

基于改进 K -means 聚类算法的风电场动态等值

吴红斌¹, 何 叶¹, 赵 波², 李 鹏², 季 宇³, 吴 鸣³

(1. 合肥工业大学电气与自动化工程学院, 合肥 230009; 2. 浙江省电力公司电力科学研究院, 杭州 310014;
3. 中国电力科学研究院, 北京 100192)

摘 要: 提出一种基于改进 K -means 聚类算法的动态等值建模方法, 将风速、输出功率及转差率作为聚类指标对风电机组进行动态分群, 然后将群内的风电机组分别进行等值, 可实现风电场的动态等值多机表征。利用算例系统进行验证, 仿真结果表明该文中提出的等值方法能更准确地反映双馈机组风电场并网点的动态特性。

关键词: 风电机组; 聚类算法; 动态模型; 等值研究

中图分类号: TM743

文献标识码: A

0 引 言

风电场并网接入必将对电网的动、静态稳定性产生影响。在研究过程中, 若搭建每台风电机组及其控制系统的详细模型, 将会极大地增加计算规模和降低运算速度, 因此有必要对风电场进行等值建模分析研究。

风电场等值分析分为单机等值和多机等值。风力发电受风速扰动、尾流效应及地形特征等因素影响, 各风力发电机处于不同的运行点, 当运行状态相差甚远时, 等值成一台发电机则无法完全表征等值前的运行状态^[1]。文献[2]提出一种针对桨距角动作情况划分的动态等效建模方法, 文献[3]提出将风电机组的转速作为单一分群指标进行多机分群等值, 文献[4]提出基于风电机组暂态电压特性的聚类分群方法。这些方法虽然可克服传统单机等值方法的局限性, 但采用单一状态变量作为聚类指标, 实际等值效果也不甚理想。

为建立更为合理的风电场等值模型, 本文提出一种基于改进 K -means 聚类算法的动态等值建模方法。将风电机组的风速、输出功率及转差率作为聚类指标, 计算聚类中心点, 先采用改进的 K -means 聚类算法将风电机组进行分群, 然后将群内的风电

机组分别进行等值, 从而得到风电场的多机动态等值模型。最后利用 Matlab/Simulink 软件进行仿真, 验证该方法的可行性和有效性。

1 改进的 K -means 聚类分析

1.1 改进的 K -means 聚类算法

传统 K -means 聚类方法是通过迭代计算把数据对象划分到不同的群中, 同群内对象的相似度较高, 而不同群间对象的相似度较小^[5,6]。该算法需事先确定 k 个聚类个数和初始聚类中心, 初始参数选择的好坏, 将直接影响聚类结果的准确性。特别是对于 k 个初始聚类中心的选择, 传统上采用随机选取方式, 极易导致算法陷入局部最优^[7]。

因此, 针对传统 K -means 算法存在的不足, 在选择初始聚类中心上, 本文采用欧式距离作为变量之间的聚类目标函数, 为防止某些大值属性左右样本间的距离, 先对样本数据进行处理, 即减去均值, 除以标准差。

先计算 n 个样本 $X_1, X_2, X_3, \dots, X_n$ 两两之间的欧式距离, 计第 i 个样本为 $X_i = [x_{i1}, x_{i2}, \dots, x_{iq}]$, 则第 i 个样本和第 j 个样本之间的欧式距离为:

$$d_{ij} = \left[\sum_{k=1}^q (X_{ik} - X_{jk})^2 \right]^{1/2} \quad (1)$$

收稿日期: 2016-08-20

基金项目: 国家重点研发计划(2016YFB0900400; 2016YFB0900404); 国家电网公司科技项目(5211DS150015)

通信作者: 吴红斌(1972—), 男, 博士、教授、博士生导师, 主要从事分布式发电微电网技术方面的研究。hfwuhongbin@163.com

再筛去 m 个与其他样本之间欧式距离和最大的样本,然后从剩下的数据集中选出 2 个欧式距离最大的点作为 2 个不同群的聚类中心,接着从其余样本中找出已挑选出来的所有聚类中心的距离和最大的点为另一个聚类中心,每个聚类中心点表示对应的群中聚类指标的平均特征,直到选出 k 个聚类中心。这样得到的初始聚类中心不受样本的输入顺序影响。

1.2 K-means 聚类指标

在改进的 K-means 算法聚类过程中,设 $X_1, X_2, X_3, \dots, X_n$ 为 n 个样本, $z_1, z_2, z_3, \dots, z_k$ 为 k 个群的中心点集,簇间距离为所有群的中心距离之和 D ,即:

$$D = \sum_{1 \leq j \leq i \leq k} d(z_j, z_i) \quad (2)$$

设定簇间距离的加权平均值为类簇指标,类簇指标对 k 的取值十分敏感,当假设的分群数目等于或者高于真实的分群数目时,该指标会缓慢下降,而少于真实数目的分群时,该指标会急剧下降^[8]。

为评价聚类效果的好坏,结合群内距离和群间距离 2 种因素,引入轮廓系数 $S(i)$ 。

$$S(i) = \begin{cases} 1 - \frac{a(i)}{b(i)}, & a(i) < b(i) \\ 0, & a(i) = b(i) \\ \frac{b(i)}{a(i)} - 1, & a(i) > b(i) \end{cases} \quad (3)$$

式中, $a(i)$ ——第 i 个样本到所有它属于的群中其他点的距离; $b(i)$ ——第 i 个样本到所有非本身所在群的点的平均距离。

由式(3)可见,轮廓系数 $S(i)$ 的值介于 $[-1, 1]$, 越趋近 1 代表内聚度和分离度都相对较优,聚类效果越好,反之亦然。

2 风力发电机组的聚类

2.1 风力发电机组的仿真模型

当风以一定的速率吹向风力机时,风轮上产生的力矩驱使风轮转动,风力驱动的机械能可按空气动力学原理计算得到风力机的输出功率 P_w 为:

$$P_w = \frac{1}{2} \rho A C_p(\lambda, \beta) v_w^3 \quad (4)$$

式中, ρ ——空气密度; A ——叶片扫风横截面积; $C_p(\lambda, \beta)$ ——风能利用系数; λ ——叶尖速比;

β ——叶片桨距角; v_w ——风速。

双馈风力发电机组风轮将风能转变为转动能量,经过齿轮箱增速驱动发电机,可由式(5)所示的微分方程^[9,10]表示:

$$\frac{dT_r}{dt} = \frac{1}{T_d}(T_m - T_r) \quad (5)$$

$$\frac{ds}{dt} = \frac{1}{T_j}(T_e - T_r) \quad (6)$$

式中, s ——转差率; T_m ——风力机转矩; T_r ——转子机械转矩; T_e ——发电机电磁转矩; T_d ——传动系统时间常数; T_j ——发电机惯性时间常数。

双馈风力发电机的暂态模型用暂态电压源 $E' = E'_d + E'_q$ 和暂态电抗 X'_s 来表示,交换到同步旋转坐标系的状态方程为^[11,12]:

$$\begin{cases} \frac{dE'_d}{dt} = -\frac{1}{T'_0} \left(E'_d + \frac{X_m^2}{X_r + X_m} i_{qs} \right) + \\ \quad s\omega_e E'_q - \frac{X_m}{X_r + X_m} \omega_e u_{qr} \\ \frac{dE'_q}{dt} = -\frac{1}{T'_0} \left(E'_q - \frac{X_m^2}{X_r + X_m} i_{ds} \right) - \\ \quad s\omega_e E'_d + \frac{X_m}{X_r + X_m} \omega_e u_{dr} \\ E'_d = u_{ds} + X'_s i_{qs} \\ E'_q = u_{qs} - X'_s i_{ds} \end{cases} \quad (7)$$

式中, X_s ——定子电抗; X_m ——励磁电抗; X_r ——转子电抗; T'_0 ——暂态时间常数; ω_e ——电网速度; i_s ——定子电流, $i_s = i_{ds} + j i_{qs}$; u_s ——机端电压, $u_s = u_{ds} + j u_{qs}$; i_r ——转子电流, $i_r = i_{dr} + j i_{qr}$; u_r ——转子电压, $u_r = u_{dr} + j u_{qr}$ 。

2.2 风力机聚类指标的选取

当风电场的风速发生变化时,风电机组的控制系统发出变桨距指令,通过变桨距系统改变风轮叶片桨距角,最大限度捕获风能,控制输出功率。

当风力发电机组系统侧发生短路故障时,发电机端电压的变化将引起电磁转矩的变化。在不平衡转矩作用下,发电机的转速增加,导致发电机等值阻抗减小,功率因数也会下降,因此异步发电机转差率可以表征风电机组的运行特性。

考虑到故障期间风电机组的运行特性和异步发电机转差率有关,风电场不同区域的风速发生变化影响控制系统。故选用风速、输出功率、转差率作为分群的指标。

2.3 风力机的分群聚类

因 K -means 算法具有一定的随机性,所以针对每个 k 值,重复执行多次,并计算轮廓系数 $S(i)$,取平均值作为最终评价标准。

基于 K -means 算法的风力机分群聚类步骤如下:

- 1) 收集风电场参数,将风速、输出功率和转差率作为样本数据进行标准化处理;
- 2) 计算样本间欧式距离及类簇指标;
- 3) 筛掉距离最大的对象,从剩余样本中选取 k 个初始聚类中心点;
- 4) 根据样本与聚类中心的相似度,分别将样本分配给与其距离最相近的群,迭代计算。
- 5) 重复步骤 3) 和步骤 4),直至所有样本都不能分配,目标函数小于算法允许误差 ε 并不在变化为止;
- 6) 计算轮廓系数,若轮廓系数 $S(i)$ 不能满足条件,重新选取 k 个初始聚类中心点,直到 $S(i)$ 满足条件。若所有样本轮廓系数 $S(i)$ 都不满足条件,根据类簇指标,重新对 k 取值。

3 风电场的动态等值

3.1 发电机参数

$$\begin{cases} S_{eq} = mS, P_{eq} = \sum_{i=1}^m P_i, x_{m-eq} = \frac{x_m}{m} \\ x_{1-eq} = \frac{x_1}{m}, x_{2-eq} = \frac{x_2}{m}, r_{1-eq} = \frac{r_1}{m}, r_{2-eq} = \frac{r_2}{m} \end{cases} \quad (8)$$

式中, S 、 P ——风力机的容量和有功功率; m ——等值前同群的风力机数量; x_m ——发电机励磁电抗; x_1 、 x_2 ——定子电抗和转子电抗; r_1 、 r_2 ——定子电阻和转子电阻; 下标 eq 代表等值参数。

3.2 轴系参数

$$H_{g-eq} = \sum_{i=1}^m H_g, H_{l-eq} = \sum_{i=1}^m H_l, K_{eq} = \sum_{i=1}^m K_i, D_{eq} = \sum_{i=1}^m D_i \quad (9)$$

式中, H_g 、 H_l ——发电机和风力机的惯性时间常数; K 、 D ——轴系刚度系数和轴系阻尼系数。

3.3 电容器和变压器参数

$$C_{eq} = \sum_{i=1}^m C_i, S_{T-eq} = mS_T, Z_{T-eq} = \frac{Z_T}{m} \quad (10)$$

式中, C ——机端补偿电容; S_T ——机端变压器容

量; Z_T ——机端变压器阻抗。

3.4 集电系统参数

根据等值损耗功率法,等效风电机组接入点电压等于等效前同群中所有风电机组接入点电压的加权平均值,其权重为风电机组的输出功率,可得到等值电缆的阻抗^[13] Z_{eq} 为:

$$Z_{eq} = \frac{\sum_{i=1}^m (P_{zi}^2 Z_{li})}{P_{zs}^2} \quad (11)$$

式中, Z_{li} ——第 i 台机组的线路阻抗; P_{zi} ——流过阻抗 Z_{li} 的总功率; P_{zs} ——流过等值阻抗 Z_{eq} 的总功率。

3.5 等值模型误差分析

为了对风电场多机等值模型进行误差分析,以风电场详细模型的仿真结果作为参考,采用均方差定义风电场等值模型的有功功率、无功功率误差评价指标 E 为:

$$E = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{z_i - z_{fi}}{z_i} \right)^2} \quad (12)$$

式中, z_i 、 z_{fi} ——分别为风电场详细模型、等值模型在风电场出口的电气量; n ——采样点数。

4 算例分析

4.1 算例介绍

利用 Matlab/Simulink 仿真平台,搭建由 30 台额定功率为 1.5 MW 的双馈式风力发电机组成的详细模型如图 1 所示。

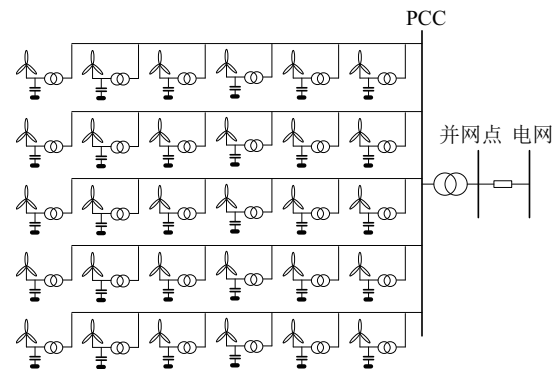


图1 风电场系统单线图

Fig. 1 Single line diagram of wind farm

风电机组机端电压为 690 V,由一机一变的单元接线方式就地升压至 35 kV,场内每 6 台变压器

由架空线路连接,输送至 35 kV/220 kV 变电站并输送至外部电网。双馈风力发电机、变压器主要参数参见文献[14],详细模型的各台参数如表 1 所示。

表 1 风电机组初始状态参数
Table 1 Initial state parameters of wind turbines

风力机 编号	$v_w/$ $\text{m}\cdot\text{s}^{-1}$	$P_w/$ kW	s	风力机 编号	$v_w/$ $\text{m}\cdot\text{s}^{-1}$	$P_w/$ kW	s	风力机 编号	$v_w/$ $\text{m}\cdot\text{s}^{-1}$	$P_w/$ kW	s
1	12.32	1405.52	-0.2	11	9.49	778.17	-0.196	21	8.72	872.59	-0.11
2	8.22	705.97	-0.197	12	8.07	685.32	-0.086	22	7.66	558.52	-0.2
3	6.95	445.79	-0.2	13	6.99	458.32	-0.086	23	10.10	1006.38	-0.2
4	10.05	1066.75	-0.2	14	10.58	1205.7	-0.2	24	7.07	489.90	-0.085
5	9.59	970.58	-0.196	15	9.39	929.39	-0.2	25	8.59	870.59	-0.2
6	8.21	728.97	-0.085	16	10.95	1277.67	-0.2	26	9.21	928.90	-0.196
7	6.75	1185.30	-0.2	17	9.76	994.63	-0.2	27	7.25	460.61	-0.085
8	10.51	969.75	-0.2	18	8.59	742.36	-0.2	28	10.51	1185.6	-0.2
9	9.45	1174.03	-0.2	19	7.49	538.95	-0.11	29	8.46	869.72	-0.2
10	10.65	958.63	-0.2	20	11.35	1159.45	-0.2	30	10.65	1174.05	-0.2

注： v_w —风速； P_w —输出功率； s —转差率。

4.2 动态聚类结果分析

使用 K-means 算法,分别对风速、输出功率和转差率 3 个状态变量所形成的矩阵进行标准化处理。

在 K-means 算法聚类过程中,设定群间距离的加权平均值为类簇指标,类簇指标对 k 的取值十分敏感,当假设的分群数目等于或高于真实的分群数目时,该指标会缓慢下降,而少于真实数目的分群时,该指标会急剧下降^[8]。图 2 是 k 从 2 到 6 时类簇指标的变化曲线,从图 2 可知,当 k 取值 5 时,类簇指标的变化趋势最明显,故 k 的最佳取值为 5。

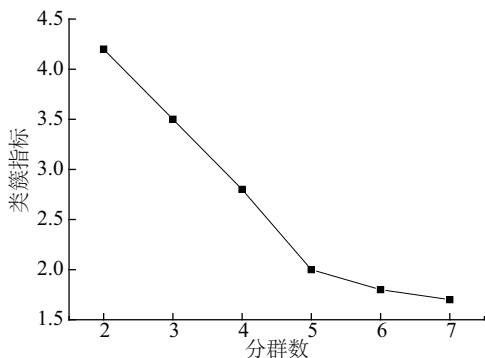


图 2 类簇指标变化曲线
Fig. 2 Varying curve of cluster index

此时系统的平均轮廓值如图 3 所示,随着迭代次数的增加,轮廓值越趋近 1 说明聚类效果越好,

进一步验证了本例取 5 个分群时效果最佳。

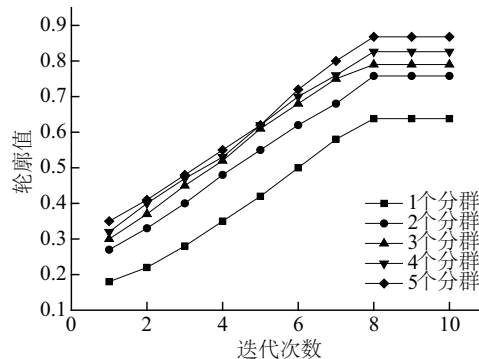


图 3 轮廓值的变化曲线
Fig. 3 Varying curve of contour value

根据风力机分群的 3 个指标,对应的三维图如图 4 所示,此时风力机的分群结果如表 2 所示。

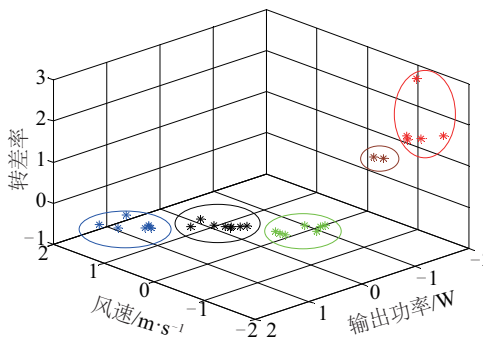


图 4 聚类结果分布图
Fig. 4 Distribution map of clustering results

表2 风电机组聚类结果

Table 2 Cluster results of wind turbines

聚类结果	风力机编号	聚类中心	轮廓值
1	3,7,13,24,27	(-1.4749,-1.5688,1.9264)	0.6381
2	2,6,12,18,21,25,29,30	(-1.0716,-1.1886,1.1328)	0.7585
3	4,5,9,11,15,17,23,26	(-0.4082,-0.3009,-0.504)	0.7903
4	19,22	(0.3863,10.3563,-0.5852)	0.8264
5	1,8,10,14,16,20,28	(1.2240,1.1801,-0.5340)	0.8681

4.3 多机等值模型的比较分析

为比较详细模型、单机等值模型、多机等值模型的效果,本文在风速变化、系统侧故障的2种情况下进行对比分析。

4.3.1 风速发生变化时的对比分析

设置风速在仿真时间3~4 s期间以2 m/s的速度持续下降后不再发生变化。图5给出了此时单机等值模型,多机等值模型和详细模型在风电场并网点的有功、无功功率变化曲线。

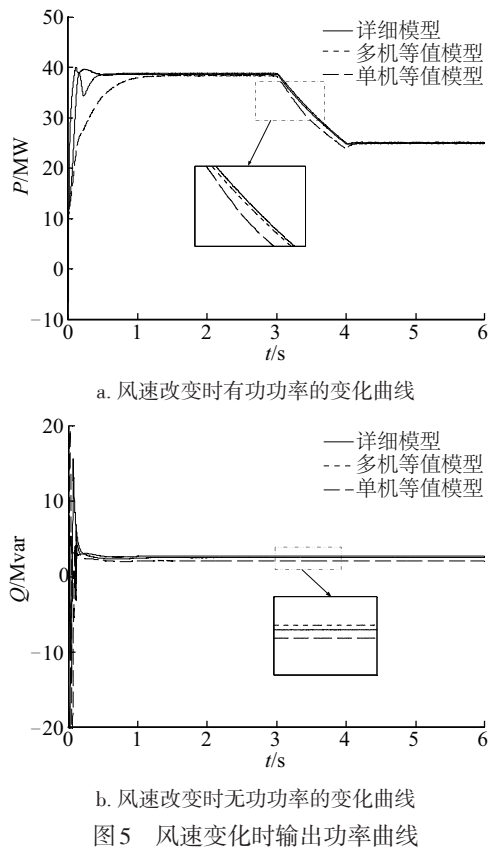


图5 风速变化时输出功率曲线

根据式(12)分别计算此时单机等值模型、多机等值模型的有功功率和无功功率误差指标,结果如

表3所示。

表3 等值模型误差指标

Table 3 Error indices of equivalent model

评价指标	有功功率误差/%	无功功率误差/%
单机模型	0.86	0.45
多机模型	0.48	0.33

4.3.2 系统侧故障时的对比分析

设置风电场在2 s时并网点发生三相短路故障,2.1 s后故障切除。图6给出了此时单机等值模型,多机等值模型和详细模型在风电场并网点的有功、无功功率变化曲线。

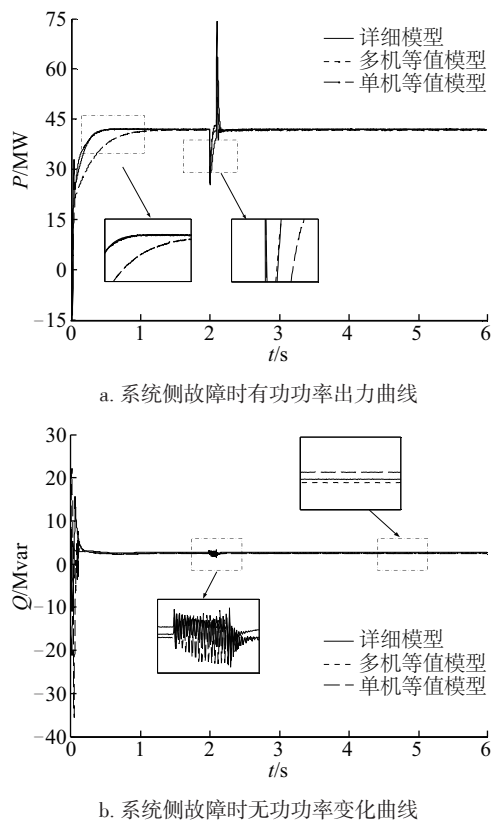


图6 系统侧故障时输出功率曲线

Fig. 6 Output power curve under system fault

由图 5、图 6 和表 3 可看出,相比于传统单机模型,本文中根据聚类分群结果建立的多机等值模型的功率出力曲线同详细模型相似度更高,误差更小,因此该多机等值模型能更准确地描述双馈机组风电场的实际情况,适用于大型风电场的等值建模。

5 结 论

针对双馈风力发电机组动态特性的差异性及等值问题的复杂性,提出一种基于改进 K -means 聚类算法的动态等效建模方法,该方法将状态参数作为聚类指标对风电机组进行动态分群,再分别进行群内等值。仿真结果表明,与传统单机等值建模相比,该多机等值模型克服了计算规模大、结果不够精确等缺点,能够准确地反映出双馈风力发电机组的动态运行特性,从而实现了双馈风力发电机组的动态等值,为风电机组并网分析提供了一种较好的建模方法。

[参考文献]

- [1] 郭志,陈洁,黄净,等.基于改进遗传KM聚类算法的风电场机群划分方法[J].可再生能源,2016,34(2):238—243.
- [1] Guo Zhi, ChenJie, Huang Jing, et al. Study on equivalent model of wind farms based on improved genetic K -means algorithm[J]. Renewable Energy Resources, 2016, 34(2): 238—243.
- [2] 米增强,苏勋文,余洋,等.双馈机组风电场动态等效模型研究[J].电力系统自动化,2010,34(17):72—77.
- [2] Mi Zengqiang, Su Xunwen, Yu Yang, et al. Study on dynamic equivalence model of wind farms with wind turbine driven doubly fed induction generator[J]. Automation of Electric Power Systems, 2010, 34(17): 72—77.
- [3] 朱乾龙,韩平平,丁明,等.基于聚类-判别分析的风电场概率等值建模研究[J].中国电机工程学报,2014,34(28):4770—4780.
- [3] Zhu Qianlong, Han Pingping, Ding Ming, et al. Probabilistic equivalent model for wind farms based on clustering-discriminant analysis[J]. Proceedings of the CSEE, 2014, 34(28): 4770—4780.
- [4] 徐玉琴,王娜.基于聚类分析的双馈机组风电场动态等值模型的研究[J].华北电力大学学报:自然科学版,2013,40(3):1—5.
- [4] Xu Yuqin, Wang Na. Study on dynamic equivalence of wind farms with DFIG based on clustering analysis[J]. Journal of North China Electric Power University: Natural Scitnce Edition, 2013, 40(3): 1—5.
- [5] 吴凤慧,成颖,郑彦宁,等. K -means算法研究综述[J].现代图书情报技术,2011,27(5):28—35.
- [5] Wu Suhui, Cheng Ying, Zheng Yanning, et al. Survey on K -means algorithm[J]. New Technology of Library and Information Service, 2011, 27(5): 28—35.
- [6] 王千,王成,冯振元,等. K -means聚类算法研究综述[J].电子设计工程,2012,20(7):21—24.
- [6] Wang Qian, Wang Cheng, Feng Zhenyuan, et al. Review of K -means clustering algorithm[J]. Electronic Design Engineering, 2012, 20(7): 21—24.
- [7] 连凤娜,吴锦林,唐琦.一种改进的 K -means聚类算法[J].电脑与信息技术,2008,16(1):38—40.
- [7] Lian Fengna, Wu Jinlin, Tang Qi. An improved algorithm of K -means[J]. Computer and Information Technology, 2008, 16(1): 38—40.
- [8] 拉贾拉曼,厄尔曼.大数据:互联网大规模数据挖掘与分布式处理[M].王斌,译.北京:人民邮电出版社,2012,187—189.
- [8] Rajarama A, Ullman J D. Mining of massive datasets [M]. Translation by Wang Bin. Beijing: Posts Telecom Press, 2012, 187—189.
- [9] 陈树勇,王聪,申洪,等.基于聚类算法的风电场动态等值[J].中国电机工程学报,2012,32(4):11—19.
- [9] Chen Shuyong, Wang Cong, Shen Hong, et al. Dynamic equivalence for wind farms based on clustering algorithm[J]. Proceedings of the CSEE, 2012, 32(4): 11—19.
- [10] 孙建锋,焦连伟,吴俊玲,等.风电场发电机动态等值问题的研究[J].电网技术,2004,28(7):58—61.
- [10] Sun Jianfeng, Jiao Lianwei, Wu Junling, et al. Research on multi machine dynamic aggregation in wind farm [J]. Power System Technology, 2004, 28(7): 58—61.
- [11] 李辉,王荷生,史旭阳,等.基于遗传算法的风电场等值模型的研究[J].电力系统保护与控制,2011,39(11):1—8.
- [11] Li Hui, Wang Hesheng, Shi Xuyang, et al. Study on equivalent model of wind farms based on genetic algorithm[J]. Power System Protection and Control,

- 2011, 39(11): 1—8.
- [12] 米增强, 苏勋文, 杨奇逊, 等. 风电场动态等值模型的多机表征方法[J]. 电工技术学报, 2010, 25(5): 162—169.
- [12] Mi Zengqiang, Su Xunwen, Yang Qixun, et al. Multi-machine representation method for dynamic equivalent model of wind farms[J]. Transactions of China Electrotechnical Society, 2010, 25(5): 162—169.
- [13] 苏勋文, 米增强, 王 毅. 风电场常用等值方法的适用性及其改进研究[J]. 电网技术, 2010, 34(6): 175—180.
- [13] Su Xunwen, Mi Zengqiang, Wang Yi. Applicability and improvement of common-used equivalent methods for wind farms[J]. Power System Technology, 2010, 34(6): 175—180.
- [14] 王 聪. 基于聚类算法的风电场动态等值问题研究[D]. 吉林: 东北电力大学, 2012.
- [14] Wang Cong. The dynamic equivalence for wind farms based on clustering algorithm[D]. Jilin: Northeast Electric Power University.

RESEARCH ON DYNAMIC EQUIVALENT OF WIND FARM BASED ON IMPROVED *K*-MEANS CLUSTERING ALGORITHM

Wu Hongbin¹, He Ye¹, Zhao Bo², Li Peng², Ji Yu³, Wu Ming³

(1. School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China;

2. Zhejiang Electric Power Corporation Research Institute, Hangzhou 310014, China;

3. China Electric Power Research Institute, Beijing 100192, China)

Abstract: In the study of grid connection of wind farm, computing scale and calculating speed will be greatly increased by simulating on each wind power generator, so it is necessary to study the equivalence modeling of wind farm. A dynamic equivalent modeling method based on improved *K*-means clustering algorithm is proposed, the method is based on three state parameters of the wind speed, output power and the slip as the clustering index to cluster wind power generator dynamically, and then the wind power generators in the same group are equivalent as a wind generator, in order to realize the dynamic equivalent multi-machine representation of the wind farm. By textting the example system to verify, the simulation results show that the proposed method is more accurate to describes the dynamic characteristics of the wind farm of the doubly fed induction generator at the grid connection point.

Keywords: wind turbines; clustering algorithm; dynamic models; equivalence research