1-14.(2024-05-06)[2024-06-04]. http://kns.cnki.net/KCMS/detail/detail.aspx?filename=JSKS20240428002&dbname=CJFD&dbcode=CJFQ.
WANG C G, QIN H R, YANG C Y, et al. Prediction of frozen soil strength and analysis of influencing factors based on XGBoost algorithm[J/OL]. China Industrial Economics,2024:1-14.(2024-05-06)[2024-06-04]. http://kns.cnki.net/KCMS/detail/detail.aspx?filename=JSKS20240428002&dbname=CJFD&dbcode=CJFQ.(in Chinese)

[25] 王燎原,孙玉利,张桂冠,等. 基于Taguchi-GA协同的磁性磨料抛光性能预测及制备工艺参数寻优[J]. 湖南大学学报(自然科学版),2024,51(4):43-53.
WANG L Y, SUN Y L, ZHANG G G, et al. Predicting polishing performance of magnetic abrasive and optimizing its preparation process parameters based on Taguchi-GA synergy[J]. Journal of Hunan University (Natural Sciences), 2024, 51(4): 43-53.(in Chinese)

[26] NGUYEN V L, SANG W H, KIHAK L, Genetic algorithm hybridized with extreme gradient boosting to predict axial compressive capacity of CCFST columns [J]. Composite Structures, 2021,278:114733.

[27] TOMAL J H, WELCH W J, ZAMAR R H. Robust ranking by ensembling of diverse models and assessment metrics[J]. Journal of Statistical Computation and Simulation, 2023, 93(1):77-102.