

基于周期调整负荷成分指数的行业 用电大数据价值挖掘

严玉婷¹, 薛冰¹, 方力谦¹, 黄国权², 张勇军²

(1.深圳供电局有限公司, 广东 深圳 518001; 2.华南理工大学电力学院, 广东 广州 510641)

摘要:能源数字经济背景下为展示行业用户的用电行为,探索地区和经济景气情况,本文借鉴股市成分指数提出一种周期调整负荷成分指数。首先,按照给定规则筛选出具有代表性的若干位行业用户;然后,利用周期分解算法提取所选用户历史日电量的周期分量,以计算调整日电量,并提出工作强度系数,再考虑其所涉及的行业及其个体差异提出多项指标,并基于模糊专家评价法等方式计算其权重以对调整日电量加权整合;最后,选定基期展示负荷指数曲线。分析可得,工作强度系数有助于联系实际生产活动,指数能够及时反映地区日用电行为模式,且去除气温影响后与经济指标强相关,能够阐释投资、产出与生产间的关系以及表征社会主体经济的动向。

关键词:行业用户;负荷成分指数;周期分解;工作强度;格兰杰因果检验

DOI:10.19781/j.issn.1673-9140.2022.06.021 中图分类号:TM-9 文章编号:1673-9140(2022)06-0181-09

Big data mining of industry power consumption based on component index about seasonal-adjusted load

YAN Yuting¹, XUE Bing¹, FANG Liqian¹, HUANG Guoquan², ZHANG Yongjun²

(1.Shenzhen Power Supply Co., Ltd., Shenzhen 518001, China; 2.School of Electric Power Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract: Under the background of energy digital economy, in order to display and analyze power consuming behavior, and to explore regional economic trends, this paper proposes a periodic adjusted load component index by referring to the stock market index. Firstly, some representative enterprise users are selected as samples based on given rules. Then, the periodic components of the selected user's historical daily electricity quantity are extracted by STL. Hence, the adjustment of the daily electricity quantity can be calculated, and the working strength coefficient is proposed. Then, according to the industry and individual differences, multiple indices are proposed, and the cycle adjusted daily load is weighted and integrated by fuzzy expert evaluation method. After that, based on a selected day's value, the load trend can be displayed. Finally, the analysis shows that the working strength coefficient is helpful to link with the actual production activities, and the load index can reflect regional daily electricity consumption behavior. Furthermore, after neglecting the influence of temperature, the index has a strong correlation with economic indicators, which can explain the relationship between investment, output and production, and represent the economy of social subject.

收稿日期:2021-06-18;修回日期:2021-11-19

基金项目:深圳供电局有限公司科技项目(SZKJXM20190594)

通信作者:黄国权(1996—),男,硕士研究生,主要从事电力大数据、负荷特性分析等研究;E-mail:907918377@qq.com

Key words: industry power users; component index; cycle decomposition; working strength; Granger causality test

中国能源革命不断向纵深推进,期间数字经济加入并助推能源高效转型,两者的深度融合催生了一种新型经济发展形态——能源数字经济^[1-3]。发展能源数字经济,数字化是战略性抓手,数据挖掘是关键出路。而作为社会能源供应的枢纽,电网企业应当充分发挥电力大数据实时监控、直观反馈的特点,分析负荷特性变化进而挖掘其深层价值。

近年来,基于电力大数据的用电分析研究逐渐深入,主要思路为①对负荷曲线聚类以提取典型负荷模式;②提出多个指标捕捉负荷的特定变化;③基于数据驱动方式分析用电行为。文献[4]总结了负荷模式提取常用的对负荷曲线进行聚类的方法特点;文献[5]提出了一种基于深度学习的 YLP 特征提取,捕获负荷每日和季节性变化,并通过负荷聚类图像观察其全年特征;文献[6-7]计算不同时间尺度下(年/月/日)的负荷率、最小负荷率和峰谷差率等时序特征指标,用以对比描述负荷的特性变化;文献[8]采用三相电压、电流、功率因数等建立多维数据特征,基于 Graph 模型聚类分析用电用户行为并对用户分类;文献[9]借鉴股市指数计算京电指数,从电力交易角度分析用电特点。

但是,上述研究与用户实际生产行为变化的联系并不紧密,具体表现:一方面上述所提指标均为负荷特性指标,与用户调休、增产等实际生产活动无关;另一方面,仅对精细时间尺度下的日负荷曲线进行聚类,并未对受工作日效应影响的周期波动和受气温影响的随机波动进行分离研究。因此,对于行业大用户用电大数据的挖掘分析,关注其实际生产行为变动能更好挖掘该类用户用电数据的潜在价值。

在证券交易中,股市指数被用于反映股市行情变动,按所包含股票样本数可分为 2 类^[10]:按一定标准选取部分上市股票作为样本的成分指数(如沪深 300 等)和以全部上市股票作为样本的综合指数(如上证综指等)。在电力大数据环境下,借鉴成分指数,通过科学客观的方法挑选出少量最具代表性的样本用户来降低数据广度,并分析行业大用户群体的用电行为变化是值得探索的方向。

鉴于此,本文旨在构造一种负荷指数,展示行业

大用户群体的负荷波动情况,为分析行业用户生产经营状态以及未来经济走势提供数据基础。首先,根据地区实际情况筛选出若干位具有代表性的行业大用户作为样本对象;然后,利用 STL 算法(seasonal-trend decomposition procedure based on lo-ess)获取样本历史日电量的周期分量,用以计算调整日电量,并提出工作强度系数指标,联系用户生产实际,再按样本所属行业及其个体地位,基于模糊专家评价法等方式赋予多项权重值,综合上述构造企业用户调整负荷成分指数;最后,以深圳市为案例分析负荷指数应用效果。

1 负荷指数样本基础

电力用户成千上万,每日产生大量数据,然而价值密度较低^[11-12],通过筛选其中部分用户的数据来降低数据广度可以提高价值密度。

自改革开放以来,中国经济总量快速攀升的同时发展不平衡问题也越发突出,利润分配在不同行业、不同规模企业上严重失衡,此现象在经济发达地区更为严重。以深圳市为例,2019 年地区生产总值(gross regional product, GDP)约为 2.7 万亿元,而仅中国平安保险集团的净利润已将近 0.15 万亿元,占地区 GDP 超 5%,营收总额排名前 10 位企业的净利润总和更是超过地区 GDP 的 20%。在此意义上,少量的大型企业的经营状态即可一定程度上代表地区经济发展的景气程度。因此,分析少量行业大用户群体的负荷特性变化在宏观上具备地区经济表征性。

借鉴股市成分指数,考虑筛选地区若干位具有代表性的企业用户作为样本对象来计算负荷指数。

2 周期调整负荷成分指数构造方法

前述提及,用户实际生产行为变动少有关注,而利润分配严重失衡,这使得简单加和样本用户用电量并不足以支撑深入分析。因此,需要针对上述情况相应处理样本用户日电量。

2.1 负荷指数基础指标

不同于小企业或个体户等,多数大企业尤其是工业大企业一般会实行单休或双休制度,且对于一些大规模的工业企业,其工作计划性强,在季度或月度等中短时间段内每日的工作量基本稳定。因此,部分大企业用户用电量存在着工作日效应,即一周内工作日负荷率明显高于周末日负荷率^[13]。这一工作日效应带来的固有周期变化使得用户负荷存在显著的“正常”波动,而这种波动会影响对负荷受扰波动的辨别以及对异常波动严重程度的判断。因此,需要区分开工作日效应这种固有波动的影响,但也不能忽略这部分电量的实际效益。

本文将用户实际日用电量表示成 $N \times 7$ 维矩阵 \mathbf{D} , 拆解为

$$\mathbf{D} = \mathbf{F} + \mathbf{I} \cdot \mathbf{T} \quad (1)$$

式中 \mathbf{F} 为 $N \times 7$ 维的非周期分量矩阵; \mathbf{I} 为 $N \times 1$ 维的全 1 矩阵; \mathbf{T} 为 1×7 维的周期分量矩阵。

\mathbf{T} 通过 STL 算法^[14] 对历史数据进行周期分解得到。STL 是一个迭代的非参数回归过程,是将鲁棒局部加权回归作为平滑方法的时间序列分解方法,能够得到稳健的周期项,而不会被数据中的异常行为扭曲,并且适用于有缺失值的时间序列。

STL 利用局部加权散点平滑 (locally weighted scatterplot smoothing, LOESS) 对离散序列点进行平滑处理。对于一个时间序列 $y_i (i=1, 2, \dots, n)$, 选取平滑参数 $q \leq n$, 由与时间点 x 最近的 q 个序列时间点 x_j 构造时间点 x 上的邻近权重:

$$W_j(x) = \begin{cases} \left[1 - \left(\frac{|x_j - x|}{\lambda_q}\right)^3\right]^3, & |x_j - x| < \lambda_q \\ 0, & |x_j - x| \geq \lambda_q \end{cases} \quad (2)$$

式中 λ_q 为 x 与距其第 q 个远的序列值 x_λ 的距离。

以此对 y_i 局部加权线性回归的优化目标为

$$\min f_i = \sum_{j=1}^n W_j(x_i) \cdot [y_j - \theta_i \cdot x_j]^2 \quad (3)$$

基于 LOESS, STL 包含内、外 2 个循环, 其中内循环过程如下:

1) 去趋势, 即减去上一轮结果的趋势分量, $\mathbf{D} - \mathbf{T}^{(k)}$;

2) 周期子序列平滑, 对每个子序列做 LOESS

回归, 并向前、后各延展一个周期, 得到 $C^{(k+1)}$;

3) 周期子序列低通滤波, 对 $C^{(k+1)}$ 依次进行窗口长度为 7 (周期长度)、7、3 的滑动平均, 再 LOESS 回归, 得到 $H^{(k+1)}$;

4) 周期子序列去趋势, 即 $S^{(k+1)} = C^{(k+1)} - H^{(k+1)}$;

5) 去周期, 即减去周期分量, $\mathbf{D} - S^{(k+1)}$;

6) 趋势平滑, 对 $\mathbf{Y} - S^{(k+1)}$ 进行 LOESS 平滑, 得到趋势分量 $T^{(k+1)}$ 。

外循环根据上述所得周期项 S 和趋势项 T 计算余项 $R = \mathbf{D} - S - T$, 并由此对时间点 x 赋予稳健权重:

$$\kappa = \begin{cases} (1 - u^2)^2, & 0 \leq u < 1 \\ 0, & u \geq 1 \end{cases} \quad (4)$$

$$u = \frac{|R|}{6 \cdot \text{median}(|R|)} \quad (5)$$

式中 $\text{median}(\cdot)$ 表示取给定数值集合的中值。

进一步地, 为保留周期分量电量带来的实际效益, 将 \mathbf{T} 均摊到周一至周日, 以日用电量调整值 \mathbf{D}' 代替实际值 \mathbf{D} 作为负荷指数基础指标:

$$\mathbf{D}' = \mathbf{D} - \xi \cdot (\mathbf{T} - \bar{\mathbf{T}}) \quad (6)$$

式中 ξ 为 $N \times 1$ 维的工作强度系数矩阵; $\bar{\mathbf{T}}$ 为 1×7 维的矩阵, 其每个元素值均相等, 为周期分量序列的均值。

企业用户虽有固定工作日效应, 但由于每周任务量的差异会存在工作强弱区别, 以 ξ 描述此种工作强度, 其元素 ξ_i 的目标确定如下:

$$\min g_m = \sigma(\mathbf{D}_m - \xi_m \cdot \mathbf{T}) \quad (7)$$

式中 \mathbf{D}_m 为由 \mathbf{D} 中第 m 行元素组成的行矩阵; $\sigma(\cdot)$ 来计算给定数值集合的标准差。

2.2 样本用户权重确定

利润分配失衡问题使得各行业各企业对主体经济的推动力不一。因此, 考量企业用户所属行业及其个体的差异, 为样本用户分配多项权重, 并据此对其调整日电量进行加权整合。

2.2.1 行业层面

1) 地位重要度权重。

相较于一些传统行业的企业用户, 社会或是政府更关心一些高新技术行业的发展, 同时这些高新技术和行业也汇聚了更多的社会资源投入, 具有更

高的经济产值。因此,对于这部分用户用电量波动应给予更多的重视。在指数中,结合一种经典的主观赋权法——模糊专家评价法^[15]对样本涉及行业的相对重要性划分 5 个等级,并以梯形模糊数 $M = (l, m, n, r)$ 量化专家评判意见,其中, l, m, n, r 分别为模糊数的下界值、上临值、下临值和上界值,取值如表 1 所示。

表 1 梯形模糊数取值

Table 1 Trapezoidal fuzzy value

相对重要性评判结果	对应梯形模糊数 M
略微重要	(0.00,0.05,0.15,0.25)
一般重要	(0.15,0.25,0.35,0.45)
重要	(0.35,0.45,0.55,0.65)
比较重要	(0.55,0.65,0.75,0.85)
非常重要	(0.75,0.85,0.95,1.00)

首先,邀请 k 位专家评判样本涉及行业的地位相对重要性,将所有专家评判结果对应的模糊数以等权重线性加权方式进行整合,得到第 i 个行业的重要程度判别模糊数:

$$M_i = \frac{1}{k} \sum_{q=1}^k M_{iq} \quad (8)$$

式中 M_{iq} 为第 q 位专家对第 i 个行业的评判结果对应的模糊数。

然后,计算梯形模糊数的重心:

$$c(M_i) = \frac{\int x \mu_M(x) dx}{\int \mu_M(x) dx} = \frac{n_i^2 + r_i^2 + n_i r_i - m_i^2 - l_i^2 - m_i l_i}{3(n_i + r_i - m_i - l_i)} \quad (9)$$

式中 $\mu_M(x)$ 为梯形模糊数 M 对应的隶属度函数,取值如下:

$$\mu_M(x) = \begin{cases} 1, & x \in [m, n] \\ \frac{x-l}{m-l}, & x \in [l, m) \\ \frac{x-r}{n-r}, & x \in (n, r] \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

最后,对 $c(M_i)$ 归一化即可得到样本涉及行业的地位重要性权重 s_i 。

2) 行业贡献率权重。

由于地区的发展规划因地制宜,各有行业侧重,使得不同行业对地区经济发展的贡献度不一,因此,

对高贡献行业应给予更多关注。定义行业贡献率 r_C 为各行业增加值增量与地区生产总值增量之比,即

$$r_{Ci} = \frac{\Delta E_i}{\Delta Z_{GDP}} \quad (11)$$

式中 ΔE_i 为第 i 个行业的地区行业增加值增量; ΔZ_{GDP} 为地区生产总值增量。

3) 能耗产出关系权重。

高能耗并不与高经济收益完全对等,相比于传统的高能耗产业,高产出的技术型行业更为重要。对此,行业能耗产出关系反映了其对主体经济的贡献度,根据用户所属行业对其赋予经济贡献度权重,该权重计算方式如下:

$$\rho_i = \frac{E_i}{P_i} \bigg/ \sum_{i=1}^n \frac{E_i}{P_i} \quad (12)$$

式中 E_i 为第 i 个行业的地区行业增加值; P_i 为第 i 个行业的样本电量总和; n 为样本涉及行业总数。

2.2.2 个体层面

同行业内企业用户间的规模对比很大程度上在其用电量水平上有所体现,因此无需额外考量该因素并赋权。而值得关注的是,企业用户业扩潜力与该用户的发展过程有直接的关系,相比于发展到稳定期用户,日渐蓬勃的企业更是经济走势的潜力股,应当着重分析。对此提出年化业扩比例 k_p 和容量饱和度 $S_{S,p}$ 指标,判断第 p 个用户的业扩潜力:

$$k_p = \sqrt[3]{\frac{S_p^{(0)}}{S_p^{(-3)}}} \quad (13)$$

$$S_{S,p} = \begin{cases} 1, & \eta_p = 0 \\ 1 + \exp(1 - \frac{1}{\eta_p}), & \eta_p \neq 0 \end{cases} \quad (14)$$

式中 $S_p^{(y)}$ 为第 p 个用户距当前年份 y 年前的容量; η_p 为第 p 个用户的容量使用比例,采用表 2 所示的分级靠档方式获得,以使计算指数权重的容量饱和度保持相对稳定。

组合 $k_p, S_{S,p}$ 这 2 项指标,计算第 p 个用户的业扩潜力权重:

$$e_p = \frac{k_p \cdot S_{S,p}}{\sum_{p \in \Omega_i} k_p \cdot S_{S,p}} \quad (15)$$

式中 Ω_i 为第 i 个行业的用户集合。

表 2 容量使用比例分级靠档

Table 2 Category-weighted chart of the capacity usage ratio %

比例		比例	
实际	计算	实际	计算
<50	0	(70,80]	80
[50,60]	60	>80	90
(60,70]	70		

2.2.3 权重组合

组合上述业扩潜力权重,得到第 i 个行业内第 p 个用户的综合权重,即

$$\omega_p = (\alpha s_i + \beta r_{C_i} + \theta \rho_i) \cdot e_p, p \in \Omega_i \quad (16)$$

式中 α, β, θ 为组合系数,此处取 $\alpha = \beta = \theta = 1/3$ 。

2.3 构造调整负荷成分指数

综上,在样本用户调整日电量 \mathbf{D}' 的总和值 D'_s 基础上,利用组合权重求取所选用户用电量的加权平均值,并选取自然年中的某日 t_0 作为基期,以该日的负荷加权均值作为基准值,构造指数如式(16)。进而每日发布前日指数值,生成指数曲线,以便分析大企业用户的日电量相对波动情况,并及时反馈异常。

$$I_t = \frac{\sum_{i=1}^n \sum_{p \in \Omega_i} \omega_p D'_{sp}}{\sum_{i=1}^n \sum_{p \in \Omega_i} \omega_p D'_{sp0}} \times 1000 \quad (17)$$

式中 D'_{sp} 为第 p 个用户的调整日电量矩阵总和值; D'_{sp0} 为第 p 个用户于 t_0 日的调整日电量矩阵总和值。

2.4 剔除环境因素影响

环境因素是影响负荷变化的另一主要因素,尤其是气温因素。通过分析负荷与气温的相关性,剔除指数受气温因素的影响。

协变量可以用来解释时间序列中的一些变化,考虑气温因素作为具有时变系数的协变量,其时变系数随时间平稳变化,但不呈现任何周期性。将指数拆解为

$$I_t = L_t + \tau_t A_t + \zeta_t + \varepsilon \quad (18)$$

式中 L_t 为平滑的趋势项; A_t 为协变量,即气温值; τ_t 为协变量的系数; ζ_t 为随机项; ε 为拟合误差。

剔除气温因素协变量的影响,即可得到负荷指数的趋势曲线。

3 调整负荷成分指数应用示例

以深圳地区为例,所有用户按 2019 年总用电量排序,以最大用电量作为基准值的对比结果如图 1 所示,其中第 1 位用户与第 173 位用户的差异已超百倍,与第 201 位用户差异更是将近千倍。因此,剔除用电量排序 200 名外的用户,再参考行业挑选出深圳市 100 位企业用户作为样本构造 2020 年指数,电量数据规模大小由全量用户的 365×10^6 维矩阵缩减至 100 位用户的 365×100 维矩阵,样本涉及行业及其代号对照如表 3 所示。

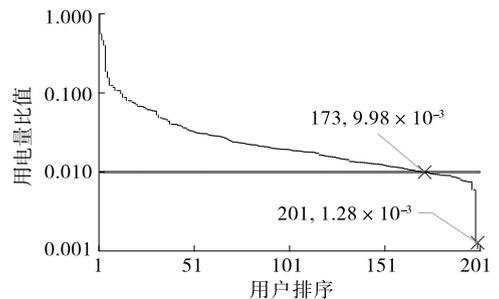


图 1 用电量排序结果

Figure 1 Power consumption sorting result

表 3 样本涉及行业及其代号对照

Table 3 Industries of samples and its code

样本涉及行业	代号	用户数
制造业	T01	53
租赁和商务服务业	T02	7
交通运输、仓储和邮政业	T03	9
信息传输、软件和信息技术服务业	T04	11
金融业	T05	3
房地产业	T06	9
电力、燃气及水的生产和供应业	T07	6
公共管理、社会保障和社会组织	T08	2
总计	—	100

首先,对样本用户前一自然年(2019年)的历史日电量进行 STL 分解,获取样本用户历史日电量的周期分量,并据此对用户 2020 年实际日电量进行调整,得到调整日电量 \mathbf{D}' 。某用户调整前、后的电量曲线如图 2 所示,可见该用户调整负荷值虽仍有细微波动,但与调整前负荷的波动性质已截然不同,不再表现出工作日效应,即不再固定于休息日大幅跌

落至极小值。同时,由此得到的 100 位用户的工作强度系数矩阵如表 4 所示。

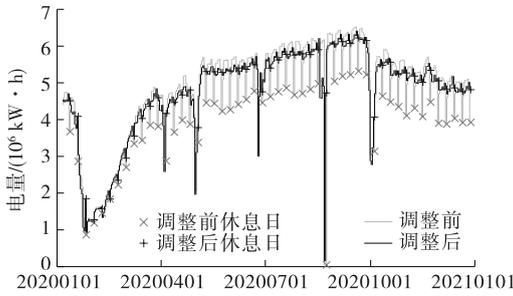


图 2 某用户调整前、后用电量对比

Figure 2 Power consumption before and after adjustment

表 4 工作强度系数矩阵

Table 4 Matrix of working strength coefficient

周数	用户 1	用户 2	用户 3	...	用户 99	用户 100
1	3.097	0.822	0.566	...	0.786	0.462
2	0.155	1.202	0.688	...	0.750	0.436
3	0.229	0.965	0.728	...	2.564	0.333
4	0.000	0.083	0.311	...	1.532	0.000
5	0.097	0.123	0.000	...	0.317	0.206
6	0.000	0.000	0.327	...	1.284	0.263
7	0.000	0.123	0.096	...	0.000	0.213
...
51	0.000	0.872	0.193	...	0.000	0.504

然后,确定各样本用户涉及行业的权重。在深圳市统计局网站上查询行业增加值增量及样本电量总和等数据,如表 5、6 所示(由于 2019 年统计数据不全,就近取 2018 年数据作为代替计算)。

表 5 2018 年样本涉及行业增加值增量及电量总和

Table 5 Increment of value added by the industries of samples and total power consumption in 2018

涉及行业	行业增加值 $E_i/(10^{11}$ 元)			样本电量总和 $D_i/(10^9 \text{ kW} \cdot \text{h})$
	2017 年	2018 年	增量 ΔE_i	
T01	8.266	8.813	0.547	11.482
T02	0.931	1.100	0.169	1.086
T03	0.706	0.738	0.033	2.536
T04	1.941	2.248	0.307	1.766
T05	3.160	3.352	0.192	0.363
T06	1.937	2.169	0.232	1.540
T07	0.237	0.310	0.072	1.010
T08	0.534	0.571	0.036	0.114

表 6 2018 年地区生产总值增量

Table 6 Increment of the regional GDP in 2018

年份	地区 GDP/(10^{11} 元)
2017	23.280
2018	25.266
增量 GDP	1.986

采用问卷调查的方式咨询专家意见,根据供电局客服经验丰富的 5 位专家意见结果,按式(8)~(10)计算模糊数重心大小,得出行业重要度权重,如表 7 所示。结合表 5、6,按式(11)、(12)可计算行业贡献率、能耗产出关系权重;根据 100 位企业用户近 3 年的容量扩装情况计算其业扩潜力权重。

表 7 行业权重值

Table 7 The weights of industries of the samples

涉及行业	s_i	r_{Ci}	ρ_i	$as_i + \beta r_{Ci} + \theta \rho_i$
T01	0.145	0.345	0.040	0.177
T02	0.096	0.106	0.053	0.085
T03	0.106	0.021	0.015	0.047
T04	0.165	0.193	0.066	0.141
T05	0.140	0.121	0.479	0.247
T06	0.093	0.146	0.073	0.104
T07	0.140	0.045	0.016	0.067
T08	0.115	0.023	0.259	0.132

最后,整合上述权重并结合式(16)对企业 2020 年调整负荷进行计算,以 2020 年 1 月 6 日作为基期绘制指数,如图 3 所示,进一步地,可去除气温因素影响的指数。

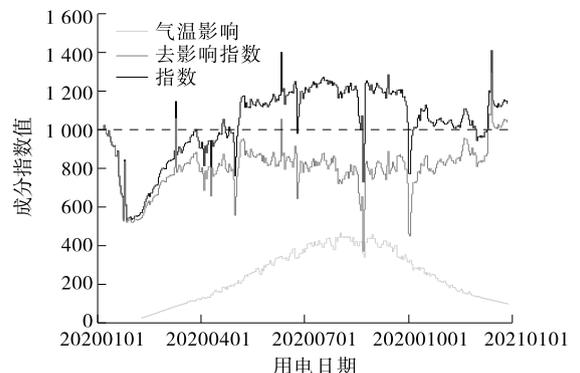


图 3 2020 年指数及气温影响情况

Figure 3 Index and its temperature impact in 2020

4 调整负荷成分指数示例分析

4.1 用户工作强度分析

基于上述所得 100 位用户的工作强度系数矩阵(表 4)绘制热力图,如图 4 所示(图中色度于 0~1 为均匀分布,于 1~10 为对数分布),其中,每行代表一周,每列代表一位用户,每个色块的颜色表示用户某一周的工作强度系数,系数值越大,说明用户当周工作量越大。

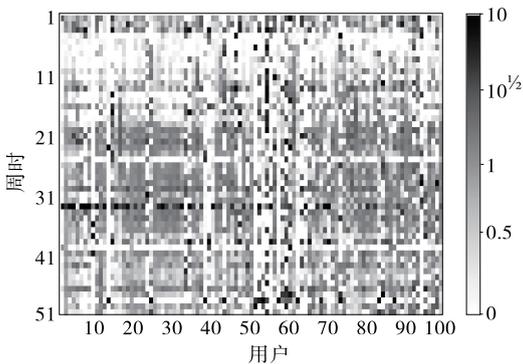


图 4 2020 年各周用户工作强度情况

Figure 4 Weekly working strength in 2020

由图 4 中可见,第 4 至第 8 周(2 月份)基本所有用户工作强度系数接近 0,当时处于春节期间,且受疫情影响直到 3 月才陆续复工复产,能看出部分企业疫情调休,复工复产后仍有几周无工作日效应;第 14、16、24、39 周同样出现了大部分用户工作强度系数降至 0 的情况,其中,第 14、16、24 周分别为劳动节/端午节前后周,均于周末调休补班,而第 39 周为国庆节当周,假期工作暂停。

综上,所提工作强度系数在分析用户复工复产、工作安排、调休补班等生产经营行为上具有重要作用。具体地,当工作强度系数接近 1 时,该用户当周处于正常生产经营状态;当接近 0 时,该用户当周处于调休补办或假期状态;当远大于 1 时,该用户当周生产经营活动较平常更为活跃。

4.2 地区日用电行为模式分析

当去除工作日特性的短期规律波动后,指数的趋势更为明显,可以清晰辨别用户电量的整体水平和异常波动。如:2020 年大部分时间下去气温影响指数值低于 1 000,以电量角度来看,地区发展整体处于低迷状态,直至年底有所恢复;再如:1 月底春

节期间指数跌至 550 左右,之后受疫情影响直至 3 月中旬指数恢复至 1 月水平,说明大企业才基本实现复工复产;又如:指数在春节、劳动节、端午节、国庆节 4 个法定节假日前后明显下跌,其中下降幅度为春节>国庆>劳动节>端午节,这方面指数也侧面反映了对节假日的重视程度。

此外,将 2020 年指数值与该年全社会总电量(以 1 月 1 日值标准化)对比,如图 5 所示,可见原始指数与全社会总电量具有相似的变化趋势,具有表征整体用户用电行为的作用。

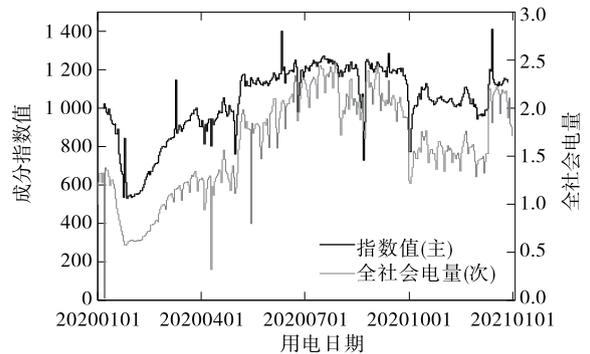


图 5 2020 年指数值与全社会总电量对比

Figure 5 Index and total electricity consumption in 2020

综上,通过对比指数值可比较地区电量变动情况,一定程度上能反映地区发展态势,当指数长期处于增长状态则说明地区生产经营行为趋于活跃,发展向好,相反则说明地区发展较为低迷。

4.3 指数与经济关系格兰杰检验

为进一步验证指数对地区经济的表征性,本文探讨指数与地区 GDP、固定资产投资(investment in fixed assets, FAI)、规模以上工业总产值(gross output value of industry above designated size, IDGO)以及居民消费价格指数(consumer price index, CPI)等指标间的关系。

计量经济学中的格兰杰因果检验^[16]可用于判断 2 个时间序列变量在统计学上的因果关系。利用格兰杰因果检验可分析去除气温影响后的指数与多个经济指标间的关系。基于卡方分布对指标与指数间关系假设进行检验,在 1% 的显著性水平下作三阶最大滞后数检验,并与未经加权处理的指数检验作对比,取显著性水平最小的滞后阶数结果,如表 8 所示;GDP 与指数的相关性如表 9 所示。

表 8 Granger 因果检验结果

Table 8 Granger causality test results

指数	假设	滞后阶数	显著性水平	结论
加 权	FAI 不是指数的 Granger 原因	3	0.000 9	拒绝
	指数不是 FAI 的 Granger 原因	2	0.247 7	接受
	IDGO 不是指数的 Granger 原因	3	0.010 9	接受
	指数不是 IDGO 的 Granger 原因	3	0.002 8	拒绝
	CPI 不是指数的 Granger 原因	3	0.000 0	接受
	指数不是 CPI 的 Granger 原因	3	0.000 0	拒绝
不 加 权	FAI 不是指数的 Granger 原因	3	0.001 8	拒绝
	指数不是 FAI 的 Granger 原因	2	0.414 6	接受
	IDGO 不是指数的 Granger 原因	3	0.020 4	接受
	指数不是 IDGO 的 Granger 原因	3	0.007 3	拒绝
	CPI 不是指数的 Granger 原因	3	0.000 0	拒绝
	指数不是 CPI 的 Granger 原因	3	0.000 0	拒绝

表 9 GDP 与指数的相关系数

Table 9 Correlation coefficient between GDP and index

计算方法	加权指数	不加权指数
Pearson	0.915 9	0.616 3
Kendall	1.000 0	0.666 7
Spearman	1.000 0	0.800 0

格兰杰检验结果显示,指数是引起 IDGO、CPI 等变化的 Granger 原因,并且均具有三阶的滞后,故负荷指数所表征的用电行为变动可在一定程度上反映生产总值以及消费水平的变化;FAI 是引起指数三阶滞后变化的 Granger 原因,说明了固定生产投资对生产经营行为起到规划作用。在 GDP 与指数的相关性分析中,多种方法均说明指数与 GDP 高度相关。同时,在指数加权与否的对比中发现,加权指数与经济指标关系检验结果的显著性水平更低,即结果更为可信,且加权指数与 GDP 的相关性也优于不加权指数的相关性,由此证明了加权方法的有效性。此外,指数与经济值的关系也印证了样本选取具备良好的表征性。

5 结语

本文借鉴股价指数,构造了调整负荷成分指数,旨在展示行业大用户的电力负荷发展情况,指数大于 1 000 时说明用电量比基期有所上涨,对比两日指数值可比较用户日电量变化情况。

1) 指数构造过程所得到的工作强度系数在分析

用户复工复产、工作安排、调休补班等生产经营行为上具有重要作用;

2) 指数反映地区日用电行为模式时具有更加准确直观的可视性,具有表征整体用户用电行为的作用,且一定程度上能反映地区发展态势;

3) 去除气温影响后的指数与经济等指标具有明显的统计因果关系,能够为投入产出关系分析以及经济走势预测提供参考价值。

未来可在调整负荷成分指数的基础上研究更细化、分类别的电力负荷指数,专门针对不同行业、不同体量、不同类型的用户制定指数曲线,通过提炼指数背后的经济价值为各方决策提供有效支持。

参考文献:

- [1] ZHAO J H, WANG Y F, XIA Y B, et al. Simulation research on the role of energy revolution in the rise of China's five central provinces[J]. IOP Conference Series: Earth and Environmental Science, 2021, 680(1): 012038.
- [2] 袁斌, 张皓维, 崔萌萌. 基于深度学习的电力基建现场安全管控系统[J]. 电网与清洁能源, 2020, 36(9): 30-36.
YUAN Bin, ZHANG Haowei, CUI Mengmeng. Deep learning based security and control system in power grid construction[J]. Power System and Clean Energy, 2020, 36(9): 30-36.
- [3] 陈睿欣, 刘素蔚. 能源数字经济为经济社会带来新动能[EB/OL]. https://www.sohu.com/a/404984306_120051337, 2020-06-30.
- [4] 张铁峰, 顾明迪. 电力用户负荷模式提取技术及应用综述[J]. 电网技术, 2016, 40(3): 804-811.
ZHANG Tiefeng, GU Mingdi. Overview of electricity customer load pattern extraction technology and its application[J]. Power System Technology, 2016, 40(3): 804-811.
- [5] RYU S, CHOI H, LEE H, et al. Convolutional autoencoder based feature extraction and clustering for customer load analysis[J]. IEEE Transactions on Power Systems, 2020, 35(2): 1048-1060.
- [6] 曹华珍, 吴亚雄, 李浩, 等. 基于海量数据的多维度负荷特性分析系统开发[J]. 电力系统保护与控制, 2021, 49(6): 155-166.
CAO Huazhen, WU Yaxiong, LI Hao, et al. Development of a multi-dimensional load characteristic analysis system based on massive data[J]. Power System Protection

and Control, 2021, 49(6):155-166.

- [7] 李富鹏,沈秋英,王森,等.基于大数据和多因素组合分析的单元制配电网精细化负荷预测[J].智慧电力, 2020, 48(1):55-62.
LI Fupeng, SHEN Qiuying, WANG Sen, et al. Refined load forecasting method for unit distribution network based on big data and multiple factors[J]. Smart Power, 2020, 48(1):55-62.
- [8] 陶鹏,张洋瑞,李梦宇,等.基于 Graph 模型的海量用电数据并行聚类分析[J].电力科学与技术学报, 2020, 35(6):144-151.
TAO Peng, ZHANG Yangrui, LI Mengyu, et al. Parallel clustering analysis for power consumption data based on graph model[J]. Journal of Electric Power Science and Technology, 2020, 35(6):144-151.
- [9] 史连军,庞博,刘敦楠,等.新电改下北京电力交易中心电力市场综合指数的交易分析[J].电力系统自动化, 2019, 43(6):163-170.
SHI Lianjun, PANG Bo, LIU Dunnan, et al. Power market transaction analysis of index of beijing electric power exchange center under new electricity reform[J]. Automation of Electric Power Systems, 2019, 43(6):163-170.
- [10] 孙清岩.股票价格指数编制理论研究[D].大连:东北财经大学, 2010.
- [11] 莫文火,陈碧云.基于邻域关系矩阵的电力大数据增量式属性约简研究[J].高压电器, 2020, 56(8):192-197+204.
MO Wenhua, CHEN Biyun. Incremental attribute reduction of electric power big data based on neighborhood relation matrix[J]. High Voltage Apparatus, 2020, 56(8):192-197+204.
- [12] 华志刚,范佳卿,郭荣,等.人工智能技术在火电行业的应用探讨[J].中国电力, 2021, 54(7):198-207.
HUA Zhigang, FAN Jiaqing, GUO Rong, et al. Discussion on application of artificial intelligence technology in thermal power industry[J]. Electric Power, 2021, 54(7):198-207.
- [13] 焦筱俊,徐青山.电力系统单用户超短期负荷预测算法研究[J].电测与仪表, 2020, 57(1):30-35.
JIAO Xiaoquan, XU Qingshan. A new ultra-short-term load forecasting algorithm for single user in power system[J]. Electrical Measurement & Instrumentation, 2020, 57(1):30-35.
- [14] XIONG T, LI C, BAO Y. Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: Evidence from the vegetable market in China [J]. Neurocomputing, 2018, 275:2831-2844.
- [15] 周一凡,胡伟,闵勇,等.基于省级数据的电力发展水平动态综合评价方法[J].电力系统自动化, 2016, 40(18):76-83.
ZHOU Yifan, HU Wei, MIN Yong, et al. Dynamic comprehensive evaluation method of power industry development level based on provincial data[J]. Automation of Electric Power Systems, 2016, 40(18):76-83.
- [16] ERIC G, JONATHAN B H, KAIJI M. Testing a large set of zero restrictions in regression models, with an application to mixed frequency Granger causality[J]. Journal of Econometrics, 2020, 218(2):633-654.