

DOI: 10.12158/j.2096-3203.2026.03.005

基于 Temporal Fusion Transformer 模型的变压器 油中溶解气体预测方法

周延豪^{1,2}, 范路², 任海龙^{1,2}, 赵谡², 王亚林², 尹毅²

(1. 上海电力大学电气工程学院, 上海 200090; 2. 上海交通大学电气工程系, 上海 200240)

摘要:油中溶解气体是评估变压器运行状态的重要指标,准确预测油中溶解气体的发展趋势有助于预防电力变压器故障。为解决传统预测模型中单一变量造成的预测效率低下,文中提出一种基于 Optuna 超参数优化的 Temporal Fusion Transformer (TFT)模型。通过引入变压器组别、绕组相别、气体类别等静态变量以及可解释性的多头注意力机制,实现多组变压器油中溶解气体的同步预测,提升变电站运维系统的预警效率。相比于传统预测模型,文中模型预测的平均相对误差仅为 0.306%,较 Transformer 模型降低了 66.7%,且在短期和长期预测时均具有更高的预测准确度。此外,文中模型的训练时间仅为 Transformer 模型的 1/4,更契合当前智能预警平台中多组别设备同步预测的发展趋势。模型中的多头注意力机制表明氢气和甲烷之间以及二氧化碳和甲烷之间具有强相关关系,其与油纸绝缘裂解的产气规律相一致,进一步表明文中模型具有良好的可解释性,可为多组别设备同步预测提供技术保障。

关键词:电力变压器;油中溶解气体;同步预测;Temporal Fusion Transformer (TFT)模型;时间序列;注意力机制

中图分类号:TM854

文献标志码:A

文章编号:2096-3203(2026)03-0037-09

0 引言

电力变压器油中溶解气体分析(dissolved gas analysis, DGA)是监测评估变压器内部状态的关键技术手段^[1-3],在电力变压器发生电学故障或热故障时,油纸绝缘将降解生成多种气体分子。通过探究油中溶解气体的发展规律,预测油中溶解气体的浓度发展趋势,能够实现变压器潜在故障的及时预警^[4-8],从而为变压器的健康管理和维护决策提供科学依据^[9-12]。

目前主要采用统计方法和人工智能算法预测变压器油中溶解气体浓度。统计方法依赖相邻数据间的关联性来预测未来的发展趋势^[13],常用的统计模型包括移动平均模型和自回归积分滑动平均模型等^[14]。变压器的油色谱受自然环境、运行负荷以及放电故障等多重因素的影响,数据发展的长期平稳性较差,相邻数据间关联性低,因此统计模型在长期预测时的跟随性较差。人工智能算法以其良好的非线性拟合能力实现对数据非平稳波动的及时响应^[15-17]。文献[18]采用长短期记忆(long short-term memory, LSTM)网络显著提升了气体浓度预测的准确性和跟随性。文献[19]在 LSTM 模型的基础上引入了特征和时序双重注意力机制,通过挖掘特征变量与时间依赖之间的关联关系,进一

步提高了不同波动状态下的预测性能。文献[20]基于 Transformer 编码器构建了一种多阶段暂态稳定评估方法,通过多时刻信息的特征构建将电力系统暂态失稳的漏判率降低了 22%,注意力权重的可视化增强了模型的可解释性。文献[21]结合卷积神经网络与 Transformer 构建预测模型,所提模型在电力系统暂态分量受到扰动时仍能准确预测频率稳定指标,且在时效性和准确性上均优于传统方法。尽管 LSTM 和 Transformer 等深度模型在短期时间序列预测中表现优异,但当序列长度增大时,历史信息的长期累积效应使得模型难以实现明确的输入归因,从而削弱了预测结果的精度。

为提高非平稳气体序列的预测精度,不少学者提出了“分解-预测-重构”的组合模型框架。该类方法先对原始气体序列进行信号分解以削弱非平稳性,再预测各气体分量,最后将预测结果叠加重构^[22-24]。文献[25]将变分模态分解(variational mode decomposition, VMD)与自回归积分滑动平均模型(auto-regressive integrated moving average model, ARIMA)相结合,并采用智能优化算法调整参数,实现了对油中溶解气体序列的高效预测;文献[26]利用黏菌算法优化的 VMD,结合门控循环单元,使氢气浓度预测的平均绝对百分比误差低至 0.636%,显著提高了预测精度。另外,改进的自适应噪声完备集合经验模态分解(improved complete ensemble empirical mode decomposition with adaptive noise, ICE-

收稿日期: 2025-07-30; 修回日期: 2025-10-16

基金项目: 国家自然科学基金资助项目(52407024)

EMDAN)结合粒子群优化算法优化的极限学习机模型,以及集合经验模态分解(complete ensemble empirical mode decomposition, CEEMD)结合时间卷积网络的模型,也在油中气体浓度预测中表现出优异的精度和可靠性^[27-28]。这些多模块组合方法流程复杂、参数众多,模型构建与参数调优困难。相比之下,文中模型无需任何信号分解预处理环节,即可取得相当或更优的预测精度,大大简化了预测流程。

在实际电力运维场景中,通常需要同时监测多台变压器的状态,并对群体运行趋势进行综合评估。然而,现有研究多着眼于单台设备预测,难以满足多组变压器油中溶解气体同步预测的迫切需求。为此,文中提出一种由 Optuna 算法优化的 Temporal Fusion Transformer (TFT) 预测模型。TFT 作为一种融合了注意力机制的深度学习架构,具有同时建模多组别时间序列并行动态的能力,并通过引入静态协变量编码等专门机制,使模型能够同时捕捉油中溶解气体序列的静态特征和动态演变规律,实现多组变压器油中溶解气体的同步高精度预测。借助 Optuna 超参数优化框架,模型的参数配置得以自动调优,进一步提升了训练收敛效率和预测精度。在河北某变电站的油色谱预测任务中,文中模型的训练时长仅为传统模型的 1/4,且 30 d 预测的平均相对误差为 0.306%。

1 基于 Optuna 优化的 TFT 模型架构

TFT 模型作为 Transformer 模型的改进架构,通过整合多种类型的输入数据(如静态特征、历史数据和已知未来特征)对时间序列进行建模^[29]。TFT 模型引入多头注意力机制进行序列信息的全局关注,并通过门控残差网络(gated residual network, GRN)对特征进行快速处理,确保信息在多层之间的有效传递。GRN 在加速 TFT 模型收敛中起到关键作用,负责对模型的输入特征进行选择性地处理,并根据需要抑制不必要的非线性贡献,从而使模型能够根据数据的特点进行自适应处理。

GRN 模块接收 \mathbf{a} 、 \mathbf{c} 两个输入矩阵并输出 ξ , 其传递函数 F_{GRN} 为:

$$\xi = F_{GRN}(\mathbf{a}, \mathbf{c}) = F_{LayerNorm}(\mathbf{a} + F_{GLU,W}(\eta_1)) \quad (1)$$

$$\eta_1 = W_{1,\omega}\eta_2 + b_{1,\omega} \quad (2)$$

$$\eta_2 = F_{ELU}(W_{2,\omega}\mathbf{a} + W_{3,\omega}\mathbf{c} + b_{2,\omega}) \quad (3)$$

$$F_{GLU,W}(\eta_1) = \sigma(W_{4,\omega}\eta_1 + b_{4,\omega}) \odot (W_{5,\omega}\eta_1 + b_{5,\omega}) \quad (4)$$

式中: $\sigma(\cdot)$ 为 S 型生长曲线函数; $F_{GLU,W}$ 为 \mathbf{W} 权重矩阵的门控线性单元; $F_{ELU}(\cdot)$ 为指数线性单元激活函数; η_1 、 η_2 为经 F_{ELU} 线性激活函数处理的中间特征

变量; $F_{LayerNorm}$ 为标准层归一化函数; $\mathbf{W}_{y,\omega}$ 为第 y 个线性映射的第 ω 组权重矩阵, $y=1,2,3,4,5$; $\mathbf{b}_{y,\omega}$ 为第 y 个线性映射的第 ω 组偏置变量列矩阵; \odot 为哈达玛积。

GRN 通过残差连接将初始输入与经过非线性处理的特征进行加权组合,最后通过 $F_{LayerNorm}$ 函数进行标准化处理。TFT 的网络架构如图 1 所示,其主要包括变量选择网络、静态协变量编码器、序列编码器、时序解码器等部分。 X_t 、 X_{t+1} 分别为 t 和 $t+1$ 时刻的动态变量; c_s 为变量选择网络得到的时间特征筛选结果变量; c_c 、 c_h 为时间特征的局部处理变量; c_e 为用静态信息丰富时间特征变量。图中虚线表示变量之间的间接关系,模型采用分位数预测的形式预测 $t+\tau$ 时不同分位数 q 的输出 $\hat{y}_{t+\tau}(q)$, 其中 τ 为未来时间步的预测长度。

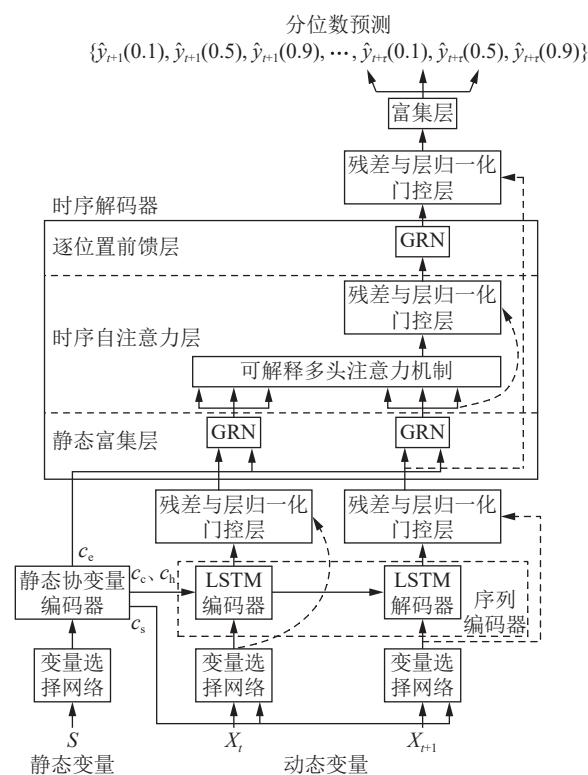


图 1 TFT 网络架构
Fig.1 Network architecture of TFT

1.1 变量选择网络

在多变量时间序列预测中,并非所有特征对于模型的预测都有相同的重要性。变量选择网络通过在每个时间步上对输入变量进行加权,帮助模型选择最有价值的特征。变量选择网络结构如图 2 所示,其中, $\xi_{i,j}$ 、 $\hat{\xi}_{i,j}$ 分别为 t 时刻第 j 个特征向量和经过 GRN 模块处理后的特征向量; Ξ_t 为 $\xi_{i,j}$ 的集合向量; \mathbf{v}_{x_i} 为第 x_i 个变量的选择权重; $\hat{\xi}_t$ 为输入特征向量。

对于每一个 t 时刻的变量, TFT 模型利用 GRN

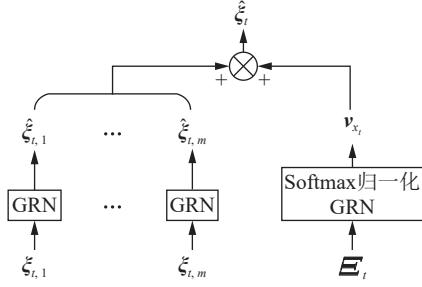


图2 变量选择网络

Fig.2 Variable selection network

模块进行非线性处理, 计算式为:

$$\hat{\xi}_{t,j} = F_{\text{GRN}}(\xi_{t,j}) \quad (5)$$

所有 ξ_t 与 c_s 通过 Softmax 归一化指数函数处理后生成第 x_t 个变量的选择权重 v_{x_t} :

$$v_{x_t} = \text{Softmax}(F_{\text{GRN}}(\xi_t, c_s)) \quad (6)$$

将所有处理后的特征根据变量选择权重 v_{x_t} 进行加权组合, 得到输入特征向量 $\hat{\xi}_t$ 。

$$\hat{\xi}_t = \sum_{j=1}^m v_{x_t} \hat{\xi}_{t,j} \quad (7)$$

式中: m 为模型中的特征向量数量。

1.2 静态协变量编码器

在 TFT 模型中, 静态协变量编码器用于处理不随时间变化的静态特征(如变压器组别、绕组相别等)。静态协变量编码器使用 GRN 编码器生成 4 个不同的静态协变量, 具体为 c_s 、 c_c 、 c_h 和 c_e 。这些静态协变量通过 GRN 与时间序列中的动态特征结合, 充分利用静态特征所包含的信息。

1.3 可解释性多头注意力机制

在 TFT 模型中, 通过引入可解释性多头注意力机制提升模型的解释性。相比传统多头注意力机制, 常见的注意力机制通过获取查询变量 Q 和键变量 K 之间的相关性权重确定注意力强度 $F_{\text{Attention}}(Q, K, V)$ 。

$$F_{\text{Attention}}(Q, K, V) = A(Q, K)V \quad (8)$$

$$A(Q, K) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_{\text{att}}}}\right) \quad (9)$$

式中: $A(\cdot)$ 为归一化函数; V 为键值; d_{att} 为 K 的维度。

TFT 通过引入多头注意力机制, 对每个注意力头的输出权重进行加权求和, 使得该模型能够并行处理不同特征子空间, 从而增强其特征选择的能力。多头注意力机制的计算式为:

$$\beta(t, n) = \tilde{H}W_H \quad (10)$$

其中:

$$\tilde{H} = \frac{1}{H} \sum_{h=1}^H F_{\text{Attention}}(QW_{Q,h}, KW_{K,h}, VW_{V,h}) \quad (11)$$

式中: $\beta(t, n)$ 为 t 时刻对第 n 个特征的注意力输出; $W_{Q,h}$ 、 $W_{K,h}$ 、 $W_{V,h}$ 分别为 Q 、 K 、 V 的第 h 个注意力头的线性映射权重矩阵, 每个头使用独立的权重矩阵学习特征表示; H 为注意力头数。最后, 通过注意力头数为 H 的注意力矩阵 W_H 将各个注意力头的输出集进行线性组合。

1.4 序列编码器

序列编码器的主要组成部分是 LSTM 编码器-解码器。LSTM 编码器接收的输入向量序列为: $\xi_{t-k:t} = [\xi_{t-k} \ \xi_{t-k+1} \ \cdots \ \xi_t]$, 其中 t 为当前时刻, k 为历史回溯窗口长度。编码器生成的隐状态变量 h_t 传递给解码器, 解码器将其与未来时间步的输入 $\xi_{t+1:t+\tau}$ 结合, 生成预测结果。更新方法如下:

$$h_t = F_{\text{LSTMEncoder}}(\xi_{t-k:t}) \quad (12)$$

$$\hat{\xi}_{t+1:t+\tau} = F_{\text{LSTMDecoder}}(h_t, \xi_{t+1:t+\tau}) \quad (13)$$

式中: $F_{\text{LSTMEncoder}}$ 、 $F_{\text{LSTMDecoder}}$ 分别为 LSTM 编码器与解码器程序; $\hat{\xi}_{t+1:t+\tau}$ 为未来时间步的预测结果。

TFT 在序列编码层和时序融合解码器层之间引入了门控层, 用于控制输入特征的选择性激活:

$$\hat{\phi}(t, n) = F_{\text{LayerNorm}}(\hat{\xi}_{t+n} + F_{\text{GLU}, \phi}(\phi(t, n))) \quad (14)$$

式中: $\phi(t, n)$ 、 $\hat{\phi}(t, n)$ 分别为 t 时刻第 n 个特征的中间表示和门控融合后的特征输出; $F_{\text{GLU}, \phi}$ 为 $\phi(t, n)$ 的门控线性单元。

1.5 时序解码器

在 TFT 模型中, 时序融合解码器通过整合静态富集层和时序自注意力层等模块, 学习数据集中不同时间步长的时序关系, 并采用分位数预测的形式预测不同分位数 q 的输出 $\hat{y}_{t+\tau}(q)$, 其传递函数为:

$$\theta(t, n) = F_{\text{GRN}}(\hat{\phi}(t, n), c_e) \quad (15)$$

$$\delta(t, n) = F_{\text{LayerNorm}}(\theta(t, n) + F_{\text{GLU}, \delta}(\beta(t, n))) \quad (16)$$

$$\psi(t, n) = F_{\text{GRN}}(\delta(t, n)) \quad (17)$$

$$\hat{y}_{t+\tau}(q) = W_q \psi(t, \tau) + b_q \quad (18)$$

式中: $\theta(t, n)$ 为 t 时刻对第 n 个特征的处理结果; $\delta(t, n)$ 为经过 $F_{\text{LayerNorm}}$ 程序处理后的 t 时刻对第 n 个特征的处理结果; $F_{\text{GLU}, \delta}$ 为 $\delta(t, n)$ 的门控线性单元; $\psi(t, n)$ 为 GRN 模块在 t 时刻对 $\delta(t, n)$ 的处理结果; W_q 为第 q 个预测任务的权重矩阵, 用于加权输入特征; b_q 为第 q 个预测任务的偏执项。通过分位数预测方法, 模型不仅能够提供预测值, 还能够提供对应的置信区间, 从而增加预测结果的准确度。

1.6 Optuna 超参数优化

文中模型采用 Optuna 算法实现 TFT 模型的超参数优化。Optuna 算法通过动态构建和调整超参

数搜索空间,实现在复杂的优化任务中自适应地选择合适的超参数,从而避免传统方法中静态定义搜索空间的局限。Optuna 算法在优化过程包含以下关键步骤:(1)利用 Define-by-Run 编程范式,在更新期间对超参数的搜索空间进行动态构建和调整,提升搜索效率;(2)利用树结构帕尔森估计器(tree-structured Parzen estimator, TPE)和协方差矩阵自适应进化策略(covariance matrix adaptation evolution strategy, CMA-ES)等优化策略加速搜索效率;(3)引入自适应同步分层自动(adaptive synchronous hierarchical auto, ASHA)机器学习剪枝机制^[30],在每次更新中,Optuna 实时评估中间结果,并通过剪枝机制提前淘汰表现不佳的超参数组合,避免无效计算。文中模型的初始超参数设置见 3.1 节。

2 实验设置

文中将两台 ODFPS-1 000 000 kV·A/1 000 kV 型变压器的实际油色谱气体监测数据作为样本,变压器容量为 1 000/1 000/334 MV·A,电压等级为 1 050/(525±4×1.25%)/110 kV,数据时间范围为 2023 年 3 月 22 日至 2024 年 2 月 16 日,共计 300 d。油色谱数据包括氢气、甲烷、乙烷、乙烯、总烃、一氧化碳和二氧化碳。这两台变压器在运行过程中均出现过较明显的 DGA 异常波动,具有代表性故障特征,适合用于算法的预测及可解释性验证。文中模型将前 270 d 的气体含量作为训练集,后 30 d 的气体含量作为测试集。

鉴于原始数据量有限,文中未划分独立验证集,而是依赖 Optuna 框架内置的 ASHA 机制来进行超参数选择。Optuna 提供的 Optuna SearchCV 等接口支持基于交叉验证的超参数搜索,可在不减少训练数据量的前提下评估模型性能^[30]。同时,Optuna 的剪枝机制能够在每轮超参数评估过程中动态划分内部训练集,并对性能不佳的参数组合提前中止训练,可在无需额外验证集的情况下充分利用全部数据进行模型训练,从而有效提高超参数选择的精度并抑制过拟合风险。

2.1 数据预处理

为了同时预测多台变压器及其三相油色谱信息,文中将变压器组别与绕组相别统一视为静态协变量并嵌入 TFT 模型。表 1 为各类输入特征变量及其处理方式。变压器组别、绕组相别和气体类别等静态特征均采用 one-hot 编码进行处理,二者匹配时,编码值为 1;若不匹配,则编码值为 0。这有助于提高模型对周期性波动特征的捕捉能力。对

于气体含量、星期、月份等动态变量,则通过将其归一化到(-1,1)区间,在一定程度上放大周期性信号的区分度,从而使网络更易捕捉序列可能存在的周期波动特性。

表 1 变量处理方式
Table 1 Variable handling mode

变量类型	变量	处理方式
静态变量	变压器组别	one-hot编码
	绕组相别	one-hot编码
	气体类别	one-hot编码
时序变量	各种气体含量	归一化
	节假日	归一化
	星期	归一化
	月份	归一化

2.2 评估标准

为评估 Optuna-TFT 算法的预测效果,文中选用平均绝对误差、平均相对误差、最大相对误差与决定系数作为评价指标。

$$e_{MAE} = \frac{1}{n_c} \sum_{i=1}^{n_c} |y_i - \hat{y}_i| \quad (19)$$

$$e_{MAPE} = \frac{1}{n_c} \sum_{i=1}^{n_c} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (20)$$

$$e_{MRE} = \max_i \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (21)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n_c} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_c} (y_i - \bar{y})^2} \quad (22)$$

式中: e_{MAE} 为平均绝对误差值; e_{MAPE} 为平均相对误差值; e_{MRE} 为最大相对误差值; R^2 为决定系数; n_c 为输入的变量个数; y_i 、 \hat{y}_i 分别为第*i*个变量的真实值和预测值; \bar{y} 为输入变量的平均值。

实验采用 Intel Xeon(R) Platinum 8255C CPU 和 RTX 2 080 Ti (22 GB) GPU 作为硬件平台,模型构建基于 Python 3.9 和 Pytorch 平台。

3 算例分析

为验证文中模型的优势,文中对单一气体预测和多组别变压器油色谱的同步预测结果进行分析。

3.1 Optuna 超参数优化

文中采用 Optuna 算法进行超参数优化,其中超参数包含学习率、隐藏层大小、注意力头数量和丢弃率。以 1 号变压器 A 相的总烃气体含量预测为例,图 3 反映不同超参数对 e_{MAE} 值的影响。动态优化后的超参数组合获得最低的 e_{MAE} 值,此时学习率

为 0.01、隐藏层大小为 64、注意力头数为 4、丢弃率为 0.1。文中模型选用该超参数组合进行后续预测。

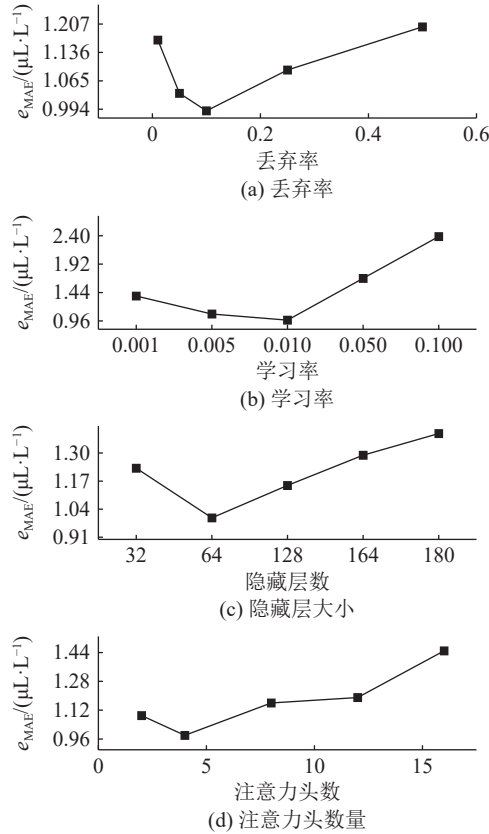


图 3 不同超参数对 e_{MAE} 的影响

Fig.3 Influence of different hyperparameters on e_{MAE}

3.2 单一气体预测结果

以 1 号变压器 A 相为例, 不同模型的预测结果如表 2 所示, 文中模型对所有特征气体的预测误差均低于对比模型, 表明 Optuna-TFT 模型具备更高的预测准确性。其中 Optuna-TFT 模型氢气浓度的预测误差为 Transformer 模型的 1/3。

图 4 为总烃含量的预测曲线, 可以看到文中模型能够更为准确地捕捉原始数据的变化趋势, 表明文中模型在学习时间序列的长短期依赖关系方面具有显著优势, 能够更有效地适应溶解气体浓度的动态变化。

上述结果表明 Optuna-TFT 模型通过动态变量选择、静态协变量编码及全局上下文搜索, 实现了多种复杂气体特征的精准预测, 其在变压器的油中溶解气体预测方面优于传统模型。

3.3 多组别变压器同步预测结果

3.3.1 同步预测准确性与时效性

不同组别的油色谱变化规律呈现出不同的增长或波动模式, 若分别对其训练, 将导致预测模型效率较低, 且无法掌握不同组别间的关联性。文中模型将变压器组别和绕组相别信息作为静态变量,

表 2 不同模型预测误差对比

Table 2 Comparison of prediction errors of different models

特征气体	$e_{MAE}/\%$			
	文中模型	未优化TFT	Transformer	LSTM
氢气	1.105	2.036	3.574	4.782
甲烷	0.976	1.342	2.817	3.712
乙烷	0.872	1.103	1.331	2.243
乙烯	2.435	2.724	2.812	4.422
总烃	0.499	1.032	1.421	1.326
一氧化碳	1.056	1.583	2.681	4.632
二氧化碳	1.132	2.067	2.559	3.764

特征气体	$e_{MRE}/\%$			
	文中模型	未优化TFT	Transformer	LSTM
氢气	0.057	0.117	0.143	0.356
甲烷	0.046	0.103	0.157	0.564
乙烷	0.053	0.079	0.096	0.742
乙烯	0.073	0.167	0.195	1.754
总烃	0.051	0.081	0.115	0.096
一氧化碳	3.043	6.234	9.543	2.653
二氧化碳	4.057	6.384	7.655	2.873

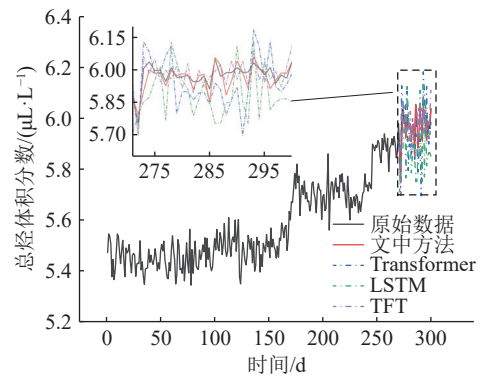


图 4 总烃预测曲线

Fig.4 Total hydrocarbon prediction curve

对两台变压器的所有油色谱数据进行同步预测, 其曲线如图 5(a)所示。以二氧化碳为例, 2 号主变 B 相二氧化碳体积分数呈上升趋势, 而 1 号主变 A 相二氧化碳体积分数因负载波动呈现周期性起伏。图 5(b)—(d)为文中模型对多组别变压器二氧化碳含量的预测结果, Optuna-TFT 模型在不同变化趋势的预测任务中, 能够有效捕捉目标气体体积分数的变化特征, 其预测结果展现了良好的拟合性和跟随性。

表 3 为不同模型训练多组别气体所需时间, 文中模型的计算时间较 Transformer 模型减少了 75.2%, 较 LSTM 模型减少了 81.6%, 表明文中模型在大规模电网故障快速响应层面具有优异的应用潜力。

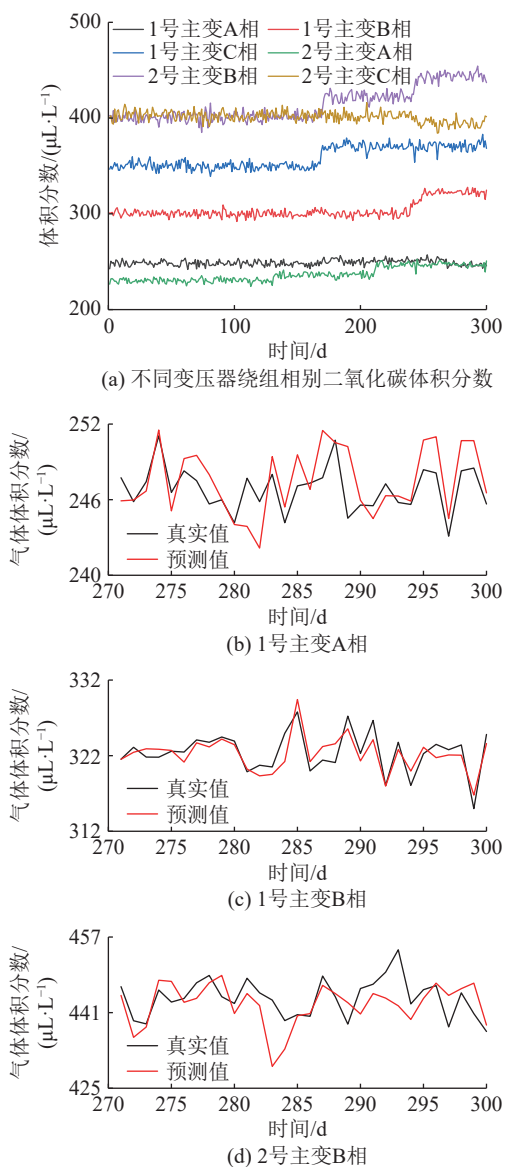


图5 不同变压器二氧化碳体积分数及预测曲线
Fig.5 Carbon dioxide concentration and prediction curves of different transformers

表3 不同模型训练与测试时间

Table 3 Different model training and testing time

模型	训练时间/s	测试时间/s
文中模型	750	41
未优化TFT	1 042	51
Transformer	3 030	55
LSTM	4 073	62

文中进一步分析了 Optuna-TFT 模型针对不同变压器及其绕组相别的油中溶解气体同步预测效果,结果如表4、表5所示,文中模型在短期和长期预测上的效果均优于传统模型。当预测时长为30 d时,Optuna-TFT模型的 e_{MAPE} 相较LSTM和Transformer模型分别降低了66.7%和80.8%;相较未经Optuna超参数优化的TFT模型降低了57.9%。由

决定系数 R^2 可知,即使预测时长为30 d,文中模型的 R^2 值依然接近1,表明该模型在长期预测中仍具有优异的跟随性。上述结果表明,在保证预测精度的同时,文中模型实现了对多组变压器油中溶解气体的高效同步预测。

表4 不同时间长度下各模型 e_{MAPE} 值
Table 4 The e_{MAPE} values of each model under different time lengths

时间长度/d	$e_{MAPE}/\%$			
	文中模型	未优化TFT	Transformer	LSTM
7	0.244	0.427	0.671	1.035
15	0.285	0.523	0.847	1.356
30	0.306	0.727	0.918	1.591

表5 不同时间长度下各模型 R^2 值
Table 5 The R^2 values of each model under different time lengths

时间长度/d	R^2			
	文中模型	未优化TFT	Transformer	LSTM
7	0.987	0.891	0.851	0.802
15	0.987	0.843	0.783	0.765
30	0.976	0.802	0.737	0.701

为了进一步发挥模型的工程应用价值,将所提 Optuna-TFT 算法作为核心模块,部署至河北某变电站变压器在线监测系统,并在系统中设计针对溶解气体浓度预测值的多级阈值预警机制。当模型输出的未来气体浓度预测值达到不同预警等级的阈值时,实时通知运维人员进行后续检查与处置。

3.3.2 模型注意力机制中的物理含义

图6为 Optuna-TFT 模型在预测二氧化碳体积分数时对静态变量和动态变量的注意力权重排序。对于静态变量,变压器组别的注意力权重占比为0.55,高于气体类别和绕组相别,表明文中模型认为不同运行条件对油中溶解气体含量的变化规律影响更大,而同一变压器内部三相之间的差异性较低。

对于动态变量,在预测二氧化碳体积分数时,甲烷和氢气对二氧化碳体积分数预测的影响最大,权重占比分别0.33和0.23。在实际工况下的绝缘材料裂解过程中,二氧化碳与甲烷、氢气的生成机理存在内在关联。绝缘纸(主要成分为纤维素)在高温下热解会产生大量一氧化碳和二氧化碳气体^[31-32];与此同时,处于相同故障环境中的绝缘油中长链碳氢化合物也会发生裂解,释放出丰富的甲烷和氢气等小分子气体^[33-34]。在实际工况下,绝缘纸和绝缘油会同步分解,其产物二氧化碳常常伴随着甲烷和氢气的升高共同出现,这3种气体在DGA

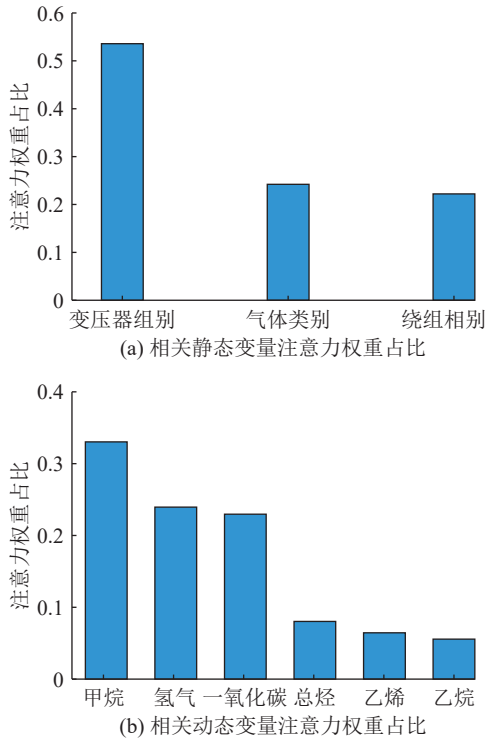


图6 二氧化碳预测中各变量注意力权重占比
Fig.6 Attention weight ratio of each variable in carbon dioxide prediction

数据中表现出强正相关关系。

上述内在机理解释了模型在预测二氧化碳浓度时对甲烷和氢气赋予更高的注意力权重,模型有效捕捉到了这些气体在故障中的相互关系,从而使其对故障状态的判断具有实际物理意义。注意力机制揭示了输入特征与目标气体之间的因果关联,提升了模型结果的可信度。

图7进一步展示了所有特征气体对预测气体的注意力权重,其中预测气体自身并不参加注意力权重计算。

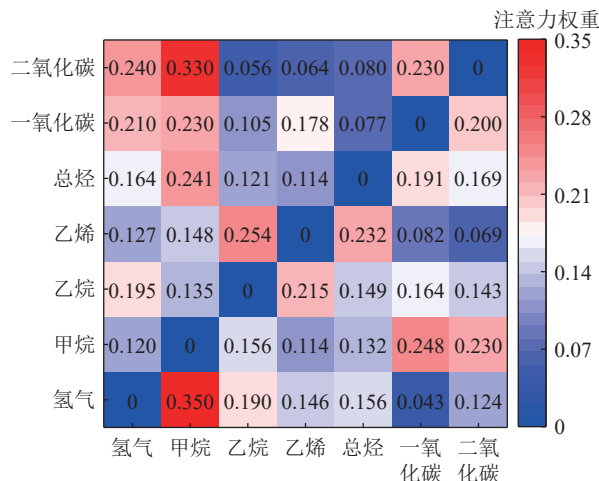


图7 变压器各故障气体关联关系热度值
Fig.7 Association heat values of transformer fault gas correlation

在各个特征气体的关联层面,氢气和甲烷以及二氧化碳和甲烷在热度图中呈强相关关系,乙烯与乙烷之间呈强相关关系,一氧化碳与二氧化碳之间呈强相关关系。上述权重占比符合绝缘油中油纸绝缘裂解期间的产气规律,表明文中模型的多头注意力机制能够识别油中溶解气体间的关联性,可以为后续研究变压器故障机理提供技术支撑。

4 结论

文中针对油中溶解气体的变化趋势预测问题,提出了 Optuna-TFT 模型,该模型在变压器油中溶解气体的预测任务中性能表现优越,具有较强的应用价值。基于以上工作,文中得出以下结论:

(1) 文中首次采用基于 Optuna 优化的 TFT 算法预测油中溶解气体的变化趋势,其中 Optuna 超参数优化有效避免了人工调参的盲目性和低效性。

(2) Optuna-TFT 模型在单一气体预测与多组别同步预测任务中均表现出显著优势。在多组别同步预测任务中,模型综合考虑多台变压器及其相别间的特性差异。在 30 d 的预测时长下,Optuna-TFT 的 R^2 依然达到 0.976,表明其在复杂多变量的长期同步预测中具备优异的跟随性。

(3) Optuna-TFT 采用可解释性的多头注意力机制发掘了油中溶解气体的相关性,其中二氧化碳与甲烷、氢气之间的强关联性与油纸绝缘裂解的产气规律高度一致,可为智能预警平台的精准诊断提供可靠的技术支撑。

尽管文中模型在训练样本中取得了优异效果且已载入预警平台,但受限于样本数量限制,距离推广应用于其他变压器仍存在差距。一方面,多物理量(如光、热、声、电、化学等)的状态感知技术已广泛应用于变压器的监测和诊断,后续可将多组别状态量通过动态变量和静态变量的方式引入模型,实现变压器故障预警的多维度评估和智能决策。另一方面,针对故障训练样本占比较小这一难题,后续可构建电力变压器全寿命仿真平台用于模拟多种故障发展规律,增强故障类型数据体量。

参考文献:

- [1] 杨金鑫,廖才波,胡雄,等. 基于 DGA 与 TPE-LightGBM 的变压器故障诊断[J]. 电力科学与技术学报, 2024, 39(4): 70-77.
YANG Jinxin, LIAO Caibo, HU Xiong, et al. Transformer fault diagnosis based on DGA and TPE-LightGBM[J]. Journal of Electric Power Science and Technology, 2024, 39(4): 70-77.
- [2] 何宁辉,吴旭涛,沙伟燕,等. 电力变压器油中溶解气体在线监测数据修复方法[J]. 高压电器, 2024, 60(11): 37-48.
HE Ninghui, WU Xutao, SHA Weiyuan, et al. On-line monitor-

- ing data repair method for dissolved gas in power transformer oil[J]. *High Voltage Apparatus*, 2024, 60(11): 37-48.
- [3] 阎对丰, 刘昌林, 李元超, 等. 基于 CTGAN-CRS 与改进卷积神经网络的变压器故障诊断方法[J]. *高压电器*, 2025, 61(6): 120-130,137.
- YAN Duifeng, LIU Changlin, LI Yuanchao, et al. Fault diagnosis method of transformer based on CTGAN-CRS and improved convolutional neural network[J]. *High Voltage Apparatus*, 2025, 61(6): 120-130,137.
- [4] MANOJ T, RANGA C, ABU-SIADA A, et al. Analytic hierarchy processed grey relational fuzzy approach for health assessment of power transformers[J]. *IEEE Transactions on Dielectrics and Electrical Insulation*, 2024, 31(3): 1480-1489.
- [5] 江军, 张文乾, 李波, 等. 电力变压器油中溶解气体离群值识别和数据重构[J]. *电工技术学报*, 2024, 39(17): 5521-5533.
- JIANG Jun, ZHANG Wenqian, LI Bo, et al. Outlier detection and data reconstruction of dissolved gas in oil for power transformers[J]. *Transactions of China Electrotechnical Society*, 2024, 39(17): 5521-5533.
- [6] 李元, 李星辉, 孙渭薇, 等. 基于多模型级联的油浸式电力变压器故障诊断方法[J]. *智慧电力*, 2023, 51(6): 86-92.
- LI Yuan, LI Xinghui, SUN Weiwei, et al. Fault diagnosis method of oil-immersed power transformer based on multi-model cascade fusion[J]. *Smart Power*, 2023, 51(6): 86-92.
- [7] 王子凌, 汪科, 柴卫健, 等. 基于 PCA 和 Fast-MCD 的油中溶解气体在线监测异常数据识别[J]. *浙江电力*, 2025, 44(3): 100-107.
- WANG Ziling, WANG Ke, CHAI Weijian, et al. Identification of abnormal data in online dissolved gas monitoring in oil based on PCA and Fast-MCD[J]. *Zhejiang Electric Power*, 2025, 44(3): 100-107.
- [8] 王娜娜, 栗文义, 李小龙. 基于不均衡小样本 DGA 数据与改进 CatBoost 决策树的油浸式变压器故障诊断方法[J]. *电力系统保护与控制*, 2024, 52(23): 167-176.
- WANG Nana, LI Wenyi, LI Xiaolong. An oil-immersed transformer fault diagnosis method based on DGA unbalanced limited sample processing and improved CatBoost[J]. *Power System Protection and Control*, 2024, 52(23): 167-176.
- [9] 范志远, 杜江. 基于相关变分模态分解和 CNN-LSTM 的变压器油中溶解气体体积分数预测[J]. *高电压技术*, 2024, 50(1): 263-273, I0020.
- FAN Zhiyuan, DU Jiang. Prediction of dissolved gas volume fraction in transformer oil based on correlation variational mode decomposition and CNN-LSTM[J]. *High Voltage Engineering*, 2024, 50(1): 263-273, I0020.
- [10] 曹旺, 万元, 唐伟, 等. 变压器油中溶解气体在线监测系统的设计[J]. *变压器*, 2023, 60(5): 31-37.
- CAO Wang, WAN Yuan, TANG Wei, et al. Design of on-line monitoring system for dissolved gas in transformer oil[J]. *Transformer*, 2023, 60(5): 31-37.
- [11] 杜振铭. 变压器油中溶解乙烯气体检测用气体传感器响应特性研究[D]. 武汉: 华中科技大学, 2024.
- DU Zhenming. Study on response characteristics of gas sensor for detecting dissolved ethylene gas in transformer oil[D]. Wuhan: Huazhong University of Science and Technology, 2024.
- [12] 罗文萱. 基于胶囊神经网络的电力变压器故障诊断方法研究[J]. *高压电器*, 2024, 60(5): 92-98.
- LUO Wenxuan. Research on fault diagnosis method of power transformer based on capsule networks[J]. *High Voltage Apparatus*, 2024, 60(5): 92-98.
- [13] 路海阳. 基于机器学习的变压器油中溶解气体含量预测及运行状态评估[D]. 杭州: 浙江大学, 2023.
- LU Haiyang. Prediction of dissolved gas content in transformer oil and evaluation of transformer's operating state based on machine learning methods[D]. Hangzhou: Zhejiang University, 2023.
- [14] 曹正江, 付文龙, 文斌, 等. 基于 TVFEMD 和多模型融合的变压器油中溶解气体体积分数预测方法[J]. *高压电器*, 2024, 60(8): 156-166.
- CAO Zhengjiang, FU Wenlong, WEN Bin, et al. Prediction method of dissolved gas volume fraction in transformer oil based on TVFEMD and multi-model fusion[J]. *High Voltage Apparatus*, 2024, 60(8): 156-166.
- [15] 李佳, 邓科, 侯玉莲, 等. 基于 GRA-CEEMDAN-BiLSTM 的变压器油中溶解气体浓度预测[J]. *变压器*, 2022, 59(6): 42-47.
- LI Jia, DENG Ke, HOU Yulian, et al. Prediction of dissolved gas concentration in transformer oil based on GRA-CEEMDAN-BiLSTM[J]. *Transformer*, 2022, 59(6): 42-47.
- [16] SAMSON MOGOS A, LIANG X D, CHUNG C Y. Enhancing transformer health index prediction using dissolved gas analysis data through integration of LightGBM and robust EM algorithms[J]. *IEEE Access*, 2024, 12: 108472-108483.
- [17] 谢明浩, 张林鎰, 董小刚, 等. 基于 OVMD-HWOA-KELM 模型的变压器油中溶解气体体积分数预测方法[J]. *高电压技术*, 2024, 50(8): 3793-3807.
- XIE Minghao, ZHANG Linxuan, DONG Xiaogang, et al. Prediction method of dissolved gas volume fraction in transformer oil based on OVMD-HWOA-KELM model[J]. *High Voltage Engineering*, 2024, 50(8): 3793-3807.
- [18] 王科, 苟家萁, 彭晶, 等. 基于 LSTM 网络的变压器油中溶解气体浓度预测[J]. *电子测量技术*, 2020, 43(4): 81-87.
- WANG Ke, GOU Jiaqi, PENG Jing, et al. Prediction of dissolved gas concentration in transformer oil based on LSTM network[J]. *Electronic Measurement Technology*, 2020, 43(4): 81-87.
- [19] 崔宇, 侯慧娟, 胥明凯, 等. 基于双重注意力机制的变压器油中溶解气体预测模型[J]. *中国电机工程学报*, 2020, 40(1): 338-347.
- CUI Yu, HOU Huijuan, XU Mingkai, et al. A prediction method for dissolved gas in power transformer oil based on

- dual-stage attention mechanism[J]. *Proceedings of the CSEE*, 2020, 40(1): 338-347.
- [20] 房佳妹, 刘崇茹, 苏晨博, 等. 基于自注意力 Transformer 编码器的多阶段电力系统暂态稳定评估方法[J]. *中国电机工程学报*, 2023, 43(15): 5745-5758.
- FANG Jiashu, LIU Chongru, SU Chenbo, et al. Multi-stage transient stability assessment of power system based on self-attention Transformer encoder[J]. *Proceedings of the CSEE*, 2023, 43(15): 5745-5758.
- [21] 杨雨昕, 姚伟, 邓贤哲, 等. 基于 Koopman 时序延拓和 CNN-Transformer 模型的频率稳定指标预测[J]. *电网技术*, 2025, 49(8): 3474-3485.
- YANG Yuxin, YAO Wei, DENG Xianzhe, et al. Frequency stability index prediction based on Koopman time series extension and CNN-Transformer model[J]. *Power System Technology*, 2025, 49(8): 3474-3485.
- [22] 汤健, 侯慧娟, 陈洪岗, 等. 基于 BI-GRU 改进的 Seq2Seq 网络的变压器油中溶解气体浓度预测方法[J]. *电力自动化设备*, 2022, 42(3): 196-202, 217.
- TANG Jian, HOU Huijuan, CHEN Honggang, et al. Concentration prediction method based on Seq2Seq network improved by BI-GRU for dissolved gas in transformer oil[J]. *Electric Power Automation Equipment*, 2022, 42(3): 196-202, 217.
- [23] LUO D S, FANG J, HE H Y, et al. Prediction for dissolved gas in power transformer oil based on TCN and GCN[J]. *IEEE Transactions on Industry Applications*, 2022, 58(6): 7818-7826.
- [24] WANG N N, LI W Y, LI J Q, et al. Prediction of dissolved gas content in transformer oil using the improved SVR model[J]. *IEEE Transactions on Applied Superconductivity*, 2024, 34(8): 9002804.
- [25] XING Z K, HE Y G, WANG X, et al. VMD-IARIMA-based time-series forecasting model and its application in dissolved gas analysis[J]. *IEEE Transactions on Dielectrics and Electrical Insulation*, 2023, 30(2): 802-811.
- [26] 杨童亮, 胡东, 唐超, 等. 基于 SMA-VMD-GRU 模型的变压器油中溶解气体含量预测[J]. *电工技术学报*, 2023, 38(1): 117-130.
- YANG Tongliang, HU Dong, TANG Chao, et al. Prediction of dissolved gas content in transformer oil based on SMA-VMD-GRU model[J]. *Transactions of China Electrotechnical Society*, 2023, 38(1): 117-130.
- [27] 李长云, 杨静雨, 连鸿松, 等. 基于 ICEEMDAN-IPSO-ELM 的硅油溶解气体浓度组合预测方法[J]. *高电压技术*, 2023, 49(9): 3887-3897.
- LI Changyun, YANG Jingyu, LIAN Hongsong, et al. Combined prediction method of dissolved gas concentration of silicone oil based on ICEEMDAN-IPSO-ELM[J]. *High Voltage Engineering*, 2023, 49(9): 3887-3897.
- [28] 张文乾, 刘金凤, 江军, 等. 基于 CEEMDAN 和 TCN 的变压器油中溶解气体含量预测[J]. *电力工程技术*, 2024, 43(3): 192-200, 233.
- ZHANG Wenqian, LIU Jinfeng, JIANG Jun, et al. Prediction of concentration for dissolved gas in oil based on CEEMDAN and TCN[J]. *Electric Power Engineering Technology*, 2024, 43(3): 192-200, 233.
- [29] LIM B, ARİK S Ö, LOEFF N, et al. Temporal Fusion Transformers for interpretable multi-horizon time series forecasting[J]. *International Journal of Forecasting*, 2021, 37(4): 1748-1764.
- [30] 曹秀梅, 赵文, 王志国, 等. 基于 Optuna-XGBoost 的砂土地层盾构渣土改良剂预测[J]. *东北大学学报(自然科学版)*, 2024, 45(12): 1769-1777.
- CAO Xiumei, ZHAO Wen, WANG Zhiguo, et al. Prediction of soil conditioners for sandy soil shield based on Optuna-XGBoost[J]. *Journal of Northeastern University (Natural Science)*, 2024, 45(12): 1769-1777.
- [31] 国家能源局. 变压器油中溶解气体分析和判断导则: DL/T 722—2014[S]. 北京: 中国电力出版社, 2015.
- National Energy Bureau of the People's Republic of China. Guide to the analysis and the diagnosis of gases dissolved in transformer oil: DL/T 722-2014[S]. Beijing: China Electric Power Press, 2015.
- [32] 高胜友, 王昌长, 李福祺. 电力设备的在线监测与故障诊断[M]. 2 版. 北京: 清华大学出版社, 2018: 78-81.
- GAO Shengyou, WANG Changchang, LI Fuqi. On-line monitoring and diagnosis for power equipment[M]. 2nd ed. Beijing: Tsinghua University Press, 2018: 78-81.
- [33] 刘昌, 邵山峰, 廖瑞金, 等. 特高压换流变压器绝缘油老化特性研究[J]. *高电压技术*, 2019, 45(3): 730-736.
- LIU Chang, SHAO Shanfeng, LIAO Ruijin, et al. Study on the aging characteristics of insulation oil of UHVDC converter transformer[J]. *High Voltage Engineering*, 2019, 45(3): 730-736.
- [34] WANG H, WU R, FAN Z, et al. Decomposition characteristics and mechanism of mineral transformer oil under arc and spark discharge[C]//2024 IEEE International Conference on Plasma Science (ICOPS). Beijing, China. IEEE, 2024: 1.

作者简介:



周延豪

周延豪(2001), 男, 硕士在读, 研究方向为变压器状态预测与故障诊断(E-mail: zhouyanhao@mail.shiep.edu.cn);

范路(1997), 男, 博士, 研究方向为表面电位测量及仿真分析、电力设备在线监测等;

任海龙(2001), 男, 硕士在读, 研究方向为变压器状态预测与电缆油含水量分析。

An anti-interference fault location simulation for XLPE cables based on improved SSTDR

WANG Yi^{1,2}, LAI Guopeng¹, LUO Zhang³, HE Jiajun³, CHEN Quanfeng³, HUA Zhilei⁴

(1. School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; 2. Chongqing Key Laboratory of Mobile Communication Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; 3. State Grid Chongqing Electric Power Company Marketing Service Center, Chongqing 400014, China; 4. Yangzhong Electric Power Research Center, North China Electric Power University, Zhenjiang 212200, China)

Abstract: Complex laying environments and external electromagnetic interference have significantly impacted the accuracy of cable fault location, particularly in modern power grids with the widespread application of cross-linked polyethylene (XLPE) power cables. In order to solve the low accuracy of existing fault location methods in high-noise environments, an improved fault positioning method based on spread spectrum time domain reflectometry (SSTDR) is proposed. Firstly, by introducing a multi-period signal modeling method using M-sequences, combined with windowed filtering and adaptive wavelet denoising, the reflected signals are optimized. Subsequently, a quadratic cross-correlation weighting method based on the smoothed coherence transform (SCOT) weighting function is employed to process the incident signals and the filtered reflected signals for each period, obtaining the correlation coefficients for each period. The final location information is derived through arithmetic averaging. Finally, Gaussian smoothing filtering is applied to enhance the peak characteristics of the fault location, and the fault position is determined using a peak detection algorithm. Simulation experiments conducted on the MATLAB/Simulink platform demonstrate that under a low signal-to-noise ratio of -20 dB, the average location error for short-circuit, open-circuit faults at different positions, and multi-branch faults is less than 0.062 m, with an average relative error below 0.25% . Compared to traditional methods and existing methods, the proposed method exhibits effective localization under low signal-to-noise ratio conditions, with the relative localization error improving by at least 0.01% . Furthermore, the experimental results validate that the method proposed in this paper has robust noise-resistant fault localization performance.

Keywords: cable fault location; spread spectrum time domain reflectometry (SSTDR); M-sequence; multi-period signal modeling; wavelet denoising; Gaussian smoothing filter; peak detection

(编辑 方晶)

(上接第 45 页)

Prediction method for dissolved gas in transformer oil based on the Temporal Fusion Transformer model

ZHOU Yanhao^{1,2}, FAN Lu², REN Hailong^{1,2}, ZHAO Su², WANG Yalin², YIN Yi²

(1. College of Electrical Engineering, Shanghai University of Electric Power, Shanghai 200090, China;
2. Department of Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: Dissolved gas analysis in transformer oil is regarded as an important indicator for evaluating the operational status of transformers. Accurate prediction of trends in dissolved gases in oil is beneficial for preventing power transformer failures. A Temporal Fusion Transformer (TFT) model, optimized via Optuna hyperparameter tuning, is proposed to address the technical challenge of low prediction efficiency inherent in traditional models that rely on a single variable. Static variables including transformer group, winding phase, and gas type are introduced into the model, and an interpretable multi-head attention mechanism is integrated as well. Synchronous prediction of all dissolved gases in the oil of multiple transformers is thereby achieved, improving the early warning efficiency of substation operation and maintenance systems. An average relative error of only 0.306% is achieved by the proposed model, representing a 66.7% reduction relative to the Transformer baseline model. Higher predictive accuracy is also demonstrated in both short-term and long-term forecasting. In addition, the model's training time is only one quarter that of the Transformer baseline model. This efficiency aligns with the current trend toward simultaneous prediction across multiple device groups in intelligent early-warning platforms. Strong correlations between hydrogen and methane and between carbon dioxide and methane are indicated by the model's multi-head attention mechanism. These correlations are consistent with the gas generation patterns of oil-paper insulation degradation, further demonstrating the model's good interpretability and providing technical support for synchronous prediction in multiple device groups.

Keywords: electric power transformer; dissolved gas in transformer oil; synchronous prediction; Temporal Fusion Transformer (TFT) model; time series; attention mechanism

(编辑 陈雯嘉)