

基于改进聚类算法的电力系统运行方式识别方法

陈典¹, 杨楠², 孙晓强², 王蒙², 陆润钊¹, 刘东浩¹

(1. 电网安全与节能国家重点实验室(中国电力科学研究院有限公司), 北京 100192;

2. 国家电网有限公司西北分部, 西安 710000)

摘要:对于电力系统计算分析,运行方式选取的准确性和合理性是决定计算质量的关键。随着“双碳”目标的提出,清洁能源接入比例将逐步提高,传统的人工选取方法不再适用,如何结合生产模拟得到的运行方式样本集自动提取典型运行方式是双碳目标下新型电力系统规划工作亟待解决的科学技术问题。本文首先提出双碳目标下电力系统运行方式提取需求分析;接着提出了基于K-means++算法和改进聚类有效性指标的自动方式提取算法;随后设计了基于人工和自动算法结合的方式提取模式;最后基于西北电网实际数据开展算例分析,从方式特征和安全性校核两方面与人工选取的运行方式对比体现本文算法的有效性和合理性,验证了本文算法在提升科学规划水平方面的贡献。

关键词:新型电力系统;聚类;运行方式提取

Identification Method of Power System Operation Mode Based on Improved Clustering Algorithm

CHEN Dian¹, YANG Nan², SUN Xiaoqiang², WANG Meng², LU Runzhao¹, LIU Donghao¹

(1. China Electric Power Research Institute State Key Laboratory of Power Grid Safety and

Energy Conservation, Beijing 100192, China; 2. Northwest Branch of the State Grid

Corporation of China, Xi'an 710000, China)

Abstract: For the calculation analysis of power system, the correctness and rationality of operation of operation mode are the key to determine calculation quality. With the proposal of *two carbon target*, the access ratio of clean energy will increase gradually and the traditional artificial selection method is not suitable any longer. How to combine the the operation mode sample set which is obtained by the production simulation to automatically extract the typical operation mode is the scientific and technological problem to be solved urgently of the planning work of new power system under the two carbon target. In this paper, the extraction need analysis of the operation mode of power system under two carbon target is proposed firstly; then the automatic extraction algorithm based on K-means++ algorithm and improved clustering validity index is proposed. After that, the extraction mode based on the combination of artificial and automatic algorithms is designed. Finally, the calculation analysis is developed based on the actual data of China Northwest Power Grid and compared with artificially selected operation mode from such two aspects as mode characteristic and safety so to reveal the effectiveness and rationality of the algorithm in this paper, and the contribution of the algorithm in this paper to the improvement of scientific planning level is verified.

Keywords: new power system; clustering; operation mode extraction

0 引言

随着我国能源战略“双碳”目标的提出,电力系统将从传统的同步机为主导的模式向新能源高比例接入的模式转变,对电力系统分析带来了新的挑

战^[1]。电力系统安全稳定计算分析是电力系统规划和运行工作的重要基础,新版《电力系统安全稳定导则》要求保证电力系统运行的稳定性,维持电力系统频率、电压的正常水平,保证系统有足够的静态稳定储备和有功功率、无功功率备用容量。《导

则》强调,合理的电网结构和电源结构是电力系统安全稳定运行的基础。在电力系统规划设计阶段,应统筹考虑,合理布局,系统运行方式的安排应注重电网结构和电源开机的合理性^[2]。

在电力系统安全稳定计算分析中,基础运行方式的选择是计算分析顺利开展和准确分析的重要前提,合理的运行方式选取能够充分发现和反映电网在实际运行中遇到的典型和极端安全稳定问题,所选方式要保障电网规划运行的适应性,能合理的反映电网的极端运行情况^[3-8]。

电力系统运行方式提取和机器学习算法在电力系统中的应用已有国内外学者开展研究和分析。文献[9]针对智能用电数据面临的数据量大的问题,利用 map-reduce 模型提升计算效率,并采取改进 K-means 算法进行聚类提取关键用户模式。文献[10]采用多种方法开展聚类分析,识别生成典型用电曲线。文献[11]提出结合降维技术的电力负荷曲线集成聚类算法。文献[12]采用划分聚类和层次聚类相结合的集成聚类算法,并对比了多种降维方法对负荷曲线聚类的影响。文献[13]提出基于 SPARK 的并行 K-means 算法辨识电力系统不良数据的新方法,能够有效提高状态估计的准确性,提升计算效率。文献[14]提到运行方式的生成难点在于多源数据的整合,文章基于模式识别和机器学习提出一种运行断面估计算法。文献[15]提出一种基于云模型和模糊聚类的电力负荷模式提取方法。文献[16]提出了数据驱动的电力系统运行方式及其形态变化分析方法,以甘肃电网实际算例验证了文中数据驱动方法的有效性。文献[17]利用聚类技术对 SCADA 系统的潜在攻击进行检测。文献[18]介绍了数据驱动方法在电池电量状态估计中的应用。文献[19]对负荷数据进行多维尺度分析降维后,输入高斯混合模型聚类算法中,实现大规模负荷数据集的分类。文献[20]提出了一种分层聚类技术选取典型工作日,捕捉参数的重要统计特征(如负荷和可再生能源可用性)。文献[21]提出一种基于粒子群优化的 K-means 马尔科夫-蒙特卡洛风电时间序列建模新方法。文献[22]综述了利用人工智能技术进行电力系统暂态响应分析和控制决策的方法。

综上所述,如何自动提取典型运行方式是新型电力系统规划供需平衡工作亟待解决的科学问题。为此,本文提出了基于聚类算法及改进聚类有效性指标的方式自动提取方法,通过与人工方式对比提前本文方法的有效性。本文结构如下:第1节提出基于 K-means++ 算法和改进聚类有效性指标的方式

提取算法;第2节提出了适应双碳目标下电力系统规划工程实际的的关键方式提取方法;第3节基于西北电网规划数据开展算例分析;第4节为结论。

1 基于 K-means++ 算法和改进最佳聚类有效性指标的方式提取算法

1.1 聚类算法比选

聚类算法是机器学习算法中的无监督算法^[23-24],通过预先制定的规则对数据进行迭代计算,最终形成优化聚集分类。聚类算法根据其不同的算法逻辑可以分为以下几种^[25-27]:

1) 基于层次。主要包含凝聚和分裂两种方式,根据不同的簇间距离计算划分为不同的类型;

2) 基于密度。密度聚集算法(DBSCAN)是一种典型的密度聚类算法,将聚类样本分为核心点、噪声点和边界点,并根据特定的规则开展样本聚类;

3) 基于模型。如层次型聚类算法(BIRCH),先叫每个样本点作为单独的聚类,自下而上不断合并,形成聚类划分;

4) 基于划分。典型的如 K-means 算法,通过聚类中心选择和样本距离划分,直观、高效的给出聚类结果,工程应用效果佳,便于理解。

新型电力系统运行方式提取要求分类清晰,方式具有明确的划分和特点,关键特征指标具有可解释性和物理意义。基于此,本文选择基于划分的 K-means 算法及其改进算法。传统 K-means 算法在应用于电力系统运行方式提取有两个核心问题,一是聚类中心选取问题,随机选取聚类中心可能会导致聚类结果不合理;二是难以确认合理的聚类数量选择。1.2 和 1.3 节将针对此问题提出解决方法。

1.2 K-means++ 算法

在选择聚类中心时,一方面,直观想象如果选取的聚类中心过于靠近,聚类得到的结果很可能出现不合理的情况,如图1所示;另一方面,作为电力系统运行方式的选取,如果单纯选择距离最远的节点可能导致选取到一些极端方式不具有代表性。

K-means++ 算法的思路是在选取聚类中心时,在选取第 $n+1$ 个聚类中心时:距离当前 n 个聚类中心越远的点会有更高的概率被选中,在本文算法中,其概率与其距离当前 n 个聚类中心中心点的距离成正比,即:

$$P(c_{n+1} = x_i) \propto \sqrt{\left(x_i - \frac{1}{n} \sum_{i=1}^n c_i\right)^T \left(x_i - \frac{1}{n} \sum_{i=1}^n c_i\right)} \quad (1)$$

式中: x_i 代表数据集中的第 i 个元素; c_i 代表第 i 个聚类中心。

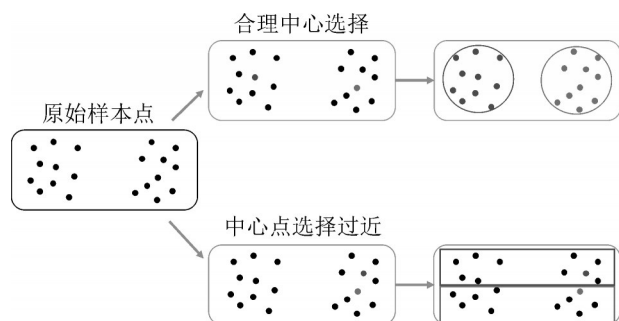


图1 中心点选择对比

Fig. 1 Comparison of center point selection

1.3 聚类有效性评价指标

确定合理的聚类数量是提高聚类质量的关键因素。常用的聚类有效性指标包含CH指标、DB指标、BWP指标、Sil指标等,这些指标都是基于簇内相似度和簇间相似度综合形成的评价指标^[28-30]。以Sil指标为例,其定义如下:

假设 n 个样本对象被聚类为 c 个簇,定义第 c 个簇的第 i 个样本的最小簇间距离 $b(j,i)$ 为该样本到其他每个簇中样本平均距离的最小值,簇内距离为该样本到第 $w(j,i)$ 簇中其他所有样本的平均距离,即

$$b(j,i) = \min_{1 \leq l \leq k, l \neq j} \left\{ \frac{1}{n_l} \sum_{p=1}^{n_l} \|x_p^{(l)} - x_i^{(j)}\|^2 \right\} \quad (2)$$

$$w(j,i) = \frac{1}{n_j - 1} \sum_{q=1, q \neq i}^{n_j} \|x_q^{(j)} - x_i^{(j)}\|^2 \quad (3)$$

式中: $x_q^{(j)}$ 表示第 j 簇的第 q 个样本; n_j 表示样本个数。

结合 $b(j,i)$ 和 $w(j,i)$, 定义簇间簇内划分指标 $Sil(j,i)$ 为

$$b'(j,i) = \min_{1 \leq l \leq k, l \neq j} \left\{ \|z^{(l)} - x_i^{(j)}\|^2 \right\} \quad (4)$$

式中 $z^{(l)}$ 代表簇 C_l 内样本的算术平均值。

使用 $b'(j,i)$ 替换 $b(j,i)$ 得到改进 Sil 指标,即:

$$\text{impSil}(j,i) = \frac{b'(j,i) - w(j,i)}{\max\{b'(j,i), w(j,i)\}} \quad (5)$$

2 电力系统规划工作工程实际的关键方式提取方法

传统人工选取运行方式的方法极度依赖人工经验,其弊端总结有以下几点:

1) 人工选取方式只能针对少量典型时段,不能做到全年场景全覆盖;

2) 极度依赖人工经验,难以适应电力系统运行特征由水火电联合运行向新能源为主体运行转变;

3) 容易产生季节、时段错位,不能反映空间和时序特性。

比较而言,自动聚类算法具有以下优点:

1) 可以综合考虑多不确定性边界条件下,多场景运行模拟得到的大量运行方式时间断面。综合考虑所有场景,根据选取特征找到相似方式之前的关联性,对海量方式进行聚类 and 特征提取;

2) 不依赖人工经验,通过数据特征驱动分类选取运行方式,尤其对于新进人员或对系统特性没有了解的人员能够给予很好的指导作用,分类的结果易于理解发掘高比例新能源接入后电力系统规划运行方式的特点和意义,发掘新的运行模式;

3) 选取的运行方式天然保持了各省之间的时序一致性,保证拼接方式取自同一个时间断面,避免出现理解偏差和时序错乱的现象;

4) 新能源接入后,电网特性复杂,本项目提出的自动算法能够通过聚类有效性指标给出选取类别的数量,并相应开展分类工作,提取出来的运行方式能够有效指导工作人员理解改变后的电力系统特性,并基于此完善形成最终计算的方式。

聚类算法虽然有上述优点,但受本身算法局限性不能完全符合实际工程应用的需求,本文将关键指标选取、方式特征分析和方式选取部分划分为人工执行,将方式聚类、最佳聚类数确定、方式特征提取部分由自动算法计算,提出了基于自动聚类算法和人工选取结合的关键方式生成提取方法。主要步骤如下:

步骤 1: 开展基于多种不确定性因素的多场景生产模拟计算分析,采集样本得到多时间断面的运行方式基础数据集;

步骤 2: 人工选取关键指标作为每个样本的多维特征向量;

步骤 3: 调用 K-means++ 算法对数据集开展聚类;

步骤 4: 计算聚类有效性指标,确定最佳聚类数即最佳方式数;

步骤 5: 根据自动算法划分聚类指标的季节和时段特性;

步骤 6: 人工开展聚类方式的特征分析;

步骤 7: 结合需求确定关注的关键指标,从每个聚类中提取相应数量的运行方式,完成关键方式的提取。具体流程如图 2 所示。

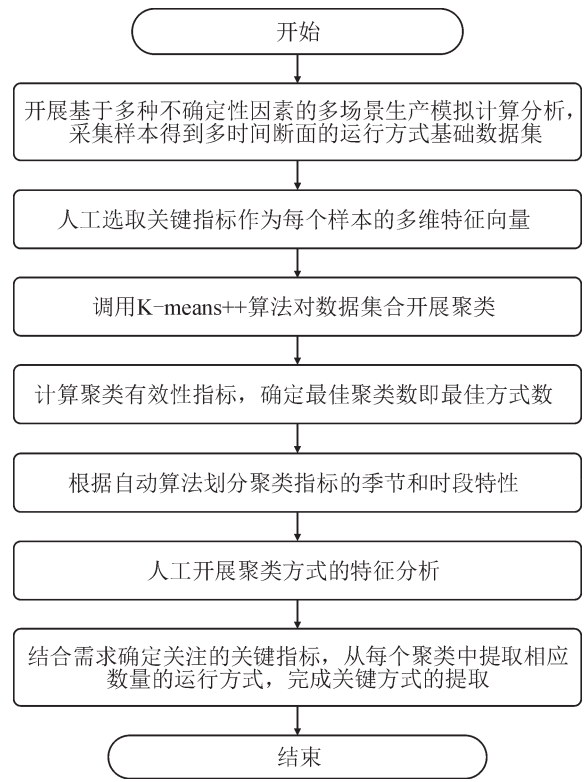


图2 关键方式提取流程
Fig. 2 Key mode extraction process

3 算例分析

本文以西北区域电网“十四五”规划水平年数据为基础开展算例分析。

3.1 计算边界条件

1) 全社会负荷预测

预计2025年,西北电网最大负荷预测值约1.58亿千瓦,电量预测值约11 190亿千瓦时,“十四五”年均增长率约6.69%。西北全社会负荷和用电量情况见表1。

表 1 西北全社会负荷和用电量情况						
Table 1 Load and electricity consumption of the whole society in northwest of China						
地区	全社会用电量/亿千瓦时			全社会最大负荷/万千瓦		
	2020年	2025年	增速	2020年	2025年	增速
西北	8 087	9 424	6.78%	11 352	15 792	6.83%

预计2025年,西北电网最大负荷预测值约1.58亿千瓦,电量预测值约11 190亿千瓦时,“十四五”年均增长率约6.69%。

2) 电源规划情况

西北地区电源发展遵循积极开发水电,大力发展风电、光伏、光热等新能源,清洁高效发展煤电,推进煤电大基地大通道建设,促进清洁有序发展,扩大西电东送规模,适度建设抽水蓄能电站、电化学储能电站,调整电源结构,提升系统灵活性的原则。

表2 西北电网电源“十四五”电源装机安排
Table 2 Northwest power grid power supply 145th power installation arrangement

万千瓦							
总计	水电	火电	风电	光伏	光热	生物质	气电
46 561	4 416	22 823	10 353	8 483	192	139	154

可以看到,“十四五”期间西北风电装机占比达22.2%,光伏装机占比达18.2%,新能源装机总占比达40%以上,构成高比例新能源电力系统。

3) 直流外送规模

“十四五”期间,西北新增4回特高压直流,部分已建直流提升至设计能力,2025年全部通道满送,预测西北外送直流2025年达约9 500万千瓦。

4) 计算软件平台

采用某电科院自主研发的源网荷一体化生产模拟计算软件(PSD-PEBL)支撑生产模拟计算分析和聚类样本集的生成。

3.2 考虑风、光、水电规模和特性不确定性的多场景生成

在电网规划工作中,应根据发电资源特点构建不同方案。针对西北电网的特点,本文主要考虑根据水电的水文特性和风、光不确定性规模构建基础数据集。

1) 水文特性曲线

水文特性一般分为枯水年、平水年和丰水年,水文特性的不同将显著影响水电的电量 and 调节空间,图3-4展示了不同省份不同水电机组的丰、枯、平水文特性曲线,可以看出每个电厂的水文特性曲线根据其自身所处的流域、环境等因素差别较大。

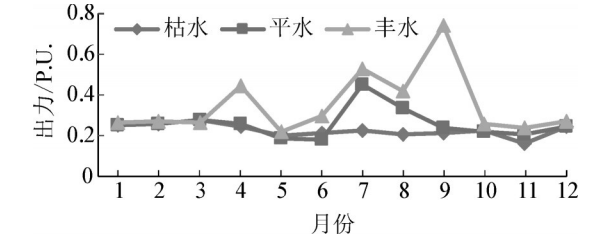


图3 A水电站各水文年逐月水电平均出力特性曲线
Fig. 3 Monthly average output characteristic curves of A Hydropower Station

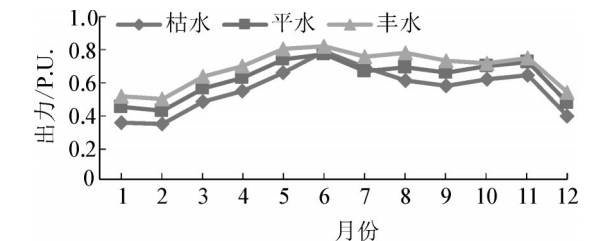


图4 B水电站各水文年逐月水电平均出力特性曲线
Fig. 4 Monthly average output characteristic curves of B Hydropower Station

开展灵敏度分析发现,一方面,水文特性会影响水电的调节特性和新能源消纳的空间,如表3所示;另一方面,水文特性的变化将显著影响省间断面的交换情况,以青海电网为例,丰水期外送有比较明显的提升,受入需求较平水年有比较明显的降低,枯期受水电小发的影响,枯期的电力受入需求有明显的增加,如图5所示。

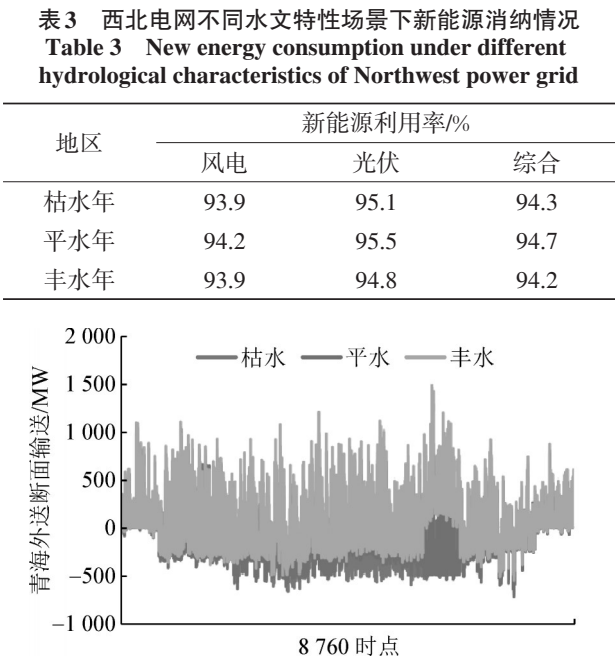


图5 青海外送断面功率8760曲线
Fig. 5 Qinghai external transmission section power 8760 curve

2) 新能源规模和出力不确定性

风、光新能源的不确定性也是影响生产模拟及规划电网方式的重要因素,西北电网不同新能源场景下新能源消纳情况见表4。通过表4可以看出,新能源大发场景下,新能源的电量占比有显著提高,但火电机组开机减少,新能源综合利用率降低。新能源小发场景下,新能源的电量占比降低,但火电机组开机增多,新能源综合利用率有所增长。

表4 西北电网不同新能源场景下新能源消纳情况
Table 4 New energy consumption in different new energy scenarios of Northwest power grid

地区	新能源利用率/%		
	风电	光伏	综合
新能源装机2亿千瓦	91.7	92.5	92.0
新能源装机1.8亿千瓦	94.2	95.5	94.7
新能源装机1.6亿千瓦	95.7	97.0	96.2

综上,考虑3种水电场景和3种新能源场景,两两组合形成9套场景,对每种场景分别开展生产

模拟计算形成总计78 840个数据样本形成数据集,作为后续运行方式提取的基础数据集。

3.3 关键方式生成实例分析

综合考虑运行实际选取的运行方式数量和程序计算效率等原因,选择聚类数的上限为30,采用K-means++算法,利用改进Sil指标评价聚类有效性选择最佳聚类,有效性指标结果如图6所示。

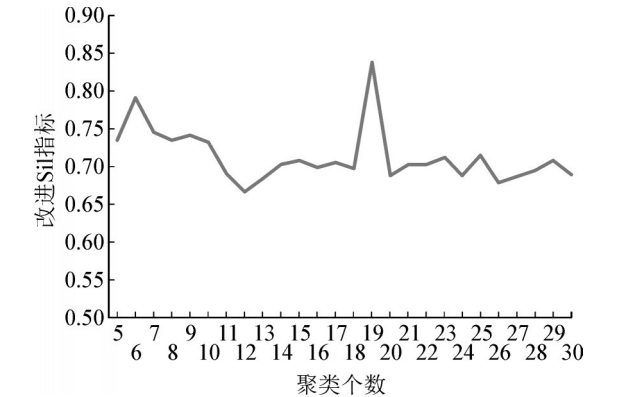


图6 聚类有效性指标计算结果
Fig. 6 Calculation results of cluster validity index

可以看到,聚类数为19的聚类有效性相对最佳,选择该组聚类结果作为提取方式分类的结果,整理聚类结果的时序特性可以得到各个聚类包含方式的季节和时序特性如图7所示,由图7可以看出,每个聚类具有鲜明的时段和季节特征,证明了本文聚类方法的有效性。

通过图7可以看出,按照传统人工经验按照时段和季节划分,每种聚类具有比较显著的季节和时段特征,主导因素主要是新能源出力、负荷特性和联络线交换。聚类的特性总结见表5。

表5 聚类季节和时段特性分类
Table 5 Clustering season and period characteristic classification

方式	春季	夏季	秋季	冬季
晚高峰大方式	4、7	11、12	4、6	12、13
凌晨早间小方式	14、15	9、10、17	5、14、17	1
午间腰方式	2、3、8	11、18	8、19	13、16

3.4 自动提取方式与人工方式对比

为了验证本文提出方式提取方法的有效性,结合“十四五”规划人工安排的方式开展方式对比,通过方式分析以及安全性校核对比验证本项目自动提取方式和人工安排方式的异同。

1) 提取方式特征对比

以西北冬季典型大方式为例,西北2025“十四五”规划冬大方式表见表6。通过表6对比可以看出,自动提取方式在负荷,直流外送方面基本与人

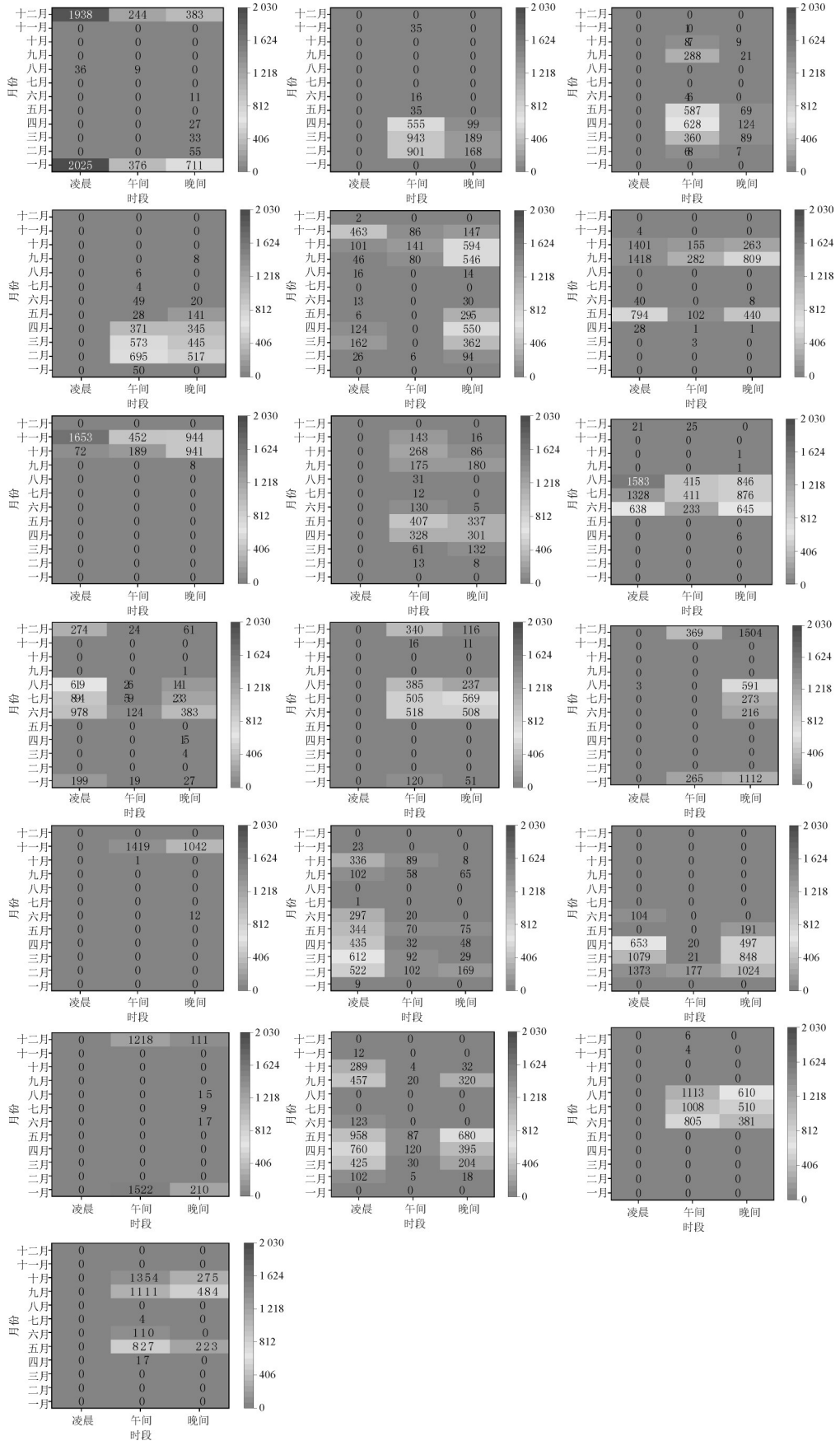


图7 聚类特征图
Fig. 7 Characteristic of clustering

表6 西北2025“十四五”规划冬大方式
Table 6 Northwest 2025 14th five-year operation mode for winter evening

项目	人工安排						自动选取					
	陕西	甘肃	青海	宁夏	新疆	西北	陕西	甘肃	青海	宁夏	新疆	西北
火电	4 627.7	3 775.1	299.4	4 045.5	8 300.5	21 048.2	4 725	3 229	1 137	2 695	5 922	17 708
水电	58.9	476.8	1185.7	41.2	332.6	2 095.2	258	477	766	14	520	2 035
风电	0	422.8	124.8	166.1	193.1	906.8	41	320	83	76	1 992	2 512
光伏	0	0	0	0	0	0	—	—	22	0	115	137
抽蓄	140	0	0	0	0	140	140	—	—	—	—	140
直流	1 210	1 600	920	2 200	2 800	8 730	1 211	1 600	860	2 200	2 800	8 669
外送	-400	400	-800	300	500	0	-78	-151	7	130	80	-12

工安排方式保持一致,风电出力比人工安排方式相对偏大,新疆的风电出力在自动提取方式中显著高于人工安排方式,出力率接近 0.5,而其他各省的风电出力率在 10% 左右,说明在同一时刻由于不同省之间地理位置,自然资源等特性不同,风电的出力特性具有差异。

传统人工选取方式时,一般安排各省按照同样的出力率安排新能源机组的出力,由于西北地域辽阔,不同省的资源特性各不相同,这种选取方式存在一定的缺陷,而自动提取方式能够保证选取方式的一致性。在区域交流联络方面可以看出,冬季大方式下,基本形成了各省的自平衡,与人工安排方式区别较大,体现了生产模拟计算中优先自平衡的特征。

2)提取方式安全性校核对比

以陕甘断面陕送甘重载问题为例,根据工作经验,陕西电网送出受制于同杆线路 N-2 故障后暂态功角问题,寻找反映该问题的典型方式应选择断面送出需求较大,常规机组开机较小的方式,根据所选需求从自动提取方式中选出了特定方式,针对陕西电网特性指标与人工选择的冬季小方式对比,如表 7 所示。

表7 选取方式关键指标对比
Table 7 Comparison of key indicators of selection mode

项目	万千瓦						
	火电	水电	风电	光伏	直流	外送	负荷
人工安排	3 353	140	834	0	980	600	2 778
自动提取	1 547	140	1 109	1 201	781	700	2 516

可以看出,自动提取的方式,负荷较小交流断面外送需求更大,同时常规机组开机显著降低,新能源机组出力占比显著提高。

计算陕西乾县至甘肃平凉同杆异名相 N-2 故

障,对比送端陕西电网机组对受端甘肃机组的功角稳定情况(见图 7)。可以看出,自动方式下,陕西同步机组开机占比显著降低,新能源出力较大,同步稳定性支撑较弱,发生同杆异名相短路故障后,机组功角摆开幅度更大,说明自动方式下暂态稳定水平较低,本文方法能够准确找到反映陕送甘断面重载时的稳定特性的方式,说明了合理性。

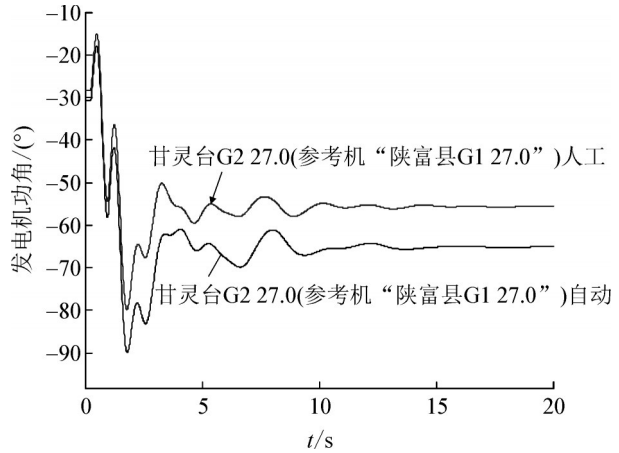


图7 乾县-平凉同杆异名相N-2故障机组功角差对比
Fig. 7 Comparison of power angle difference of N-2 fault units with the same bar and different phases

4 结语

随着“双碳”目标的提出,新能源为主体的电力系统逐渐取代传统同步机组为主导的电力系统,对电力系统的安全运行和仿真分析带来了新的挑战,电力系统运行方式方面,传统的人工选取典型方式的方法遇到瓶颈,多样化的运行方式需要新的自动化算法进行方式提取。本文提出了基于机器方式提取技术,开展算例分析从方式特征和与人工安排方式的对比体现本文算法的有效性。主要结论如下:

1)“双碳”目标下,新能源接入比例不断增高,传统的人工选取典型方式的方法不再适用,使用基于机器学习的聚类算法能够有效确定典型方式分

类,体现了“双碳”目标下电力系统运行的多样性及自动提取方式的必要性;

2)分析自动提取算法和人工安排方式的异同发现,人工安排方式存在时段拼接错位、负荷和新能源特性设置单一、难以反映各省资源和地理特性特点等问题,自动方法更真实的反映实际的运行状态和电力流向,具有优越性和合理性;

3)通过安全性校核分析发现,自动提取的方式能够真实反映实际的机组开机出力 and 电力流向情况,自动提取方式能够正确反映电力系统薄弱环节存在的安全稳定问题,证明了本文所提方法的有效性。

参考文献

- [1] 王康,张青蕾,王泽,等.高比例风电系统的爬坡备用需求评估[J].电网与清洁能源,2022,38(8):94-101.
WANG Kang, ZHANG Qinglei, WANG Ze, et al. Evaluation of ramping reserve requirement for high-proportion wind power systems[J]. Power System and Clean Energy, 2022, 38(8): 94-101.
- [2] 电力系统安全稳定导则:GB 38755—2019[S]. 2019.
Code on security and stability for power system:GB 38755—2019[S]. 2019.
- [3] HOU Qingchun, DU Ershun, ZHANG Ning, et al. Impact of high renewable penetration on the power system operation mode: a data-driven approach[J]. IEEE Transactions on Power Systems, 2020, 35(1): 731-741.
- [4] 吴继岩,张怀垠,马鑫,等.双碳量约束下主动配电网网架规划方法设计[J].电网与清洁能源,2022,38(4):117-123.
WU Jiyan, ZHANG Huaiyin, MA Xin, et al. Design of Active distribution network planning method under double carbon constraints[J]. Power System and Clean Energy, 2022, 38(4): 117-123.
- [5] 矫真,王兆军,李骁,等.新能源汽车充电站浪涌侵入波防护分析[J].电瓷避雷器,2022(1):103-106.
JIAO Zhen, WANG Zhaojun, LI Xiao, et al. Lightning surge protection for the new energy vehicle charging station[J]. Insulators and Surge Arresters, 2022(1): 103-106.
- [6] 张亚迪,陈钊,卢杰.考虑组间间隙与体表场强的特高压直流线路进入等电位路径安全评估研究[J].高压电器,2022,58(10):164-172.
ZHANG Yadi, CHEN Zhao, LU Jie. Study on safety assessment of entering isopotential path of UHVDC line considering combined clearance and human surface field strength[J]. High Voltage Apparatus, 2022, 58(10): 164-172.
- [7] 黎寿涛,夏成军,钟明明,等.基于故障链聚类算法的电网关键线路辨识[J].电力工程技术,2022,41(1):84-92.
LI Shoutao, XIA Chengjun, ZHONG Mingming, et al. Critical line identification of power grid based on fault chain clustering algorithm[J]. Electric Power Engineering Technology, 2022, 41(1): 84-92.
- [8] 杨彬,徐伟杰,韩文,等.750 kV输电线路下地面金属附着物感应电压分析研究[J].电网与清洁能源,2023,39(1):52-57.
YANG Bin, XU Weijie, HAN Wen, et al. Analysis of the induced voltage of ground metal attachment under 750 kV transmission lines[J]. Power System and Clean Energy, 2023, 39(1): 52-57.
- [9] 赵莉,候兴哲,胡君,等.基于改进K-means算法的海量智能用电数据分析[J].电网技术,2014,38(10):2715-2720.
ZHAO Li, HOU Xingzhe, HU Jun, et al. Improved K-means algorithm based analysis on massive data of intelligent power utilization[J]. Power System Technology, 2014, 38(10): 2715-2720.
- [10] 王志勇,曹一家.电力客户负荷模式分析[J].电力系统及其自动化学报,2007,19(3):62-65.
WANG Zhiyong, CAO Yijia. Electric power system load profiles analysis[J]. Proceedings of the CSU-EPSC, 2007, 19(3): 62-65.
- [11] 张斌,庄池杰,胡军,等.结合降维技术的电力负荷曲线集成聚类算法[J].中国电机工程学报,2015,35(15):3741-3749.
ZHANG Bin, ZHUANG Chijie, HU Jun, et al. Ensemble clustering algorithm combined with dimension reduction techniques for power load profiles[J]. Proceedings of the CSEE, 2015, 35(15): 3741-3749.
- [12] 刘莉,王刚,翟登辉.K-means聚类算法在负荷曲线分类中的应用[J].电力系统保护与控制,2011,39(23):65-68.
LIU Li, WANG Gang, ZHAI Denghui. Application of K-means clustering algorithm in load curve classification[J]. Power System Protection and Control, 2011, 39(23): 65-68.
- [13] 孟建良,刘德超.一种基于Spark和聚类分析的辨识电力系统不良数据新方法[J].电力系统保护与控制,2016,44(3):85-91.
MENG Jianliang, LIU Dechao. A new method for identifying bad data of power system based on Spark and clustering analysis[J]. Power System Protection and Control, 2016, 44(3): 85-91.
- [14] 黄宇鹏.基于模式识别和机器学习的电网运行断面估计算法研究[D].北京:中国电力科学研究院,2019.
- [15] 宋易阳,李存斌,祁之强.基于云模型和模糊聚类的电力负荷模式提取方法[J].电网技术,2014,38(12):3378-3383.
SONG Yiyang, LI Cunbin, QI Zhiqiang. Extraction of power load patterns based on cloud model and fuzzy clustering[J]. Power System Technology, 2014, 38(12): 3378-3383.

- [16]侯庆春,杜尔顺,田旭,等.数据驱动的电力系统运行方式分析[J].中国电机工程学报,2021,41(1):1-12.
HOU Qingchun, DU Ershun, TIAN Xu, et al. Data-driven power system operation mode analysis[J]. Proceedings of the CSEE, 2021, 41(1): 1-12.
- [17]ALMALAWI A, FAHAD A, TARI Z, et al. An efficient data-driven clustering technique to detect attacks in SCADA systems[J]. IEEE Transactions on Information Forensics and Security, 2016, 11(5): 893-906.
- [18]LI Yue, CHATTOPADHYAY P, XIONG Sihan, et al. Dynamic data-driven and model-based recursive analysis for estimation of battery state-of-charge[J]. Applied Energy, 2016, 184: 266-275.
- [19]张美霞,李丽,杨秀,等.基于高斯混合模型聚类和多维尺度分析的负荷分类方法[J].电网技术,2020,44(11): 4283-4293.
ZHANG Meixia, LI Li, YANG Xiu, et al. A load classification method based on gaussian mixture model clustering and multi-dimensional scaling analysis[J]. Power System Technology, 2020, 44(11): 4283-4293.
- [20]LIU Yixian, SIOSHANSI R, CONEJO A J. Hierarchical clustering to find representative operating periods for capacity-expansion modeling[J]. IEEE Transactions on Power Systems, 2018, 33(3): 3029-3039.
- [21]黄越辉,曲凯,李驰,等.基于K-means MCMC算法的中长期风电时间序列建模方法研究[J].电网技术,2019, 43(7): 2469-2476.
HUANG Yuehui, QU Kai, LI Chi, et al. Research on modeling method of medium-and long-term wind power time series based on K-means MCMC algorithm[J]. Power System Technology, 2019, 43(7): 2469-2476.
- [22]汤奕,崔晗,李峰,等.人工智能在电力系统暂态问题中的应用综述[J].中国电机工程学报,2019,39(1):2-13.
TANG Yi, CUI Han, LI Feng, et al. Review on artificial intelligence in power system transient stability analysis[J]. Proceedings of the CSEE, 2019, 39(1): 2-13.
- [23]KAUSHIK S. An introduction to clustering and different methods of clustering [EB/OL]. (2016-11-03) [2020-05-20]. <https://www.analyticshya.com/blog/2016/11/an-introduction-to-clustering-and-different-methods-of-clustering/>.
- [24]LI Kehua, MA Zhenjun, ROBINSON D, et al. Identification of typical building daily electricity usage profiles using gaussian mixture model-based clustering and hierarchical clustering[J]. Applied Energy, 2018, 231: 331-342.
- [25]伍育红.聚类算法综述[J].计算机科学,2015,42(S1): 491-499.
WU Yuhong. General overview on clustering algorithms[J]. Computer Science, 2015, 42(S1): 491-499.
- [26]杨俊闯,赵超. K-means 聚类算法研究综述[J]. 计算机工程与应用,2019,55(23):7-14.
YANG Junchuang, ZHAO Chao. Survey on K-means clustering algorithm[J]. Computer Engineering and Applications, 2019, 55(23): 7-14.
- [27]章永来,周耀鉴.聚类算法综述[J]. 计算机应用,2019,39(7):1869-1882.
ZHANG Yonglai, ZHOU Yaojian. Review of clustering algorithms[J]. Journal of Computer Applications, 2019, 39(7): 1869-1882.
- [28]郭靖,侯苏. K-means 算法最佳聚类数评价指标研究[J]. 软件导刊,2017,16(11):5-8.
GUO Jing, HOU Su. Study on the index of determining the optimal clustering number of K-means algorithm[J]. Software Guide, 2017, 16(11): 5-8.
- [29]周世兵,徐振源,唐旭清. K-means 算法最佳聚类数确定方法[J]. 计算机应用,2010,30(8):1995-1998.
ZHOU Shibing, XU Zhenyuan, TANG Xuqing. Method for determining optimal number of clusters in K-means clustering algorithm[J]. Journal of Computer Applications, 2010, 30(8): 1995-1998.
- [30]王勇,唐靖,饶勤菲,等.高效率的 K-means 最佳聚类数确定算法[J]. 计算机应用,2014,34(5):1331-1335.
WANG Yong, TANG Jing, RAO Qinfei, et al. High efficient K-means algorithm for determining optimal number of clusters[J]. Journal of Computer Applications, 2014, 34(5): 1331-1335.

作者简介:

陈典(1990—),男,高级工程师,研究方向为电力系统安全稳定,电力系统供需平衡分析,电力市场等。E-mail: lhc7177@163.com。

杨楠(1984—),男,副高级工程师,研究方向为电力系统调度运行、电网规划、新能源消纳等。E-mail: yangn@nw.sgcc.com.cn。

孙晓强(1974—),男,正高级工程师,研究方向:大电网规划、能源规划及电网调控技术等。E-mail: sunxq@nw.sgcc.com.cn。

王蒙(1978—),男,副高级工程师,研究方向:电网规划、直流输电及短路电流控制技术。E-mail: wangmeng@nw.sgcc.com.cn。

陆润钊(1994—),男,硕士,工程师,研究方向为电力系统供需平衡分析、电力系统规划等。E-mail: lurunzhao@epri.sgcc.com.cn。

刘东浩(1995—),男,工程师,研究方向为电力系统供需平衡分析、电力系统规划等。E-mail: 695219776@163.com。