

DOI:10.3969/j.issn.2097-0706.2022.01.003

考虑风-光-荷不确定性的数据驱动型机组组合模型

Data-driven unit commitment model incorporating the uncertainty of wind-PV-load

石立宝, 翟放
SHI Libao, ZHAI Fang

(清华大学深圳国际研究生院 电力系统国家重点实验室深圳研究室, 广东 深圳 518055)
(National Key Laboratory of Power Systems in Shenzhen, Shenzhen International Graduate School,
Tsinghua University, Shenzhen 518055, China)

摘要:传统的概率分布模型在不确定性建模方面主要依赖于模型的选择,但假定的单一模型往往无法准确刻画随机量的复杂变化规律。除此之外,参数概率模型对随机量之间时间、空间相关性的描述不够,而基于连接函数等相关性建模方法在描述多个随机量的相关性方面又过于复杂,给模型实际应用增加了难度。通过采用一种数据驱动的不确定性建模方法,提出了一种数据驱动下的两阶段机组组合模型,基于非参数狄利克雷过程高斯混合模型(DPGMM)和变分贝叶斯推断(VBI)方法来描述风电、光伏和负荷的不确定性,考虑了多风电场和各节点负荷间的相关性,并采用传统数学优化方法对所构建的机组组合模型进行求解。最后,在IEEE-30节点算例系统上进行了仿真验证,结果表明,采用DPGMM模型能够较好地拟合随机量的概率分布并描述随机量之间的相关性。

关键词:风电;光伏;负荷;不确定性;随机量;数据驱动;机组组合;狄利克雷过程;高斯混合模型

中图分类号:TK 01*9:TM 732 **文献标志码:**A **文章编号:**2097-0706(2022)01-0018-08

Abstract: The traditional probability distribution model mainly relies on the choice of model in uncertainty modelling, but the assumed single model often cannot accurately describe the complex variations of random quantities. In addition, the parametric probability model does not adequately describe the temporal and spatial correlations between random variables, while the correlation modelling methods like Copula functions are too complicated in describing the correlations of multiple random variables, which adds difficulty to the practical applications of the model. Taking a data-driven uncertainty modelling method, a data-driven two-stage unit commitment model is proposed. The non-parametric Dirichlet process Gaussian mixture model (DPGMM) and the variational Bayesian inference (VBI) method are used to describe the uncertainty of wind power, photovoltaics and load. Taking the correlations between multiple wind farms and the load of each node into consideration, the traditional mathematical optimization method is applied to solve the established unit commitment model. Finally, the simulation verification is carried out on the IEEE-30 node test system. The results show that the DPGMM model can fit the probability distribution of random quantities and describe the correlation between them.

Keywords: wind power; PV; load; uncertainty; random variable; data-driven; unit commitment; DP; GMM

0 引言

相较于经济调度问题,机组组合(UC)问题在电力系统发电计划编制中应用更加广泛。UC问题包含机组启停优化和经济调度2个基本问题,综合考虑了机组启停成本、机组启停时间约束、爬坡约束乃至网络安全约束。由于存在时间和空间上的双重优化,UC问题的优化难度相对较大,是电网优化

调度领域研究的重点问题之一。UC中不确定性的来源有风电和光伏出力、机组和线路故障、电价和负荷波动,不确定性UC问题的主要建模方法分成2大类,即随机优化的UC(SUC)和鲁棒优化的UC(RUC)。其中,基于概率方法的SUC模型常利用蒙特卡罗法生成相应的不确定性场景,有研究基于风电概率分布生成风电场景^[1-2],还有研究则从负荷预测误差服从正态分布出发建立了负荷不确定性的场景^[3]。随机优化可提供一个相对准确的优化结果,但场景数量和求解精度需要进行一定的平衡。

除此之外,SUC模型很容易忽视极端场景对调度的影响,而且基于概率方法建模所要求的信息量很大,不利于多种不确定性的分析。

随着电网量测技术的进步和可再生能源穿透率的提高,电网调度侧积累了大量发电和负荷数据,针对电力大数据的应用研究已比较广泛^[4-6],特别是针对数据驱动下UC问题的研究方法已有许多新的进展^[7-10]。这些数据驱动下的UC问题优化模型可以分成数据驱动下的SUC模型、数据驱动下的RUC模型以及分布式鲁棒优化模型。相较于传统的随机优化模型,数据驱动的SUC模型能够基于不确定量的预测结果以及实测数据构造更为准确的数学模型,从而提高随机优化方法的准确性,进而增强调度方案的实用性。已有学者对如何利用数据驱动方法构造更准确的概率分布模型进行了细致的研究。其中,数据驱动的SUC模型多通过采用非参数模型、随机过程模型或人工智能方法生成场景来描述随机量。数据驱动下的SUC模型相较于传统SUC模型,虽然一定程度上提高了结果的安全性和经济性,但仍存在实用性低、求解复杂等问题。

本文提出了一种数据驱动下的两阶段机组组合模型。首先,采用非参数狄利克雷过程高斯混合模型(DPGMM)对风、光、荷等数据的概率分布进行拟合,并考虑随机量之间的相关性,使用变分贝叶斯推断(VBI)方法对该模型进行参数估计和采样;接下来,建立数据驱动下的两阶段机组组合模型,在场景法的基础上采用传统数学优化方法并结合商业求解器对该模型进行求解;最后,将所构建的优化模型在IEEE-30节点算例系统上进行仿真验证分析。

1 数据驱动下风-光-荷不确定性建模

1.1 非参数狄利克雷过程高斯混合模型

随机过程可以简单地理解为随机分布的分布。换句话说,从狄利克雷过程(DP)中每次抽取的样本是一个分布,而不是一个随机向量^[11]。作为一种随机过程,狄利克雷过程在贝叶斯推断中有着广泛的应用,通常用来描述随机变量分布的先验知识,即随机变量按照一个或另一个特定分布来分布的可能性有多大。它被称为狄利克雷过程则是因为边缘分布为狄利克雷分布。DPGMM是一个基于断棒过程表达的DP非参数贝叶斯模型,它可以提供无穷多个组成成分来分析变量的概率分布。给定一个 D 维的随机变量 $x = [x_1, x_2, \dots, x_D]$,其联合概率密度分布(PDF)为^[12]

$$p(x|\beta, \Omega) = \sum_{m=1}^{\infty} \beta_m p_m(x|\theta_m), \quad (1)$$

$$\beta = \{\beta_m\}_{m=1}^{\infty}, \Omega = \{\theta_m\}_{m=1}^{\infty}, \quad (2)$$

式中: $p_m(\cdot)$ 为第 m 个混合成分的 D 维PDF,也是多元高斯分布; x 为变量; β 为成分比例向量; Ω 为参数空间; θ_m 为可数无限的独立随机变量。

如果用高斯分布进行表达,上式变为

$$p(x|\beta, \Omega) = \sum_{m=1}^{\infty} \beta_m N_m(x|\mu_m, \Lambda_m^{-1}), \quad (3)$$

$$N_m(x|\mu_m, \Lambda_m^{-1}) = \frac{1}{(2\pi)^{D/2} |\Lambda_m^{-1}|^{1/2}} \exp \left[-\frac{1}{2} (x - \mu_m)^T \Lambda_m (x - \mu_m) \right], \quad (4)$$

$$\theta_m = \{\mu_m, \Lambda_m\}, \Omega = \{\mu_m, \Lambda_m\}_{m=1}^{\infty}, \quad (5)$$

式中: $N_m(\cdot)$ 为高斯分布的概率密度函数; μ_m 为第 m 个高斯成分的均值向量; Λ_m 为第 m 个高斯成分的协方差矩阵。

如果数据集 $X = \{x^{[n]}\}_{n=1}^N$,那么基于数据集 X 的DPGMM联合概率密度表达式为

$$p(X|\beta, \Omega) = \prod_{n=1}^N \left\{ \sum_{m=1}^{\infty} \beta_m N_m(x^{[n]}|\mu_m, \Lambda_m^{-1}) \right\}. \quad (6)$$

设 $z_m^{[n]}$ 为二进制变量且仅在数据点 $x^{[n]}$ 来自第 m 个高斯成分时取值为1。则在示性变量 $Z = \{z_m^{[n]}\}_{n=1, m=1}^{N, \infty}$ 下的条件联合概率密度表达式为

$$p(X|Z, \Omega) = \prod_{n=1}^N \prod_{m=1}^{\infty} N_m(x^{[n]}|\mu_m, \Lambda_m^{-1})^{z_m^{[n]}}. \quad (7)$$

传统高斯混合模型(GMM)中高斯成分的个数对模型的整体表现影响很大,实际应用中靠观察数据集后由人为指定,成分个数设置不合理,模型的表现会显著下降。DPGMM中高斯成分的个数不再依赖于人为指定,而是直接从真实数据中学习得到,相较于GMM具有很大优势。在考虑不确定性因素的相关性方面,DPGMM同GMM一样依赖于协方差矩阵对相关性的刻画。由于具备自我学习的能力,当概率分布模型与设定的Copula函数形状差距较大时,DPGMM对变量拟合的表现将超过Copula函数。

建立DPGMM模型后,下一步则需要估计模型中的超参数和潜在变量 $W = \{Z, \Omega, Y, \Phi\}$ 的后验概率分布,其中 $Y = \{v_m\}_{m=1}^{\infty}, \Phi = \{\phi_m\}_{m=1}^{\infty}$ 表示用于描述DP断棒过程的贝塔分布参数。假设潜在变量之间两两独立,本文采用VBI方法来实现,具体步骤如下。

(1) $\forall j \in \{1, 2, \dots, \bar{M}\}$ ($\bar{M}$ 为给定的截断数),初始化变分分布 $q_j(W_j)$ 。

(2) 循环开始,令 $\varphi(q)$ 为潜在变量 W 和数据集 X

真实的后验分布和基于 VBI 方法近似的变分分布 $\varphi(q)$ 的 Kullback-Leibler (KL) 散度, 即 $\varphi(q) = E_{q(W)}[\ln p(W, X)] - E_{q(W)}[\ln q(W)]$ (式中: E 表示期望运算)。如果 $\varphi(q)$ 未收敛, 则转到步骤(3), 否则转到步骤(5)。

(3) 对于每一个 $j \in \{1, 2, \dots, \bar{M}\}$, 根据下式计算 $q_j(W_j)$

$$\ln q_j(W_j) = E_{q(W)_{i \neq j}}[\ln p(W, X)] + \text{constant},$$

式中: constant 为所有除 $q_j(W_j)$ 以外的常数项。

(4) 由 $q(W) = \prod_{j=1}^{\bar{M}} q_j(W_j)$ 计算 $q(W)$, 并计算 $\varphi(q)$ 。

(5) 循环结束, 返回最优的 $q^*(W)$ 。

1.2 数据驱动下的两阶段 UC 模型

1.2.1 目标函数

通过对数据驱动下风-光-荷的不确定性建模和抽样, 可以得到风电、光伏和负荷的出力场景。因此, 本节建立了一个数据驱动下的两阶段 UC 模型, 其中第 1 阶段先优化机组的启停计划, 第 2 阶段优化机组的出力。因此, 目标函数表达为

$$\min \sum_{t=1}^T \left[\sum_{i \in N_G} S'_{c, G(i)}(u_{G(i)}^t, v_{G(i)}^t) + E_s[Q(o_{G(i)}^t, P_{G(i)}^{t,s}, \xi^{t,s})] \right], \quad (8)$$

$$C'_{G(i)}(u_{G(i)}^t, v_{G(i)}^t) = c'_{o, G(i)} u_{G(i)}^t + c'_{c, G(i)} v_{G(i)}^t, \quad (9)$$

$$\xi^{t,s} = \left\{ P_{W(j)}^{t,s}, j \in N_W, P_{R(r)}^{t,s}, r \in N_R, P_{d(b)}^{t,s}, b \in N_B \right\}, \quad (10)$$

式中: $S'_{c, G(i)}$ 为 t 时刻第 i 个火电机组的启停成本函数; E_s 为第 s 个场景下的期望成本; t 为时间; T 为最大调度时间; N_G 为火电机组集合; N_W 为风电场集合; N_R 为光伏电站集合; N_B 为母线集合; S 为场景的个数, $s \in S$; $o_{G(i)}^t, u_{G(i)}^t, v_{G(i)}^t$ 为二进制变量; 当 t 时刻第 i 个火电机组处于运行状态时, $o_{G(i)}^t=1$, 反之为 0; 当 t 时刻第 i 个火电机组开启时, $u_{G(i)}^t=1$, 反之为 0; 当 t 时刻第 i 个火电机组关停时, $v_{G(i)}^t=1$, 反之为 0; $C'_{G(i)}$ 为 t 时刻第 i 个火电机组的启停成本函数; $c'_{o, G(i)}$ 为 t 时刻第 i 个火电机组的开启成本; $c'_{c, G(i)}$ 为 t 时刻第 i 个火电机组的关停成本; $P_{G(i)}^{t,s}$ 为 t 时刻第 s 个场景下第 i 个火电机组的出力; Q 为第 2 阶段的燃料成本函数; $\xi^{t,s}$ 为 t 时刻第 s 个场景下不确定量的集合; $P_{W(j)}^{t,s}$ 为 t 时刻第 s 个场景下第 j 个风电场的出力; $P_{R(r)}^{t,s}$ 为 t 时刻第 s 个场景下第 r 个光伏电站的出力; $P_{d(b)}^{t,s}$ 为 t 时刻第 s 个场景下节点 b 的负荷。

相较于机组出力, 机组的启停需要在第 1 阶段进行优化, 到了第 2 阶段再针对不同场景优化机组的出力, 而总成本函数则由启停成本和期望成本构成。基于 S 个风-光-荷出力场景, 期望成本表达为

$$E_s[Q(o_{G(i)}^t, P_{G(i)}^{t,s}, \xi^{t,s})] = \frac{1}{S} \sum_{s=1}^S Q(o_{G(i)}^t, P_{G(i)}^{t,s}, \xi^{t,s}), \quad (11)$$

$$Q(o_{G(i)}^t, P_{G(i)}^{t,s}, \xi^{t,s}) = F_{c, G(i)}^{t,s}(o_{G(i)}^t, P_{G(i)}^{t,s}) + C_{W,+}^{t,s} P_{W,+}^{t,s} + C_{R,+}^{t,s} P_{R,+}^{t,s}, \quad (12)$$

$$F_{c, G(i)}^{t,s}(o_{G(i)}^t, P_{G(i)}^{t,s}) = o_{G(i)}^t [a_i (P_{G(i)}^{t,s})^2 + b_i P_{G(i)}^{t,s} + c_i], \quad (13)$$

式中: $C_{W,+}^{t,s}$ 为 t 时刻第 s 个场景下风电的弃风成本系数; $C_{R,+}^{t,s}$ 为 t 时刻第 s 个场景下光伏的弃光成本系数; $P_{W,+}^{t,s}$ 为 t 时刻第 s 个场景下的总弃风功率; $P_{R,+}^{t,s}$ 为 t 时刻第 s 个场景下的总弃光功率; a_i, b_i, c_i 为第 i 个火电机组的燃料成本系数; $F_{c, G(i)}^{t,s}$ 为 t 时刻第 s 个场景下第 i 个火电机组的燃料成本。

需要说明的是, 第 2 阶段各机组出力优化结果与不同场景相关联, 最终总成本为期望成本, 因此, 不同于确定性的机组组合模型, 这里无法给出针对第 1 阶段得到的处于开启状态机组所对应时刻的具体出力值。

1.2.2 第 1 阶段约束条件

(1) 火电机组最小开启和关停时间约束

$$\begin{aligned} -o_{G(i)}^{t-1} + o_{G(i)}^t - o_{G(i)}^k &\leq 0, \\ \forall k: 1 \leq k - (t-1) \leq T_{\text{up}, G(i)}, \\ \forall i \in N_G, \forall t = 1, \dots, T, \end{aligned} \quad (14)$$

$$\begin{aligned} o_{G(i)}^{t-1} - o_{G(i)}^t + o_{G(i)}^k &\leq 1, \\ \forall k: 1 \leq k - (t-1) \leq T_{\text{down}, G(i)}, \\ \forall i \in N_G, \forall t = 1, \dots, T, \end{aligned} \quad (15)$$

式中: $T_{\text{up}, G(i)}, T_{\text{down}, G(i)}$ 分别为第 i 个火电机组最小开启时间和关停时间。

(2) 二进制变量约束

$$\begin{aligned} -o_{G(i)}^{t-1} + o_{G(i)}^t - u_{G(i)}^k &\leq 0, \\ \forall i \in N_G, \forall t = 1, \dots, T, \end{aligned} \quad (16)$$

$$\begin{aligned} o_{G(i)}^{t-1} - o_{G(i)}^t - v_{G(i)}^k &\leq 0, \\ \forall i \in N_G, \forall t = 1, \dots, T, \end{aligned} \quad (17)$$

$$\begin{aligned} o_{G(i)}^t - o_{G(i)}^{t-1} = u_{G(i)}^t - v_{G(i)}^t &\leq 0, \\ \forall i \in N_G, \forall t = 1, \dots, T, \end{aligned} \quad (18)$$

$$\begin{aligned} o_{G(i)}^t, u_{G(i)}^t, v_{G(i)}^t &\in \{0, 1\}, \\ \forall i \in N_G, \forall t = 1, \dots, T. \end{aligned} \quad (19)$$

1.2.3 第 2 阶段约束条件

(1) 火电出力上下限约束

$$\begin{aligned} u_{G(i)}^t L_{G(i)} &\leq P_{G(i)}^{t,s} \leq u_{G(i)}^t U_{G(i)}, \\ \forall i \in N_G, \forall t = 1, \dots, T, \forall s \in S, \end{aligned} \quad (20)$$

式中: $L_{G(i)}, U_{G(i)}$ 分别为第 i 个火电机组最小、最大出力。

(2) 火电机组爬坡功率约束

$$\begin{aligned} o_{G(i)}^t (P_{G(i)}^{t,s} - P_{G(i)}^{t-1,s}) &\leq P_{\text{max}, G(i)}^{\text{up}}, \\ \forall i \in N_G, \forall t = 1, \dots, T, \forall s \in S, \end{aligned} \quad (21)$$

$$o_{G(i)}^t (P_{G(i)}^{t-1,s} - P_{G(i)}^{t,s}) \leq P_{\text{max}, G(i)}^{\text{down}},$$

$$\forall i \in N_G, \forall t = 1, \dots, T, \forall s \in S, \quad (22)$$

式中: $P_{\max, G(i)}^{\text{up}}$ 和 $P_{\max, G(i)}^{\text{down}}$ 分别为第 i 个火电机组最大上行爬坡功率和最大下行爬坡功率。

(3) 功率平衡约束

$$\sum_{i \in N_G} P_{G(i)}^{t,s} + \sum_{j \in N_W} P_{W(j)}^{t,s} + \sum_{r \in N_R} P_{R(r)}^{t,s} - P_{\text{loss}}^{t,s} = \sum_{b \in N_B} P_{d(b)}^{t,s}, \quad (23)$$

$$\forall t = 1, \dots, T, \forall s \in S,$$

式中: $P_{\text{loss}}^{t,s}$ 为 t 时刻第 s 个场景下电网的网损。

(4) 旋转备用约束

$$\sum_{b \in N_B} (1 + r_m) P_{d(b)}^{t,s} \leq \sum_{i \in N_G} P_{rs, G(i)}^{t,s}, \quad (24)$$

$$P_{rs, G(i)}^{t,s} = \min \left\{ o_{G(i)}^t (U_{G(i)}^t - P_{G(i)}^{t,s}), o_{G(i)}^t P_{\max, G(i)}^{\text{up}}, o_{G(i)}^t P_{\max, G(i)}^{\text{down}} \right\}, \quad (25)$$

式中: r_m 为预设的比例系数; $P_{rs, G(i)}^{t,s}$ 为 t 时刻第 s 个场景下第 i 个火电机组的旋转备用。

(5) 网络安全约束

$$-S_{l,k}^{\max} \leq \sum_{b \in N_B} K_{l,k}^b \left[\sum_{i \in \Lambda_G^b} P_{G(i)}^{t,s} + \sum_{j \in \Lambda_W^b} P_{W(j)}^{t,s} + \sum_{r \in \Lambda_R^b} P_{R(r)}^{t,s} - P_{d(b)}^{t,s} \right] \leq S_{l,k}^{\max}, \quad (26)$$

$$\forall l \in N_B, \forall k \in N_B, \forall t = 1, \dots, T, \forall s \in S,$$

式中: $K_{l,k}^b$ 为在节点 b 注入功率, 连接在节点 l 和节点 k 上传输线的线路转移因子; $S_{l,k}^{\max}$ 为连接节点 l 和节点 k 的支路视在功率上限; Λ_G^b 为接入在节点 b 的火电机组集合; Λ_W^b 为接入在节点 b 的风电场集合; Λ_R^b 为接入在节点 b 的光伏电站集合。

2 求解方法

在利用场景法的基础上, 上述模型已转化为一个混合整数线性调度运行优化模型, 因此本节采用传统数学优化方法对该模型进行求解, 其中求解软件为 MATLABTM 环境下 YALMIP 工具箱, 调用 Gurobi 商业求解器 (9.1.1 版本) 进行求解。由于 DPGMM 模型可以有效地对数据进行滚动处理, 因此该 UC 模型可以实现数据驱动和滚动优化, 其求解流程如图 1 所示。

3 算例分析

选取 IEEE-30 节点算例系统对本文提出的模型进行仿真验证分析, 系统接线如图 2 所示。其中: 在节点 5, 6, 11, 15 处各接入一个风电场, 节点 17 处接入一个光伏电站, 节点 12, 21 处负荷具有不确定性 (概率描述)。风电、光伏和负荷均利用 DPGMM 模型对概率分布进行描述并采样, 数据来源于下节的建模案例。模型中, 弃风成本系数 $C_{W,+}^{t,s}$ 和弃光成

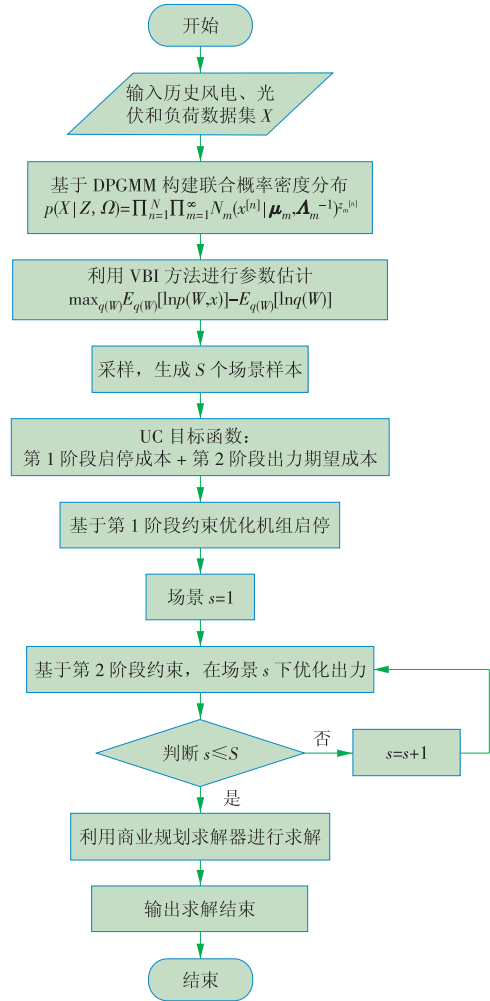


图 1 数据驱动下的 UC 模型求解流程

Fig. 1 Solution of the data-driven UC model

本系数 $C_{R,+}^{t,s}$ 为 2 美元/(MW·h) ($\forall t=1, \dots, T, \forall s \in S$)。

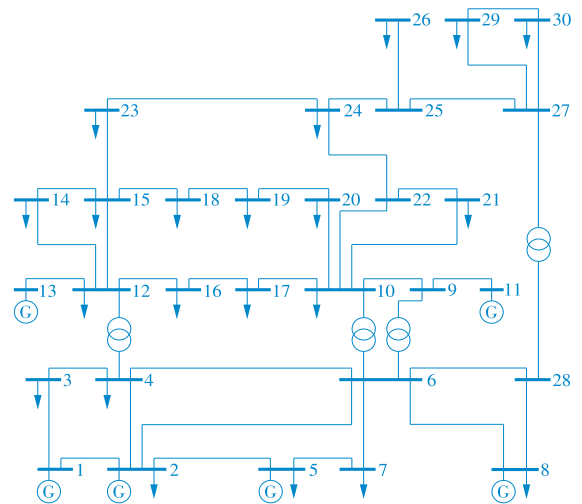


图 2 IEEE-30 节点算例系统

Fig. 2 IEEE-30 node test system

3.1 基于 DPGMM 的风-光-荷不确定性建模案例

本节以美国国家风电数据集提供的风电数据^[13]、美国国家太阳能辐射数据库^[14]提供的太阳辐

照度和光伏出力数据以及纽约市电力系统独立运营商提供的负荷数据^[15]为基础,分别给出基于DPGMM模型的单变量不确定性及不确定性相关性建模案例。

首先,选取某风场2006年6月10日、11日的风速数据并基于DPGMM模型进行拟合,每5 min记录1个风速数据并同Gamma分布、Weibull分布、Lognorm分布和一、二、三、四阶GMM进行比较。

为研究方便,风速数据已进行归一化处理。6月10日、11日风速概率拟合对比分别如图3、图4所示。

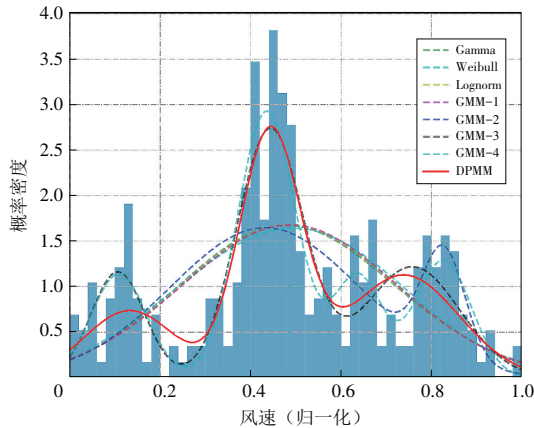


图3 6月10日风速概率拟合对比

Fig. 3 Fitting curves of the wind speed probability on June 10th

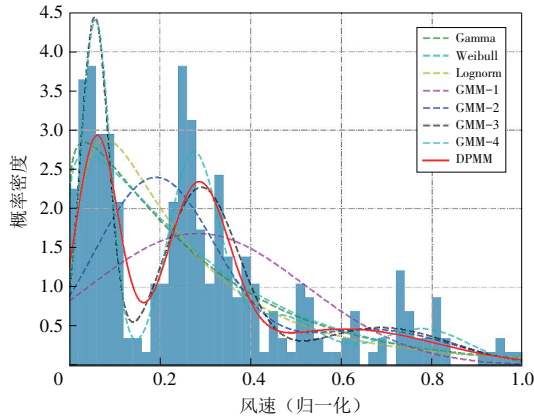


图4 6月11日风速概率拟合对比

Fig. 4 Fitting curves of the wind speed probability on June 11th

为了量化比较各概率模型的有效性,选取对数似然函数值(lnL)和卡方拟合优度(GoF)2个指标进行比较。需要注意的是,lnL值在概率分布模型中越大,代表模型对数据的拟合度越高;而GoF值则刚好相反,模型的GoF值越小其拟合程度越高。6月10日和6月11日风电场1风速拟合的lnL值和GoF值对比结果分别见表1、表2。

光伏电站的功率大小与太阳总辐射(GHI)紧密相关,该指标是指太阳照射在地面上某一处水平面上的总辐射能量。基于GHI值,可以方便地计算出

表1 不同概率模型对风电场1的风速拟合lnL值
Table 1 Wind speed fitting lnL values of wind farm 1 made by different probability models

概率模型	2006-06-10	2006-06-11
Gamma	58.06	214.01
Weibull	61.48	215.12
Lognorm	58.29	198.14
GMM-1	58.44	198.14
GMM-2	72.75	144.04
GMM-3	108.20	247.10
GMM-4	107.48	231.95
DPGMM	115.53	251.63

表2 不同概率模型对风电场1的风速拟合GoF值
Table 2 Wind speed fitting GoF values of wind farm 1 made by different probability models

概率模型	2006-06-10	2006-06-11
Gamma	134.78	81.80
Weibull	129.60	81.08
Lognorm	135.00	98.45
GMM-1	134.81	234.95
GMM-2	111.03	122.70
GMM-3	30.13	35.44
GMM-4	20.80	27.91
DPGMM	18.57	20.09

太阳能电站的出力^[16-17]。选取2006年8月21日的GHI数据,基于DPGMM模型的GHI概率拟合对比如图5所示。负荷则选取纽约市某节点2006年9月1日的数据进行拟合,对比结果如图6所示。

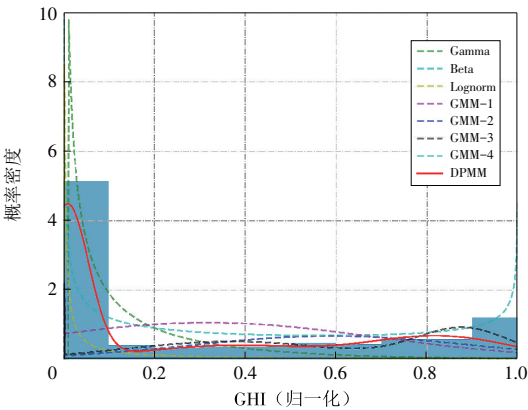


图5 8月21日GHI概率拟合对比

Fig. 5 Fitting curves of GHI probability on August 21st

从以上结果可以看出,在单个不确定因素建模方面,DPGMM模型的优化结果明显好于Gamma分布、高斯分布等传统概率分布模型。

在不确定性相关性建模方面,选取临近的2个风电场,其风速三维直方图如图7所示,由图7可以看出,2个风电场风速间的相关性很高。

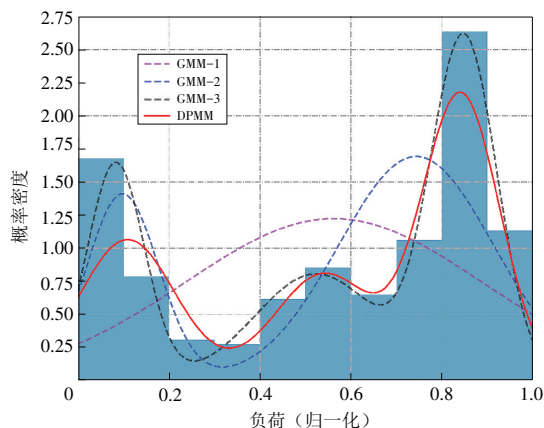


图6 9月1日负荷概率拟合对比

Fig. 6 Fitting curves of the load probability on September 1st

基于DPGMM模型,可以得到2个风电场风速联合概率密度(如图8a所示),其 $\ln L$ 值为420.49。与此同时,应用Copula理论进行拟合,其最优结果为Frank Copula, Copula参数为16.81, Kendall相关系数为0.78, $\ln L$ 值为279.19, 相较DPGMM模型小了很多,其拟合出的概率密度图(如图8b所示)也与直方图相差较大。而当不确定性因素增加到3个以上时, Copula函数则需要引入藤Copula,其处理思路和

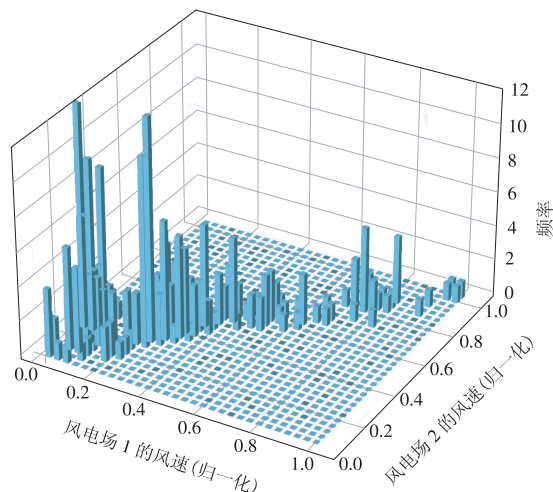


图7 风电场1,2的风速三维直方图

Fig. 7 3D histogram of the wind speed in wind farm 1 and wind farm 2

求解方法也很复杂。选取4个临近风电场的风速数据(2006年6月11日), 基于R藤Copula和C藤Copula理论,在R语言Vine Copula包中优化得到该风电场的Copula概率密度,其 $\ln L$ 值分别为657.44和626.82, 而基于DPGMM模型得到的 $\ln L$ 值为847.60, 显然, DPGMM的拟合效果更好。

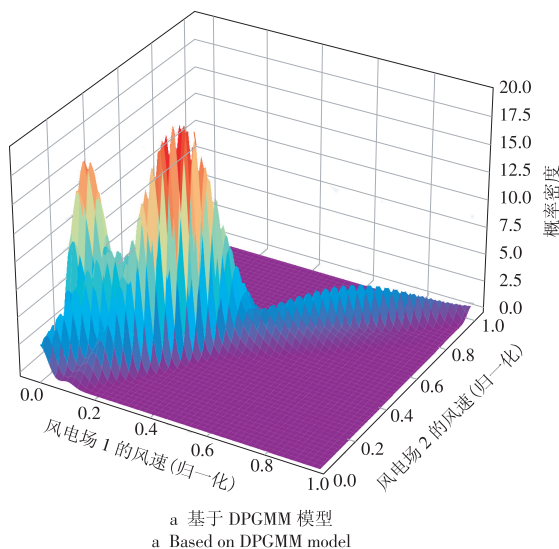
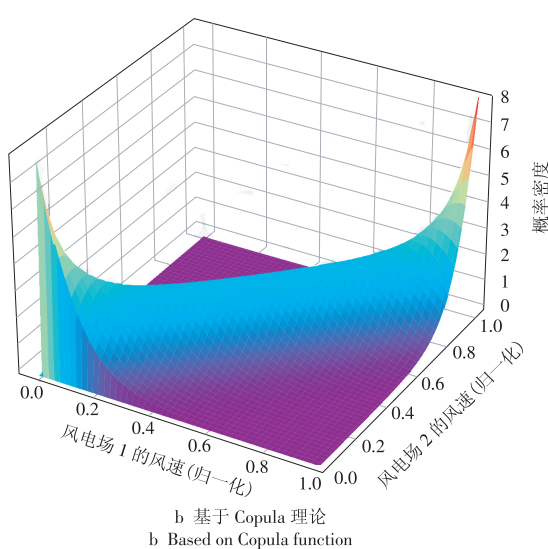
a 基于 DPGMM 模型
a Based on DPGMM modelb 基于 Copula 理论
b Based on Copula function

图8 风电场1,2的风速联合概率密度

Fig. 8 Joint probability density of the wind speed in wind farm 1 and wind farm 2

对于Copula理论来说,其相关性的表达较为明确,但随着随机量的增加,藤结构的数量将非常庞大,计算相对复杂。除此之外,选择不同的藤结构对于最终的结果也不尽相同,这实际上增加了应用难度。在DPGMM模型中,虽然相关性描述依赖于线性矩阵,但多个线性矩阵的描述结果并不一定比基于Copula理论的计算结果差,而且基于DPGMM的方法利于滚动更新,可操作性更强。

3.2 不同风光穿透率和负荷波动率下的调度结果

以新能源穿透率为20%、负荷波动率(负荷波动占总负荷比重)为10%为例,求解出机组的启停计划(见表3,表中:0表示停机;1表示开机)。针对不同的新能源穿透率和负荷波动率,其优化后的成本见表4。

优化结果显示,当新能源穿透率上升时,相同负荷波动率下调度总成本有所下降,其主要原因是低成本的新能源替代了传统火力发电;而当负荷波

表 3 优化求解后的机组启停计划

Table 3 Startup and shutdown schedule of the units after optimization

机组编号	启停计划(01:00—24:00)
G1	11111111111111111111111111111111
G2	11111111111111111111111111111111
G3	11111111111111111111111111111111
G4	000000000000000000000000000000
G5	0011111111111111111111111111000
G6	000000000000000000000000000000

表 4 新能源渗透率和负荷波动率对调度总成本的影响

Table 4 Impacts of new energy penetration level and load fluctuation rate on the total dispatch cost 美元

渗透率/%	负荷波动率/%		
	10	15	20
10	$1.198\ 7 \times 10^4$	$1.283\ 9 \times 10^4$	$1.356\ 4 \times 10^4$
20	$1.029\ 3 \times 10^4$	$1.112\ 0 \times 10^4$	$1.193\ 7 \times 10^4$
30	$9.042\ 0 \times 10^3$	$9.850\ 1 \times 10^3$	$1.056\ 0 \times 10^4$
40	$8.348\ 0 \times 10^3$	$9.117\ 2 \times 10^3$	$9.743\ 0 \times 10^3$

动率增加时,相同新能源渗透率下调度总成本均有所增加,其主要原因则是负荷波动导致弃风和弃光成本增加。

4 结论

本文提出了一种数据驱动下的两阶段 UC 模型,其中采用 DPGMM 和 VBI 方法来描述风电、光伏和负荷的不确定性,并进行拟合和参数估计;在此基础上运用 Gurobi 商业求解器来进行求解。通过算例仿真分析发现:与采用已有的单一概率分布模型(如 Gamma 分布、Weibull 分布等)以及 GMM 拟合模型比较,DPGMM 模型在拟合数据概率上具有更好的表象,而在考虑变量相关性描述中相较于 Copula 模型的表现更优;此外,当新能源渗透率提高时,系统调度总期望成本有所下降,而负荷波动率提高时,调度总期望成本相应增加。

需要指出的是,本文所采用的 DPGMM 模型仅讨论了风电和负荷的空间相关性,且仅刻画了线性相关关系,未来将对风电和负荷的时空相关性和非线性相关性建模展开研究。

参考文献:

[1] TUOHY A, MEIBOM P, DENNY E, et al. Unit commitment for systems with significant wind penetration [J]. IEEE Transactions on Power Systems, 2009, 24(2): 592-601.

[2] WANG J, SHAHIDEHPOUR M, LI Z. Security-constrained unit commitment with volatile wind power generation [J]. IEEE Transactions on Power Systems, 2008, 23(3): 1319-1327.

[3] POZO D, CONTRERAS J. A chance-constrained unit commitment with an n -K security criterion and significant wind generation [J]. IEEE Transactions on Power systems, 2013, 28(3): 2842-2851.

[4] 胡杰,唐静,谢仕义,等.基于大数据技术的电厂设备状态评估和预警应用研究[J].华电技术,2020,42(2):1-6.

HU Jie, TANG Jing, XIE Shiyi, et al. State assessment and early warning application for power plant equipment based on big data technology [J]. Huadian Technology, 2020, 42(2):1-6.

[5] 王凤,周铁梁.基于大数据的大型装备全生命周期管理研究[J].华电技术,2020,42(2):12-16.

WANG Feng, ZHOU Tieliang. Explore the value of data for lifecycle management of large equipment [J]. Huadian Technology, 2020, 42(2):12-16.

[6] 张徐亮,万里冰,钱伟中,等.基于区块链的电力大数据安全保障体系[J].华电技术,2020,42(8):68-74.

ZHANG Xuliang, WAN Libing, QIAN Weizhong, et al. Security assurance system for electric power big data based on blockchain technology [J]. Huadian Technology, 2020, 42(8):68-74.

[7] NING C, YOU F. Data-driven adaptive robust unit commitment under wind power uncertainty: A bayesian nonparametric approach [J]. IEEE Transactions on Power Systems, 2019, 34(3):2409-2418.

[8] 鲁卓欣,徐潇源,严正,等.不确定性环境下数据驱动的电力系统优化调度方法综述[J].电力系统自动化,2020,44(21):172-183.

LU Zhuoxin, XU Xiaoyuan, YAN Zheng, et al. Overview on data-driven optimal scheduling methods of power system in uncertain environment [J]. Automation of Electric Power Systems, 2020, 44(21):172-183.

[9] 赵鑫,郑文禹,侯智华,等.基于粒子群优化算法的多能互补系统经济调度研究[J].华电技术,2021,43(4):14-20.

ZHAO Xin, ZHENG Wenyu, HOU Zhihua, et al. Research on economic dispatch of multi-energy complementary system based on Particle Swarm Optimization [J]. Huadian Technology, 2021, 43(4):14-20.

[10] ZHENG X, CHEN H. Data-driven distributionally robust unit commitment with Wasserstein metric: Tractable formulation and efficient solution method [J]. IEEE Transactions on Power Systems, 2020, 35(6):4940-4943.

[11] ZHANG C, BÜTEPAGE J, KJELLSTRÖM H, et al. Advances in variational inference [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 41

- (8): 2008–2026.
- [12] SUN W, ZAMANI M, HESAMZADEH M R, et al. Data-driven probabilistic optimal power flow with nonparametric Bayesian modeling and inference [J]. IEEE Transactions on Smart Grid, 2019, 11(2): 1077–1090.
- [13] DRAXL C, CLIFTON A, HODGE B M, et al. The wind integration national dataset (WIND) toolkit [J]. Applied Energy, 2015, 151: 355–366.
- [14] SENGUPTA M, XIE Y, LOPEZ A, et al. The national solar radiation data base (NSRDB) [J]. Renewable and Sustainable Energy Reviews, 2018, 89: 51–60.
- [15] NYISO. New York independent system operator: Load data [EB/OL]. [2021-07-20]. <https://www.nyiso.com/load-data>.
- [16] ULLAH Z, WANG S, RADOSAVLJEVIĆ J, et al. A solution to the optimal power flow problem considering WT and PV generation [J]. IEEE Access, 2019, 7: 46763–46772.
- [17] BISWAS P P, SUGANTHAN P, AMARATUNGA G A. Optimal power flow solutions incorporating stochastic wind and solar power [J]. Energy Conversion and Management, 2017, 148: 1194–1207.

(本文责编:刘芳)

作者简介:

石立宝(1971),男,副教授,博士,从事多能源系统互补与协同调度技术、电力信息物理系统风险评估、电力系统连锁故障及韧性评估、人工智能技术在智能电网中的应用等方面的研究,shilb@sz.tsinghua.edu.cn。

“能源电力数字化转型关键技术 with 优选路径” 专刊征稿启事

“碳达峰”和“碳中和”目标驱动下,能源革命与数字革命相融并进已成为必然趋势,能源电力行业升级转型的重要性和迫切度持续升高。国家“十四五”规划纲要明确提出“以数字化转型整体驱动生产方式、生活方式和治理方式变革”。能源电力数字化转型,即实现数字技术与以电力为中心的能源生产、传输、消费、交易、管理等环节的深度融合。近年来,云计算、大数据、人工智能、区块链、数字孪生等新兴技术层出不穷,发展迅猛,数据采集、传输、分析、管理、决策、执行等过程将迎来变革式创新,为能源电力行业创造新价值、培育新动能。值得注意的是,数字化转型是一个长周期、系统性的过程,而现阶段数字技术融合应用存在瓶颈,运营与监管体系尚不完善,信息物理安全风险不明。因此,亟须围绕能源电力数字化转型,科学设计顶层架构与总体目标,综合分析现阶段水平与未来发展趋势,全面梳理基础理论与关键技术,统筹规划发展路径与示范应用,进而整体提升能源电力行业的资源配置能力、安全保障能力和智能互动能力,为落实“碳达峰”和“碳中和”目标提供重要支撑。

鉴于当前的问题与机遇,《综合智慧能源》(《原《华电技术》)拟推出“能源电力数字化转型关键技术 with 优选路径”专刊,特邀请浙江大学李知艺研究员、东北林业大学班明飞副教授和长沙理工大学施星宇讲师担任特约主编,共同探讨能源电力数字化转型相关的前沿成果、关键技术、工程案例以及未来发展趋势,欢迎业内同仁踊跃投稿。

一、征稿范围(包括但不限于)

- (1)能源电力数字化转型的顶层设计、路径优化。
- (2)数字化转型下的能源电力系统投资、规划模式。
- (3)数字化转型下的能源电力系统运营、交易技术。
- (4)能源电力大数据的智能分析技术与应用。
- (5)能源电力数字化转型的监管体系与标准化工作。
- (6)能源电力系统的信息物理安全风险分析与防护。
- (7)能源电力系统的预测、运行控制、故障诊断技术。
- (8)能源电力系统数据要素的价值挖掘与隐私防护技术。
- (9)人工智能、区块链、数字孪生等技术的融合应用。
- (10)智慧园区、车联网与需求侧主动支撑技术。

二、时间进度

专刊拟于2022年7月1日截稿,2022年择期出版。

三、征稿要求

- (1)专刊只收录未公开发表的论文,拒绝一稿多投。作者对论文内容的真实性和客观性负责。
- (2)按照《综合智慧能源》论文格式要求使用 Word 软件排版,论文模板请在网站(www.hdpower.net)首页作者中心下载。
- (3)请保留论文图片、曲线和表格原始文件,并在投稿时按规定提交。
- (4)论文作者应遵守相关学术不端规定。

四、投稿方式

- (1)在线投稿(推荐):登录在线采编系统(www.hdpower.net),完成在线全文投稿,欢迎投稿时推荐审稿人。
- (2)邮箱投稿:zhiyi@zju.edu.cn(李老师);hdjs-chd@vip.163.com(编辑部)。
- (3)咨询联系:刘芳 0371-58501060/13838002988;杨满成 010-63918755/ 13801175292。