DOI：10.13234/j.issn.2095-2805.2026.7.243 **中图分类号**：TM743 **文献标志码**：A

谐波和三相不平衡扰动源联合概率建模方法

张一帆[1,2]，潘大志[3]，张永宏[4]，辛力坚[1]，汪 颖[5]，王心茹[5]
(1.内蒙古电力(集团)有限责任公司内蒙古电力科学研究院分公司，呼和浩特 010020；2.内蒙古自治区新型电力系统智能电网企业重点实验室，呼和浩特 010020；3.内蒙古电力(集团)有限责任公司，呼和浩特 010020；4.内蒙古电力(集团)有限责任公司鄂尔多斯供电公司，鄂尔多斯 017000；5.四川大学电气工程学院，成都 610065)

摘要：谐波和三相不平衡是电力系统常见的2种扰动形式，彼此相互作用与影响，共同对电力系统造成更复杂的影响。为了更全面、准确地分析谐波和三相不平衡的分布规律与相互作用机制，对谐波和三相不平衡扰动源H-TU-DS(harmonic and three-phase unbalanced disturbance sources)的联合发射水平进行研究，提出了一种联合概率发射水平建模方法。首先，基于扰动源功率、谐波和三相不平衡数据的多元相关系数，筛选出扰动源特征数据集，实现数据降维。接着，利用模糊C均值FCM(fuzzy C-means)算法完成对特征数据的聚类划分，为扰动源分工况分析奠定基础。然后，通过非参数核密度估计方法构建H-TU-DS的联合概率模型。最后，通过仿真和实测数据对比，证明了所提特征数据集筛选方法、运行工况划分方法和联合概率建模方法的有效性和准确性。

关键词：谐波；三相不平衡；非参数核密度；概率建模

Joint Probabilistic Modeling Approach for Harmonic and Three-phase Unbalanced Disturbance Sources

ZHANG Yifan[1,2], PAN Dazhi[3], ZHANG Yonghong[4], XIN Lijian[1], WANG Ying[5], WANG Xinru[5]
(1. Inner Mongolia Power (Group) Co., Ltd., Inner Mongolia Power Research Institute Branch, Hohhot 010020, China; 2. Inner Mongolia Key Laboratory of Smart Grid of New-type Power System, Hohhot 010020, China; 3. Inner Mongolia Power (Group) Co., Ltd., Hohhot 010020, China; 4. Inner Mongolia Power (Group) Co., Ltd., Ordos Power Supply Company, Ordos 017000, China; 5. College of Electrical Engineering, Sichuan University, Chengdu 610065, China)

Abstract: Harmonics and three-phase unbalance are two common forms of disturbances in power systems, which interact with each other and jointly cause more complex impacts on power systems. In order to more comprehensively and accurately analyze the distribution law and interaction mechanism of harmonics and three-phase unbalance, the joint emission level of harmonic and three-phase unbalanced disturbance sources (H-TU-DS) is investigated, and a joint probabilistic emission level modeling method is proposed. First, based on the multivariate correlation coefficients of disturbance source power, harmonic and three-phase unbalanced data, the disturbance source characteristic dataset is screened to realize the data dimensionality reduction. Subsequently, the fuzzy C-means (FCM) algorithm is used to complete the clustering division of the feature data, which lays the foundation for the disturbance source sub-condition analysis. Then, the joint probability model of H-TU-DS is constructed by the nonparametric kernel density estimation method. Finally, the effectiveness and accuracy of the proposed feature dataset screening method, operation condition classification method and joint probability modeling method are proved by comparing and verifying the simulated and measured data.

收稿日期：2024-09-10；**修回日期**：2024-10-03；**录用日期**：2024-10-13；**网络首发日期**：2024-10-25
基金项目：2023年内蒙古电力公司科技创新“揭榜挂帅”资助项目(2023-3-2)；2023年内蒙古自治区科技资助计划(2023KYPT0020)
This work is supported by 2023 Inner Mongolia Power (Group) Co., Ltd. Science and Technology Innovation “Unveiled Marshal” Project under the grant 2023-3-2; 2023 Inner Mongolia Autonomous Region Science and Technology Program under the grant 2023KYPT0020

Keywords: Harmonics; three-phase unbalance; nonparametric kernel density; probability modeling

随着经济与技术的发展，用户对电能质量的要求越来越高，不仅要求保证电能稳定持续供应，还要求电能的安全可靠性足够高[1-4]。其中，影响用户最多的是电网谐波和三相不平衡问题。新能源机组、冶炼负荷、轧机等设备的投入运行会给电网带来谐波扰动和三相不平衡扰动，这类设备既是谐波源，也是三相不平衡扰动源。为了评估此类扰动源的扰动特性，构建谐波和三相不平衡扰动源 H-TU-DS(harmonic and three-phase unbalanced disturbance sources)的联合扰动发射水平模型是目前亟需攻克的课题。

但是，截至目前，关于电网扰动源建模考虑的电能质量因素单一，主要针对单一谐波扰动源和单一三相不平衡扰动源进行建模，而对于 H-TU-DS 的联合概率建模研究较少。现有关于谐波源的建模方法主要可以分为基于机理分析的建模方法和基于数据分析的建模方法。基于机理分析的谐波源建模一般通过分析电压和电流的伏安特性，构建一种能够描述谐波电压和谐波电流耦合关系的物理模型[5]。目前国内外研究者提出的谐波源物理模型有恒流源模型[6]、谐波耦合导纳矩阵模型[7]和诺顿等效模型[8]。基于数据分析的建模方法，主要是通过数据建立谐波源机器学习模型[9]和概率统计模型[10]。关于负荷三相不平衡建模，文献[11]提出将负序不平衡因素等效为戴维南等值电路，建立不平衡分析的数学模型，同时将不平衡源等效为三相不平衡阻抗，通过改变阻抗的大小实现三相不平衡的等效仿真模型。在工程上，常用的方法是基于用户监测数据计算三相不平衡度，进而评估不平衡源对系统不平衡发射水平的影响[12]。对于兼有谐波扰动特性和三相不平衡扰动特性的电网用户，尚无研究者明确提出其联合概率模型的构建方法。

在联合建模领域，建模方法主要分为基于参数估计的方法和基于非参数估计的方法。基于参数估计的方法有 Copula 分析方法[13]、高斯混合模型[14]和联合分布的混合模型[15]等，但上述方法模型均需预先假设待观测数据遵循某种分布，进而开展参数估计。参数估计方法依赖对输入数据的先验假设，对处理复杂且未知分布的数据具有局限性。对于本文的研究对象 H-TU-DS，其实际运行时的数据分布特征是未知的，无法对数据进行概率分布假设，故采用基于非参数估计的方法进行联合建模。在非参数估计方法中，核密度函数凭借其适应多种概率分布的数据类型且不需要参数估计的优势得到广泛运用。Hao 等[16]在研究中指出，非参数核密度模型能够准确处理光伏发电出力概率建模中的边界偏差和局部适应性问题；Hou 等[17]将非参数核密度用于解决可再生能源的不确定性问题，提出了鲁棒核密度估计的数据驱动最优调度方法，强调了非参数核密度无需数据先验假设在处理不确定性问题上的优势。以上文献均表明了非参数核密度广泛运用于电力系统不确定性问题的处理，且具有较高的准确性和鲁棒性。

为了分析 H-TU-DS 的谐波和不平衡双重扰动特性，本文提出了一种基于非参数核密度估计的 H-TU-DS 联合建模方法。首先采集 H-TU-DS 并网点功率数据、谐波数据和三相不平衡数据，提出扰动源电能质量发射水平的特征数据集构建方法；接着考虑 H-TU-DS 在不同运行状态下电能质量发射水平存在差异的情况，对 H-TU-DS 进行分工况分析，提出基于特征数据集聚类的工况划分方法；然后采用双变量非参数核密度函数对不同工况下的 H-TU-DS 构建概率分布函数模型；最后通过仿真验证了所提方法的有效性和准确性。

本文所提出的联合建模方法为深入理解 H-TU-DS 的电能质量发射特性提供了坚实的基础。通过构建的特征数据集和联合概率分布函数模型，可以更准确地描述 H-TU-DS 在电网中的动态行为及其对电网电能质量的潜在影响。文中的联合概率分布模型能够运用于配电网分析，有效评估 H-TU-DS 接入电网对电能质量的影响。同时，根据不同运行工况下 H-TU-DS 的运行特性，当其运行状态发生

变化时，可提前预测电能质量特性并采取预防措施，从而确保电网的稳定和用户的电力供应质量。文献[18]也表明，基于概率模型的方法能够提高对电力系统扰动源行为预测和评估的准确性。因此，本文方法在实际应用中具有较高价值，为电能质量的预测和评估提供了一种新途径。

1　基于多元相关系数的特征筛选

分析由电能质量监测装置采集到的 H-TU-DS 并网点处的功率数据、三相不平衡度数据和谐波数据。由于提取的各类数据均具有多维性的特点，为实现数据降维、降低计算难度、提高建模效率，本文提出了基于多元相关系数的 H-TU-DS 特征数据筛选方法，具体步骤如下。

步骤 1 数据集 $\boldsymbol{A}$、$\boldsymbol{E}$、$\boldsymbol{H}$ 获取。

功率数据集 $\boldsymbol{A}$ 可表示为

$$\boldsymbol{A}=[P_j \;\; Q_j \;\; S_j] \qquad j=1,2,\cdots,N \tag{1}$$

式中：P_j、Q_j、S_j 分别为扰动源有功功率、无功功率、视在功率在 j 时刻的采样值；N 为总采样点数。

三相不平衡度数据集 $\boldsymbol{E}$ 可表示为

$$\boldsymbol{E}=[\varepsilon_{u0,j} \;\; \varepsilon_{u2,j} \;\; \varepsilon_{i0,j} \;\; \varepsilon_{i2,j}] \tag{2}$$

式中，$\varepsilon_{u0,j}$、$\varepsilon_{u2,j}$、$\varepsilon_{i0,j}$、$\varepsilon_{i2,j}$ 分别为扰动源电压和电流零序、负序不平衡度在 j 时刻的采样值。

谐波数据集 $\boldsymbol{H}$ 可表示为

$$\boldsymbol{H}=[I_{\mathrm{T},\phi,j} \;\; U_{\mathrm{T},\phi,j}] \qquad \phi=\mathrm{A,B,C} \tag{3}$$

式中，$I_{\mathrm{T},\phi j}$、$U_{\mathrm{T},\phi j}$ 分别为 ϕ 相电流、电压总谐波畸变率在 j 时刻的采样值。

由式(1)~式(3)可知，$\boldsymbol{A}$ 包含 3 个数据指标、$\boldsymbol{E}$ 包含 4 个数据指标、$\boldsymbol{H}$ 包含 6 个数据指标。通过依次从 $\boldsymbol{A}$、$\boldsymbol{E}$ 和 $\boldsymbol{H}$ 中提取 1 个数据指标，构建不同的数据组合。根据排列组合原则，共有 72 种数据组合(3×4×6)，依次对 72 组数据进行计算分析，筛选出 H-TU-DS 特征数据集，实现数据降维。

步骤 2 特征筛选。

多元相关系数具有可以表示 3 个或 3 个以上变量之间相关性的特点，多用于探索变量之间的复杂关系[19]，取值范围为[0, 1]，其值越接近 1，数据指标之间的线性相关性越强。因此，本文采用计算各组合的多元相关系数来描述 $\boldsymbol{A}$、$\boldsymbol{E}$、$\boldsymbol{H}$ 所提取数据指标之间的关联程度。多元相关系数的具体计算式为

$$R_{A,E,H}=\sqrt{\frac{r_{AE}^{\;2}+r_{AH}^{\;2}-2r_{AE}r_{AH}r_{EH}}{1-r_{EH}^{\;2}}} \tag{4}$$

式中：$R_{A,E,H}$ 为 $\boldsymbol{A}$、$\boldsymbol{E}$、$\boldsymbol{H}$ 之间的多元相关系数指标；r_{AE}、r_{AH}、r_{EH} 分别为各数据指标两两之间的皮尔逊相关系数。

以功率数据指标与三相不平衡度数据指标为例，r_{AE} 的计算式为

$$r_{AE}=\frac{\sum_{j=1}^{N}(A_{f,j}-A_{f0})(E_{g,j}-E_{g0})}{\sqrt{\sum_{j=1}^{N}(A_{f,j}-A_{f0})^2}\sqrt{\sum_{j=1}^{N}(E_{g,j}-E_{g0})^2}} \tag{5}$$

式中：$A_{f,j}$、$E_{g,j}$ 分别为 $\boldsymbol{A}$ 的第 f(f=1, 2, 3)个数据指标和 $\boldsymbol{E}$ 的第 g(g=1, 2, 3, 4)个数据指标的第 j 个采样值；A_{f0}、E_{g0} 分别为对应数据指标的数学期望值。同理，可以依次计算出 r_{AH} 和 r_{EH}，并用 $H_{h,j}$、H_{h0} 分别表示 $\boldsymbol{H}$ 的第 h(h=1, 2, 3, 4, 5, 6)个数据指标的第 j 个采样值和对应的数学期望值。计算得到多元相关系数集合 R={$R_{f,g,h}$ | f=1, 2, 3; g=1, 2, 3, 4; h=1, 2, 3, 4, 5, 6}，其中，$R_{f,g,h}$ 为 $\boldsymbol{A}$ 的第 f 个数据指标、$\boldsymbol{E}$ 的第 g 个数据指标与 $\boldsymbol{H}$ 的第 h 个数据指标之间的多元相关系数。

步骤 3 特征数据集构建。

筛选出 R 中最大元素 max$R$=$R_{F,G,H}$，在 $\boldsymbol{A}$ 的第 F 个数据指标、$\boldsymbol{E}$ 的第 G 个数据指标与 $\boldsymbol{H}$ 的第 H 个数据指标时取得。根据式(4)与式(5)，将计算得到的 $R_{F,G,H}$ 对应的 $\boldsymbol{A}$、$\boldsymbol{E}$、$\boldsymbol{H}$ 数据指标作为 H-TU-DS 的特征数据集，用 $\boldsymbol{X}$($\boldsymbol{X}$ 是一个 $N\times3$ 的矩阵)表示。$\boldsymbol{X}$ 包含 3 列数据，分别来自 $\boldsymbol{A}$ 的第 F 列、$\boldsymbol{E}$ 的第 G 列、$\boldsymbol{H}$ 的第 H 列，分别表示功率特征数据指标、三相不平衡特征数据指标和谐波特征数据指标。$\boldsymbol{X}$ 可表示为

$$\boldsymbol{X}=[A_{F,j}\ \ E_{G,j}\ \ H_{H,j}]\quad j=1,2,\cdots,N \tag{6}$$

式中：$A_{F,j}$为 $\boldsymbol{A}$ 的第 F 个数据指标在 j 时刻的采样值；$E_{G,j}$为 $\boldsymbol{E}$ 的第 G 个数据指标在 j 时刻的采样值；$H_{H,j}$为 $\boldsymbol{H}$ 的第 H 个数据指标在 j 时刻的采样值。

2 基于模糊 C 均值算法的扰动源工况划分

本文的研究对象是同时具有谐波发射特性和三相不平衡发射特性的 H-TU-DS。在第 1 节的研究中已经得到该类扰动源的 $\boldsymbol{X}$。考虑到 H-TU-DS 在不同运行条件下所产生的电能质量问题可能不同，本文对 H-TU-DS 采用分工况分析方法。理论上，可以用聚类的方法实现工况划分目标。

$\boldsymbol{X}$ 中的特征数据具有整体呈非均匀分布但局部聚拢的特点，模糊 C 均值 FCM(fuzzy C-means)聚类算法对于处理这类数据具有较高的灵活度和准确度。因此，本文提出了一种基于 FCM 算法的扰动源工况划分方法，通过 FCM 聚类算法将 $\boldsymbol{X}$ 划分为不同的类别，并将每一个类别看作一种运行工况。在同一种运行工况下，扰动源的电能质量发射水平可视为一致。

2.1 FCM 聚类算法

$\boldsymbol{X}$ 中的采样点为离散点。FCM 根据不同样本点对聚类中心的隶属度进行类别划分[20]。FCM 聚类算法的一般计算步骤如下。

步骤 1 聚类初始化。

将 $\boldsymbol{X}$ 作为输入，并将 $\boldsymbol{X}$ 的每一行作为一个数据点 $\boldsymbol{x}_j$，$\boldsymbol{x}_j=\{X_{j,1},X_{j,2},X_{j,3}\}$。设置初始聚类中心集合 $C=\{\boldsymbol{c}_i\mid i=1,2,\cdots,k\}$，$k$ 为聚类簇数，$\boldsymbol{c}_i$为第 i 个聚类簇的聚类中心；取模糊度因子 $m=2$，初始迭代次数 $t=1$、$t_{\max}=100$；初始化隶属度矩阵 $\boldsymbol{U}_{ji}$。$\boldsymbol{X}$ 中每个数据点 $\boldsymbol{x}_j$ 到 C 的隶属度和均为 1，其简要表达式为

$$\sum_{i=1}^{k}\boldsymbol{U}_{ji}=1\qquad \boldsymbol{U}_{ji}\in[0,1] \tag{7}$$

步骤 2 构建目标函数。

FCM 算法将所有数据点 $\boldsymbol{x}_j$ 到各个聚类中心 $\boldsymbol{c}_i$ 的欧氏距离的和作为目标函数 J_m，并通过最小化 J_m 来优化 $\boldsymbol{c}_i$ 的选取。反复迭代运算，直至 J_m 收敛，此时所得 C 为 k 下的最优解，并可得对应的 $\boldsymbol{U}_{ji}$。J_m 的表达式为

$$J_m=\sum_{j=1}^{N}\sum_{i=1}^{k}\boldsymbol{U}_{ji}^{m}\left\|\boldsymbol{x}_j-\boldsymbol{c}_i\right\|^2 \tag{8}$$

式中，U_{ji}^{m}为 FCM 迭代计算的加权系数。

步骤 3 更新隶属度矩阵和聚类中心集合。

J_m 每迭代一次，均会得到新的 $\boldsymbol{U}_{ji}$ 和 C。其中，$\boldsymbol{U}_{ji}$ 与 $\boldsymbol{c}_i$ 可分别表示为

$$\boldsymbol{U}_{ji}=\sum_{z=1}^{k}\left(\frac{\left\|\boldsymbol{x}_j-\boldsymbol{c}_i\right\|}{\left\|\boldsymbol{x}_j-\boldsymbol{c}_z\right\|}\right)^{-\frac{2}{m-1}} \tag{9}$$

$$\boldsymbol{c}_i=\frac{\sum_{j=1}^{N}\boldsymbol{U}_{ji}^{m}\boldsymbol{x}_j}{\sum_{j=1}^{N}\boldsymbol{U}_{ji}^{m}} \tag{10}$$

式中，$\boldsymbol{c}_z$为第 z 个聚类簇的聚类中心。迭代收敛条件为

$$\max\{|\boldsymbol{U}_{ji}^{t+1}-\boldsymbol{U}_{ji}^{t}|\}<\varepsilon \tag{11}$$

式中，ε 为误差阈值，$\varepsilon=10^{-5}$。

如果不满足式(11)并且 $t<t_{\max}$，就转回步骤 2 继续迭代，此时 $t=t+1$；否则就转到步骤 4。

步骤 4 输出聚类结果。

输出聚类簇数为 k 时的最优聚类结果。最优聚类中心即为扰动源最优运行工况中心，扰动源运行工况总数也为 k(扰动源总运行工况数与聚类簇数意思相同，前者为专业说法、后者为算法参数说法)，每个样本点属于其隶属度值最大的运行工况。

2.2 扰动源最优工况数 K

上述 FCM 聚类算法需要提前设置 k，由于 k 的选择具有主观性，故得到的扰动源总运行工况

数不够准确。本文针对这一情况做了改善，采用簇内差异最小、簇间差异最大的原则来选取最优聚类工况数，用 K 表示 H-TU-DS 最优运行工况数。其中，聚类评价指标簇内-簇间比值指标 CWB (cluster within-between ratio)的计算步骤如下。

步骤 1 确定运行工况总数 k 的取值范围。

分析本文所研究 H-TU-DS 的运行特点，根据是否发生谐波扰动和三相不平衡扰动，可以粗略地将扰动源的运行工况划分为 4 种情况。同时考虑到在扰动源工况聚类过程中，运行工况总数一般都大于 1，所以本文中扰动源运行工况总数 k 的取值范围为{2, 3, 4}，初始聚类簇数 k=2。

步骤 2 确定簇内差异、簇间差异。

在设置运行工况总数为 k 的前提下，通过第 2.1 节的 FCM 聚类算法输出工况聚类结果。簇内、簇间差异可分别表示为

$$C_{\mathrm{CW}}(a)=\frac{1}{N_i}\sum_{m=1}^{N_i}(x_m-c_a)^2 \quad a=1,2,\cdots,k \tag{12}$$

$$C_{\mathrm{CB}}(a,b)=(c_a-c_b)^2 \quad a,b=1,2,\cdots,k \tag{13}$$

式中：N_i 为第 2.1 节聚类结果中隶属于第 i 种运行工况的数据点个数；x_m 为第 i 种运行工况的第 m 个数据点；c_a、c_b 分别为第 a 个和第 b 个聚类中心点。

步骤 3 确定聚类评价指标值。

扰动源运行工况总数为 k 时的聚类评价指标值为

$$C_{\mathrm{CWB}}(k)=\frac{\max C_{\mathrm{CW}}(a)}{\min C_{\mathrm{CB}}(a,b)} \tag{14}$$

按照上述步骤依次计算 k={2, 3, 4}对应的聚类评价指标 C_{CWB}={C_{CWB}(2), C_{CWB}(3), C_{CWB}(4)}。取 minC_{CWB} 对应的 k 作为扰动源最优工况数 K，继而将 K 代入第 2.1 节作为 FCM 聚类簇数 k，重新执行第 2.1 节完整的 FCM 迭代流程，输出最优聚类结果。基于 FCM 聚类算法的扰动源工况划分流程如图 1 所示，详细描述了 FCM 算法流程、扰动源最优工况数确定的具体步骤。

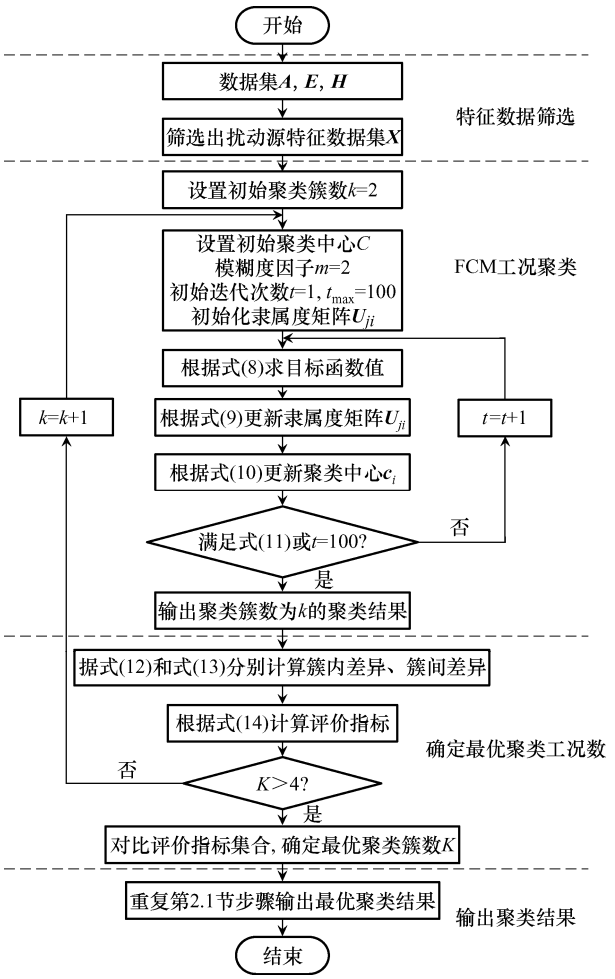


图 1　基于 FCM 聚类算法的工况划分流程

Fig. 1 Flowchart of working condition division based on FCM clustering algorithm

3　扰动源联合概率建模方法

扰动源在不同运行工况下的谐波数据和三相不平衡数据分布特征不同。为了更直观地分析 H-TU-DS 处于不同运行工况的数据分布特性，本文提出了基于非参数核密度估计的 H-TU-DS 联合概率建模方法。

3.1　联合概率密度函数构建

采用双变量非参数核密度函数对 H-TU-DS 不同运行工况下的谐波特性和三相不平衡特性进行联合概率建模分析。

将扰动源第 i 种运行工况下的谐波特征数据指标 H_{i,H,j_i} 和三相不平衡特征数据指标 E_{i,G,j_i} 作为输

入，j_i=1, 2, …, N_i；谐波数据和三相不平衡数据的联合概率值$f(H_{i,H}, E_{i,G})$作为输出，其表达式为

$$f(H_{i,H},E_{i,G})=\frac{1}{N_i\cdot h_{i,h}\cdot h_{i,e}}\cdot \sum_{j_i=1}^{N_i}K\left(\frac{\bar{H}_{i,H}-H_{i,H,j_i}}{h_{i,h}},\frac{\bar{E}_{i,G}-E_{i,G,j_i}}{h_{i,e}}\right) \tag{15}$$

式中：$h_{i,h}$、$h_{i,e}$分别为第 i 种运行工况下谐波数据指标和三相不平衡数据指标的窗宽；$\bar{H}_{i,H}$、$\bar{E}_{i,G}$分为对应数据指标的平均值；$K(\cdot)$为核函数。

3.2 参数选取

在使用非参数双变量核密度函数对谐波数据指标和三相不平衡度数据指标进行联合概率分析时，$K(\cdot)$与 $h_{i,h}$、$h_{i,e}$ 的选择均会对拟合的精度产生影响。研究表明，当 $h_{i,h}$、$h_{i,e}$ 一定时，$K(\cdot)$的选择对拟合精度的影响不大；当 $K(\cdot)$一定时，$h_{i,h}$、$h_{i,e}$ 的选择对拟合精度的影响较大，过小拟合出来的曲线平滑度差，过大拟合出来的曲线过于光滑无法反映细节[21]。

(1)核函数选择

选用二维高斯核函数。高斯核函数具有良好的光滑度和可微性，使估计的密度函数相对平滑，在H-TU-DS不同运行工况的建模中，有助于减小局部噪声的影响。具体核函数的表达式为

$$K\left(\frac{\bar{H}_{i,H}-H_{i,H,j_i}}{h_{i,h}},\frac{\bar{E}_{i,G}-E_{i,G,j_i}}{h_{i,e}}\right)=\frac{1}{2\pi}\cdot \exp\left[-\frac{(\bar{H}_{i,H}-H_{i,H,j_i})^2}{2h_{i,h}^2}-\frac{(\bar{E}_{i,G}-E_{i,G,j_i})^2}{2h_{i,e}^2}\right] \tag{16}$$

(2)窗宽选择

窗宽的不同取值对概率密度函数的影响较大，所以窗宽是一个需要优化的参数。本文通过计算估计值与实际值的均方根误差 RMSE(root mean square error)，选择使误差达到最小时的 $h_{i,h}$、$h_{i,e}$作为最优窗宽。均方根误差的计算式为

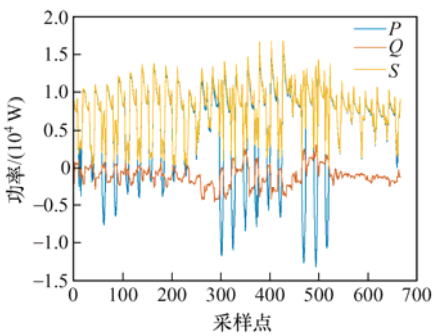
$$R_{\mathrm{RMSE}}=\sqrt{m_{\mathrm{mean}}[f_{\mathrm{emp}}(H_{i,H},E_{i,G})-f(H_{i,H},E_{i,G})]^2} \tag{17}$$

式中：$m_{\mathrm{mean}}(\cdot)$为求平均值函数；$f_{\mathrm{emp}}(H_{i,H}, E_{i,G})$为扰动源在第 i 种运行工况下的谐波和三相不平衡数据指标的实际概率密度分布值。

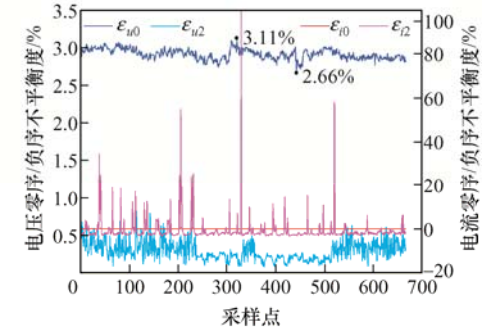
4 算例分析

4.1 算例数据

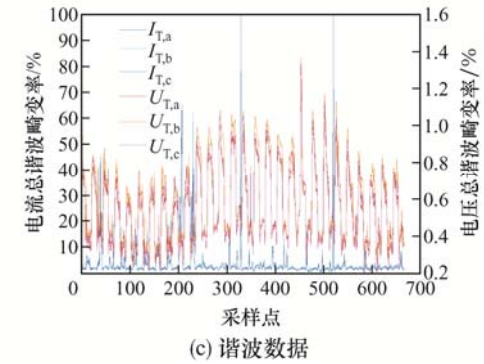
为了验证本文所提 H-TU-DS 联合概率建模方法的有效性与正确性，对中国中部某大型城市的某一交流电弧炉公共连接点 PCC(point of common coupling)的实际电能质量监测数据进行算例分析，监测点的电压等级为 35 kV、最小短路容量为 367.6 MV·A、供电设备容量为 90 MV·A。为了保证监测点数据能够反映负荷一段时间的运行状况，采样周期为一个月，采样时间间隔为 1 h，采样点总数为 720，采样点数据格式为 95%概率大值。监测点功率数据、三相不平衡度数据和谐波数据如图 2 所示。其中，图 2(b)显示的是 4 个三相不平衡数据指标。从图中可以看出，电压零序不平衡度的监测值较大，在 3.00%上下波动，波谷在第 440 个采样点取得，为 2.66%，波峰在第 320 个采样点取得，为 3.11%，波动跨值 0.45%；电压负序不平衡度的取值范围为[0, 1%]，在 240~320 采样点和 380~510 采样点之间取值接近 0 且变化平缓，剩余采样点取值变化无规则且剧烈；电流零序不平衡度监测点取值几乎无波动，在图中呈现为一条 0.5 附近的直线；电流负序不平衡度的波动范围最大，波动跨值 100%。此外，可以从图中清晰地观察出，取值突变点最能体现交流电弧炉的运行特点与变化。下述实验结果也表明，电流负序不平衡度最终被认定为三相不平衡度的特征数据指标。



(a) 功率数据



(b) 三相不平衡数据



(c) 谐波数据

图 2　某 35 kV 变电站 PCC 点监测数据

Fig. 2 PCC point monitoring data of a 35 kV substation

4.2　结果 1：扰动源特征数据集

利用式(4)和式(5)计算 ***A***、***E***、***H*** 的多元相关系数，获得 ***X***。多元相关系数计算结果见表 1，第 1 列为 ***A***、***E***、***H*** 各自包含的数据指标的 72 种组合，第 2、3、4 列分别对应 ***A*** 的第 f 个数据指标、***E*** 的第 g 个数据指标与 ***H*** 的第 h 个数据指标，第 5 列为每种组合形式的多元相关系数值 $R_{A,E,H}$。此算例中，$R_{A,E,H}$ 的最大值为 0.637，在第 71 组数据中取得，所以 ***X*** 由 ***A*** 的第 3 列(视在功率 S)、***E*** 的第 4 列(电流负序不平衡度 ε_{i2})和 ***H*** 的第 5 列(B 相电压总谐波畸变率 $U_{\mathrm{T,b}}$)组成，记为 $\boldsymbol{X}=[A_{3,j}\quad E_{4,j}\quad H_{5,j}]$。

表 1　多元相关系数结果

Tab. 1 Results of multiple correlation coefficients

序号	f	g	h	$R_{A,E,H}$
1	1	1	1	0.277
2	1	1	2	0.231
3	1	1	3	0.227
⋮	⋮	⋮	⋮	⋮
70	3	4	4	0.633
71	**3**	**4**	**5**	**0.637**
72	3	4	6	0.635

为验证所选特征数据集优于其他数据集，计算运用不同数据集进行联合概率建模的 RMSE。从表 1 中选取 6 个从小到大排列的 $R_{A,E,H}$，对应的组合分别为 33、2、34、51、64、71，并计算其 RMSE。图 3 为取不同 $R_{A,E,H}$ 时对应的 H-TU-DS 概率模型误差值 RMSE。从图中可以看出，$R_{A,E,H}$ 越大，RMSE 越小、概率模型精度越高。案例中选用的特征数据为第 71 组，RMSE 明显小于其余组合，表明所选特征数据集优于其他数据集。

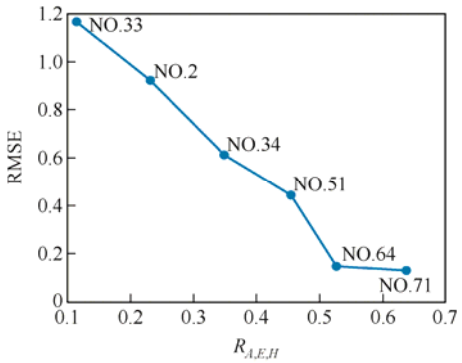


图 3　不同多元相关系数概率模型误差

Fig. 3 Probability model errors with different multivariate correlation coefficients

4.3　结果 2：运行工况划分

运用 FCM 算法对 ***X*** 进行划分，得到该交流电弧炉的运行工况。设置运行工况总数 k 分别为 2、3、4，通过计算每种情况的 CWB 聚类评价指标，选取该交流电弧炉最优运行工况数。不同运行工况数的 CWB 见表 2。从表中可以看出，在 k 取 2 时，CWB 取得最小值 0.095 9，工况划分效果最好。因此最优运行工况数为 2，该交流电弧炉存在 2 种典型运行工况，记为工况 1、工况 2。

表 2　不同运行工况总数的 CWB

Tab. 2 CWB for different total numbers of operating conditions

k	CWB
2	**0.095 9**
3	0.181 6
4	0.215 8

图 4 为当 k=2 时，该交流电弧炉的运行工况聚类结果。从图中可以看出：工况 1 和工况 2 最主要

的区别在于 S，工况 1 的 S 大于工况 2；工况 1 的 ε_{i2} 分布较集中、$U_{T,b}$ 分布较分散，工况 2 则相反。分析可得，该交流电弧炉工况 1 的谐波问题较为突出、工况 2 的三相不平衡问题较为突出。

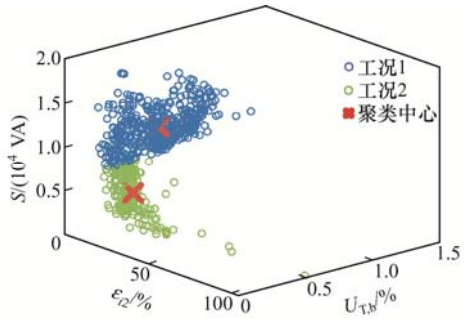


图 4 交流电弧炉工况划分结果

Fig. 4 AC arc furnace working condition classification results

图 5 和图 6 分别为工况 1 和工况 2 的实际联合概率密度值，$f(H, E)$表示谐波和三相不平衡度的实际联合概率密度值。由图 5 可知，工况 1 中柱形分布比较均匀，但谐波和三相不平衡仍存在一定的区别，ε_{i2} 集中分布在区间[0.80%, 1.40%]，呈现中间高、两边低的形态；$U_{T,b}$ 均匀分布在区间[1%, 3%]内。由图 6 可知，工况 2 中柱形分散，集中分布在 ε_{i2} 的区间[0.24%, 0.42%]内、$U_{T,b}$ 的区间[0, 26%]内。工况 1 和工况 2 的实际联合概率密度柱形图存在明显差异，证明了所提聚类方法的有效性。

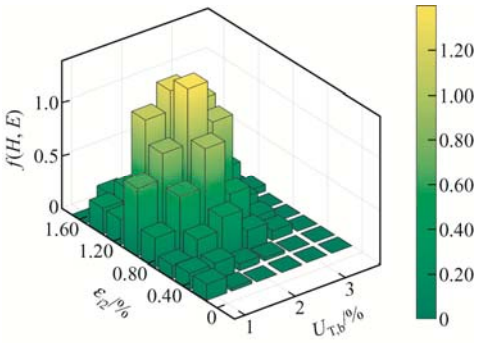


图 5 工况 1 实际联合概率密度

Fig. 5 Actual joint probability density for Case 1

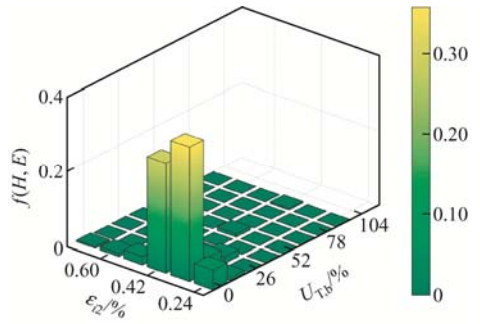


图 6 工况 2 实际联合概率密度

Fig. 6 Actual joint probability density for Case 2

4.4 结果 3：联合概率建模

运用双变量非参数核密度函数分别对工况1 和工况 2 的联合概率密度进行建模分析，建模结果分别如图 7 和图 8 所示。图 7 和图 8 建模图形的大致走向分别与图 5 和图 6 一致，建模的精度见表 3。其中 h_h、h_e 分别为谐波特征最优窗宽、三相不平衡特征最优窗宽。工况 1 的 RMSE 为 0.013，工况 2 为 0.244。由此可以看出，所采用的双变量非参数核密度拟合方法具有较高的精确性。

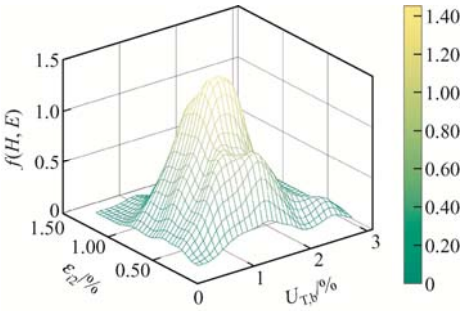


图 7 工况 1 联合概率密度建模结果

Fig. 7 Case 1 joint probability density modeling results

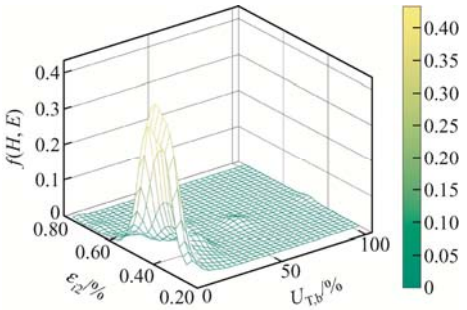


图 8 工况 2 联合概率密度建模结果

Fig. 8 Case 2 joint probability density modeling results

表 3 建模误差值和窗宽取值

Tab. 3 Modeling error values and window width values

工况	RMSE	h_h	h_e
工况 1	0.013	0.015 7	1.716 8
工况 2	0.244	0.055 7	0.197 8

4.5 时序扰动特性

本文所构建的 H-TU-DS 联合概率模型依赖于数据区间划分，反映了 H-TU-DS 的谐波指标和三相不平衡指标取不同数据区间大小时的概率密度值，清晰地显示了谐波和三相不平衡的相互关系。同时，本文也考虑了建立基于时序的扰动源扰动特征刻画的方法，如图 9 所示，刻画了连续 28 天的交流电弧炉扰动特性。从图中看出，该案例中交流电弧炉 $U_{\mathrm{T,b}}$ 在多数时间段大于 ε_{i2}，谐波问题占主导地位。ε_{i2} 波动幅度大，较易产生严重的电能质量事故，需要进行重点防范。图 9 从时间维度上分析扰动源的扰动特性，主要反映某时刻何种扰动占主导地位，以及持续运行时不同扰动的变化情况。

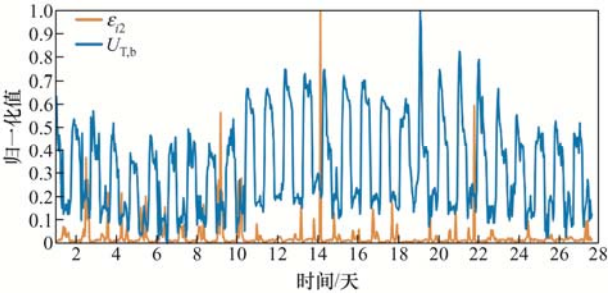


图 9　交流电弧炉时序扰动特征

Fig. 9 Temporal Disturbance Characteristics of the AC arc furnace

扰动源时序模型和本文提出的概率密度模型分别从时间和数据区段的角度对扰动源进行分析，从横向、纵向 2 个维度对扰动源扰动特征进行刻画。横向表明谐波和三相不平衡数据之间的联合关系，纵向表明扰动源发生扰动的时序变化特征。两者结合，更能准确且全面地刻画扰动源运行特性，对谐波和三相不平衡的治理更具指导意义。

5　结论

本文提出了一种谐波和三相不平衡扰动源联合概率建模方法，并采用实测数据进行仿真，验证了所提方法的有效性和准确性。

本文的主要工作分成 3 部分：①采集扰动源并网点数据，并通过计算各数据类型之间的多元相关系数，选出扰动源特征数据集；②运用 FCM 算法对扰动源特征数据集进行聚类划分，每个划分对应扰动源的每种工况，继而实现扰动源分工况分析；③采用双变量非参数核密度函数对扰动源每种运行工况进行联合概率密度建模。

所提方法揭示了 H-TU-DS 在不同工况下的概率分布，为分析该类扰动源的电能质量发射特性提供了基础。此外，将所建模型接入配电网有利于预测和评估扰动源对电能质量的影响，进而通过提前治理有效预防电能质量问题的发生。

参考文献：

[1] 付勉, 刘志涵, 宋振浩, 等. 配电网络电能质量综合治理设备优化配置策略[J]. 电源学报, 2024, 22(2): 336- 344.
Fu Mian, Liu Zhihan, Song Zhenhao, et al. Optimal allocation strategy for comprehensive control equipment of power quality in distribution area [J]. Journal of Power Supply, 2024, 22(2): 336-344. (in Chinese)

[2] 熊亮雳, 游力, 韩刚, 等. 基于统一不平衡/谐波电压-电流下垂的构网型变流器电能质量控制[J]. 电源学报, 2025, 23(4): 186-199.
Xiong Liangli, You Li, Han Gang, et al. Power quality control for grid-forming converters based on unified unbalanced/harmonic voltage-current droop [J]. Journal of Power Supply, 2025, 23(4): 186-199. (in Chinese)

[3] 李建闽, 曹远远, 姚文轩, 等. 基于自适应移频滤波的电力系统谐波分析方法[J]. 电工技术学报, 2024, 39(13): 4015-4024.
Li Jianmin, Cao Yuanyuan, Yao Wenxuan, et al. Power system harmonic analysis method based on adaptive frequency-shift filtering [J]. Transactions of China Electrotechnical Society, 2024, 39(13): 4015-4024. (in Chinese)

[4] 范鹏程, 李明哲, 庄坤, 等. 含高比例分布式光伏的用户侧共享储能容量优化配置方法[J]. 电源技术, 2025, 49(2): 443-453.
Fan Pengcheng, Li Mingzhe, Zhuang Kun, et al. Capacity optimization configuration method of user side shared energy storage for high-proportion distributed photovoltaic [J]. Chinese Journal of Power Sources, 2025, 49(2): 443-453. (in Chinese)

[5] Abdelrahman S, Milanović J V. Practical approaches to assessment of harmonics along radial distribution feeders [J]. IEEE Transactions on Power Delivery, 2019, 34(3): 1184-1192.

[6] Hiyama T, Hammam M S A A, Ortmeyer T H. Distribution system modeling with distributed harmonic sources [J]. IEEE Transactions on Power Delivery, 1989, 4(2): 1297-1304.

[7] Sun Yuanyuan, Xie Xiangmin, Zhang Linghan, et al. A voltage adaptive dynamic harmonic model of nonlinear home appliances [J]. IEEE Transactions on Industrial Electronics, 2020, 67(5): 3607-3617.

[8] Shu Qin, Wu Yayue, Xu Fangwei, et al. Estimate utility harmonic impedance via the correlation of harmonic measurements in different time intervals [J]. IEEE Transactions on Power Delivery, 2020, 35(4): 2060-2067.

[9] Chen C I, Chen Y C. A neural-network-based data-driven nonlinear model on time- and frequency-domain voltage-current characterization for power-quality study [J]. IEEE Transactions on Power Delivery, 2015, 30(3): 1577-1584.

[10] Zhao Yuqi, Milanović J V. Equivalent modelling of wind farms for probabilistic harmonic propagation studies [J]. IEEE Transactions on Power Delivery, 2022, 37(1): 603-611.

[11] Jayatunga U, Perera S, Ciufo P. Voltage unbalance emission assessment in radial power systems [J]. IEEE Transactions on Power Delivery, 2012, 27(3): 1653-1661.

[12] Wen He, Cheng Da, Teng Zhaosheng, et al. Approximate algorithm for fast calculating voltage unbalance factor of three-phase power system [J]. IEEE Transactions on Industrial Informatics, 2014, 10(3): 1799-1805.

[13] Zhang Xiaoqin, Zhu Hongbin, Li Bo, et al. Power transformer diagnosis based on dissolved gases analysis and copula function [J]. Energies, 2022, 15(12): 4192.

[14] Kloubert M L. Assessment of generation adequacy by modeling a joint probability distribution model [J]. Electric Power Systems Research, 2020, 189: 106803.

[15] Merino Z D, Farmer J, Jacobs D J. Probability density estimation through nonparametric adaptive partitioning and stitching [J]. Algorithms, 2023, 16(7): 310.

[16] Hao Wenbin, Meng Zhigao, Zeng Peng, et al. A method for modeling the output power of photovoltaic power supplies based on non-parametric kernel density estimation [C]// 2022 5th International Conference on Renewable Energy and Power Engineering (REPE). Beijing, China, 2022: 407-413.

[17] Hou Wenting, Yi Longxian, Miao Huaibin, et al. Robust kernel density estimation based data-driven optimal scheduling for power systems considering data errors and uncertainties of renewable energy [J]. IEEE Access, 2024, 12: 81329-81337.

[18] Pukhrem S, Basu M, Conlon M F. Probabilistic risk assessment of power quality variations and events under temporal and spatial characteristic of increased PV integration in low-voltage distribution networks [J]. IEEE Transactions on Power Systems, 2018, 33(3): 3246-3254.

[19] 陈斌，高彦梅．excel 在统计分析中的运用[M]．北京：清华大学出版社，2019.
Chen Bin, Gao Yanmei. The Use of Excel In Statistical Analysis [M]. Beijing: Tsinghua University Press, 2019. (in Chinese)

[20] Zhong Min, Zhou Yucai. Short-term PV power prediction based on SAGA-FCM and XGBoost [J]. Journal of Physics: Conference Series, 2023, 2562(1): 012024.

[21] 饶显杰, 徐忠林, 刘翔宇, 等. 基于反射系数与核函数构建的新型电缆缺陷诊断方法[J]. 电工技术学报, 2024, 39(7): 2184-2192.
Rao Xianjie, Xu Zhonglin, Liu Xiangyu, et al. A new cable defect diagnosis method based on reflection coefficient and kernel function construction [J]. Transactions of China Electrotechnical Society, 2024, 39(7): 2184- 2192. (in Chinese)



张一帆

作者简介：

张一帆(1984—)，女，本科，高级工程师。研究方向：电力系统，电能质量。E-mail：zyf_dky@163.com。

潘大志(1977—)，男，本科，高级工程师。研究方向：电力系统及其自动化。E-mail：321694976@qq.com。

张永宏(1975—)，男，硕士，高级工程师。研究方向：电网运维管理。E-mail：zhangyh0438@sina.com。

辛力坚(1978—)，男，本科，正高级工程师。研究方向：电力系统。E-mail：1261049227@qq.com。

汪颖(1982—)，女，博士，教授。研究方向：电能质量与优质供电。E-mail：769429505@qq.com。

王心茹(2000—)，女，通信作者，博士研究生。研究方向：电能质量与优质供电。E-mail：290172429@qq.com。

(责任编辑：耿越)