

基于变分贝叶斯学习的光伏功率波动特性研究

李 芬¹,李春阳¹,闫全全²,赵晋斌¹,段善旭³

(1. 上海电力学院 电气工程学院,上海 200090;2. 上海市电力公司检修公司,上海 200063;
3. 华中科技大学 强电磁工程与新技术国家重点实验室,湖北 武汉 430074)

摘要: 光伏出力波动严重影响电力系统稳定运行。对光伏出力爬坡率进行分析,建立光伏出力爬坡率的高斯混合模型,并用变分贝叶斯学习算法估计模型参数。某光伏电站大量实测数据检验表明,在进行光伏功率波动特性研究方面,在不同时间尺度和天气类型下,变分贝叶斯学习算法比单一分布及基于最大期望算法的方法具有更好的拟合效果。

关键词: 光伏功率波动; 变分贝叶斯学习; 高斯混合模型; 爬坡率

中图分类号: TM 615

文献标识码: A

DOI: 10.16081/j.issn.1006-6047.2017.08.013

0 引言

光伏发电是一个多变量耦合的高度非线性随机过程,受天气、气候等气象环境因素的影响显著。随着目前光伏并网渗透率的提升,光伏功率的不确定性和波动性给电网电能质量、规划、运行和调度等方面带来很大影响和挑战。而短时间内大幅度有功功率增减,会对负载和电网平衡产生巨大危害,对此,国家电网公司专门对并网光伏电站功率波动制定了企业标准^[1]。

针对光伏功率波动问题,文献[2]采用功率变化率控制的策略对储能进行控制,从而控制光伏出力的爬坡率 RR(Ramp Rate)。文献[3]通过太阳辐射数据及建立的光伏出力模型,模拟多地光伏电站小时级光伏功率数据,用来评估光伏电站的总爬坡率,并根据卫星资料分析特殊天气类型下光伏爬坡率。文献[4]分析了不同时间尺度下的光伏功率波动。文献[5]使用秒级太阳辐射度数据模拟光伏出力数据,来分析光伏电站每天的最大波动量。文献[6]建立了光伏功率变化的指数分布,用于评估短期功率波动幅值。文献[7]采用光伏功率变化率来分析不同时空尺度下光伏功率波动,并用 t location-scale 分布求解波动变化率的置信区间,来制定自动发电控制(AGC)备用容量。文献[8]分析了光伏出力波动对电网的影响,采用光伏出力一阶差分量来刻画功率波动,并运用 3 种单一分布研究不同时间尺度下光伏功率波动的随机分布特性。现有文献对光伏

功率波动特性的研究仅限于单一分布^[6-8],还没有关于混合分布的研究,但在风电功率波动方面,关于混合分布的研究已很成熟,所使用的混合分布模型有混合高斯分布^[9-10]和混合 t location-scale 分布^[11]等,并运用最大期望 EM(Expectation Maximization)算法或最大似然估计 MLE(Maximum Likelihood Estimation)求解混合模型参数。由于使用 EM 算法或 MLE 求解参数时,无法确定混合分布的混合度,上述文献只是研究了固定混合度的混合分布。

本文基于武汉某屋顶分布式光伏电站 2 年 5 min 级实测数据,首先给出爬坡率的定义以及分析不同时间尺度下光伏爬坡率的变化;其次建立光伏功率爬坡率的高斯混合模型,运用变分贝叶斯学习算法估计混合模型参数,并通过在不同时间尺度和天气类型下,分别与单一分布和 EM 算法求取的混合模型参数进行对比检验。

1 爬坡率定义

10min 时间间隔内的光伏功率波动主要影响电网频率,进而影响电能质量;大于 10min 时间间隔的光伏功率波动需要增加或减小备用容量或通过调度消除^[6],因此需要用爬坡率等指标^[12]量化光伏功率波动。文献[12]定义 Δt 时间间隔的端点功率差值作为光伏功率爬坡率,但是这样并不能表示 Δt 时间间隔内功率的最大波动,因此,定义 Δt 时间间隔内功率爬坡率 $\delta_{RR}(t)$ 为 Δt 时间间隔内最大功率和最小功率差值(最大波动量)与光伏电站额定功率的比值,即:

$$\delta_{RR}(t) = \begin{cases} \frac{\max[P(t, t+\Delta t)] - \min[P(t, t+\Delta t)]}{P_N} & t_{P_{\max}} > t_{P_{\min}} \\ \frac{\min[P(t, t+\Delta t)] - \max[P(t, t+\Delta t)]}{P_N} & t_{P_{\max}} < t_{P_{\min}} \end{cases} \quad (1)$$

其中, $P(t, t+\Delta t)$ 为 $t \sim t+\Delta t$ 时间段内的所有瞬时功

收稿日期:2017-03-01;修回日期:2017-07-03

基金项目:国家自然科学基金青年项目(51307105);上海绿色能源并网工程技术研究中心(13DZ2251900);上海市经济和信息化委员会专项资金资助项目(沪 CXY-2016-012)

Project supported by the Young Scientists of National Natural Science Foundation of China(51307105),Shanghai Engineering Research Center of Green Energy Grid-Connected Technology(13DZ2251900) and Shanghai Economic and Information Committee Special Funds(CXY-2016-012)

率; $\max[P(t, t+\Delta t)]$ 和 $\min[P(t, t+\Delta t)]$ 分别为 $t \sim t+\Delta t$ 时间段内的最大功率和最小功率; $t_{p_{\max}}$ 和 $t_{p_{\min}}$ 分别为最大功率和最小功率出现的时刻; P_N 为光伏装机容量。

在 Δt 时间间隔内, 当最大功率出现的时刻晚于最小功率出现的时刻, 为负爬坡率, 反之为正爬坡率; 负爬坡率对电网危害较大, 需要通过增加备用容量或调度补充因光伏功率跌落而缺失的功率。

图 1 给出了 3 月 17 日的光伏出力, 以及 5 min 和 15 min 时间尺度下光伏爬坡率的波形, 图中采样点间隔 5 min。可以看出, 15 min 时间尺度下光伏电站最大正负爬坡率均达额定功率的 40%, 且随着时间尺度的增加, 爬坡率逐渐增加。为了定量分析不同时间尺度下光伏电站爬坡率的分布情况, 便于备用容量整定或调度, 需要对光伏电站不同时间尺度下的爬坡率建立模型。

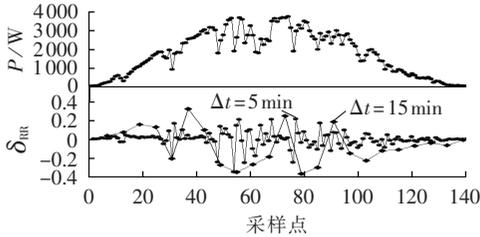


图 1 日光伏出力与不同时间尺度下爬坡率对比
Fig.1 Comparison between daily PV output and ramp rate, for different time scales

2 高斯混合模型建立

本文采用有限阶高斯混合模型, 即有限个高斯分布的线性叠加, 来对光伏爬坡率进行建模。收集待评估光伏电站历史出力数据, 计算光伏爬坡率, 建立如式(2)所示的高斯混合模型。

$$p(\delta_{RR}(t)|\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{i=1}^K \pi_i f(\delta_{RR}(t)|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \quad (2)$$

其中, K 为高斯混合模型的混合度; $\boldsymbol{\pi}$ 为混合系数, 即每个高斯分布在混合模型中的权重; $\boldsymbol{\mu}$ 和 $\boldsymbol{\Sigma}$ 分别为高斯分布的均值和协方差矩阵; $\pi_i, \boldsymbol{\mu}_i$ 和 $\boldsymbol{\Sigma}_i$ 分别为高斯混合模型中第 i 个高斯分布的混合系数、均值和协方差; $f(\cdot)$ 为高斯分布的概率密度函数。

对于建立的高斯混合模型, 需要估计模型参数 $\{K, \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\}$ 。采用变分贝叶斯学习算法, 也称为变分贝叶斯期望最大 VBEM (Variational Bayesian Expectation Maximization) 算法^[13] 估计混合模型参数。VBEM 算法利用样本数据的先验信息及自带模型复杂度惩罚函数^[14], 可自适应地估计模型阶数, 即混合度 K , 同时避免了传统 EM 算法和经典 MLE 算法过拟合的情形, 及蒙特卡洛算法计算量大且收敛较慢的问题, 进而提高模型精度。

定义样本集 $\boldsymbol{X} = \{\delta_{RR}(t_1), \delta_{RR}(t_2), \dots, \delta_{RR}(t_N)\}$ 表示

以 Δt 为时间间隔的光伏爬坡率, N 为样本长度, 隐变量 $\boldsymbol{Z} = \{z_{1k}, z_{2k}, \dots, z_{Nk}\} (k=1, 2, \dots, K)$ 表示光伏功率爬坡率在高斯混合模型的归属情况, \boldsymbol{Z} 为 $N \times K$ 阶矩阵, \boldsymbol{Z} 中任一元素 $z_{nk} \in \{0, 1\}$, 且 $\sum_{k=1}^K z_{nk} = 1, z_{nk} = 1$ 表示第 n 时刻的光伏功率爬坡率属于第 k 个高斯分布。

根据共轭指数域内先验和后验分布规律的一致性定律^[13], 可以通过高斯混合模型参数的先验分布计算混合参数。混合系数 $\boldsymbol{\pi}$ 的共轭先验分布是 Dirichlet 分布^[13-14]:

$$p(\boldsymbol{\pi}) = \text{Dir}(\boldsymbol{\pi}|\boldsymbol{\alpha}_0) = C(\boldsymbol{\alpha}_0) \prod_{k=1}^K \pi_k^{\alpha_0 - 1} \quad (3)$$

其中, $\text{Dir}(\boldsymbol{\pi}|\boldsymbol{\alpha}_0)$ 表示以 $\boldsymbol{\alpha}_0$ 为参数、 $\boldsymbol{\pi}$ 为自变量的 Dirichlet 分布概率密度函数; $C(\boldsymbol{\alpha}_0)$ 为 Dirichlet 分布归一化常数; $\boldsymbol{\alpha}_0$ 可以理解为观测样本的有效先验数, 与混合模型的每个参数都有关系。

定义 $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}^{-1}$, 均值 $\boldsymbol{\mu}$ 和精度矩阵 $\boldsymbol{\Lambda}$ 的共轭先验分布为 Wishart 分布与高斯分布的乘积^[12], 即:

$$p(\boldsymbol{\mu}, \boldsymbol{\Lambda}) = p(\boldsymbol{\mu}|\boldsymbol{\Lambda})p(\boldsymbol{\Lambda}) = \prod_{k=1}^K f(\boldsymbol{\mu}_k | m_0, (\beta_0 \boldsymbol{\Lambda}_k)^{-1}) \boldsymbol{W}(\boldsymbol{\Lambda}_k | W_0, \nu_0) \quad (4)$$

其中, ν_0 为自由度; \boldsymbol{W} 为正定阵; m_0, β_0 和 W_0 均为超参数。

为了避免多变量情况下对边缘似然函数 $p(\boldsymbol{X})$ 的复杂积分运算, 根据混合模型参数的先验信息 $p(\boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda})$, 用一种相对容易计算的分布 $q(\boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda})$ 逼近混合模型参数的真实后验分布 $p(\boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}|\boldsymbol{X}, \boldsymbol{\pi})$ 。通过最大化变分似然函数的下界, 估计混合模型参数。边缘似然函数^[14]可等价表示为:

$$\begin{aligned} \ln p(\boldsymbol{X}|\boldsymbol{\pi}) = & \iiint q(\boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) \ln \frac{p(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}|\boldsymbol{\pi})}{q(\boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda})} d\boldsymbol{\mu} d\boldsymbol{\Lambda} d\boldsymbol{Z} + \\ & \left(- \iiint q(\boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) \ln \frac{q(\boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda})}{p(\boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}|\boldsymbol{X}, \boldsymbol{\pi})} d\boldsymbol{\mu} d\boldsymbol{\Lambda} d\boldsymbol{Z} \right) = \\ & F(q) + \text{KL}(q||p) \end{aligned} \quad (5)$$

其中, $F(q)$ 为变分下限; $\text{KL}(q||p) \geq 0$ 为 $q(\boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda})$ 与 $p(\boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}|\boldsymbol{X}, \boldsymbol{\pi})$ 间的 K-L 散度。为逼近给定函数 $\ln p(\boldsymbol{X}|\boldsymbol{\pi})$, 只需使变分下限 $F(q)$ 最大化, 即最小化 $\text{KL}(q||p)$ 。由式(5)知, $\text{KL}(q||p) = 0$ 即 $q(\boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) = p(\boldsymbol{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}|\boldsymbol{X}, \boldsymbol{\pi})$ 时, 变分下限 $F(q)$ 达到最大化。

VBEM 算法和 EM 算法一样, 需要经过循环迭代才能得到最优解, 即需要经过求取变分贝叶斯期望和使变分贝叶斯最大化 2 步。

a. 变分贝叶斯期望步骤: 根据隐变量 \boldsymbol{Z} 的近似后验分布计算隐变量分布参数 r_{nk} , 其中隐变量的近似后验分布为:

$$q(\boldsymbol{Z}) \propto \prod_{n=1}^N \prod_{k=1}^K \rho_{nk}^{z_{nk}} \quad (6)$$

其中, ρ_{nk} 可表示为^[13]:

$$\ln \rho_{nk} = E[\ln \pi_k] + \frac{1}{2} E[\ln |A_k|] - \frac{1}{2} \ln(2\pi) - \frac{1}{2} E_{\mu_k, A_k}[(x_n - \mu_k)^T A_k (x_n - \mu_k)] \quad (7)$$

对 ρ_{nk} 进行标准化:

$$r_{nk} = \rho_{nk} / \sum_{i=1}^K \rho_{ni} \quad (8)$$

b. 变分贝叶斯最大化步骤: 根据参数的先验分布,混合系数 π_k 的后验分布 $q(\pi_k)$ 为 Dirichlet 分布, μ 和 Λ 的后验分布 $q(\mu_k, \Lambda_k)$ 为高斯-威沙特分布。

$$q(\pi_k) = \text{Dir}(\pi_k | \alpha_k) \quad (9)$$

$$q(\mu_k, \Lambda_k) = f(\mu_k | m_k, (\beta_k \Lambda_k)^{-1}) \mathbf{W}(\Lambda_k | W_k, v_k) \quad (10)$$

其中,超参数 $\alpha_k, \beta_k, m_k, v_k$ 和 W_k 计算公式见文献[13]。在最大化步骤中,根据更新的超参数及式(3)、(4)和式(6)–(10)计算变分下界 $F(q)$ 。

$$F(q) = E[\ln p(\mathbf{X} | \mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) + \ln p(\boldsymbol{\mu}, \boldsymbol{\Lambda}) + \ln p(\mathbf{Z} | \boldsymbol{\pi}) + \ln p(\boldsymbol{\pi})] - E[\ln q(\boldsymbol{\mu}, \boldsymbol{\Lambda}) + \ln q(\mathbf{Z}) + \ln q(\boldsymbol{\pi})] \quad (11)$$

在迭代过程中,对数据解析能力较弱的混合分量的混合系数逐渐减小,因此可以通过设定阈值剔除。当式(11)变分下界达到最大化时迭代停止,具体迭代步骤如图 2 所示。根据迭代停止后输出的最佳超参数 $\alpha_k, \beta_k, m_k, v_k$ 和 W_k , 计算高斯混合模型混合系数为 $\pi_k = N_k / \sum_{k=1}^K N_k$ (N_k 为第 k 个高斯分布的 r_{nk} 之和)、均值 μ_k 和协方差 Σ_k 。

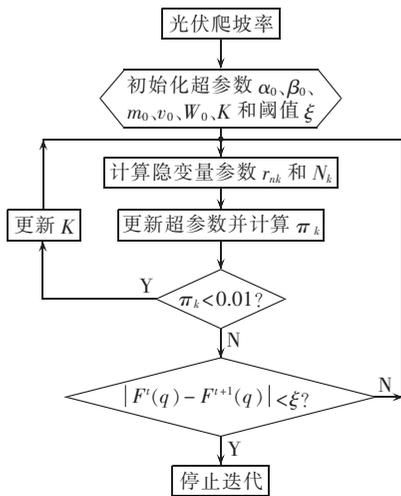


图 2 VBEM 算法参数求解流程图

Fig.2 Flowchart of VBEM method for parameter estimation

3 实例分析

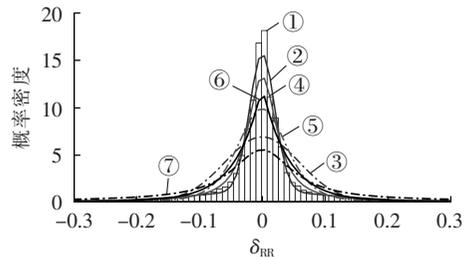
选取武汉某屋顶分布式光伏电站 2 年 5 min 实测有功功率数据作为样本进行分析。由于光伏发电固有的周期性及日变化等特征,本文根据日出日落时角值,截取太阳辐射有效时段对应的光伏电站出

力数据分析。

3.1 与单一分布的对比检验

选取正态分布、 t location-scale 分布、Laplace 分布、Logistic 分布和柯西分布^[8-11] 5 种单一分布作为光伏功率爬坡率的概率密度函数模型。为了更直观地展现光伏功率爬坡率的变化情况,选用概率密度函数直方图进行研究,并运用 Moore 公式^[16]作为直方图组数的选择标准。使用可决系数 R^2 和均方根误差 RMSE(Root Mean Squared Error)作为判断理论概率分布与观测数据是否一致的标准^[17]。

5 种单一分布及运用 VBEM 算法求解的高斯混合模型拟合对比如图 3 所示,其拟合误差如表 1 所示。从表 1 可看出,运用 VBEM 算法求解的高斯混合模型的混合度为 2,其 R^2 为 0.9798,大于 5 种单一分布;RMSE 为 0.0421,小于 5 种单一分布,说明高斯混合模型更适合描述光伏功率爬坡率的概率分布。单一分布中, t location-scale 分布的 R^2 为 0.9494, RMSE 为 0.0666,两拟合指标均优于其他 4 种单一分布,柯西分布拟合效果最差。



① 频率直方图, ② VBEM 算法
③ 正态分布, ④ t location-scale 分布
⑤ Logistic 分布, ⑥ Laplace 分布, ⑦ 柯西分布

图 3 单一模型与混合模型对比结果

Fig.3 Comparison among single distribution models and mix distribution model

表 1 单一模型与混合模型拟合指标

Table 1 Fitting indexes of single distribution models and mix distribution model

分布类型	RMSE	R^2
VBEM 算法	0.0421	0.9798
正态分布	0.1791	0.6340
t location-scale 分布	0.0666	0.9494
Logistic 分布	0.1262	0.8184
Laplace 分布	0.1067	0.8702
柯西分布	0.1939	0.5709

根据林德伯格-列维定理,对于一个独立同分布且数学期望和方差有限的随机变量序列,其标准化后服从标准正态分布。文献[9]认为,若标准化后的随机变量序列服从标准正态分布,则这个序列可以用单一分布较好地描述,反之则不能全局最优。对所采用的数据样本进行标准化,图 4 给出了 5、10、15 min 3 种时间尺度下标准化后的概率密度函数曲

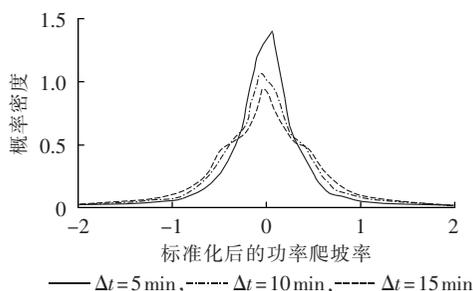


图 4 不同时间尺度标准化后的爬坡率概率密度曲线
Fig.4 Normalized ramp rate probability density curves of different time scales

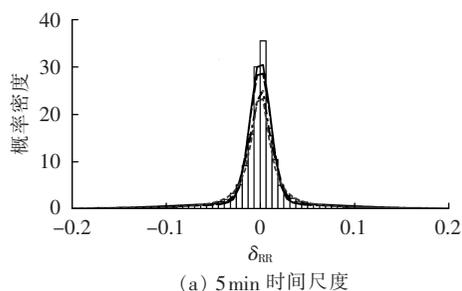
线。可以看出,标准化后的光伏功率爬坡率主要在 $-1 \sim 1$ 之间变化,且随着时间尺度的增加,概率分布曲线逐渐变宽,峰度^[18]逐渐减小,且不同时间尺度下的概率密度函数曲线形状也有差异。样本在 3 种时间尺度下并不是严格地遵守标准正态分布,说明样本不是同分布的,理论上不能用单一分布描述光伏功率爬坡率的概率分布。

3.2 不同时间尺度的对比检验

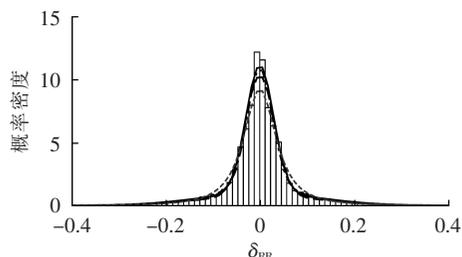
光伏出力主要与太阳辐照度有关,太阳辐照度在不同时间尺度上具有不同的相关性,时间尺度越小,前后 2 个时间间隔的太阳辐照度相关性越强,辐照度变化越小,光伏功率波动越小。为了验证算法在不同时间尺度下对光伏功率爬坡率概率密度函数的拟合效果,选择 5、15、30 min 时间尺度下的光伏功率爬坡率作为样本,分别进行 2、3 阶高斯混合模型和 t location-scale 分布的对比。由于 EM 算法对初值要求严格,因此,本文采用 K 均值聚类确定 EM 算法的初值,以提高算法精度。

图 5 给出了 3 种情况下的拟合图形,拟合误差如表 2 所示。随着时间尺度的增加,光伏功率爬坡率概率密度函数直方图从“高而窄”逐渐变为“矮而宽”,峰度逐渐减小,“拖尾”逐渐严重。低峰度倾向于在均值附近有较小的峰值,直观上看,随着时间尺度的增加,直方图在爬坡率 0 附近的概率密度值逐渐减小,如 5 min 时间尺度时,爬坡率 0 附近的峰值是 30 min 时间尺度时峰值的 5 倍,且主要变化范围从 $-0.1 \sim 0.1$ 增加到 $-0.4 \sim 0.4$,95% 置信区间从装机容量 8% 增加到 22.8%,较大爬坡率所占比重也逐渐增加。由图 5 可见,由 EM 算法求解的高斯混合模型与由 VBEM 算法求解的高斯混合模型在爬坡率较小时,拟合结果相差不大,但是在 0 附近,两者有明显的差异,且后者拟合效果较前者更好。

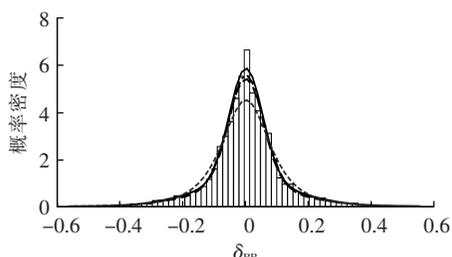
由表 2 可见,随着时间尺度增加,根据 VBEM 算法求取的高斯混合模型最佳混合度由 3 阶降为 2 阶。当混合度超过模型最佳混合度时,会使模型变得复杂,而复杂模型容易受到伪细节结构的影响,即求取的模型参数并非全局最优,使得在 15 min 和 30 min



(a) 5 min 时间尺度



(b) 15 min 时间尺度



(c) 30 min 时间尺度

▮ 频率直方图, — VBEM 算法
--- 3 阶高斯, - - - 2 阶高斯, ····· t location-scale

图 5 不同时间尺度下光伏功率爬坡率拟合结果
Fig.5 Results of PV output ramp rate fitting for different time scales

表 2 3 种时间尺度拟合指标

Table 2 Fitting indexes for three time scales

时间尺度/min	误差类型	VBEM 算法	2 阶高斯	3 阶高斯	t location-scale
5	RMSE	0.0673	0.1039	0.0724	0.1050
	R ²	0.9765	0.9626	0.9394	0.9223
15	RMSE	0.0301	0.0350	0.0364	0.0588
	R ²	0.9848	0.9797	0.9781	0.9505
30	RMSE	0.0220	0.0340	0.0344	0.0524
	R ²	0.9854	0.9826	0.9812	0.9359

时间尺度下 EM 算法求解的 3 阶高斯混合模型的拟合效果差于 2 阶。VBEM 算法通过计算模型参数和隐变量的后验分布,并运用贝叶斯推理中的均值估计模型参数,而 EM 算法只是运用模型参数的众数进行参数估计,不能运用参数的后验分布,因此,同一阶数下,VBEM 算法求取的高斯混合模型拟合效果好于 EM 算法求解的混合模型。 t location-scale 分布的拟合效果差于高斯混合模型。VBEM 算法是基于变分贝叶斯框架,边缘似然函数 $\ln p(X|\pi)$ 被认为隐含复杂度惩罚函数,散度 $KL(q||p)$ 被认为是模型复杂度惩罚函数^[14],即模型越复杂散度 $KL(q||p)$

值越大,在参数上的表现是某些类的 N_k 很小,使得这些类的混合系数 π_k 很小,从而通过剔除过小的 π_k 达到自动调节模型混合度的目的。EM 算法混合度需人为设定,但光伏功率爬坡率所服从的高斯混合模型的混合度不恒定,因此,运用 EM 算法求取高斯混合模型对光伏功率爬坡率进行拟合时,需要反复尝试才能确定最佳混合度,且在最佳混合度下拟合效果也差于 VBEM 算法。

3.3 不同天气类型下的算法检验

根据文献[19]中所使用的光伏出力曲线类别划分的方法,对原始光伏功率数据进行不同天气类型划分,划分为 4 类,即晴天、多云、阴天和转折天气。图 6 为不同天气类型下光伏功率爬坡率散点图,可以看出不同天气类型下,光伏功率爬坡率分布情况差异很大。晴天时,由于受太阳辐射(天文)周期性变化的影响,光伏功率曲线比较光滑,具有明显的日(内)变化特征,在 MATLAB 的 Curve fitting 工具箱中对 5 min 时间尺度下光伏功率爬坡率进行 Fourier 曲线拟合,拟合曲线为正弦曲线,拟合结果如图 6(a) 所示,其他天气类型下的波动无明显规律;多云时,由于云层等的随机性,使得光伏出力较大的时刻功率爬坡率差异较大,爬坡率在散点图中表现为较分散;阴天时,云层的覆盖较为均匀,光伏出力水平较低,光伏功率爬坡率相对较小;转折天气时,天气发生突变,此时光伏功率会突然上升或下降,爬坡率较大,对电网的稳定运行影响较大。4 种天气类型下,95% 置信区间分别为装机容量的 4.7%、11.4%、4.4% 和 14.5%。基于以上原因,选择不同天气类型的光伏功率爬坡率对模型算法进行检验。选择 EM 算法求取的 2 阶高斯混合模型与 VBEM 算法求解的高斯混合模型作对比。

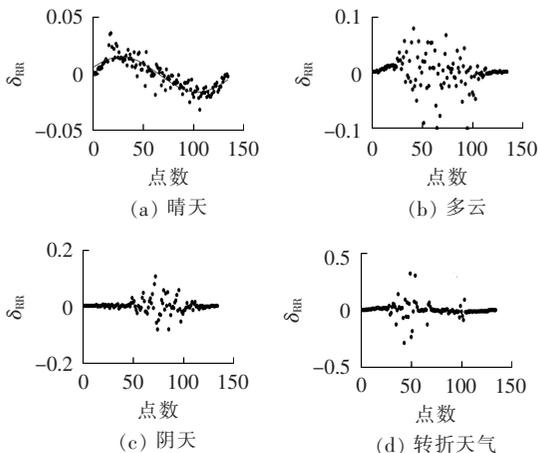


图 6 不同天气类型下光伏功率爬坡率

Fig.6 PV output ramp rate for different weather types

图 7 为晴天和转折天气类型下光伏功率爬坡率概率密度直方图拟合图形,拟合误差如表 3 所示。

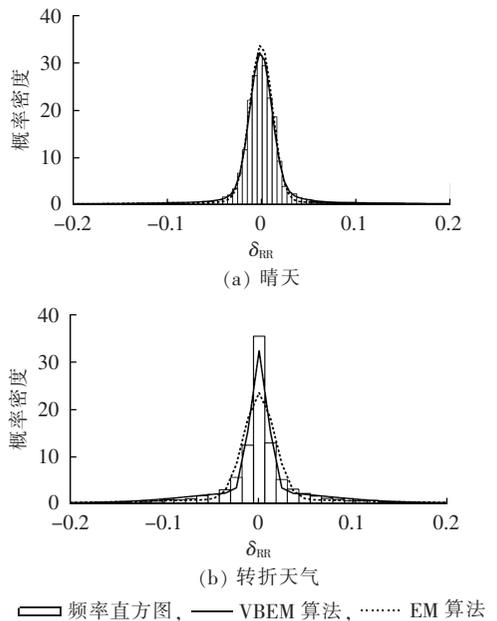


图 7 晴天和转折天气拟合结果

Fig.7 Fitting results for sunny day and transition weather

表 3 晴天和转折天气拟合误差

Table 3 Fitting errors for sunny day and transition weather			
天气类型	误差类型	VBEM 算法	EM 算法
晴天	RMSE	0.0449	0.0652
	R ²	0.9952	0.9898
转折天气	RMSE	0.0741	0.1904
	R ²	0.9773	0.8502

不同天气类型下,光伏功率爬坡率直方图不同,晴天时,“拖尾”较小,而转折天气时,呈尖峰状,“拖尾”较长。晴天时,光伏功率爬坡率主要在 $-0.05 \sim 0.05$ 区间内,直方图各组概率密度变化较缓和,而且组距较窄,而转折天气时波动范围为 $-0.1 \sim 0.1$,直方图在 $-0.05 \sim 0.05$ 范围内各组概率密度值变化较大,在 0 附近为单条纹,组距较宽。2 种天气类型下直方图左右对称,说明正负爬坡率概率基本相同。

由表 3 可知,2 种天气类型下,EM 算法求解的高斯混合模型拟合效果差于 VBEM 算法,说明变分贝叶斯求取的高斯混合模型参数更适合描述不同天气类型的功率爬坡率概率密度函数。根据 VBEM 算法求取的不同天气类型下光伏功率爬坡率概率密度函数最佳混合度不同,晴天时最佳混合度为 2 阶,转折天气时为 3 阶。由于 EM 算法求解高斯混合模型时,不能求解混合模型的最佳混合度,这严重影响高斯混合模型的应用,而 VBEM 算法可以解决这一问题。

4 结论

本文研究了基于变分贝叶斯学习的光伏功率波动特性及爬坡率的概率模型,并用武汉某屋顶分布

式光伏电站的实际运行数据进行了验证。分析得到结论如下。

a. 运用 VBEM 算法求解高斯混合模型参数, 从而自适应地确定模型混合度, 解决了高斯混合模型拟合光伏功率爬坡率时混合度无法确定的问题。随着时间尺度的增加, 光伏功率爬坡率概率密度函数的混合度从 3 阶降为 2 阶, 且不同天气类型时光伏功率波动特性不同, 需要用不同阶数的高斯混合模型拟合。相对于 EM 算法的方法和单一分布, VBEM 算法的方法具有更高的拟合精度。

b. 随着时间尺度的增加, 光伏功率波动逐渐增加, 95% 置信区间下 5 min 时间尺度功率爬坡率约占装机容量的 8%, 30 min 时间尺度时约占 22.8%。不同天气类型下波动情况不同, 5 min 时间尺度下, 晴天时 95% 置信区间的爬坡率约占装机容量的 4.7%, 转折天气时约占 14.5%。

参考文献:

- [1] 国家电网公司. 光伏电站接入电网技术规定: Q/GDW 617—2011[S]. 北京: 中国电力出版社, 2011.
- [2] ALAM M J E, MUTTAQI K M, SUTANTO D. A novel approach for ramp-rate control of solar PV using energy storage to mitigate output fluctuations caused by cloud passing[J]. IEEE Transactions on Energy Conversion, 2014, 29(2): 507-518.
- [3] JAMALY M, BOSCH J L, KLEISSL J. Aggregate ramp rates of distributed photovoltaic systems in San Diego County[J]. IEEE Transactions on Sustainable Energy, 2013, 4(2): 519-526.
- [4] MURATA A, YAMAGUCHI H, OTANI K. A method of estimating the output fluctuation of many photovoltaic power generation systems dispersed in a wide area[J]. IEEE Transactions on Power & Energy, 2007, 127(5): 645-652.
- [5] MARCOS J, MARROYO L, LORENZO E, et al. From irradiance to output power fluctuations: the PV plant as a low pass filter [J]. Progress in Photovoltaics Research & Applications, 2011, 19(19): 505-510.
- [6] MARCOS J, MARROYO L, LORENZO E, et al. Power output fluctuations in large scale PV plants: one year observations with one second resolution and a derived analytic model[J]. Progress in Photovoltaics Research & Applications, 2011, 19(2): 218-227.
- [7] 陈逍潇, 张粒子, 杨萌, 等. 考虑光伏发电功率波动性的 AGC 备用容量分析方法[J]. 电力系统自动化, 2015, 39(22): 16-21. CHEN Xiaoxiao, ZHANG Lizi, YANG Meng, et al. A method for AGC reserve capacity analysis considering photovoltaic power fluctuation characteristics[J]. Automation of Electric Power Systems, 2015, 39(22): 16-21.
- [8] 吴振威, 蒋小平, 马会萌, 等. 多时间尺度的光伏出力波动特性研究[J]. 现代电力, 2014, 31(1): 58-61. WU Zhenwei, JIANG Xiaoping, MA Huimeng, et al. Study on fluctuations characteristics of photovoltaic power output in different time scales[J]. Modern Electric Power, 2014, 31(1): 58-61.
- [9] 崔杨, 杨海威, 李鸿博. 基于高斯混合模型的风电场群功率波动概率密度分布函数研究[J]. 电网技术, 2016, 40(4): 1107-1112. CUI Yang, YANG Haiwei, LI Hongbo. Probability density distribution function of wind power fluctuation of a wind farm group based on the Gaussian mixture model[J]. Power System Technology, 2016, 40(4): 1107-1112.
- [10] 李剑楠, 乔颖, 鲁宗相. 大规模风电多尺度出力波动性的统计建模研究[J]. 电力系统保护与控制, 2012, 40(19): 7-13. LI Jiannan, QIAO Ying, LU Zongxiang. Research on statistical modeling of large-scale wind farms output fluctuations in different special and temporal scales[J]. Power System Protection and Control, 2012, 40(19): 7-13.
- [11] 杨茂, 董骏城. 基于混合分布模型的风电功率波动特性研究[J]. 中国电机工程学报, 2016, 36(增刊 1): 69-78. YANG Mao, DONG Juncheng. Study on characteristics of wind power fluctuation based on mixed distribution model[J]. Proceedings of the Chinese Society of Electrical Engineering, 2016, 36(Supplement 1): 69-78.
- [12] VAN HAAREN R, MORJARIA M, FTTHENAKIS V. Empirical assessment of short-term variability from utility-scale solar PV plants[J]. Progress in Photovoltaics Research & Applications, 2014, 22(5): 548-559.
- [13] BISHOP C M, NASRABADI N M. Pattern recognition and machine learning[M]. New York, USA: Springer, 2007: 10.
- [14] ATTIAS H. A variational Bayesian framework for graphical models[J]. Advances in Neural Information Processing Systems, 2000, 12: 209-215.
- [15] 徐定杰, 沈忱, 沈锋. 混合高斯分布的变分贝叶斯学习参数估计[J]. 上海交通大学学报, 2013, 47(7): 1119-1125. XU Dingjie, SHEN Chen, SHEN Feng. Variational Bayesian learning for parameter estimation of mixture of Gaussians[J]. Journal of Shanghai Jiaotong University, 2013, 47(7): 1119-1125.
- [16] 谢晓洁. 普通统计学[M]. 北京: 北京大学出版社, 2004: 29.
- [17] 吕晓禄, 梁军, 负志皓, 等. 风电场出力的纵向时刻概率分布特性[J]. 电力自动化设备, 2014, 34(5): 40-45. LÜ Xiaolu, LIANG Jun, YUN Zhihao, et al. Longitudinal instant probability distribution of wind farm output power[J]. Electric Power Automation Equipment, 2014, 34(5): 40-45.
- [18] 卞海红, 郑维高, 林章岁, 等. 集群风电场出力统计指标建模与应用[J]. 电力自动化设备, 2015, 35(12): 21-27. BIAN Haihong, ZHENG Weigao, LIN Zhangsui, et al. Modeling and application of statistical indicators for wind farm cluster[J]. Electric Power Automation Equipment, 2015, 35(12): 21-27.
- [19] 孙英云, 侯建兰, 何国庆, 等. 一种考虑备用影响的光伏发电碳减排量计算方法[J]. 电力系统自动化, 2014, 38(17): 177-182. SUN Yingyun, HOU Jianlan, HE Guoqing, et al. A carbon emission reduction calculation method of photovoltaic considering influence of reserve[J]. Automation of Electric Power Systems, 2014, 38(17): 177-182.

作者简介:



李芬

李芬(1984—), 女, 湖北天门人, 副教授, 博士, 研究方向为太阳能资源评估、光伏发电并网控制及功率预报技术、电力电子技术, 在清洁能源分布式发电技术中应用等 (E-mail: beckyhust@163.com);

李春阳(1989—), 男, 河南周口人, 硕士研究生, 主要研究方向为光伏功率波动对电网的影响 (E-mail: 851439317@qq.com)。

(下转第 122 页 continued on page 122)

Short circuit current calculation and fault analysis method of DFIG wind-farm groups

YIN Jun, LI Yanbin, XIONG Junhua, YAN Zhekun

(School of Electric Power, North China University of Water Resources and Electric Power, Zhengzhou 450045, China)

Abstract: The transient characteristics of different wind farms in wind-farm groups are quite different and the coupling relation between each wind farm and grid is quite strong, which make the failure analysis of the grid with wind-farm groups more complicate. An equivalent inverter model is established based on the external characteristics of its inputs and outputs, a DFIG transient model considering the influence of inverter control is provided, the influence of LVRT control strategy on the short circuit current is analyzed, and a short circuit current calculation model of DFIG wind unit is built. The interaction among wind farms during a fault is analyzed and a method for calculating the short circuit current of DFIG wind-farm groups is proposed. A physical experimental platform with DFIG wind unit controller is established based on RTDS to verify the correctness of the proposed short circuit current calculation method. The fault analysis method of the grid with DFIG wind-farm groups is discussed and analyzed.

Key words: DFIG wind-farm groups; wind farms; low-voltage ride-through; short circuit currents; failure analysis

(上接第 104 页 continued from page 104)

Photovoltaic output fluctuation characteristics research based on variational Bayesian learning

LI Fen¹, LI Chunyang¹, YAN Quanquan², ZHAO Jinbin¹, DUAN Shanxu³

(1. College of Electrical Engineering, Shanghai University of Electric Power, Shanghai 200090, China;

2. Maintenance Company of SMEPC, Shanghai 200063, China; 3. State Key Laboratory of Advanced Electromagnetic Engineering and Technology, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract: The fluctuation of PV (PhotoVoltaic) output influences the stable operation of power system significantly. The ramp rate of PV output is analyzed, its Gaussian mixture model is built, and the variational Bayesian learning method is applied to estimate the model parameters. The test based on the massive measured data of an PV station shows that, the variational Bayesian learning method has better fitting effect than the single distribution method and EM algorithm method in researching the characteristics of PV output fluctuation for different time scales and weather types.

Key words: PV output fluctuation; variational Bayesian learning; Gaussian mixture model; ramp rate