

基于LSSVM的电力变压器异常状态实时评价研究

变压器作为电力系统关键设备，对其健康状态进行快速准确评价具有重大意义。针对现有变压器异常识别误判率较高的问题，提出一种通过例行试验结果校正在线监测数据趋势，并利用最小二乘支持向量机（LSSVM）开展变压器异常实时评价的方法。通过三次指数平滑法提取油中溶解气体在线监测数据趋势特征，利用最小二乘法将在线监测数据趋势与例行试验历史数据相融合建立在线监测数据修正模型，最后在专家打分的基础上，通过LSSVM实现变压器异常实时判断。实例分析结果表明，该方法同时吸收了在线监测与例行试验的优点，可以有效识别变压器异常状态，具有广阔的工程应用前景。

满玉岩¹ 董艳唯¹ 祁麟² 张弛¹ / 1. 国网天津市电力公司电力科学研究院
2. 国网天津市电力公司城南供电分公司

0 引言

近年来，特高压电网在社会经济效益与新技术装备的双重推动下获得跨越式发展，特高压电网的快速发展在优化我国能源配置、推进清洁能源发展的同时也对电网设备安全可靠运行提出更高的要求。同时，随着社会经济水平的不断提升，供电可靠性的要求日益增高，计划停电愈发困难，抢修时间愈发紧迫，电网安全压力进一步加大。为了保证设备安全可靠运行的同时降低设备停运率，亟待提升不停电检测技术在电网的应用效益。

变压器在线监测技术具有无须人工干预、实时测量等优点，有助于及时跟踪发现设备缺陷或故障的早期特征，但由于装置运行不稳定、故障率高等原因，在线监测数据存在缺失、离群值和漂移等问题，基于在线监测的变压器异

常判别是在线监测应用研究的难点与热点之一。研究人员围绕变压器异常征兆识别基于模糊评估^[1]、证据理论^[2-3]、可拓分析法^[4]、人工神经网络^[5]、灰靶理论^[6]和贝叶斯网络^[7]等开展了大量研究工作。文献[8]利用Canopy高维聚类模型识别油中溶解气体特征模式，结合异常事件库，提出基于变压器状态的异常诊断方法，该方法可以有效处理在线监测数据的缺失与波动问题，并给出异常原因，但受异常事件库准确性与完整性影响，判断准确性难以保证；文献[9]将时间序列与滑动窗口相结合筛选判断变压器状态的多维参量，并利用k-means聚类方法识别变压器异常，该方法综合利用了油温、负荷和环境温度等多维数据，判据较全面，但没有考虑在线监测数据问题导致的误判。笔者认为，目前由于在线监测数据质量问题，单独利用在线监测数据难以有效评估设备状态，有必要将例行试验结果与在线监测数据相结合以准确判断变压器状态。

支持向量机（Support Vector Machine, SVM）是机器学习中有监督学习模型的重要学习算法，可用于分类、回归或异常值检测等任务。在用于分类时，SVM可以看作一个二元线性（非线性）分类器，对于给定的一组已标记的二元分类训练样本，SVM将通过学习建立一个具有空间最优的二元分类模型。除了线性分类之外，SVM可以建立基于内核函数的非线性分类模型，并将输入样本映射到高维特征空间中。最小二乘支持



满玉岩/工程师

关键词/Keywords

变压器 ·
油中溶解气体分析 ·
状态评价 ·
三次指数平滑 ·
最小二乘支持向量机 ·

向量机 (Least Square Support Vector Machine, LSSVM) 把原方法的不等式约束变为等式约束, 极大简化了Lagrange乘子的求解。

本文提出了一种通过例行试验结果校正在线监测数据趋势, 并基于LSSVM开展变压器异常实时评价的方法, 该方法在提取在线监测有效信息的同时能实时评估变压器状态, 可及时发现变压器缺陷或故障前期征兆, 为变压器运维检修提供有效依据, 具有广阔的应用前景。

1 变压器异常识别模型建立

工程实践表明, 较在线监测数据而言, 油中溶解气体例行试验数据更准确, 而在线监测由于采样时间短, 且数据自动处理的特点, 在趋势追踪方面具有较强优势。因此, 本文首先利用缺失值填充技术拟合油中溶解气体在线监测空白数据, 然后利用时间序列分析中的趋势分析方法提取油中溶解气体在线监测数据变化趋势, 并与例行试验数据累加修正油中溶解气体变化量, 最后基于修正结果利用LSSVM方法评价变压器状态。

1.1 数据预处理

由于通信异常、传感器故障及外界干扰等问题, 油中溶解气体在线监测数据会出现缺失, 需要采用数学方法进行填充, 本文采用基于k最近邻 (k-Nearest Neighbor, kNN) 的缺失值填充方法。

首先, 计算缺失样本与其他样本间的欧式距离, 查找与缺失样本距离最小的k个最近样本, 然后, 采取反距离加权的手段对筛选出的k个距离最近样本进行加权, 远距离对应低权值, 近距离对应高权值。这些最近位置的加权平均值即缺陷样本的估计值。该方法简单、易操作, 在缺失值补充方面获得了广泛应用。该算法计算步骤为: ① 计算缺失样本与训练样本之间的距离; ② 按照从小到大的顺序对距离进行排序; ③ 筛选出距离缺失样本最近的k个点; ④ 计算这k个点所属分类的出现频率; ⑤ 选取k个点中频度最高的类别作为缺失样本的加权基值; ⑥ 利用前k个点与缺失样本的距离, 按照反距离加权方法确定前k个点的权值; ⑦

将加权基值与前k个点权值相乘后除以k即为缺失样本估计值。

上述距离仅指欧式距离, 即

$$d(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2} \quad (1)$$

1.2 在线监测数据趋势提取

为了校正在线监测数据, 首先需要在线监测数据趋势进行提取。时间序列分析法具有挖掘时序数据隐含信息与模式的优点, 本文采用Holt-Winters时间序列分析方法进行油中溶解气体在线监测数据趋势提取。

变压器油中溶解气体含量由前端装置采集完成, 数据一般按照时序排列, 因此, 可以认为各特征气体监测量形成一系列的时间序列数据, 如矩阵X所示

$$X = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ X_{m1} & X_{m2} & \cdots & X_{mn} \end{pmatrix} \quad (2)$$

式中, X_{mn} 为第m类特征气体在第n时段的数据。

在变压器内部由于过热、局部放电等原因时, 其乙炔 (C_2H_2)、氢气 (H_2) 和总烃等特征气体含量一般呈现缓慢且有规律的上升趋势, 为了捕捉此趋势, 本文采用Holt-Winters方法对此趋势进行拟合。Holt-Winters指数平滑法公式如下

$$y_{t+k} = (a_t + b_t k) c_{t+k} \quad (3)$$

式中, y_{t+k} 为平滑后的序列; a_t 为截距; b_t 为t时刻的趋势项, 表示指数平滑的平均值; c_{t+k} 为乘法模型的季节因子; $t = 1, 2, \dots, T$; k 为向后平滑期数, $k > 0$ 。其中

$$a_t = \alpha \frac{y_t}{c_{t-s}} + (1-\alpha)[a_{t-1} + b_{t-1}] \quad (4)$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1} \quad (5)$$

$$c_t = \gamma \frac{y_t}{a_t} + (1-\gamma)c_{t-k} \quad (6)$$

式中, α 、 β 和 γ 为阻尼因子, 介于0~1之间; s 为季节周期长度, 此处取365。

未来某日油中溶解气体在线监测数据趋势预

测模型即为

$$y_{T+k} = (a_T + b_T k) c_{T+k-s} \quad (7)$$

1.3 例行试验预测

假设油中溶解气体在线监测数据趋势预测模型较准确地反映了变压器油特征气体含量变化趋势，需要将所得趋势与例行试验结果相结合以建立特征气体含量预测模型，因此，仅提取式(3)趋势项，该模型如下

$$\hat{X}_{m(t+k)} = b_t k + \varphi \quad (8)$$

式中， φ 为在线监测测量误差。

为了保证计算结果与历史例行试验结果偏差最小，本文采用最小二乘法进行计算模型参量 ϕ 。最小二乘法通过寻找已有模型不同参数下的最小方差以获得最优函数。

假设同时期例行试验油中溶解气体特征气体结果为

$$Y = \begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1n} \\ Y_{21} & Y_{22} & \cdots & Y_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ Y_{m1} & Y_{m2} & \cdots & Y_{mn} \end{pmatrix} \quad (9)$$

式中， Y_{mn} 为第 m 类特征气体在第 n 时段的数值。

则例行试验结果与模型误差为

$$e_t = \hat{X}_{mt} - Y_{mt} \quad (10)$$

定义平方损失函数为

$$Q = \sum_{t=0}^n e_t^2 = \sum_{t=0}^n (\hat{X}_{mt} - Y_{mt})^2 = \sum_{t=0}^n [b_t k + \phi - Y_{mt}]^2 \quad (11)$$

把 ϕ 看作函数 Q 的自变量，对式(11)取极值求导即可得变量 ϕ 。

1.4 变压器状态实时评价

变压器状态评价从机器学习的角度可以看作对变压器状态的分类问题。由于例行试验数据样本较少，适于LSSVM小样本学习方法，因此，本文基于例行试验预测结果利用LSSVM对变压器状态开展实时评价计算。

首先给定例行试验训练样本集 (x_i, y_i) ， $i = 1, 2, \dots, l$ ， $x \in R$ ， $y \in \{\pm 1\}$ ，假设超平面为 $(w \cdot x) + b = 0$ ，为实现所得超平面对所有样本均能够正确分类且距离最小，应满足以下条件

$$y_i [(w \cdot x) + b] \geq 1 \quad i = 1, 2, \dots, l \quad (12)$$

由此可得，分类间隔为 $2/\|w\|$ ，因此问题转换为在上述约束下求

$$\min \Phi(w) = 1/2 \|w\|^2 = 1/2 (w \cdot w') \quad (13)$$

为了求解该问题，需引入拉格朗日函数

$$L(w, b, a) = 1/2 \|w\|^2 - a \{y[(w \cdot x) + b] - 1\} \quad (14)$$

式中， $a > 0$ 为拉格朗日乘数。约束最优解必须在鞍点处对 w 和 b 的偏导为0，将该二次规划问题转化为对偶问题，即

$$\max Q(a) = \sum_{j=1}^l a_j - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j (x_i \cdot x_j) \quad (15)$$

$$\text{s.t.} \quad \sum_{j=1}^l a_j y_j = 0 \quad a_j \geq 0, j = 1, 2, \dots, l \quad (16)$$

计算得最优解为

$$a^* = (a_1^*, a_2^*, \dots, a_l^*)^T \quad (17)$$

计算 w^* 和 b^* 分别为

$$w^* = \sum_{j=1}^l a_j^* y_j x_j \quad (18)$$

$$b^* = y_i - \sum_{j=1}^l a_j^* y_j a_j (x_j \cdot x_i) \quad (19)$$

因此，最优分类函数为

$$\begin{aligned} f(x) &= \text{sgn}\{(w^* \cdot x) + b^*\} \\ &= \text{sgn}\left\{\left(\sum_{j=1}^l a_j^* y_j a_j (x_j \cdot x_i)\right) + b^*\right\}, x \in R^n \end{aligned} \quad (20)$$

2 算例分析

为了验证本算法对变压器油中溶解气体例行试验结果预测与变压器异常评价结果的准确性，本文收集了某500 kV变电站2016年1~7月的变压器油中溶解气体在线监测数据与例行试验数据，对本算法进行分析验证。计算步骤如下。

1) 选取历史数据。将历史数据分为训练样本与测试数据。

2) 数据预处理。利用kNN方法对各特征气体含量测量空缺值拟合填充。

3) 建立在线监测数据趋势模型，计算模型参数。

4) 根据步骤3)趋势模型建立在线监测数据校正模型。

5) 基于例行试验数据计算LSSVM模型参数, 建立变压器异常实时评价模型。

本文收集到174组变压器油中溶解气体在线监测数据与4条例行试验数据, 选取3条例行试验数据作为训练样本, 剩余的1组数据作为测试数据。下面以 C_2H_2 气体为例进行介绍。

首先利用kNN算法填补在线监测数据集空缺项, 该数据集缺失情况如图1所示。可见该数据集共有18项缺失值, 而且缺失项均为 C_2H_2 气体, 因此仅有一种缺失模式。

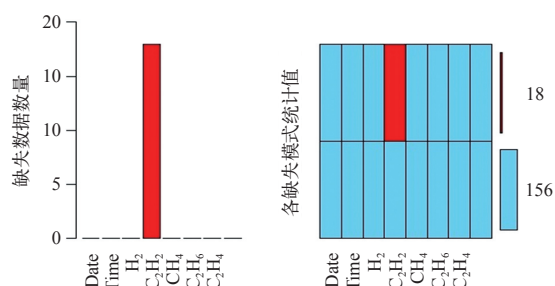


图1 缺失数据数量与缺失模式统计值

通过kNN填补后, 对填补误差进行分析, 结果如表1所示。可见平均绝对百分比误差小于0.05, 填补结果较好。

表1 kNN填补误差

误差项	标准误差(RMSE)	平均绝对百分比误差(MAPE)
误差值	0.034 5	0.048 6

然后分析该在线监测数据集趋势模型, C_2H_2 气体在线监测结果如图2所示, 容易看出, 从2016年1~7月, C_2H_2 气体含量有明显增大趋势, 且6、7

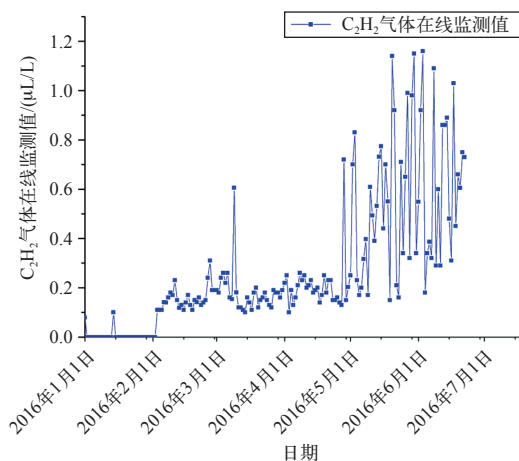


图2 C_2H_2 气体在线监测结果

月份, 部分 C_2H_2 气体含量甚至超过规程阈值。

通过估计式(3)中 a_i 、 b_i 初值, 即可计算式(3)模型参数, 求得 $\alpha = 0.187$, $\beta = 0.179$, 趋势提取结果如图3所示。

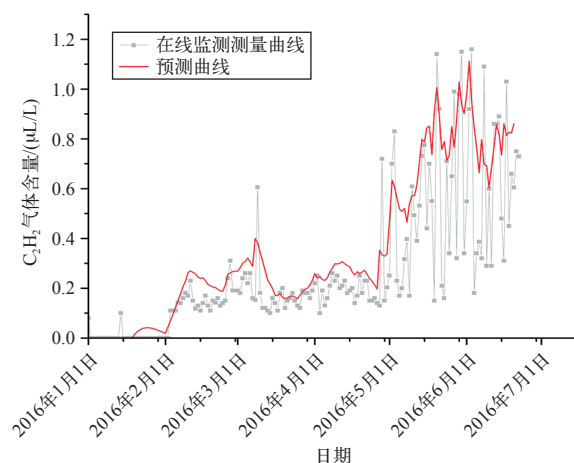


图3 C_2H_2 气体含量在线监测数据趋势提取结果

接着, 利用最小二乘法将例行试验数据与在线监测趋势模型相叠加, 结果如表2所示。

表2 C_2H_2 例行试验数据

(单位: $\mu\text{L/L}$)

日期	C_2H_2 气体含量
2016年1月7日	0.035
2016年3月19日	0.108
2016年5月8日	0.234
2016年7月3日	0.681

在线监测数据修正结果如图4所示。可见, 与原始数据相比, 修正结果根据真实测试值进行了偏移, 使得监测结果更接近真实情况。结果误差分析

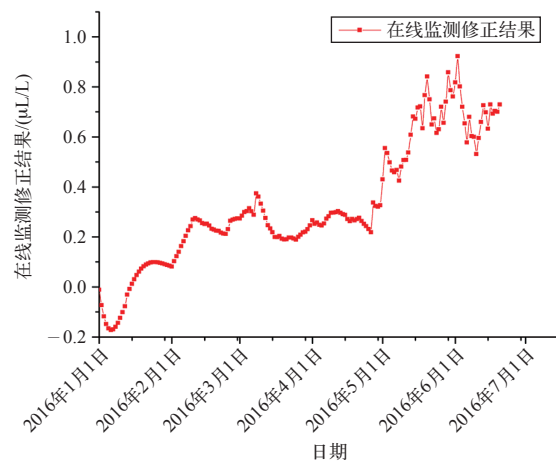


图4 C_2H_2 气体含量在线监测数据修正结果

如表3所示。2016年7月3日C₂H₂气体含量真实测量值为0.681 μL/L，而修正模型计算结果为0.729 μL/L，两者相对误差为7.05%，可见，与在线监测真实监测数据相比较，本文所建模型修正结果较准确。

表3 C₂H₂在线监测数据与例行试验数据误差分析

日期	C ₂ H ₂ 气体含量 例行试验/(μL/L)	C ₂ H ₂ 气体含量 在线监测/(μL/L)	相对误差(%)
2016年7月3日	0.681	0.729	7.05

最后研究基于变压器油中溶解气体趋势提取结果实时判断变压器状态，由于趋势预测特点，本算法较适合发现具有潜伏性、渐进性特征的变压器绝缘劣化等变压器异常。

首先通过专家打分的方式根据上文油中溶解气体在线监测修正数据对变压器状态进行计分。若计分结果大于60分，则该时刻变压器设备异常；若小于等于60分，则变压器设备正常。接着，通过LSSVM算法对所得变压器油中溶解气体修正数据与专家判定结果进行联合分析，SVM的核函数采用Gauss径向基函数。识别结果如表4所示，按识别结果正确与否分为四类，包括识别判断正确的模式1和模式4以及识别错误的模式2和模式3，其中模式4代表变压器异常正确识别结果，在2016年6月1日~7月1日间，共发现8次异常状态，经现场检测与专家判断，该变压器可能局部存在高温过热，与算法判断结果一致。可见基于LSSVM算法的变压器异常判断准确率为89.5%，具有较好的识别能力。

表4 变压器异常识别准确率

序号	样本数据	识别结果	数量	准确率(%)
1	正常	正常	30	100
2	异常	正常	4	0
3	正常	异常	0	—
4	异常	异常	8	100
准确率	—	—	—	89.5

3 结束语

1) 本方法利用例行试验数据修正在线监测结果并基于LSSVM识别变压器异常，在解决在线监测数据质量问题的同时实现变压器异常的快速准

确识别。

2) 本文仅针对变压器有中溶解气体在线监测数据开展了算例分析，下一步对其他监测类型应进一步分析验证。

3) 基于单一状态量的变压器异常识别较易受数据波动性影响，下一步可考虑融合油温、环境温度等多状态量开展变压器异常综合识别。

参考文献

- [1] 骆思佳, 廖瑞金, 王有元, 等. 带变权的电力变压器状态模糊综合评判 [J]. 高电压技术, 2007, 33 (8): 107-108.
- [2] Tang W H, Spurgeon K, Wu Q H, et al. An evidential reasoning approach to transformer condition assessments [J]. IEEE Transactions on Power Delivery, 2004, 19 (4): 1696-1703.
- [3] 朱承治, 郭创新, 孙旻, 等. 基于改进证据推理的变压器状态评估研究 [J]. 高电压技术, 2008, 34 (11): 2332-2337.
- [4] 廖瑞金, 张懿议, 黄飞龙, 等. 基于可拓分析法的电力变压器本体绝缘状态评估 [J]. 高电压技术, 2012, 38 (3): 521-526.
- [5] 阮羚, 谢齐家, 高胜友, 等. 人工神经网络和信息融合技术在变压器状态评估中的应用 [J]. 高电压技术, 2014, 40 (3): 822-828.
- [6] 郑蕊蕊, 赵继印, 吴宝春, 等. 基于加权灰靶理论的电力变压器绝缘状态分级评估方法 [J]. 电工技术学报, 2008, 23 (8): 60-66.
- [7] 白翠粉, 高文胜, 金雷, 等. 基于3层贝叶斯网络的变压器综合故障诊断 [J]. 高电压技术, 2013, 39 (2): 330-335.
- [8] 荣智海, 齐波, 张鹏, 等. 基于油中溶解气体Canopy高维模型的变压器异常状态快速识别方法 [J]. 中国电机工程学报, 2018 (13): 1-10.
- [9] 严英杰, 盛戈皞, 刘亚东, 等. 基于滑动窗口和聚类算法的变压器状态异常检测 [J]. 高电压技术, 2016, 42 (12): 4020-4025.

(收稿日期: 2017-11-13) **EA**