

# 基于多维特征融合的居民电力消费异质性模式识别研究

刘向向<sup>1</sup>,卢 婕<sup>1</sup>,周 琦<sup>2</sup>,赵文辉<sup>3</sup>,冯 颖<sup>1</sup>

(1. 国网江西省电力有限公司 供电服务管理中心,南昌 330001;2. 国网江西省电力有限公司,南昌 330077;3. 北京理工大学 管理与经济学院,北京 100081)

## The research of pattern recognition of heterogeneous residential power consumption based on multi-dimensional features

LIU Xiangxiang<sup>1</sup>, LU Jie<sup>1</sup>, ZHOU Qi<sup>2</sup>, ZHAO Wenhui<sup>3</sup>, FENG Yin<sup>1</sup>

(1. Power Supply Service Management Center, State Grid Jiangxi Electric Power Co., Ltd., Nanchang 330001, China;2. State Grid Jiangxi Electric Power Co., Ltd., Nanchang 330077, China;3. School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China)

**摘要:**基于海量数据的居民电力消费行为特征分析与模式识别研究,对电网负荷预测以及需求响应潜力挖掘等有着至关重要的作用,而居民电力消费大规模数据与多维特征的涌现,成为居民电力消费异质性模式识别的难点。研究基于大规模居民电力消费数据,首先采用电力负荷特征分解技术构建特征工程;其次通过因子分析对所构建的多维特征进行融合,并采用聚类算法对居民电力消费模式进行识别;最后以江西省居民电力消费数据为例,分别在居民家庭与台区维度进行实证分析,分别得到4种典型居民电力消费模式,可为电网公司个性化与差异化政策制定,进一步拓展服务的深度和广度提供科学支撑。

**关键词:**经验模态分解;特征工程;行为模式挖掘;多维特征融合;聚类算法

**Abstract:** Research of pattern recognition on resident power consumption behavior based on massive data plays a crucial role in power grid load prediction and demand response potential development. However, the emergence of large-scale data and multi-dimensional features of residents' electricity consumption has become a difficulty in identifying the heterogeneous pattern of residents' electricity consumption. Firstly, based on the large - scale residential electricity consumption data, the power load characteristic decomposition technology is used to construct the feature project. And then, the multi-dimensional features constructed are fused by factor analysis, and the clustering algorithm is adopted to identify the residential electricity consumption pattern. Finally, taking Jiangxi province residents' electricity consumption data as an example, an empirical analysis is made in both households and community level, and four typical residents' electricity consumption patterns have been obtained, which can provide scientific support for the grid companies to make tailored and differentiated policies and further expand the depth and breadth of services.

**Key words:**empirical mode decomposition;feature engineering;behavior pattern minin;multi-dimensional feature fusion;clustering algorithm

## 0 引言

近年来,随着居民家庭生活水平的整体提升,居民用电负荷快速增长,居民负荷已成为电网尖峰负荷的重要组成部分,对电网的调控能力和安全稳定运行提出了新的挑战。因此,有必要通过分析居民用户的用电行为,挖掘居民用户参与需求响应的潜力<sup>[1-2]</sup>。

智能电能表已在电网公司得到推广普及应用,用户与电网间通过智能电能表相连,可以通过智能电能表收集用户侧的海量电力消费数据<sup>[3-4]</sup>,从而

分析居民用户的用电行为,电力大数据也成为众多学者研究的热点<sup>[5-6]</sup>。

目前,对居民用户用电行为的分析通常采用模式识别、聚类分析等数据挖掘技术与方法。文献[7]提出一种基于聚类的用户用电行及其影响因素分析方法,对某行业典型用户进行了分析,但未对海量居民用电数据进行分析;文献[8]以传统行业为基础,采用聚类分析方法,得出与电网公司目录电价相似的结果,未考虑用户之间的差异;文献[9]通过采用云计算方式对居民用电行为进行聚类分析,但未涉及大规模用户细分策略;文献[10]提出一种基于负荷分解的居民用电行为特性分析方法,对某一典型智能小区用户夏季负荷进行分解分析,但结论不够精细,不具备普适性;文献[11]提出一种基于现有特征的用户用电行为优选策略分析方法,但未对多维特征进行分析。

收稿日期:2020-05-11;修回日期:2020-07-13

基金项目:国网江西省电力有限公司科技项目(52182018001D)

This work is supported by Science and Technology Project of State Grid Jiangxi Electric Power Co., Ltd.(No.52182018001D)

本文基于大规模居民电力消费数据,采用多维特征融合分析方法,对家庭电力消费特征进行分析与模式识别,并结合江西省居民用户实测数据,验证了方法的正确性和有效性。

## 1 分析架构

基于海量数据的居民电力消费行为特征分析与模式识别研究方法实现流程如图1所示。首先,对所采集的大规模居民电力消费数据进行数据预处理;其次,基于特征工程开展特征构建与遴选,针对时间序列数据构建多维特征,将数据属性转换为数据特征,数据属性代表了数据的维度,在数据建模时,如果仅对原始数据的所有属性进行学习,并不能很好的找到数据的潜在趋势与特征,而通过特征工程对数据进行预处理之后,算法模型能够减少受到噪声的干扰,更好的找出趋势与结构特征;最后,对于构建出的多维特征进行因子分析,获取共同因子对居民电力消费行为进行聚类,得到居民电力消费典型特征,进而对居民电力消费模式进行识别。

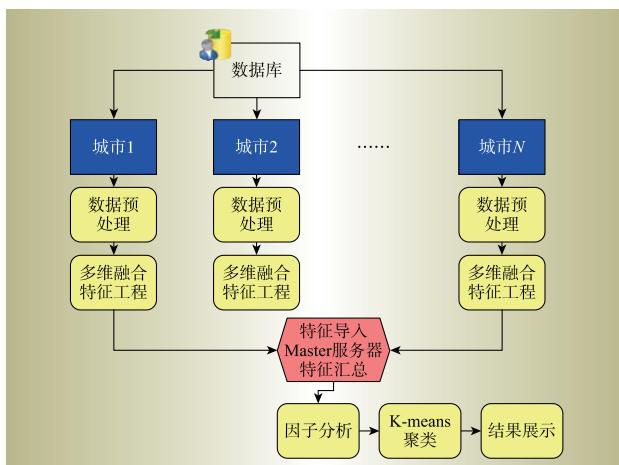


图1 基于多维特征的模式识别实现流程图

Fig. 1 Flow chart of pattern recognition based on multi-dimensional features

## 2 特征工程构建

特征工程是指通过一系列工程化的方式从原始数据中筛选出数据特征,以提升模型的训练(聚类)效果。特征工程包括特征构建、特征选择与降维环节。

考虑到进行数据插值有一定的偏差,同时由于样本数据量大,部分随机缺失数据的删除不会影响样本的总体代表性。因此,本文在数据清洗与预处理确定数据样本过程中,剔除了某些月份存在缺失值的用户,仅选择每月均有电量的用户作为聚类样本,并对数据进行标准化。

### 2.1 特征构建

基于时序数据特征挖掘方法与经验模态分解技

术(empirical mode decomposition, EMD)构建特征工程。经验模态分解依据数据自身的时间尺度特征来进行信号分解,无须预先设定任何基函数。经验模态分解可以将复杂居民用电波形分解为有限个本征模函数(intrinsic mode function, IMF),分解出来的有限个本征模函数包含了原始用电信号的不同时间尺度的局部特征,得到具有物理意义的频率。

经验模态分解将一个频率不规则的波化为多个单一频率的波与残波的和的形式。具体实现过程如下:

- (1) 找出原用电序列  $X(t)$  的极大值点,用三次样条插值函数拟合形成原数据的上包络线;
- (2) 找出原始用电序列极小值点,并将所有的极小值点通过三次样条插值函数拟合形成数据的下包络线;上包络线和下包络线的均值记作  $m_1$ ;
- (3) 将原用电时间序列数据  $X(t)$  减去  $m_1$ ,得到一个新的数据序列  $h_1$ ,

$$h_1 = X(t) - m_1 \quad (1)$$

- (4) 若  $h_1$  极值点个数大于 2,则返回第一步,开始新一轮的分解;否则分解结束。

本文中,将在原始用电负荷信号进行多层次分解后,应用EMD技术对所得到的各个本征模函数分量和原始振动信号及各自频谱,分别提取时域特征和频域特征,将提取的特征作为进一步聚类的特征单元,进而进行特征筛选与降维,用于模式聚类。

### 2.2 特征选择与降维

由于不同的特征对聚类效果的影响程度不同,需要选择一些有效特征,移除与问题相关性小的特征。本文采用方差法进行特征选择,通过计算各个特征的方差,保留方差大于阈值的特征,剔除方差接近于 0 的特征,即

$$\sigma^2 = \frac{\sum(X - \mu)^2}{N} \quad (2)$$

式中:  $\sigma^2$  为特征方差;  $X$  为各个用户特征值;  $\mu$  为该特征均值;  $N$  为用户数量。

针对以上分析流程,本文针对居民用电数据,在特征工程构建过程中,采用时序数据特征挖掘方法与经验模态分解技术,筛选出与居民用能模式高度相关、并具有代表性的特征,基于居民月度用电量,构建了用电时间,年度电量均值(mean),年度方差(std),偏度(skew),峰度(kurt),用电长、中、短期趋势的频率,周期以及拐点(拐点,拐点发生与否,上升与下降判断)等 42 个特征。进而在因子分析过程中,对所筛选出的 42 个特征进行分析,包括因子方差贡献率计算、因子得分以及综合得分计算等。

最后,对所得到的因子通过 K 均值聚类算法(K-means clustering algorithm)进行聚类。模式识别过程中,将大规模数据导入模型,根据模型训练参数对用户用电模式进行判断,对每一个用户的用电模

式设置标签,实现居民用电典型模式识别。

### 3 居民电力消费模式实证分析

将单个居民用户作为研究对象,采用多维特征融合分析方法,对居民电力消费模式进行识别。以微观家庭为单位的用能模式聚类分析,为每一类家庭的个性化用能方案制定、科学用能行为引导提供科学依据。本文采集样本为江西省 29 万居民客户近 3 年月度用电量数据。基于多维特征融合的方法,最终得到居民用户的用电模式可分为 4 类:冬夏均衡型、夏季偏好型,夏季敏感型、冬季敏感型,4 类的聚类结果如图 2 所示。

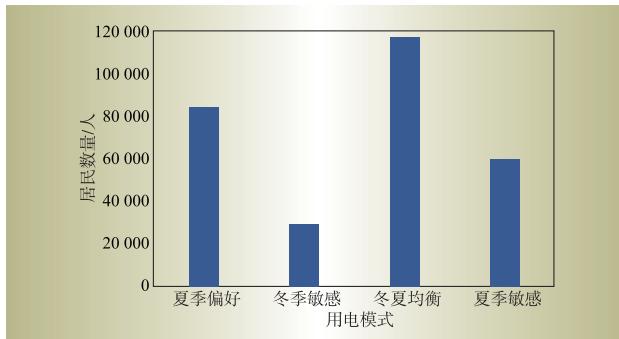


图 2 江西省 4 类典型用电模式居民数量

Fig. 2 The number of households in each typical power consumption patterns in Jiangxi province

4类典型用电模式中,冬夏均衡型、夏季偏好型、夏季敏感型、冬季敏感型占比分别为 40%、29%、21%、10%。

#### 3.1 冬夏均衡型

冬夏均衡型用电模式的典型特征为:用户四季用电温和波动、春秋季节用电平稳、冬夏用电呈现温和增长且用电峰值接近对等,该类用户多为普通刚性电力需求用户,95%置信区间下该类用户用电模式聚类结果如图 3 所示。

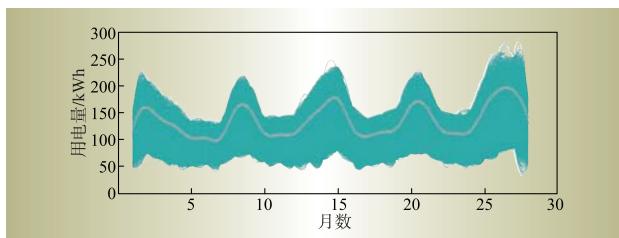


图 3 冬夏均衡型居民用电聚类结果

Fig. 3 Residential electricity clustering results of winter and summer balance mode

由图 3 可知,多数居民月均用电量在 50~240 kWh 区间波动,且四季波动稳定,波幅相对固定。基于波形分解方法,对该类型的用电模式进行长期、中期和短期趋势分解,得出波形分解图如图 4 所示。

由图 4 可知,从长期趋势来看,冬夏均衡型居民用电模式呈现长期波动规律,波长半径为 20 个月,

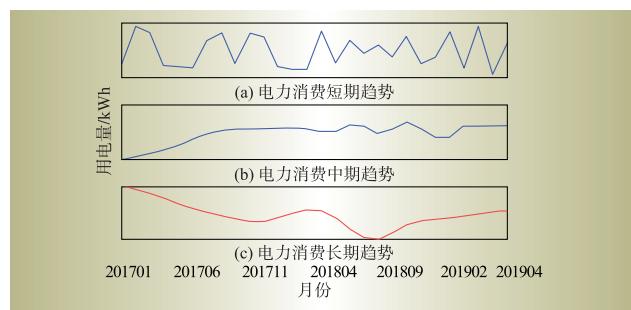


图 4 冬夏均衡型居民用电模式波形分解

Fig. 4 Electrical waveform decomposition of winter and summer balance residents

从中期趋势特征来看,呈现平稳的发展趋势;从短期趋势来看,该类用户的短期波动较大,是造成总体趋势波动的主要原因。总体来看,该类用户存在较小的长期趋势特征,未来将呈现缓慢的上升趋势。

#### 3.2 夏季偏好型

夏季偏好型用电模式的典型特征为:四季用电温和波动、夏季用电峰值略高于同年冬季用电峰值、呈现较明显的夏季用电偏好属性,该类用户表现出较高的电力消费意愿以适应高温天气。95%置信区间下该类用户用电模式聚类结果如图 5 所示。

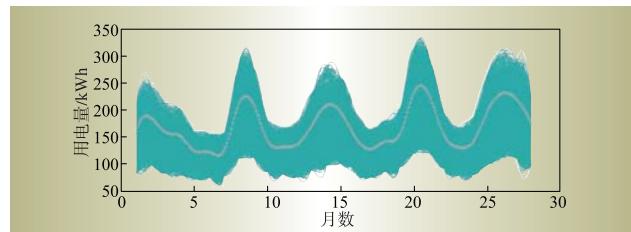


图 5 夏季偏好型居民用电聚类结果

Fig. 5 Residential electricity clustering results of summer preference mode

由图 5 可知,多数居民的用电量在 80~300 kWh 区间波动,冬季波幅相对固定,夏季有波形放大现象。基于波形分解方法,对该类型的用电模式进行长期、中期和短期趋势分解,得出波形分解图如图 6 所示。

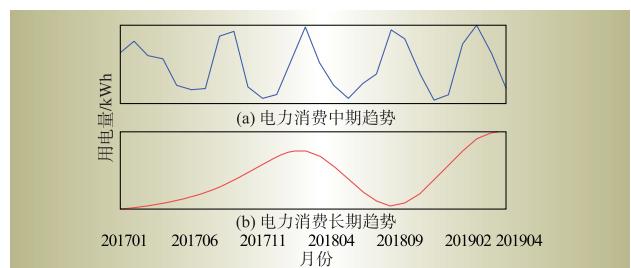


图 6 夏季偏好型居民用电模式波形分解

Fig. 6 Electrical waveform decomposition of summer preference residents

由图 6 可知,从长期趋势来看,夏季偏好型居民用电模式呈现长期波动规律,波长直径为 20 个月,2017 年 1 月至 2018 年 2 月呈现上升趋势,2018 年 2 月至 8 月,呈现下降趋势,2018 年 8 月至 2019 年 4 月

呈现上升趋势；从中期趋势特征来看，呈现稳定的正态型波动模式，周期为6个月，该类用户不存在短期趋势。总体来看，该类用户存在明显的波动趋势特征，波动较为规整，存在较小的长期趋势特征，呈现缓慢的上升趋势，且这种时间趋势在夏季被明显放大，未来仍将呈现缓慢的上升趋势。

### 3.3 夏季敏感型

夏季敏感型用电模式的典型特征为：夏季用电量波动巨大、峰值比谷值高出一倍以上、比冬季峰值高出50%以上，该类用户多为夏季温度舒适偏好型电力需求用户，对高温极度敏感。95%置信区间下该类用户用电模式聚类结果如图7所示。

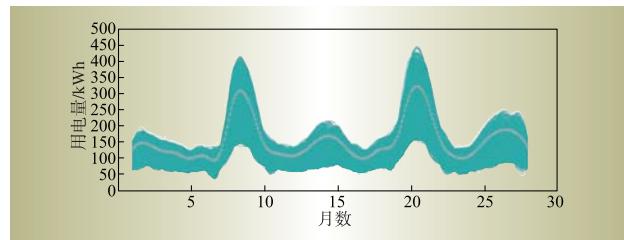


图7 夏季敏感型居民用电聚类结果

Fig. 7 Residential electricity clustering results of summer sensitive mode

由图7可知，多数居民的用电量在50~450 kWh区间波动，冬季波幅相对固定，夏季有明显波形畸形放大现象。基于波形分解方法，对该类型的用电模式进行长期、中期和短期趋势分解，得出波形分解图如图8所示。

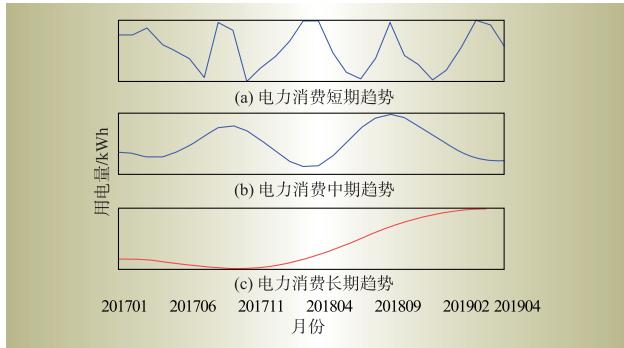


图8 夏季敏感型居民用电模式波形分解

Fig. 8 Electrical waveform decomposition of summer sensitive residents

由图8可知，从长期趋势来看，夏季敏感型居民用电模式呈现长期波动规律，波长半径为17个月，2017年1月至9月呈现下降趋势，2017年9月至2019年4月，呈现上升趋势；从中期趋势特征来看，呈现正态放大型波动模式，周期为12个月，该类用户存在显著的短期趋势，波幅剧烈。总体来看，该类用户存在明显的波动趋势特征，中期波动较大，是造成夏季显著波动的主要原因，在长期和中期趋势的叠加作用下，长期时间趋势较弱，上升趋势不明显，该类用户未来呈现平稳发展趋势。

### 3.4 冬季敏感型

冬季敏感型用电模式的典型特征为：和夏季敏感型用电模式相似，同样存在巨大的季节波动，冬季峰值比夏季峰值高出50%以上，该类用户多为冬季采暖偏好型电力需求用户，对低温极度敏感。95%置信区间下该类用户用电模式聚类结果如图9所示。

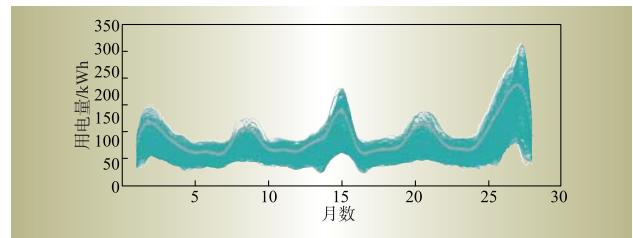


图9 冬季敏感型居民用电聚类结果

Fig. 9 Residential electricity clustering results of winter sensitive mode

由图9可知，多数居民的用电量在50~200 kWh区间波动，夏季波幅相对固定，冬季有明显波形畸形放大现象。基于波形分解方法，对该类型的用电模式进行长期、中期和短期趋势分解，得出波形分解图如图10所示。

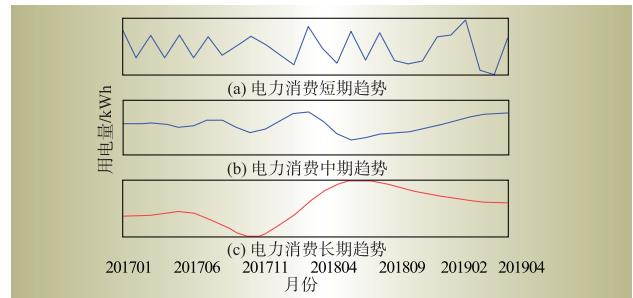


图10 冬季敏感型居民用电模式波形分解

Fig. 10 Electrical waveform decomposition of winter sensitive residents

由图10可知，从长期趋势来看，冬季敏感型居民用电模式波形不规整，不呈现周期性，从中期趋势特征来看，呈现平台波动模式，波幅较小，该类用户存在显著的短期趋势，波幅剧烈。总体来看，在3类波的叠加作用下，该类用户存在明显的波动趋势特征，短期波动在冬季放大，是造成冬季显著波动的主要原因；在长期和中期趋势的叠加作用下，该类用户未来呈现平稳发展趋势。

综上分析，4类典型居民用电模式如图11所示。

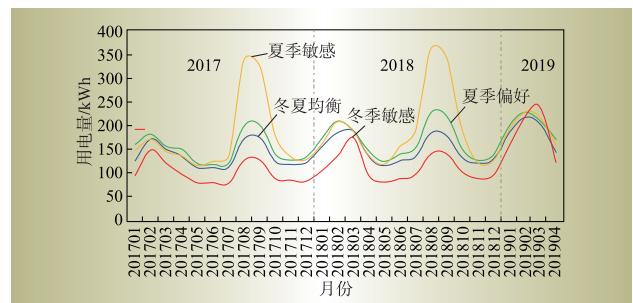


图11 江西省4类典型居民用电模式

Fig. 11 Four types of typical residential electricity consumption patterns in Jiangxi province

由图11可知,冬夏均衡型的冬季和夏季月份的用电量相较于春秋季节有一定的上升,幅值相差小,波形比较平缓;夏季偏好型的冬季和夏季月份用电量相较于春秋季节有一定的上升,但夏季峰值略高于冬季;夏季敏感型在夏季月份的用电量明显高于其他季节,夏季的峰度值也明显大于其他用电模式类型;冬季敏感型在冬季的用电量明显高于其他季节,冬季峰值明显大于其他用电模式类型。

#### 4 台区用电模式实证分析

以台区为单位,分析江西省各地区的用电模式,识别区域性的用电模式差异。以中观层面的台区为分析单元进行用能模式聚类分析,有利于电网管理侧从中观角度对整个省份/城市总体用电模式和典型类别有更宏观的把控。基于电力大数据的用电模式分析结果,台区用电模式可分为4类:平稳型、波动-低用电型、波动-中等用电型和波动-高用电型,4种用电模式占比分别为66%、21%、10%和3%,如图12所示。

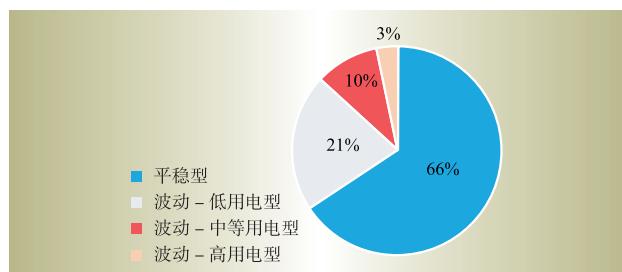


图12 台区电力消费模式占比

Fig. 12 The proportion of power consumption pattern in the community level

波动型用电模式台区(波动低用电、波动中等用电和波动高用电)的电力消费呈现长-中-短期差异性波动特征如图13所示。波动-低用电模式、波动-中等用电模式以及波动-高用电模式的长期、中期和短期的波动特征图如图14—图15所示。长期、中期和短期的波动特征差异明显,其中长期用电行为的波动相对较为平缓,中期用电行为的波动较为陡峭,短期用电行为的波动最为明显。对比可以发现,波动-高用电模式的长期波动幅度最为明显,呈现明显的峰谷特征;波动-中等用电模式的长期波动特征明

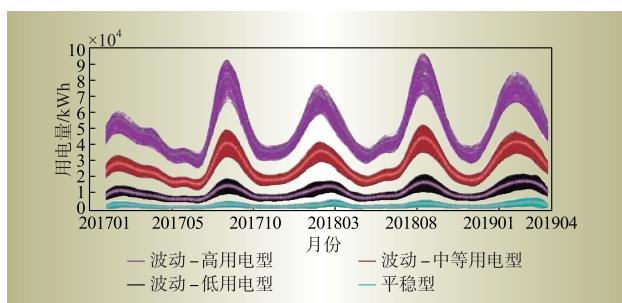


图13 台区用电模式图

Fig. 13 Electricity consumption modes of community level

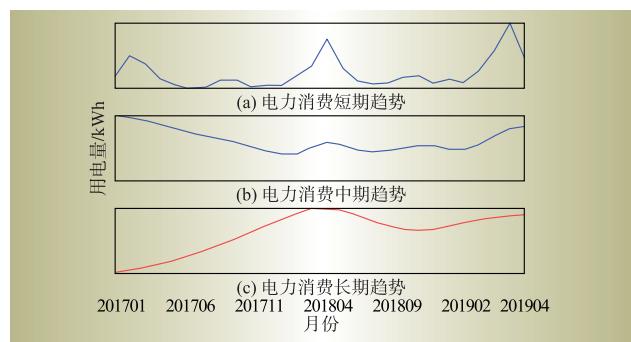


图14 波动-低用电模式波形分解

Fig. 14 Waveform decomposition of fluctuation-low power consumption mode

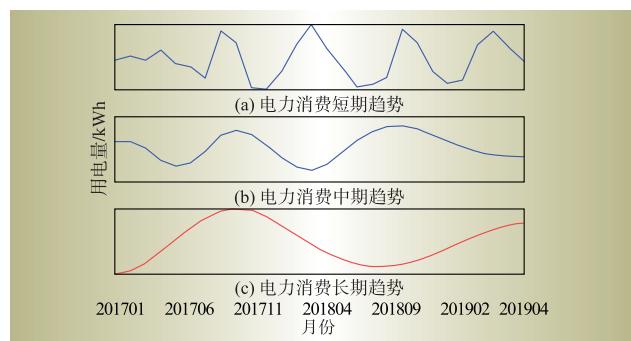


图15 波动-高用电模式波形分解

Fig. 15 Waveform decomposition of fluctuation-high power consumption mode

显区别于波动-高用电型,呈现波动上升的发展趋势;波动-低用电模式的长期波动特征介于波动-高用电型与波动-中用电模式之间,波动-低用电模式的中期波动特征明显区别于波动-中等用电模式和波动-高等用电模式。不同台区的长期、中期和短期的波动特征差异,主要由于台区空间位置的气候差异、台区的经济差异等因素所导致。通过对台区用电波形分解,识别模式间的长期-中期-短期的波动特征,对于缓解峰时用电压力具有现实意义。

#### 5 结束语

本文借助电力大数据,对居民电力消费行为特征分析与模式识别进行研究,对江西29万居民用户及所在台区近3年的用电数据进行特征分析。居民层面的聚类结果是居民家庭自身属性特征的体现,反映的是家庭属性方面的差异导致的家庭用电行为的差异;而台区层面是某些地区家庭的聚合,更多的体现的区域性的特征所导致的用电模式的差异,两者有一定的相关性,但是驱动用电行为背后的特征是不同的,故而在模式划分上也有一定的差异性,得出了以下结论:

- (1) 居民用电模式可以划分为4类:冬夏均衡型、夏季偏好型、夏季敏感型、冬季敏感型4类;
- (2) 台区用电模式可以划分为4类:平稳型、波

动-低用电型、波动-中等用电型和波动-高用电型；  
(3) 各类型用户除了在季节波动与温度敏感性差异外，在电力消费的中长期趋势上有明显差异。

基于多维特征融合的居民电力消费异质性模式识别与分析结果，可将不同家庭、台区用户按用能模式设置成不同标签，对于电网公司制定可复制、可推广的家庭智慧用能服务个性化方案，拓展用能服务的深度和广度，具有重要的技术储备价值与现实意义。由于数据量有限，本文研究仅采用了用电量单一维度数据，从不同时间尺度构建多维特征有一定的局限性。研究将家庭、台区的经济属性、社会属性、电器信息等多方面外部数据融合，采用维度更丰富的异质性数据，进行多维特征融合的居民电力消费异质性模式识别，为家庭用户智慧用能服务个性化方案制定提供更加细粒度、针对性支撑，是今后的重点研究方向。

## 参考文献：

- [1] 陆俊,陈志敏,龚钢军,等. 基于极限学习机的居民用电行为分类分析方法[J]. 电力系统自动化, 2019, 43(2):97-104.  
LU Jun, CHEN Zhimin, GONG Gangjun, et al. Classification analysis method for electricity consumption behavior based on extreme learning machine algorithm [J]. Automation of Electric Power System, 2019, 43(2):97-104.
- [2] 孙毅,刘迪,崔晓昱,等. 面向居民用户精细化需求响应的等梯度迭代学习激励策略[J]. 电网技术, 2019, 43(10):3 597-3 605.  
SUN Yi, LIU Di, CUI Xiaoyu, et al. Equal gradient iterative learning incentive strategy for accurate demand response of resident users [J]. Power System Technology, 2019, 43(10):3 597-3 605.
- [3] 胡江溢,祝恩国,杜新纲,等. 用电信息采集系统应用现状及发展趋势[J]. 电力系统自动化, 2014, 38(2):131-135.  
HU Jiangyi, ZHU Enguo, DU Xingang, et al. Application status and development trend of power consumption information collection system [J]. Automation of Electric Power Systems, 2014, 38(2):131-135.
- [4] 张新昌,周逢权. 智能电网引领智能家居及能源消费革新[J]. 电力系统保护与控制, 2014, 42(5):59-67.  
ZHANG Xinchang, ZHOU Fengquan. Smart grid leads the journey to innovative smart home and energy consumption patterns [J]. Power System Protection and Control, 2014, 42(5):59-67.
- [5] 王德文,孙志伟. 电力用户侧大数据分析与并行负荷预测[J]. 中国电机工程学报, 2015, 33(3):527-537.  
WANG Dewen, SUN Zhiwei. Big data analysis and parallel load forecasting of electric power user side [J]. Proceedings of the CSEE, 2015, 33(3):527-537.
- [6] 林顺富,郝朝,汤晓栋,等. 基于数据挖掘的楼宇短期负荷预测方法研究[J]. 电力系统保护与控制, 2016, 45(7):89-95.  
LIN Shunfu, HAO Chao, TANG Xiaodon, et al. Study of short-term load forecasting method based on data mining for buildings [J]. Power System Protection and Control, 2016, 45(7):89-95.
- [7] 李顺昕,远振海,丁健民,等. 基于聚类的用户用电行为及其影响因素分析[J]. 电力需求侧管理, 2019, 21(3):53-58.  
LI Shunxin, YUAN Zhenhai, DING Jianmin, et al. Analysis of users' electricity behavior and influencing factors based on clustering [J]. Power demand side management, 2019, 21(3):53-58.
- [8] 冯晓蒲,张铁峰. 基于实际负荷曲线的电力用户分类技术研究[J]. 电力科学与工程, 2010, 26(9):18-22.  
FENG Xiaopu, ZHANG Tiefeng. Research on electricity users' classification technology based on actual load curve [J]. Electric Power Science and Engineering, 2010, 26(9):18-22.
- [9] 张素香,刘建明,赵丙镇,等. 基于云计算的居民用电行为分析模型研究[J]. 电网技术, 2013, 37(6):1 542-1 546.  
ZHANG Suxiang, LIU Jianming, ZHAO Bingzhen, et al. Cloud computing-based analysis on residential electricity consumption behavior [J]. Power System Technology, 2013, 37(6):1 542-1 546.
- [10] 罗滇生,杜乾,别少勇,等. 基于负荷分解的居民差异化用电行为特性分析[J]. 电力系统保护与控制, 2016, 44(21):29-33.  
LUO Diansheng, DU Qian, BIE Shaoyong, et al. Analysis of differentiation residential electricity consumption characteristic based on power load decomposition [J]. Power System Protection and Control, 2016, 44(21):29-33.
- [11] 陆俊,朱炎平,彭文昊,等. 智能用电用户行为分析特征优选策略[J]. 电力系统自动化, 2017, 41(5):58-63, 83.  
LU Jun, ZHU Yanping, PENG Wenhao, et al. Feature selection strategy for electricity consumption behavior analysis in smart grid [J]. Automation of Electric Power System, 2017, 41(5):58-63, 83.

## 作者简介：

刘向向(1987),男,江西九江人,硕士,工程师,研究方向为电力需求侧管理;

卢婕(1982),女,江西南昌人,高级工程师,研究方向为电力需求侧管理;

周琪(1963),男,江西南昌人,教授级高级工程师,研究方向为电力需求侧管理;

赵文辉(1993),男,甘肃天水人,博士研究生,研究方向为大数据与因果推断的融合研究;

冯颖(1974),女,江西都昌人,高级工程师,研究方向为电费电价和电力市场。

(责任编辑 曹阳)