

doi: 10.7690/bgzd.2026.01.009

基于数据聚类的配电变压器自动容量校核方法

李金富¹, 吉云海¹, 范顺辉¹, 叶宗来², 李新飞²

(1. 云南电网有限责任公司楚雄供电局, 云南 楚雄 675000;

2. 保定众德科仪电气科技有限公司, 河北 保定 071000)

摘要: 为保证配电变压器的安全、可靠运行, 提出基于数据聚类的自动容量校核方法。通过改进模糊 C 均值聚类算法, 计算各个数据样本和所有聚类之间的隶属度, 聚类配电变压器历史电压降数据; 基于相间差分方法处理聚类后的高压侧数据, 降低高压侧数据的噪声干扰后; 以变压器的负载率和电压变化率 2 种电压降特性参数为依据, 利用时间差分方法自动校核配电变压器容量。测试结果表明: 该方法能有效完成电压器电压降数据聚类, Xie-Beni 系数计算结果均在 0.030 以下; 可极大程度减少噪声数据对变压器容量校核的影响, 可靠完成变压器容量的自动校核; 校核结果的均方根误差平均值均在 0.89 以上。

关键词: 数据聚类; 配电变压器; 容量校核; 电压降; 负载率; 电压变化率

中图分类号: TP274; TM41 **文献标志码:** A

Automatic Capacity Calibration Method of Distribution Transformer Based on Data Clustering

Li Jinfu¹, Ji Yunhai¹, Fan Shunhui¹, Ye Zonglai², Li Xinfei²

(1. Chuxiong Power Supply Bureau of Yunnan Power Grid Co., Ltd., Chuxiong 675000, China;

2. Baoding Zhongde Keyi Electrical Technology Co., Ltd., Baoding 071000, China)

Abstract: To ensure the safe and reliable operation of distribution transformer, an automatic capacity verification method based on data clustering is proposed. By improving the fuzzy C -means clustering algorithm, the membership degree between each data sample and all clusters is calculated, and the historical voltage drop data of the distribution transformer is clustered; Based on the load factor and voltage change rate of the transformer, the capacity of the distribution transformer is automatically checked by the time difference method. The test results show that the method can effectively complete the voltage drop data clustering of the transformer, and the Xie-Beni coefficient calculation results are all below 0.030. It can greatly reduce the influence of noise data on the transformer capacity check, and reliably complete the automatic check of the transformer capacity. The average root mean square error of the check results is above 0.89.

Keywords: data clustering; distribution transformer; capacity check; voltage drop; load rate; voltage change rate

0 引言

配电变压器是电力系统中用于实现交流电压变换以及交流电能传输的一种静止电器, 通常情况下, 配电变压器均安装在变电所内^[1]。安装时, 需依据电能供配需求、供配方式以及负荷情况等完成配电变压器容量选择, 该选择结果对于变压器的运行安全、使用寿命等均存在较大影响^[2]。在配电系统中, 选择配电变压器时, 均以系统中记录的容量或者变压器铭牌参数上标记的容量为主要参考, 但是在实际应用时, 会发生运行容量和标记容量参数不符合的情况^[3], 导致配电变压器安装后, 存在一定的安全隐患, 甚至不满足变电站的应用需求^[4]。相关管理部门在进行配电变压器安装时, 需采取相应的方法对变压器的实际容量进行校核, 判断该结果和记录的容量是否相符^[5], 并对错误的容量信息进行校

核, 提升配电网的管理水平, 保证变压器的安全与运行^[6]。在进行配电变压器容量校核时, 由于配电网的结构较为复杂, 导致校核的精准性较低。文献[7]为掌握变压器的应用情况, 以其历史运行数据为基础, 通过直觉模糊层次法对变压器的质量进行分析; 但是该方法在应用过程中, 无法判断变压器的质量情况, 也无法计算变压器的详细容量结果。文献[8]为保证变压器的稳定运行, 确定变压器容量情况, 基于中低频零序扫描阻抗对该容量进行检测, 获取容量结果; 但该方法在应用过程中, 不是依据变压器运行特性参数完成检测, 如果运行特性发生显著变化后, 检测结果的误差较大。数据聚类属于一种数据分析技术, 其可依据数据的内在特性, 对数据进行聚类, 实现相同特性的数据均位于一个聚合类中^[9]。数据聚类具有显著的大数据处理能力,

收稿日期: 2024-11-20; 修回日期: 2024-12-22

第一作者: 李金富(1979—), 男, 云南人。

并且可良好完成高维数据处理,除此之外,还具有降低数据中孤立点、降低数据噪声等功能,在数据挖掘、模式识别以及图像处理等多个领域中均具有较好的应用效果^[10]。笔者为实现配电变压器容量自动校核,以数据聚类为依据,提出基于数据聚类的配电变压器自动容量校核方法。

1 配电变压器自动容量校核

1.1 配电变压器自动容量校核整体流程

为实现配电变压器自动容量校核,笔者对变压器的运行特性进行充分分析,提出基于数据聚类的配电变压器自动容量校核方法。该方法以数据聚类算法为基础,实现变压器运行数据的处理,依据处理后的数据,进行容量自动校核,整个校核流程如图 1 所示。

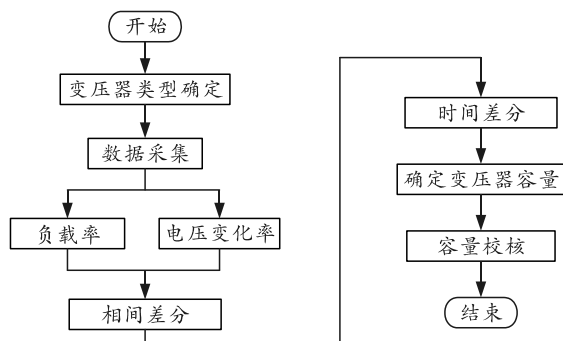


图 1 配电变压器自动容量校核整体流程

该方法在进行配电变压器容量自动校核时,以配电变压器运行的实时数据为基础,采用改进模糊 C 均值聚类算法对该数据进行聚类后,通过时间差分法完成变压器容量自动校核。

1.2 配电变压器历史电压降数据聚类

变压器容量和其铁芯的大小、绕组线圈半径大小、空载损耗、绕组电压降均存在直接关联,可依据上述任意参数对变压器容量进行校核。在实际应用过程中,铁芯的大小、绕组线圈半径大小、空载损耗等参数均需进行现场检测获取^[11];因此,笔者结合变压器容量自动校核的需求,选择电压降作为变压器的运行特性参数,进行容量自动校核。电压降主要包含变压器的负载率和电压变化率 2 方面。

在进行配电变压器运行数据采集时,由于采集的数据涵盖变压器高压侧和低压侧的所有运行数据,如果直接用于校核中,则数据的运算量较大^[12]。为确保变压器容量的校核效果,笔者采用改进模糊 C 均值聚类算法完成数据分类,获取数据中的电压降数据,依据该数据进行变压器容量自动校核。

如果配电变压器历史电压降数据样本集用 $X=\{x_1, x_2, \dots, x_n\}$ 表示, n 表示样本数量,则第 j 个数据样本的特征向量用 $x_j=[x_{j,1}, x_{j,2}, \dots, x_{j,m}]$ 表示, m 表示特征数量。采用模糊聚类算法对 X 进行分类,使其 c 个形成类别,则 X 可描述为:

$$X = X_1 \cup X_2 \cup \dots \cup X_c, X_i \cap X_j = \emptyset. \quad (1)$$

式中: $i \neq j$; X_i 为第 i 类数据的集合。

为分析各个数据样本和所有聚类之间存在的关联,采用隶属度矩阵 $G(g_{ij})$ 描述,需满足:

$$\sum_{i=1}^c g_{ij} = 1. \quad (2)$$

式中: $i=1, 2, \dots, c$; $j=1, 2, \dots, n$; g_{ij} 为数据第 i 个聚类中心的隶属度。

在进行变压器电压运行特性数据聚类时,定义目标函数为每个数据样本和全部聚类中心距离的加权平方和:

$$J_w(G, V) = \sum_{j=1}^n \sum_{i=1}^c g_{ij}^w \|x_j - v_i\|^2. \quad (3)$$

式中: v_i 为向量,对应第 i 个聚类中心; V 为矩阵,阶数为 $c \times m$; w 为加权指数。

在聚类过程中, w 的变化会导致 $J_w(G, V)$ 随之变化,且会出现一个拐点,在该拐点附近存在一个最小值,为保证算法的收敛性,计算 w :

$$w^* = \left\{ w \left| \frac{\partial}{\partial w} \left[\frac{\partial J_w(G, V)}{\partial w} \right] = 0 \right. \right\}. \quad (4)$$

通过拉格朗日惩罚对式(3)进行优化,并定义式(2)为约束条件,以此得出目标函数为局部极小值的必须条件:

$$\left. \begin{aligned} v_i &= \frac{\sum_{j=1}^n g_{ij}^w x_j}{\sum_{j=1}^n g_{ij}^w}, 1 \leq i \leq c \\ g_{ij} &= \left[\sum_{k=1}^c \left(\frac{\|x_j - v_i\|}{\|x_j - v_k\|} \right)^{2/(w-1)} \right]^{-1} \end{aligned} \right\}. \quad (5)$$

将式(5)作为迭代公式,对目标函数进行迭代,前后 2 次迭代误差用 $\Delta J_w(G, V)$ 表示,给定正数用 ε 表示,如果 $\Delta J_w(G, V) < \varepsilon$, 停止聚类,并结合 U 完成聚类分析。

1.3 变压器容量自动校核

在进行配电变压器运行数据采集时,由于变压器的运行环境中存在大量噪声以及其他干扰,导致采集的数据中存在一定的噪声^[13],聚类后的数据中

也存在一定噪声数据, 为保证变压器容量的自动校核效果, 避免噪声数据对校核结果造成干扰^[14], 笔者采用基于电压降时间-相间差分的容量校核方法。该方法整体分为: 基于相间差分的数据降噪; 基于时间差分方法实现变压器容量自动校核 2 个步骤。

1.3.1 基于相间差分的数据降噪

聚类获取的电压降数据中, 包含变压器的高压侧和低压侧的数据。中高压侧数据中, 含有大量噪声; 且该侧, 输电线路的长短存在差异^[15], 不同线路中的压降导致变压器高、低压 2 侧的电压结果, 均低于 2 侧的额定电压 U_{1N} 和 U_{2N} , 进而导致电压差值的结果和实际电压差结果之间发生显著误差。为了保证容量的自动校核效果, 利用相间差分方法对高压侧数据进行处理。

如同一时刻, 同一变压器的 3 相电压以及运行参数均一致, 为降低高压侧额定容量校核误差, 该侧电压的处理采用相间差分方法完成。

变压器电压降 K 的计算公式为:

$$K = \frac{U_{2N} - U_2}{\eta} \quad (6)$$

式中: η 为负载率; U_2 为 2 次侧电压。由于配网中 2 次侧额定电压为 220 V, 理论上 K 的取值范围为 2.2~8.8 V。

依据式(6)计算变压器 3 相电压降, 其公式为:

$$\left. \begin{aligned} K_A &= \frac{U_{2N} - U_A}{\eta_A} \\ K_B &= \frac{U_{2N} - U_B}{\eta_B} \\ K_C &= \frac{U_{2N} - U_C}{\eta_C} \end{aligned} \right\} \quad (7)$$

式中: U_A 、 U_B 、 U_C 为实际电压, 分别对应 A、B、C 3 相; 负载率分别为 η_A 、 η_B 、 η_C 。

由于同一变压器的各相之间运行参数和运行特性几乎一致, 因此得出:

$$K_A \approx K_B \approx K_C \approx K \quad (8)$$

仅以 A、B 2 个相为例, 将式(7)中的前 2 个式子相减后得出:

$$(U_A - U_B) = (\eta_B - \eta_A)K \quad (9)$$

将式(9)和(7)对比后可知, 式(7)第一个式子中的 U_{2N} 参数经过差分后消除了, 因此, 高压侧数据对 U_{2N} 的干扰被极大程度弱化。依据上述原理, 即可完成其他 2 相中高压侧数据对 U_{2N} 干扰的处理, 降低高压侧数据噪声造成的影响。

1.3.2 基于时间差分的变压器容量自动校核

通过高压侧数据处理后, 笔者采用时间差分方法进行配电变压器容量自动校核。

仍以 A、B 2 个相为例, 通过时间差分方法计算 A、B 2 个相的负载率和电压之间的关系:

$$K_A = (U_A^t - U_A^{t-1}) / (\eta_A^t - \eta_A^{t-1}); \quad (10)$$

$$K_B = (U_B^t - U_B^{t-1}) / (\eta_B^t - \eta_B^{t-1}) \quad (11)$$

式中 t 和 $t-1$ 为时刻。

在式(11)的基础上, 获取基于时间差分 and 基于相间差分的负载率和电压之间的关联:

$$K_{AB} = \frac{(U_A^t - U_A^{t-1}) - (U_B^t - U_B^{t-1})}{(\eta_A^t - \eta_A^{t-1}) - (\eta_B^t - \eta_B^{t-1})} \quad (12)$$

式中: K_{AB} 为电压降, 该结果是通过时间-相间差分后得出。

式(12)计算的结果即可理解为在不同时刻下, 同一变压器相和相之间的 η , 引起的 K 和测量得出的电压差在理论上一致, 为了降低计算的复杂程度, 式(12)转换成:

$$K_{AB} = (\Delta U_A - \Delta U_B) / (\Delta \eta_B - \Delta \eta_A) \quad (13)$$

式中: $\Delta U_A = U_A^t - U_A^{t-1}$ 为 A 相上的电压差; $\Delta U_B = U_B^t - U_B^{t-1}$ 为 B 相上的电压差; $\Delta \eta_A = \eta_A^t - \eta_A^{t-1}$ 为 A 相上的负载率差; $\Delta \eta_B = \eta_B^t - \eta_B^{t-1}$ 为 B 相上的负载率差。依据上述公式, 同理即可获取 K_{AC} 、 K_{BC} 的结果。

结合式(6)和(13)可知, 经过相间-时间差分处理后, K_{AB} 和额定容量 S_N 之间呈正比例关系。对 K_{AB} 的结果进行判断, 确定其是否满足 $2.2 \text{ V} \leq K_{AB} \leq 8.8 \text{ V}$ 条件, 以此判断变压器容量结果是否准确, 实现配电变压器容量自动校核。

2 实验结果分析

为验证笔者方法的应用效果, 以某变电站内的 30 台配电变压器作为测试对象, 该变压器的型号为 Dyn11, 该变压器的额定容量为 2 500 kva, 电源相数为 3 相, 其高压为 10 kV 低压为 0.4 kV, 空载损耗为 2 290 W, 负载损耗为 21 200 W, 空载电流为 0.4 A, 短路阻抗为 5%, 其安装尺寸为 820 mm×820 mm, 运行效率为 85%。

获取该变电站内 30 台变压器运行 1 个月的历史数据, 将该数据作为测试使用数据。

笔者在进行变压器容量自动校核时, 需先对变压器的电压降数据进行聚类。为验证笔者方法对于配电变压器电压降数据的聚类效果, 获取笔者方法

应用后，在 2 维平面内，数据的聚类结果如图 2。

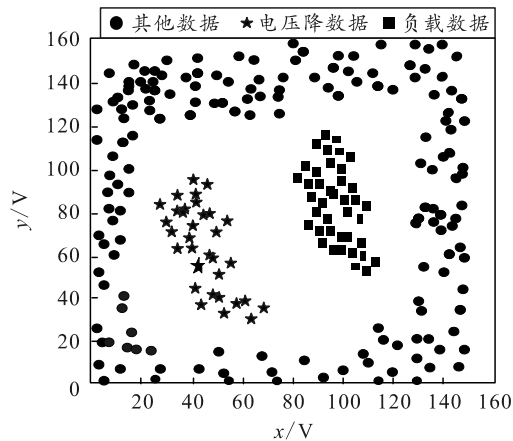


图 2 数据聚类测试结果

依据图 2 测试结果可知：应用文中方法后，能有效完成配电变压器电压降数据的聚类，同时可区分出其他类别数据；且各个类别数据之间没有发生重叠、交叉等现象。可见，该方法能够有效完成电压器电压降数据聚类，为变压器容量校核提供可靠的数据依据。

为进一步验证该方法的聚类效果，笔者采用 Xie-Beni 系数作为评价指标，该指标的结果在 0~1 之间，其值越小，表示聚类效果越佳，该指标的计算公式为：

$$\xi(G, c) = \frac{\frac{1}{n} \sum_{j=1}^c \sum_{i=1}^n g_{ij}^2 \|x_i - v_i\|}{\min_{i,j} \|v_i - v_j\|} \quad (14)$$

式中： $\min_{i,j} \|v_i - v_j\|$ 为聚类结果的紧致性；

$\frac{1}{n} \sum_{j=1}^c \sum_{i=1}^n g_{ij}^2 \|x_i - v_i\|$ 为同类别数据之间分离性。

依据式(14)计算本文中方法在不同数据类别下，随着数据量大小的逐渐增加， $\xi(U, c)$ 的计算结果，如表 1 所示。

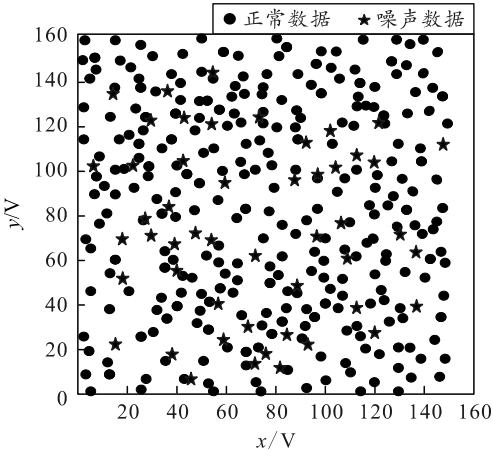
表 1 Xie-Beni 系数测试结果

数据量	类别数量		
	1	2	3
100	0.014	0.015	0.021
200	0.018	0.012	0.017
300	0.022	0.023	0.016
400	0.016	0.027	0.030
500	0.009	0.029	0.022
600	0.011	0.014	0.018
700	0.025	0.009	0.028
800	0.015	0.013	0.022
900	0.013	0.015	0.014
1 000	0.011	0.019	0.016

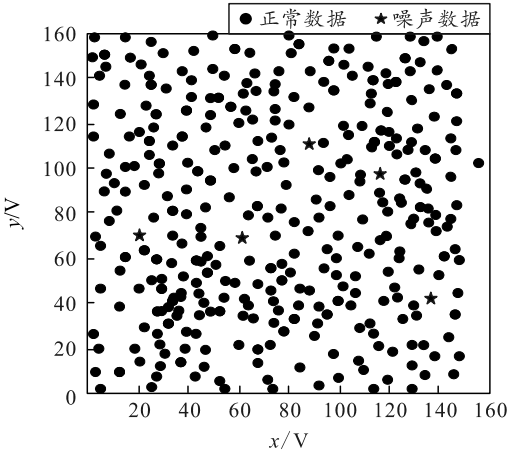
依据表 1 的测试结果可知：文中方法在进行不

同类别变压器数据聚类时，随着数据量的不断增加， $\xi(U, c)$ 的计算结果均在 0.030 以下，最小结果为 0.004；因此，本文中方法的聚类效果，能够满足变压器电压降数据的聚类需求。

本文中方法在进行变压器容量自动校核时，是依据电压降数据完成，因此，测试该方法在不同的负载率差值下，随着噪声数据的逐渐增加该方法对变压器 3 相的电压降计算结果如图 3 所示。



(a) 降噪前数据情况



(b) 降噪后数据情况

图 3 数据降噪测试结果

依据图 3 测试结果可知：采用本文中方法对数据进行处理后，能够有效去除数据中的噪声数据，极大程度减少噪声数据对变压器容量校核的影响。

为验证本文中方法对于配电变压器自动容量校核的应用效果，采用本文中方法对 30 台变压器的容量进行自动校核，获取时间-相间差分后，各个变压器相和相之间的电压降平均值，并依据该结果完成容量校核；并且获取校核结果的均方根误差平均值，该值越接近 1 表示校核效果越佳，测试结果如表 2 所示。由于篇幅有限，结果仅随机呈现 10 台变压器的校核结果。

表 2 配电变压器自动容量校核结果

变压器序号	电压降平均值/V	均方根误差平均值	校核结果
1	5.8	0.91	正确
2	5.0	0.93	正确
3	7.2	0.94	正确
4	5.5	0.92	正确
5	9.3	0.90	错误
6	7.7	0.91	正确
7	9.5	0.89	正确
8	9.5	0.92	错误
9	7.0	0.93	错误
10	6.5	0.95	正确

依据表 2 的测试结果可知：本文中方法能够获取各个变压器相和相之间的电压降平均值，并确定序号 5、7、8 变压器容量错误，电压降的平均值超过 $2.2 \text{ V} \leq K_{AB} \leq 8.8 \text{ V}$ 标准，该结果和实际结果吻合，证明本文中方法能够可靠完成变压器容量的自动校核；并且，均方根误差平均值均在 0.89 以上，极大程度接近 1。可见，本文中方法具有较好的应用效果，满足变压器容量自动校核需求。

3 结论

配电变压器作为配电网中的重要组成部分，其额定容量是相关部门依据变压器的安装位置、负荷以及过负荷承受能力等进行选择和确定，如果变压器的额定容量和实际容量不吻合，则会直接影响供电的经济效益；因此，笔者提出基于数据聚类的配电变压器自动容量校核方法。对该方法进行测试后得出：该方法具有较好的应用效果，能够完成变压器电压降数据的有效聚类，并且，能够全面处理数据中的噪声数据干扰，精准完成变压器容量校核。

参考文献：

- [1] 陈磊, 姜飞, 涂春鸣. 考虑负荷最大概率分布与故障风险的配电变压器优选方法[J]. 电网技术, 2021, 45(3): 1109-1116.
- [2] 耿俊成, 郭志民, 李晓蕾, 等. 基于 LOF 和 SVM 的配电网线变关系数据校验方法[J]. 中国测试, 2021, 47(4): 49-54.
- [3] 刘明群, 何鑫, 覃日升, 等. 基于改进 K-means 聚类 k 值选择算法的配网电压数据异常检测[J]. 电力科学与技术学报, 2022, 37(6): 91-99.
- [4] 柴大鹏, 王晓雯. 配电变压器低压无功功率自动补偿方法研究[J]. 计算机仿真, 2021, 38(5): 51-54, 198.
- [5] 范苏纯, 黄向敏, 张勇军, 等. 计及有序充电的配电变压器扩展规划建模[J]. 电力系统自动化, 2021, 45(7): 62-70.
- [6] 蔡秀雯, 陈茂新, 陈钢, 等. 考虑主动管理与需求侧管理的主动配电网分布式光伏最大准入容量计算方法[J]. 电力建设, 2021, 42(5): 38-47.
- [7] 刘振, 李林, 梁博渊, 等. 直觉模糊层次法下的配电变压器质量评估研究[J]. 中南民族大学学报(自然科学版), 2022, 41(5): 553-557.
- [8] 缪宇峰, 汪李忠, 郭强, 等. 基于中低频零序扫描阻抗的配电变压器容量检测方法[J]. 电工电能新技术, 2022, 41(7): 64-71.
- [9] 李振坤, 管琰玲, 张代红, 等. 基于年时序曲线的配电网节点光伏接纳能力评估及最佳容量研究[J]. 太阳能学报, 2021, 42(4): 200-208.
- [10] 刘健, 陈鲁鹏, 张志华. 基于本地量测信息的配电变压器静态参数估计[J]. 电力自动化设备, 2021, 41(2): 71-76.
- [11] 李汉伟. 地铁长大隧道内大功率风机启动对变压器容量选择的影响分析[J]. 城市轨道交通研究, 2022, 25(3): 178-181.
- [12] 王文森, 杨晓西, 刘阳, 等. 基于层次聚类分析的变压器油中溶解气体在线监测数据异常检测[J]. 高压电器, 2023, 59(1): 142-147.
- [13] 李克明, 江亚群, 黄世付, 等. 基于 DTW 距离和聚类分析的配电台区低压拓扑结构辨识方法[J]. 电力系统保护与控制, 2021, 49(14): 29-36.
- [14] 张悦, 宋运忠. 一种基于类内类间协同的用电数据聚类选择[J]. 武汉大学学报(工学版), 2022, 55(5): 493-502.
- [15] 李大伟, 梁得亮, 高亚晨, 等. 一种基于模糊控制的混合式配电变压器的控制策略[J]. 电工技术学报, 2021, 36(S2): 696-703, 722.