

# 基于高斯混合随机性模型的多风电场配电网概率潮流计算

王士兴<sup>1</sup>,陈树恒<sup>1</sup>,刘群英<sup>2</sup>,韩 杨<sup>1</sup>,CHEN Zhe<sup>3</sup>,胡维昊<sup>1</sup>

(1. 电子科技大学 机械与电气工程学院,四川 成都 611731;

2. 电子科技大学 自动化工程学院,四川 成都 611731;3. 奥尔堡大学 能源技术系,丹麦 奥尔堡 9220)

**摘要:**风速的随机性及风电场之间的相关性对电力系统潮流分析具有重要影响。计及风速的随机性及多风电场之间的相关性,提出一种改进的概率潮流计算方法。基于多风电场实际出力样本数据,利用 $k$ -means算法确定高斯混合模型的参数数量,并利用数据筛选过程改进高斯混合模型以提高联合分布模型的精确度;引入基于Nataf估算变换的三点估计法对所建概率分布模型进行采样,并将采样数据与电力系统潮流平衡方程结合以实现概率潮流计算。IEEE 18节点系统的算例结果表明,所提方法具有较高的计算精度和计算效率。

**关键词:**高斯混合模型; $k$ -means算法;Nataf变换;三点估计法;概率潮流

**中图分类号:**TM 614

**文献标志码:**A

**DOI:**10.16081/j.epae.202205036

## 0 引言

2021年,我国政府将实现碳达峰和碳中和目标的时间表写入政府工作报告中,大力发展清洁能源是实现该目标的重要途径之一,但风电等清洁能源规模的不断扩大将会给电网带来更多的不确定性,这些不确定性会极大地影响电力系统运行的稳定性、安全性和经济性,计及多风电场间的相关性,建立风电功率的随机性模型对进一步研究电力系统概率潮流和概率优化问题具有重要意义<sup>[1]</sup>。

近年来,国内外学者对风电功率的随机性与相关性建模问题进行了大量研究<sup>[2]</sup>。文献[3]借助线性相关因子和秩相关系数构建相依性模型,实现了对多风电场相关关系的描述,但该模型的计算精度较低。针对该问题,文献[4]在风电功率建模时引入混合Copula模型,文献[5-6]对混合滕结构的Copula模型进行改进,文献[7]将核密度估计法与Copula函数相结合以实现数据采样,这些改进均在一定程度上提高了风电功率随机性模型的性能,然而,受地理、气候等因素的影响,不同区域多风电场间的相依程度是不同的,Copula函数在描述不同随机性模型时的精度差别较大。针对该问题,文献[8]提出利用高斯混合模型GMM(Gaussian Mixture Model)对多风电场相关性进行描述,理论上该模型可以描述任意随机性分布,文献[9]进一步对GMM的定义域进行修正并设计归一化的基于截断的分布模型,虽然这些改进在一定程度上提高了联合分布模型的精度,但模型的精度仍受到样本数据分布因素的影响。

概率潮流计算是分析随机风电功率接入后电力系统潮流分布的有效方法。蒙特卡罗抽样方法是一种经典的概率潮流计算方法,但在面对较大规模的电网时,该方法的计算效率较低<sup>[10]</sup>。相比较而言,基于随机变量平均值、标准差、偏度和峰度等数据特征的 $2m+1$ 点估计法<sup>[11]</sup>利用较少的采样数据点即可求得满足精确度要求的计算结果。

针对现有研究的不足,本文进一步改进GMM并提出一种相对完整、精准的基于GMM的多风电场相关性建模方法,并将该方法应用于电力系统概率潮流计算。与传统的风电功率随机性建模方法和基本高斯混合建模方法相比,本文方法的创新点如下:通过引入补偿系数建立截断式GMM,该模型在描述多风电场相关性时具有明显的拟合精度优势;基于数据筛选过程提出一种改进的参数计算方法,该方法有效地提高了模型的精度和求解速度;在点估计法框架下,设计一种基于计算机数值求解的改进Nataf变换求解算法,该算法解决了传统Nataf变换的积分求解难题。

## 1 风电场特性和GMM

### 1.1 风电场特性

决定风电场特性的是风速或风电功率数据,受地理、气候等因素影响,不同地区风电场风速的概率分布模型不同。风电功率 $P_w$ 与风速 $v$ 直接相关,两者之间的关系如附录A式(A1)所示。

附录B图B1(a)、(b)分别为某年内每隔5 min记录的某风电场100 m高处的风速数据和归一化风电功率数据,共计105 120个数据点。受切入风速、额定风速和切出风速的影响,风电功率在0和1 p.u.的聚集性较高。然而,式(A1)没有考虑风速时序、恢复风速条件和风机状态等因素影响,实际上,基于式(A1)将风速转换形成的功率数据与实际功率数据

收稿日期:2021-07-19;修回日期:2022-04-01

在线出版日期:2022-04-13

基金项目:国家重大研发计划项目(2018YFE0127600)

Project supported by the National Key Research and Development Plan Program of China(2018YFE0127600)

之间存在一定的差距。本文将使用风电功率数据对多风电场的相关关系进行建模,以直接降低在将风速转换为风电功率的过程中产生的误差。

## 1.2 GMM

### 1.2.1 GMM基本原理

GMM 是多个高斯分布函数的叠加组合,即总体的概率分布模型由多个子高斯模型构成,高斯分布良好的数学品质使得 GMM 具有较高的应用潜能<sup>[12]</sup>。附录 A 式(A2)—(A4)给出了高斯函数的概率模型。

### 1.2.2 GMM 求解方法

通常利用极大似然估计 MLE(Maximum Likelihood Estimation)近似求解 GMM,如式(1)所示。

$$\theta = \underset{\theta}{\operatorname{argmax}} L(\theta) \quad (1)$$

式中: $\theta$ 为概率模型中的未知参数; $L(\theta)$ 为极大似然函数; $\operatorname{argmax} L(\theta)$ 表示  $L(\theta)$ 取最大值时的  $\theta$  值。

MLE 一般用于求解单变量模型的参数,对于存在隐变量的概率模型参数求解问题,在 MLE 的基础上引入迭代过程的改进期望最大化 EM(Expectation Maximization)算法得到的求解效果较好,该算法的具体求解过程如附录 A 式(A5)—(A8)和图 A1 所示。

## 2 基于 GMM 的改进建模方法

本节将结合多风电场出力联合分布特性对基于 GMM 的建模方法进行改进,包括筛选模型训练数据、改进模型概率函数和简化参数判断准则。

### 2.1 GMM 训练数据的筛选

图 1 为两风电场功率的联合分布直方图(图中风电场 1 功率和风电场 2 功率均为标么值,后同)。由图可见,模型在等边际点 0 处的分布密度过高,由于 GMM 由多个子高斯模型叠加构成,密度过高的边际点会影响子高斯模型的参数求解过程,进而影响 GMM 的拟合精度,为了提高参数求解精度,需要提前筛选 GMM 的训练数据。

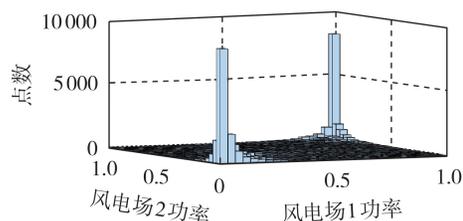


图 1 归一化的风电功率联合分布直方图

Fig.1 Histogram of joint distribution of normalized wind power

筛选边际点首先要利用核密度估计法,其是一种用于估计未知密度函数的非参数检验方法。附录 B 图 B2 为 GMM 训练数据的筛选流程,利用核密度估计法计算各点的概率密度,筛选出联合分布中概率密度过高点,对剩余点重复该筛选过程,直至筛选后

的数据满足式(2)。

$$\begin{cases} \forall p_{\text{point}} < p_{\text{max}} \\ \forall p_{\text{edge}} \geq p_{\text{max}} \end{cases} \quad (2)$$

式中: $p_{\text{point}}$ 为筛选后点的概率密度; $p_{\text{max}}$ 为模型概率密度的最大值; $p_{\text{edge}}$ 为筛选出的概率密度过高的点的概率密度。图 2 为筛选后归一化的风电功率联合分布直方图。由图可见,筛选后的联合分布直方图中聚集区分布清晰且无过高密度点,筛选出的边际点数量少且分布集中,其概率模型通过直接统计获得。

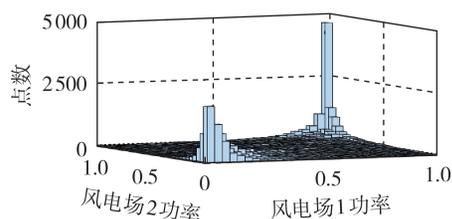


图 2 筛选后归一化的风电功率联合分布直方图

Fig.2 Histogram of joint distribution of normalized wind power after screening

引入数据筛选过程后,用于 GMM 建模的数据量与实际联合分布的数据量是不同的,需要按照式(3)在 GMM 中添加补偿系数以修正模型概率。

$$\begin{cases} p_{\text{fact}} = k_{\text{scr}} p_{\text{GMM}} \\ k_{\text{scr}} = N_{\text{all}} / N_{\text{mid}} \end{cases} \quad (3)$$

式中: $p_{\text{fact}}$ 为修正概率模型的概率密度函数; $k_{\text{scr}}$ 为筛选过程形成的概率模型修正系数; $p_{\text{GMM}}$ 为对应低密度数据区的 GMM 概率密度函数; $N_{\text{all}}$ 为原始联合分布训练数据点数; $N_{\text{mid}}$ 为筛选后的训练数据点数。

数据筛选过程极大地减少了数据聚集分布特性对 GMM 精度的影响,提高了模型精确度和适用性。

### 2.2 GMM 参数设计的改进

GMM 通过迭代过程进行求解,在处理大量样本数据时,求解该模型的计算复杂度极高。在 GMM 中,概率密度受单一子高斯模型的影响较大,求解 GMM 的关键是确定子高斯模型的数量,即聚集区数量。对于大量样本数据, $k$ -means 算法可以快速求解聚集区数量,因此,本文利用该算法判断聚集区数量,从而确定子高斯模型的数量以降低参数求解过程的复杂度。

在  $k$ -means 算法中,评价聚集程度的核心指标是误差平方和 SSE(Sum of Squared Errors)<sup>[13]</sup>:

$$v^{\text{SSE}} = \sum_{i=1}^k \sum_{x_p \in C_i} |x_p - m_i|^2 \quad (4)$$

式中: $v^{\text{SSE}}$ 为 SSE 指标; $k$ 为聚集区数量; $C_i$ 为第  $i$  个簇; $x_p$ 为  $C_i$ 中样本点的位置; $m_i$ 为第  $i$  个簇的质心。

判断样本数据聚类数的一种常见方法是手肘法。随着  $k$  的增大,样本簇的数量增加,每个簇的聚

合程度提高,  $v^{\text{SSE}}$  降低。当  $k$  小于真实聚类数时,  $k$  增大可以明显提高每个簇的聚合程度,  $v^{\text{SSE}}$  迅速降低; 当  $k$  达到真实聚类数时, 随着  $k$  的继续增大,  $v^{\text{SSE}}$  的变化趋于平缓。图3为  $v^{\text{SSE}}$ 、MLE 负值与  $k$  的函数关系。MLE 是判断 GMM 拟合程度的重要指标, 其值越大, 拟合效果越好。由图可见:  $v^{\text{SSE}}$  与  $k$  的关系呈现一个手肘形状, 肘部拐点可以在一定程度上反映真实聚类数; 随着  $k$  的增大,  $v^{\text{SSE}}$  与 MLE 负值的变化趋势相同, 通过  $v^{\text{SSE}}$  的变化可以确定子高斯模型数量。由于  $k$ -means 算法的计算复杂度远低于求解 GMM 的 EM 算法, 因此, 借助  $k$ -means 算法可以缩短计算时间。

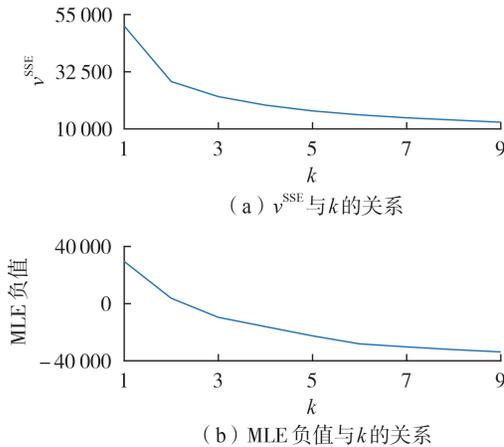


图3  $v^{\text{SSE}}$ 、MLE 负值与  $k$  的关系

Fig.3 Relationship between  $v^{\text{SSE}}$ , negative value of MLE and  $k$

### 2.3 改进的 GMM

GMM 的定义域与归一化的风电功率联合分布模型的定义域不同, 本文通过式(5)对 GMM 的概率密度函数进行改进, 设计截断式 GMM 来描述联合分布模型。

$$p_{\text{imp}} = \text{sgn}(X) \text{sgn}(1-X) p_{\text{trun}}(X|\theta) \quad (5)$$

式中:  $p_{\text{imp}}$  为改进后的 GMM 函数概率密度;  $X$  为多维随机变量;  $\text{sgn}(\cdot)$  为阶跃函数, 一维的  $\text{sgn}(\cdot)$  定义如式(6)所示;  $p_{\text{trun}}(X|\theta)$  为截断前的 GMM 概率密度函数,  $\theta$  为未知参数向量。

$$\text{sgn}(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (6)$$

式中:  $x$  为一维随机变量。

利用基于迭代的 EM 算法求解改进的 GMM 参数, 并利用式(7)增加的补偿系数, 得到最终的参数。

$$\begin{cases} p_{\text{imp}}(X|\theta) = k_{\text{trun}} p_{\text{trun}}(X|\theta) \\ k_{\text{trun}} = 1/P_{\text{trun}} \\ P_{\text{trun}} = \int_0^1 p_{\text{trun}}(X|\theta) dX \end{cases} \quad (7)$$

式中:  $k_{\text{trun}}$  为截断引入的概率补偿系数;  $P_{\text{trun}}$  为原始

概率函数。引入补偿系数的截断式 GMM 可以较好地适应归一化的风电场功率数据的分布, 其定义域准确, 描述效果好。

改进的多风电场归一化功率描述模型为:

$$p = \begin{cases} p_{\text{fact}} = k_{\text{scr}} k_{\text{trun}} p_{\text{trun}}(X|\theta) & P_{\text{point}} < P_{\text{max}} \\ p_{\text{edge}} & P_{\text{point}} \geq P_{\text{max}} \end{cases} \quad (8)$$

### 3 基于改进 GMM 的电力系统概率潮流

点估计法计算过程简单, 描述准确, 适用于独立的随机变量。结合 Nataf 变换的点估计法, 可对多个相关随机变量进行采样。本文提出一种基于 Nataf 变换的求解方法, 使点估计法可应用于更复杂的模型。

#### 3.1 三点估计法

对于一维随机变量函数  $Z = Z(x)$ , 可以在标准正态变量空间中取适量点描述数据特征,  $2m+1$  点估计法应用广泛, 计算精度高<sup>[14]</sup>。附录 C 表 C1 为基于 Gauss-Hermite 积分方法计算得到的三点估计法、五点估计法和七点估计法的独立正态随机变量的采样点与相应权重。在满足计算要求的情况下, 本文选取三点估计法来降低计算量。

#### 3.2 三点式 Nataf 逆变换

Nataf 变换可实现相关随机变量组与独立正态变量组的转换<sup>[15]</sup>, Nataf 变换过程如附录 B 图 B3 所示。在 Nataf 正变换中, 通过式(9)将  $n$  维原始随机变量  $X = [x_1, x_2, \dots, x_n]$  转换为  $n$  维相关正态随机变量  $Y = [y_1, y_2, \dots, y_n]$ , 再通过式(10)利用 Nataf 变换矩阵  $B$  将相关正态随机变量转换为  $n$  维独立正态随机变量  $Z = [z_1, z_2, \dots, z_n]$ 。其中,  $x_i (i=1, 2, \dots, n)$  为第  $i$  维随机变量,  $y_i (i=1, 2, \dots, n)$  为第  $i$  维相关正态随机变量,  $z_i (i=1, 2, \dots, n)$  为第  $i$  维不相关正态随机变量。

$$y_i = \Phi^{-1}(F_i(x_i)) \quad (9)$$

$$Z = B^{-1}Y \quad (10)$$

式中:  $\Phi(\cdot)$  为标准正态分布的累积函数;  $F_i(x_i)$  为第  $i$  维随机变量的原始分布函数。式(10)中矩阵  $B$  是通过 Cholesky 分解对称正定矩阵  $\rho_0$  得到的下三角矩阵, 如式(11)所示。

$$\rho_0 = BB^T \quad (11)$$

$\rho_0$  中元素  $\rho_{0ij}$  根据式(12)计算得到。

$$\rho_{ij} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{x_i - u_i}{\sigma_i} \frac{x_j - u_j}{\sigma_j} f_{ij}(x_i, x_j) dx_i dx_j = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{F_i^{-1}(\Phi(z_i)) - u_i}{\sigma_i} \frac{F_j^{-1}(\Phi(z_j)) - u_j}{\sigma_j} \times \phi_{ij}(z_i, z_j, \rho_{0ij}) dz_i dz_j \quad (12)$$

式中:  $\rho_{ij}$  表示  $x_i$  和  $x_j$  间的相关性;  $u_i, u_j$  分别为  $x_i, x_j$  的期望;  $\sigma_i, \sigma_j$  分别为  $x_i, x_j$  的标准差;  $\rho_{0ij}$  表示  $z_i$  和  $z_j$  间

的相关性; $\phi_{ij}(z_i, z_j, \rho_{0ij})$ 为具有相关系数的联合分布函数。

### 3.3 Nataf估算变换

Nataf变换是一种求解多变量相关关系的成熟算法。常见的变量相关关系可以利用式(12)的积分求解函数进行求解。对于复杂模型,可以利用半经验公式对其进行简化再利用式(12)进行求解。半经验公式仅能对Copula函数、多次函数、幂函数等进行变换求解,其应用范围限制了Nataf变换的求解效果,当待求解函数不能表示为上述函数之和时,基于半经验函数的Nataf变换中的相关关系系数往往不可解或者求解过程过于复杂。当随机变量使用带有离散变量的改进截断式GMM描述时,半经验公式的缺陷导致Nataf变换矩阵**B**无法求解,针对该问题,本文设计一种Nataf估算变换来替代半经验公式,形成一种具有较低计算复杂度和较高精确度的求解算法。

Nataf估算变换是对求解过程中的部分内容进行估算,式(12)中函数 $F_i^{-1}(\Phi(z_i))$ 和 $\phi_{ij}(z_i, z_j, \rho_{0ij})$ 无法求解,对于 $F_i^{-1}(\Phi(z_i))$ ,本文利用拟合多项式近似处理,则式(12)变为:

$$\left\{ \begin{aligned} \rho_{ij} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{F_{i,imp}^{-1}(\Phi(z_i)) - u_i}{\sigma_i} \frac{F_{j,imp}^{-1}(\Phi(z_j)) - u_j}{\sigma_j} \times \\ &\quad \phi_{ij}(z_i, z_j, \rho_{0ij}) dz_i dz_j \\ F_{i,imp}^{-1}(\Phi(z_i)) &= \sum_{l=1}^n w_{il} x_i^{l-1} \end{aligned} \right. \quad (13)$$

式中: $F_{i,imp}(\Phi(z_i))$ 为 $z_i$ 的边缘分布函数; $w_{il} x_i^{l-1}$ 分别为第*i*维随机变量拟合多项式的第*l*项权值和幂函数。式(13)中仅存在 $\phi_{ij}(z_i, z_j, \rho_{0ij})$ 这个复杂函数,可以对其进行积分计算。由于 $\rho_{ij}$ 可以通过积分进行求解,为已知量,借助计算机以步长0.01求解 $\rho_{0ij}$ 与 $\rho_{ij}$ 的相关关系,最终得到误差小于0.01的近似解 $\rho_{0ij}$ 。估算得到的Nataf变换方程基本可以满足数据的计算要求。

### 3.4 概率潮流计算

确定性潮流的约束方程可以表示为:

$$R = G(X_{in}) \quad (14)$$

式中:**R**为输出变量,包括发电机机端电压、线路传输有功功率 $P_{ij}$ 、线路传输无功功率 $Q_{ij}$ ;G(·)为基于节点导纳矩阵的函数<sup>[16]</sup>;X<sub>in</sub>为输入变量,包括各节点的有功功率 $P_i$ 和无功功率 $Q_i$ 、补偿无功功率 $Q_{in}$ 等。基于改进点估计法的概率潮流计算流程如附录B图B4所示。

## 4 算例分析

仿真实验平台是搭载64位Windows操作系统的计算机,CPU为Intel Core i5-8265U,1.60 GHz, RAM为8 GB,程序开发语言为MATLAB 2018b。

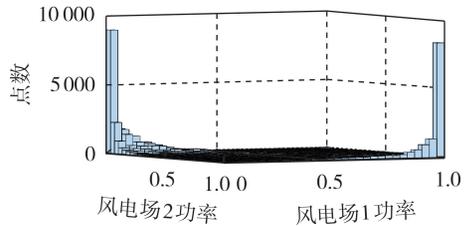
### 4.1 风电功率相关性建模分析

本文使用美国东海岸2座风电场的风电功率数据,数据来自WIND Toolkit。分别使用高斯Copula函数、t分布Copula函数<sup>[10]</sup>、传统GMM分布函数<sup>[17]</sup>和改进的GMM分布函数对归一化功率相关关系进行建模,并对各模型进行比较。

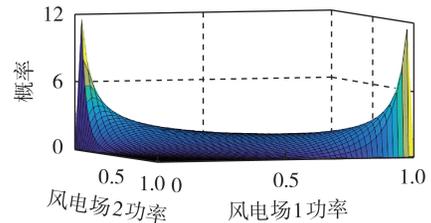
1)改进的GMM拟合精度。

图4为不同模型对2座风电场相关功率的拟合图。由图可知,在描述模型时,3种常规模型均存在较大的误差。为定量分析模型的精度,本文选用二阶欧氏距离*d*作为评价指标,如式(15)所示。

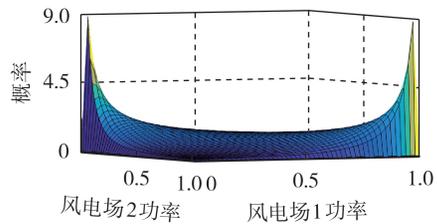
$$d = \sum_{i=1}^n (C_M^{(i)}(v_1, v_2) - C_D^{(i)}(v_1, v_2))^2 \quad (15)$$



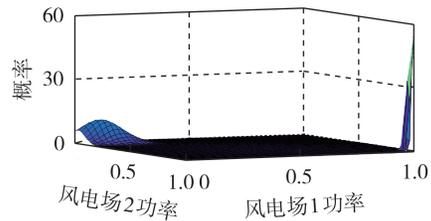
(a) 2座风电场的归一化功率分布直方图



(b) 高斯 Copula 函数模型



(c) t分布 Copula 函数模型



(d) 传统GMM

图4 不同模型对2座风电场相关功率的拟合图  
Fig.4 Fitting diagrams of correlative power of two wind farms for different models

式中:  $C_M^{(i)}(v_1, v_2)$  为模型在点  $i(v_1, v_2)$  处的概率分布函数,  $C_D^{(i)}(v_1, v_2)$  为模型在点  $i(v_1, v_2)$  处的经验概率分布,  $v_1, v_2$  分别为风电场1和风电场2的归一化功率出力点。

表1为4种模型的二阶欧氏距离。由表可知,改进的GMM精度相较于其他模型具有明显优势。

表1 4种模型的二阶欧氏距离比较

Table 1 Comparison of second-order Euclidean distance among four models

模型	二阶欧氏距离
t分布 Copula 函数模型	9.0114
高斯 Copula 函数模型	8.9574
传统 GMM	0.2403
改进的 GMM	0.0366

2) 筛选过程在提高模型精度中的效果。

为提高拟合精度,本文设计了基于核密度估计法的筛选过程。图5为数据筛选后改进的GMM拟合图,筛选点主要为(0,0)、(1,1) p.u.。由图可知,筛选过程解决了GMM受训练数据制约的问题,筛选后的数据拟合图完整反映了2座风电场功率数据分布的相关关系。对比表1的传统GMM和改进的GMM精度可知,引入筛选过程的GMM误差仅为传统GMM误差的15%。表2为数据筛选前后改进的GMM的MLE,由表可知,筛选过程显著提高了模型的精度。改进的GMM模型参数如附录C表C2所示。

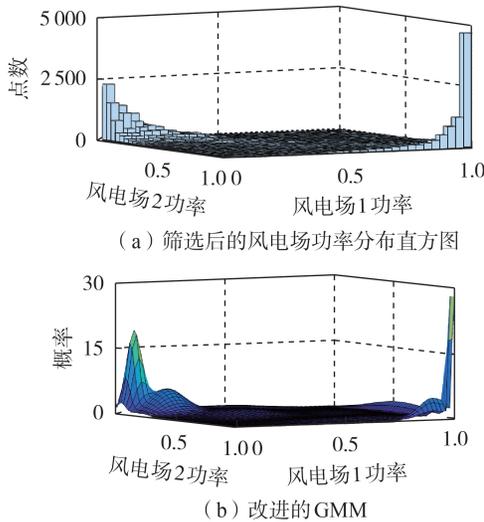


图5 数据筛选后改进的GMM拟合图  
Fig.5 Fitting diagram of improved GMM after data selection

表2 MLE比较

Table 2 Comparison of MLE

条件	求解方法	MLE
数据筛选前	EM算法	$5.2879 \times 10^4$
数据筛选后	EM算法	$6.8054 \times 10^4$

3)改进参数计算方法的优势。

求解参数须先选择合适的 $k$ 值,假设 $k$ 的取值范围为1~9,表3为基于2种计算方法求解 $k$ 值的计算时间。由表可知,基于 $k$ -means算法通过 $v^{SSE}$ 判断 $k$ 值的计算时间相较于直接求解明显缩短。

表3 2种计算方法的计算时间对比

Table 3 Comparison of calculation time between two calculation methods

计算方法	评价指标	计算时间/s
直接求解	MLE	53.653
$k$ -means算法	$v^{SSE}$	21.808

## 4.2 概率潮流数字仿真实验

利用Nataf估算变换得到三点估计法的采样结果,如附录C表C3所示。基于概率潮流数字仿真实验验证本文所提概率潮流计算方法的准确性和计算效率优势。实验基于IEEE 18节点系统,节点2、17、18各接入1座功率为40 kW的风电场。基于三点估计法与蒙特卡罗抽样方法得到的计算时间如表4所示,基于2种方法得到的节点15电压的概率分布图如图6所示。其中,蒙特卡罗抽样方法的样本点规模为10000。由表4可知,在保证实际精度要求的条件下,三点估计法大幅缩短了计算时间,提高了计算效率。图6中,受到采样点较少的影响,2种方法得到的分布图仅能大致拟合,但三点估计法基本满足实际分析需求。

表4 基于三点估计法和蒙特卡罗抽样方法得到的计算时间

Table 4 Calculation time obtained by three-point estimation method and Monte Carlo sampling method  
单位:s

潮流计算方法	采样时间	潮流计算时间	总时间
蒙特卡罗抽样方法			13 722.32
三点估计法	31.87	29.83	61.70

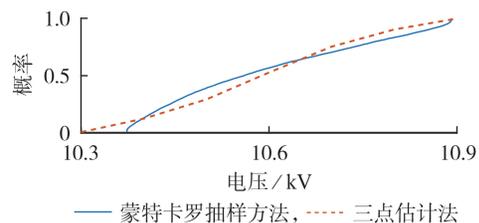


图6 基于三点估计法和蒙特卡罗抽样方法得到的节点15电压的概率分布

Fig.6 Probabilistic distribution of voltage at Bus 15 obtained by three-point estimation method and Monte Carlo sampling method

附录C表C4给出了IEEE 18节点系统中部分关键节点的电压和线路总损耗的均值和标准差。表中,三点估计法潮流计算结果的电压均值与实际值

的相对误差均小于 0.05%,电压标准差与实际值的相对误差均小于 5%,线路总损耗的均值、方差与实际值之间的相对误差均较小,可满足实际应用需求。

## 5 结论

本文在建立 GMM 的基础上,改进多风电场联合概率分布模型的计算过程,通过数字仿真实验得到如下结论:

1) 本文基于补偿系数的截断式 GMM 在函数图像数据拟合和欧氏距离精度指标拟合方面具有明显优势,欧氏距离误差指标仅为传统 Copula 函数模型的 2% 左右;

2) 本文设计的数据筛选过程解决了 GMM 受训练数据聚集特性影响的问题,使改进的 GMM 欧氏距离误差指标降低至传统 GMM 的 20% 以下;

3) 采用本文设计的参数计算方法求解 GMM 中的  $k$  值时,计算时间明显缩短,相较于直接求解的计算时间约缩短了 50%;

4) 本文所提出的基于 Nataf 估算变换的三点估计法的潮流计算方法进一步提高了计算结果的精度,与实际值相比,采用本文所提方法得到的关键节点电压以及线路总损耗误差小于 5%,满足实际应用需求。

附录见本刊网络版(<http://www.epae.cn>)。

## 参考文献:

- [1] XU Y J, KORKALI M, MILI L, et al. Risk assessment of rare events in probabilistic power flow via hybrid multi-surrogate method[J]. IEEE Transactions on Smart Grid, 2019, 11(2):1593-1603.
- [2] 苏晨博,刘崇茹,李至峪,等. 基于贝叶斯理论的考虑多维风速之间相关性的概率潮流计算[J]. 电力系统自动化, 2021, 45(3):157-165.
- SU Chenbo, LIU Chongru, LI Zhiyu, et al. Bayesian theory based calculation of probabilistic power flow considering correlation between multi-dimensional wind speed[J]. Automation of Electric Power Systems, 2021, 45(3):157-165.
- [3] 蔡德福,石东源,陈金富. 基于 Copula 理论的计及输入随机变量相关性的概率潮流计算[J]. 电力系统保护与控制, 2013, 41(20):13-19.
- CAI Defu, SHI Dongyuan, CHEN Jinfu. Probabilistic load flow considering correlation between input random variables based on Copula theory[J]. Power System Protection and Control, 2013, 41(20):13-19.
- [4] TAMRAZ M. Mixture copulas and insurance applications[J]. Annals of Actuarial Science, 2018, 12(2):391-411.
- [5] XU Y D, YUAN Y. Analysis of aggregated wind power dependence based on optimal vine copula[C]//2019 IEEE Innovative Smart Grid Technologies-Asia. Chengdu, China: IEEE, 2019:1788-1792.
- [6] XIE Z Q, JI T Y, LI M S, et al. Quasi-Monte Carlo based probabilistic optimal power flow considering the correlation of wind speeds using copula function[J]. IEEE Transactions on

Power Systems, 2018, 33(2):2239-2247.

- [7] 徐玉琴,陈坤,李俊卿,等. Copula 函数与核估计理论相结合分析风电场出力相关性的一种新方法[J]. 电工技术学报, 2016, 31(13):92-100.
- XU Yuqin, CHEN Kun, LI Junqing, et al. A new method analyzing output correlation of multi-wind farms based on combination of Copula function and kernel estimation theory[J]. Transactions of China Electrotechnical Society, 2016, 31(13):92-100.
- [8] WANG Z W, SHEN C, LIU F, et al. Chance-constrained economic dispatch with non-Gaussian correlated wind power uncertainty[J]. IEEE Transactions on Power Systems, 2017, 32(6):4880-4893.
- [9] SUN W G, ZAMANI M, ZHANG H T, et al. Probabilistic optimal power flow with correlated wind power uncertainty via Markov chain quasi-Monte-Carlo sampling[J]. IEEE Transactions on Industrial Informatics, 2019, 15(11):6058-6069.
- [10] XU J, WU W, WANG K Y, et al. C-Vine pair copula based wind power correlation modelling in probabilistic small signal stability analysis[J]. IEEE/CAA Journal of Automatica Sinica, 2020, 7(4):1154-1160.
- [11] 韩海腾,高山,吴晨,等. 基于 Nataf 变换的电网不确定性多点估计法[J]. 电力系统自动化, 2015, 39(7):28-34.
- HAN Haiteng, GAO Shan, WU Chen, et al. Uncertain power flow solved by multi-point estimate method based on nataf transformation[J]. Automation of Electric Power Systems, 2015, 39(7):28-34.
- [12] ROBERT C. Machine learning, a probabilistic perspective[J]. Chance, 2014, 27(2):62-63.
- [13] 李国庆,陆为华,李赫,等. 基于模糊 C-均值聚类的时序概率潮流快速计算方法[J]. 电力自动化设备, 2021, 41(4):116-122.
- LI Guoqing, LU Weihua, LI He, et al. Fast calculation method of time sequence probabilistic power flow based on fuzzy C-means clustering[J]. Electric Power Automation Equipment, 2021, 41(4):116-122.
- [14] MORALES J M, PEREZ-RUIZ J. Point estimate schemes to solve the probabilistic power flow[J]. IEEE Transactions on Power Systems, 2007, 22(4):1594-1601.
- [15] ZHU J Z, ZHANG Y. Probabilistic load flow with correlated wind power sources using a frequency and duration method[J]. IET Generation, Transmission & Distribution, 2019, 13(18):4158-4170.
- [16] 王聪,颜伟,户秀琼. 考虑频率波动的孤岛微电网在线三相概率潮流计算方法[J]. 电力自动化设备, 2020, 40(2):109-118.
- WANG Cong, YAN Wei, HU Xiuqiong. Online three-phase probabilistic power flow calculation method for islanded microgrid considering frequency fluctuation[J]. Electric Power Automation Equipment, 2020, 40(2):109-118.
- [17] KE D P, CHUNG C Y, SUN Y Z. A novel probabilistic optimal power flow model with uncertain wind power generation described by customized Gaussian mixture model[J]. IEEE Transactions on Sustainable Energy, 2016, 7(1):200-212.

## 作者简介:



王士兴

王士兴(1997—),男,硕士研究生,主要研究方向为电力系统运行与优化(E-mail: 201921040317@std.uestc.edu.cn);

陈树恒(1974—),男,副教授,博士,通信作者,主要研究方向为电力系统运行与优化(E-mail: chenshuheng@uestc.edu.cn)。

(编辑 王锦秀)

(下转第 91 页 continued on page 91)

## Control strategy of virtual harmonic resistance-type energy storage inverter with harmonic governance function

ZENG Jiang, FENG Jianlei, CHEN Shuliang, XIONG Taojun, CHEN Weiguo, LIU Qiwei

(School of Electric Power, South China University of Technology, Guangzhou 510640, China)

**Abstract:** The traditional energy storage inverter can only compensate the harmonic current generated by the specified nonlinear load, but cannot reduce the original harmonic content of the power grid. Therefore, the suppression effect of parallel resistance on harmonic voltage at the PCC (Point of Common Coupling) and system resonance is analyzed, and the control strategy of virtual harmonic resistance-type energy storage inverter is proposed. For fundamental wave, the energy storage inverter can still achieve normal charging and discharging function. For harmonics, the energy storage inverter is controlled to output harmonic current, whose magnitude is proportional to the magnitude of harmonic voltage and whose phase is opposite to that of harmonic voltage. In this case, the inverter is equivalent to a virtual harmonic resistance, which can absorb harmonic power and suppress harmonic voltage at the PCC. Considering the resonance, the perturbation and observation method is used to adjust the virtual harmonic conductance automatically so that the energy storage inverter can absorb the maximum harmonic power. The simulative and experimental results show that the virtual harmonic resistance-type energy storage inverter can realize good harmonic governance function while charging and discharging normally.

**Key words:** energy storage inverter; virtual harmonic resistance; harmonic governance; harmonic suppression; control strategy

(上接第 69 页 continued from page 69)

## Probabilistic power flow calculation of distribution network with multiple wind farms based on Gaussian mixture random model

WANG Shixing<sup>1</sup>, CHEN Shuheng<sup>1</sup>, LIU Qunying<sup>2</sup>, HAN Yang<sup>1</sup>, CHEN Zhe<sup>3</sup>, HU Weihao<sup>1</sup>

(1. School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China; 2. School of Automation Engineering,

University of Electronic Science and Technology of China, Chengdu 611731, China;

3. Department of Energy Technology, Aalborg University, Aalborg 9220, Denmark)

**Abstract:** The power flow analysis of power system is significantly affected by the randomness of wind speed and the correlation between wind farms. An improved probabilistic power flow calculation method is proposed considering the randomness of wind speed and the correlation between multiple wind farms. Based on the actual output sample data of multiple wind farms, *k*-means algorithm is used to determine the number of parameters of Gaussian mixture model, and the data selection process is used to improve Gaussian mixture model for improving the accuracy of joint distribution model. A three-point estimation method based on Nataf evaluation transformation is introduced for sampling of the built probabilistic distribution model, and the sampling data and power flow balance equations of power system are combined to realize probabilistic power flow calculation. The case results of IEEE 18-bus system show that the proposed method has high calculation precision and efficiency.

**Key words:** Gaussian mixture model; *k*-means algorithm; Nataf transformation; three-point estimation method; probabilistic power flow

## 附录 A

风速与风电功率特性关系为：

$$P_w = \begin{cases} 0 & 0 \leq v < v_{in} \\ f_w(v) & v_{in} \leq v < v_r \\ P_r & v_r \leq v < v_{out} \\ 0 & v \geq v_{out} \end{cases} \quad (A1)$$

式中： $v_{in}$  为切入风速； $v_{out}$  为切出风速； $v_r$  为额定风速； $f_w(\cdot)$  为从风速到风电功率的转换函数； $P_r$  为额定风速下的额定风电场功率。

### 单高斯模型的概率模型

1) 高斯模型。

对于一维样本数据  $x$ ，其满足下面概率密度函数：

$$P(x|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad (A2)$$

式中： $\mu$  为变量  $x$  的数据期望； $\sigma$  为变量  $x$  的数据标准差； $\boldsymbol{\theta}$  为概率模型参数向量。

当样本数据组为多维数据时，模型的概率密度函数为：

$$P(\mathbf{X}|\boldsymbol{\theta}) = \frac{1}{(2\pi)^{\frac{D}{2}} |\boldsymbol{\Sigma}|^2} \exp\left[-\frac{(\mathbf{X}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X}-\boldsymbol{\mu})}{2}\right] \quad (A3)$$

式中： $\boldsymbol{\Sigma}$  为数据协方差； $D$  为数据维度； $\mathbf{X}$  为多维变量； $\boldsymbol{\mu}$  为多维均值。

2) GMM。

高斯混合模型由  $K$  个子高斯模型共同构成，第  $k$  个子高斯模型对应的权值为  $\alpha_k$ ，模型的概率分布函数为：

$$\begin{cases} P(\mathbf{X}|\boldsymbol{\theta}) = \sum_{k=1}^K \alpha_k \phi(\mathbf{X}|\boldsymbol{\theta}_k) \\ \sum_{k=1}^K \alpha_k = 1 \end{cases} \quad (A4)$$

式中： $\phi(\mathbf{X}|\boldsymbol{\theta}_k)$  为第  $k$  个子模型的高斯分布密度函数；参数  $\boldsymbol{\theta}=(\mu_k, \sigma_k, \alpha_k)$ ， $\mu_k$ 、 $\sigma_k$ 、 $\alpha_k$  分别为第  $k$  个子高斯模型的期望、方差和权值。

### EM 算法的求解过程

1) 依据当前参数，计算多维数据  $\mathbf{x}_j$  来自第  $k$  个子高斯模型的可能性。

$$\gamma_{jk} = \frac{\alpha_k \phi(\mathbf{x}_j | \boldsymbol{\theta}_k)}{\sum_{s=1}^K \alpha_s \phi(\mathbf{x}_j | \boldsymbol{\theta}_s)} \quad j=1,2,\dots, N; \quad k=1,2,\dots, K \quad (\text{A5})$$

2) 计算新一轮迭代的模型参数。

$$\mathbf{u}_k = \frac{\sum_{j=1}^N \gamma_{jk} \mathbf{x}_j}{\sum_{j=1}^N \gamma_{jk}} \quad k=1,2,\dots, K \quad (\text{A6})$$

$$\boldsymbol{\Sigma}_k = \frac{\sum_{j=1}^N \gamma_{jk} (\mathbf{x}_j - \boldsymbol{\mu}_k)(\mathbf{x}_j - \boldsymbol{\mu}_k)^\top}{\sum_{j=1}^N \gamma_{jk}} \quad k=1,2,\dots, K \quad (\text{A7})$$

$$\alpha_{jk} = \frac{\sum_{j=1}^K \gamma_{jk}}{N} \quad k=1,2,\dots, K \quad (\text{A8})$$

式中： $\alpha_{jk}$  为第  $k$  个子模型的概率； $\mathbf{x}_j$  为第  $j$  个数据点； $N$  为数据点总数； $\gamma_{jk}$  为第  $j$  个数据属于第  $k$  个子模型的概率。

算法在迭代过程中不断逼近最优解，迭代收敛后即得到可行解。EM 算法的具体实现步骤如附录 A 图 A1 所示。

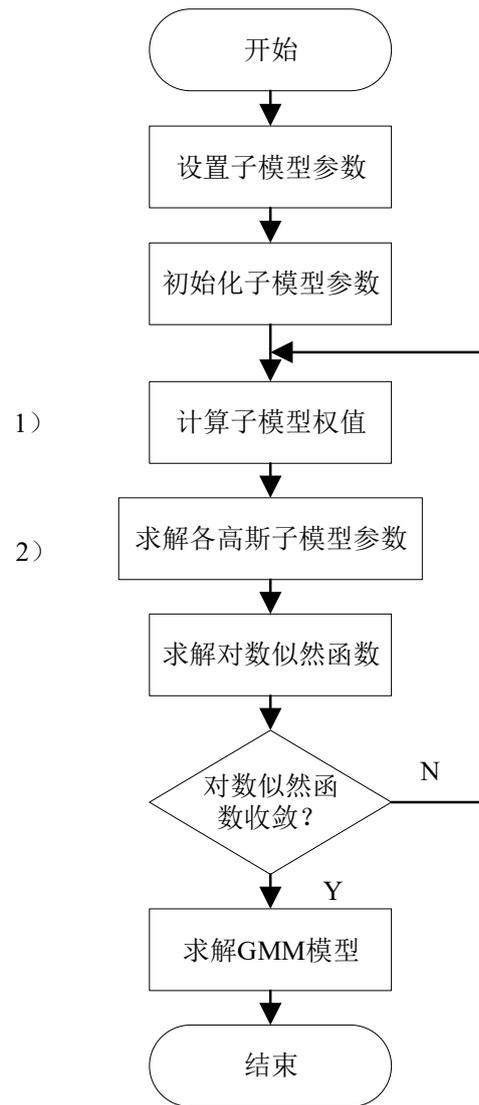
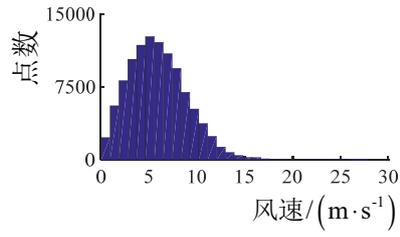
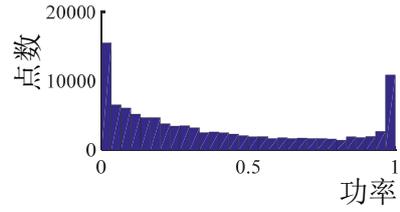


图 A1 GMM 求解流程  
Fig.A1 Solution flowchart of GMM

附录 B



(a) 风速分布直方图



(b) 风电功率分布直方图

图 B1 某地区风速与归一化功率分布直方图

Fig.B1 Histogram of wind speed and normalized power in an area

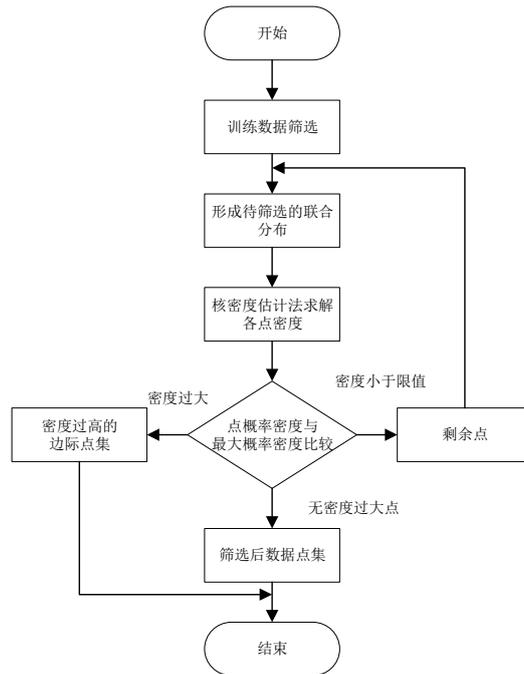


图 B2 GMM 训练数据的筛选流程

Fig.B2 Flowchart of training data screening for GMM

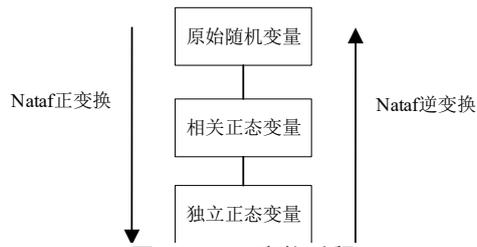


图 B3 Nataf 变换过程

Fig.B3 Process of Nataf transformation

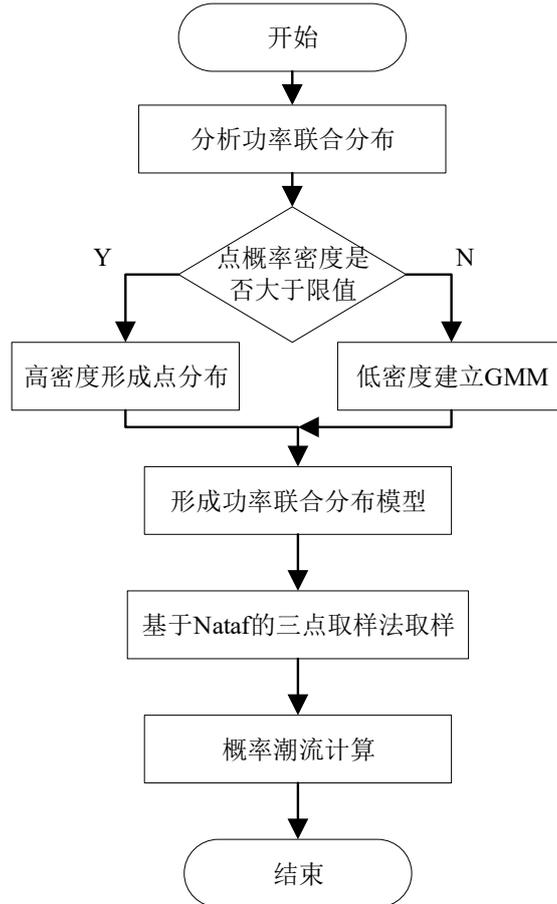


图 B4 接入多风电场电力系统概率潮流计算流程

Fig.B4 Flowchart of probabilistic power flow in power system connected with multiple wind farms

附录 C

表 C1 Gauss-Hermite 采样点与相应权重  
Table C1 Gauss-Hermite sampling points and corresponding weights

节点数	$v_j$	$P_j$
3	$\pm 1.224\ 744\ 871$	0.166 666 667
	0	0.666 666 667
5	$\pm 2.020\ 182\ 871$	0.011 257 411
	$\pm 0.958\ 314\ 808$	0.222 075 922
	0	0.533 333 333
7	$\pm 2.651\ 961\ 356$	0.000 548 269
	$\pm 1.673\ 551\ 628$	0.030 757 124
	$\pm 0.816\ 287\ 882$	0.240 123 179
	0	0.457 142 857

注： $v_j$  为正态随机变量的采样点； $P_j$  为采样点对应的权值。

表 C2 改进的 GMM 参数  
Table C2 Parameters of improved GMM

子模型	$w$	$\mu$	$\theta$
1	0.405 1	[0.567 9, 0.543 4]	[0.044 5, 0.006 8; 0.006 8, 0.047 9]
2	0.066 9	[0.974 0, 0.954 1]	[0.000 4, -0.000 2; -0.000 2, 0.001 3]
3	0.168 7	[0.079 7, 0.072 1]	[0.001 7, -0.000 8; 0.000 8, 0.001 4]
4	0.038 5	[0.995 1, 0.990 9]	$[1.599 \times 10^{-3}, -1.50 \times 10^{-2}; -1.50 \times 10^{-2}, 7.698 \times 10^{-3}]$
5	0.285 6	[0.229 9, 0.220 7]	[0.010 8, 0.003 5; 0.003 5, 0.047 9]
6	0.107 1	[0.889 9, 0.872 0]	[0.005 0, -0.001 8; -0.001 8, 0.008 1]

表 C3 三点估计法的采样点及对应权值  
Table C3 Sampling points and corresponding weights of three-point estimation method

组数	权值	采样点
1	1/6	[0.962, 0.870, 0.791]
2	0	[0.290, 0.363, 0.341]
3	1/6	[0.016, 0.084, 0.093]
4	1/6	[0.290, 0.896, 0.625]
5	1/6	[0.290, 0.074, 0.154]
6	1/6	[0.290, 0.363, 0.860]
7	1/6	[0.290, 0.363, 0.070]

表 C4 2 种方法节点电压及线路总损耗的均值和标准差  
Table C4 Mean value and standard deviation of node voltage and total line power loss for two methods

指标	标准差			均值		
	三点估计法	蒙特卡罗抽样方法	相对误差/%	三点估计法	蒙特卡罗抽样方法	相对误差/%
$V_4$	0.060	0.062	3.22	10.602	10.603	0.01
$V_7$	0.118	0.123	4.06	10.591	10.592	0.02
$V_8$	0.133	0.138	3.62	10.582	10.584	0.03
$V_{11}$	0.156	0.164	4.87	10.585	10.591	0.05
$V_{15}$	0.161	0.166	2.86	10.590	10.592	0.02
$P_{loss}$	6 005.65	5 588.18	6.95	21 471	20 847	2.90