

基于电力大数据的企业复工电力指数的应用分析

梁雪青, 杜舒明, 赵小凡, 刘超

(广东电网有限责任公司广州供电局, 广东 广州 510000)

摘要: 在社会经济发展过程中, 我国电力行业取得了突飞猛进的发展, 作为支撑现代化建设的重要公共事业, 不断优化电力行业管理模式, 借助信息技术实现电力企业的创新发展, 具有十分重要的意义。电力大数据是指与电力企业运营、销售、生产相关的数据信息, 是电力企业发展的重要内容, 可以分析行业发展现状、周期, 为电网企业预测行业用电趋势, 制定电网建设规划。随着疫情逐渐得到良好控制, 各地企业逐渐复工, 影响就业情况及经济发展。基于此, 文章立足电力大数据角度, 通过对电力指数的应用分析, 探讨企业复工情况, 阐述电力指数在企业复工中的应用思路, 并按照企业分群和企业复工情况分析两个步骤研究电力指数的实际应用, 以为相关学者及政府决策提供参考。

关键词: 大数据; 复工; 电力指数

中图分类号: TM72

文献标识码: A

文章编号: 1674-1064 (2022) 01-079-03

DOI: 10.12310/j.issn.1674-1064.2022.01.027

新冠肺炎疫情常态防控后, 虽然部分企业逐步复工复产, 但是人员流动和实际复工率相对较低, 在此情况下, 为推动国民经济稳定发展, 促使社会有序运转, 各地政府必须加强对于当地企业复工情况的了解, 为后续决策提供相应数据支持, 通过分析电力大数据, 明确企业复工电力指数, 帮助政府充分了解企业的复工情况。

1 电力指数在企业复工中的应用思路

基于实际生产经营活动与企业电力指数之间的关系, 此次研究主要针对一定区域范围内进行相应生产活动的企业, 一些已经关停或小电量的单位和企业并不在文章的讨论范围内。在此前提之下, 应用电力大数据和电力指数进行企业复工情况分析。

结合各地区衡量企业复工复产情况的标准, 主要从以下两个方面分析企业的复工情况: 一方面, 分析已复工企业在当地企业总数中的占比; 另一方面, 分析该地区复工后的电量水平占春节前电量的比例。前者被称为复工企业比例, 后者被称为复工电量比例^[1]。

在此基础上, 将这两个指数按照各50%的比例, 构建复工电力指数, 其计算公式为:

复工电力指数 = (复工企业比例 × 0.5 + 复工电量比例 × 0.5) × 100%。

2 实际应用分析

根据电力指数分析企业复工情况, 判断企业复工情况

的主要方式, 结合企业停工前的用电规律进行分析。

通常情况下, 在春节期间及春节前后, 企业用电情况包括两种。第一种企业在春节期间并未停止营业, 其在春节前后和春节期间的用电情况并不会产生较大波动。第二种企业会在春节期间停产或停工, 随着假期结束, 企业员工陆续返程, 企业会恢复生产, 用电量也会逐步增加。

在分析企业复工情况时, 由于两种企业的实际情况和用电规律呈现出不同特点, 如果采用同一个标准或方法进行分析, 就会得出错误的结论。因此, 为保障企业复工情况分析的准确性和可靠性, 要结合实际情况针对不同类型企业的特点, 进行分类讨论和分析。

2.1 企业分群分析

根据上述分析结果, 探讨企业复工情况时, 首先要做的就是企业分群。

基于春节前后不同企业的不同用电规律, 主要以春节期间停工与否为界限进行企业分群, 再通过电力数据的应用和计算, 对春节期间停工企业的后续复工情况展开分析和判断。

针对企业春节期间用电规律进行分群的方式, 属于无监督模型分群, 对于此类分群模式, 常用的分群模型包括层次聚类、高斯混合模型等多种类型。高斯混合模型作为一种聚类方式, 具有更好的实践应用效果和价值, 因此, 选择高斯混合模型作为企业分群模型。

按照上述理论分析, 对需要进行企业复工情况分析地区, 企业用电量信息和相关数据进行收集整理, 构建春节用电比例特征模型, 计算公式为:

春节用电比例特征 = 春节期间平均日电量/春节前3个

作者简介: 梁雪青 (1985—), 女, 广东广州人, 硕士, 高级工程师, 研究方向: 信息技术、项目管理、大数据应用。

月平均日电量

使用高斯混合模型，以收集到的企业用电量信息为基础，进行聚类分析。但是，在实际计算分析过程中，根据模型得到的企业聚类情况，并不一定会将所有用户划分为上文分析的两个类型。因此，在结果分析过程中，要结合模型收敛的实际情况，确定合理的聚类数目。

按照上述研究情况，文章以某地区2020年春节前后企业用电情况大数据为例，进行聚类结果分析，经计算分析后最终得到聚类数为三类。

在使用春节用电比例对企业进行分类划分后，经过统计，各类型企业在春节期间的用电比例情况如表1所示。结合当地企业实际生产经营业务和类型，将第一类群体和第二类群体合并，将其作为春节不停工企业，将第三类群体划分为春节停工企业，其划分结果和用电比例情况如表2所示。

根据表2可知，对于该地区企业复工情况而言，春节期间不停工且用电比例在0.27以上的企业，在春节后停工的概率相对较小，可认为其在春节后复工。春节期间处于停工状态的企业，其春节后复工情况需要进一步分析。

表1 聚类算法下各类型企业在春节期间的用电比例情况

企业类型划分	春节用电比例最小值	春节用电比例最大值	数量	占比
第一类	0.88	2.92	21 570	5.33%
第二类	0.27	0.87	95 480	23.54%
第三类	0.0	0.26	288 390	71.13%

表2 春节期间企业类别划分及用电比例

企业类别	春节用电比例最小值	春节用电比例最大值	数量	占比
春节不停工	0.27	2.92	117 050	28.87%
春节停工	0	0.26	288 390	71.13%

在企业分群分析中，结合春节用电规律开展分群工作的方法均为无监督模型分群，最常用的模型包括K-means模型、K-medoids模型、Mean Shift模型、层次聚类模型、DB-SCAN模型和GMM模型。

GMM属于一种高斯混合模型，相比而言，使用频率更高。该模型是运用多个高斯分布拟合数据集，数据分析统计效率较高，实践应用效果良好。GMM这种聚类方法具有概率式特征，在计算公式中，需要设 x_i ($i=1, 2, \dots, m$) 为第 i 个电力数据观察值，该样本与高斯混合分布相符。

$$\begin{cases} p(x) = \sum_{i=1}^L W_i N(x|u_i, \Sigma_i) \\ N(x|u_i, \Sigma_i) = \frac{1}{\sqrt{|\Sigma_i|} (2\pi)^{d/2}} \exp[-\frac{1}{2}(x-u_i)^T \Sigma_i^{-1} (x-u_i)] \\ \sum_{i=1}^L W_i = 1, 0 \leq W_i \leq 1 \end{cases}$$

在以上公式中， $p(x)$ 是指基于高斯混合模型的概率密度函数； $N(x|u, \Sigma)$ 是指高斯概率密度函数； W_i 是第 i 个组件的权重， u_i 是第 i 个组件的平均值， Σ_i 是第 i 个组件的平均方差矩阵。 L 代指模型中的高斯分布个数，取值通常在

3个~5个。

从GMM模型的聚类分析看，其具体步骤分为以下几步：

在初始化中，基于高斯混合分布的模型参数被定义为 W_i 、 u_i 和 Σ_i 。

准确计算样本 x_j 由各种混合成分所生成的后验概率，从观测数据 x_j 引进了一个隐变量 $z_j \in \{1, 2, \dots, k\}$ ，从而得出基于样本 x_j 的高斯分布模型。然后，从第 i 个分模型会生成概率在 $p(z_j=i|x_j)$ ，用公式表示为：

$$\frac{W_i \cdot p(x_j|u_i, \Sigma_i)}{\sum_{i=1}^k W_i \cdot p(x_j|u_i, \Sigma_i)}$$

新模型参数如下：

$$\begin{aligned} \mu_i' &= \frac{\sum_{j=1}^m \gamma_{ij} x_j}{\sum_{j=1}^m p(z=i|x_j)} \\ W_i' &= \frac{\sum_{j=1}^m \gamma_{ij}}{m} \\ \Sigma_i' &= \frac{\sum_{j=1}^m \gamma_{ij} (x_j - \mu_i')(x_j - \mu_i')^T}{\sum_{j=1}^m \gamma_{ij}} \end{aligned}$$

在以上公式中， μ_i' 是一种向量，属于新生成的均值， Σ_i' 是新生成的协方差矩阵， W_i' 是一种新混合系数。

参照新模型的参数，重复操作第二步和第三步，一直到满足最后的条件。

依据 $\lambda_i = \gamma_{ij}$ 的标准，把每个样本归入相对应的簇。与此同时，需要在GMM模型中分别代入每个样本，做好每一个模型概率的计算工作。

必须注意的是，每一个样本来自哪一个分模型，就需要将计算出的概率归入该分模型所对应的簇中，最后会得出 k 个聚类。

2.2 企业复工情况

对于春节期间停工的企业，在春节后复工过程中，这些企业的用电情况会随着实际经营情况出现波动，其主要用电曲线呈现出如下特点。春节停工期间，用电量相对较少，因此用电量处于较低水平，且呈现较为平缓的状态，但是当企业复工之后，在复工节点时，企业用电量将会骤增。

基于春节停工企业复工后其用电量变化这一特点，分析企业复工情况时，要结合这部分企业历年春节前后用电情况和特点，判断用电量拐点出现的时间，然后计算非疫情防控期间，企业在春节后达到复工状态时的用电情况，再计算复工状态下企业用电量与春节前平均用电量之间的比值，将其作为判断疫情防控期间企业复工情况的依据。

企业复工情况分析主要采用Knee point算法，计算公式为：

$$K_i(x) = \frac{f''(x)}{[1+f'(x)]^{2.5/2}}$$

其中, $f(x)$ 为任意连续函数; $K_f(x)$ 为该函数所对应的标准闭形式, 将 $f(x)$ 在任何点的曲率定义为其一阶和二阶导数的函数^[2]。

Knee point算法的具体计算流程如下:

第一, 使用多项式插值法拟合元数据, 在此过程中, 要尽量确保源数据的集合形式不被破坏, 构建平滑曲线上点的有限集。

第二, 为防止异常值对平滑曲线的曲率产生不良影响, 要先对曲线上的点进行归一处理。

第三, 确定平滑曲线出现急剧变化的点, 并构建差分数据集。

第四, 通过计算确定曲线中的拐点, 并在拐点位置计算差分曲线当中的局部最大值。

第五, 针对上一环节计算得到的每一个局部最大值, 使用相应模型敏感度参数和连续值的平均差, 定义唯一阈值。

第六, 检测上述计算得到的拐点值, 以此确保拐点计算的准确性^[3]。

此外, 电力企业为了加强电力指数管理, 提供更优质服务, 非常重视加强电力营销信息化建设, 设计完善的电力营销管理系统(该系统包括电力企业用电检查管理系统), 搭建综合性营销服务管理平台。从整体结构分析, 完善电力企业电力营销管理系统, 应做好以下四项基本工作。

第一, 优化系统业务功能。在综合性营销服务管理平台中, 企业电力营销系统业务功能主要包括销售管理、系统登录与权限管理、客户信息管理、质量服务管理、支付交易管理等。

其中, 销售管理业务细分为普通订单管理、特殊订单管理、生成订单合同管理、订单合同变更管理、订单审核管理、订单报表查询管理。系统登录与权限管理主要是为用户提供信息登记和权限设置服务, 通过合法验证确保系统安全。客户信息管理主要是完成客户信息登记工作, 包括变更信息管理, 同时, 做好客户信息安全保密工作。质量服务管理包括电表以及各种电力设备保修、旧件管理、零配件管理。支付交易管理由电费核算服务管理、电费缴纳管理、交易结算管理组成。

第二, 设计完善的营销宣传管理系统。该系统主要负责电力营销服务产品以及企业文化宣传工作, 树立电力企业良好的市场形象, 通过微信、公众号留言、微博、互联网等媒介宣传电力企业信息和营销方针策略, 促进客户和

电力企业营销人员的沟通。

第三, 设计质量服务管理系统。设计该系统时, 不仅需要纳入电力设备保修服务管理、旧件管理、零配件管理等模块, 而且应设置电力企业质量咨询管理部门, 该部门服务内容包括企业文化服务、电力科技服务、电力安全教育、用电服务体验、及时服务和营销动态需求信息服务等。

第四, 优化支付交易服务管理系统。该系统分为四大模块, 分别是电费核算管理模块、交易结算管理模块、电费缴纳管理模块和电费报销管理模块。

其中, 电费核算管理细分为用户电费核算资料收集服务、区域电费核算信息收集服务。交易结算管理模块由应收应付管理、电费交易管理、户对账管理、日记账管理共同组成。电费缴纳管理模块主要是对不同电费缴纳方式与信息的统一管理。电费报销管理模块主要负责“煤改电”和电费补贴优惠管理工作。

3 结语

综上所述, 为帮助政府在疫情防控期间充分了解企业复工情况, 电力企业要积极利用自身资源, 在电力大数据的支持下, 针对用电企业计算电力指数, 并以此确定企业复工情况, 通过对复工企业比例和复工用电量比例的计算, 判断企业复工情况。

在春节停工期间, 该企业的用电量相对较少, 因此用电量处于较低水平上, 且呈现出较为平缓的状态。但是当企业复工之后, 在复工节点时, 企业用电量将会呈现骤增情况。

随着对用电指数以及企业用电特点的深入研究和分析, 电力大数据将在企业复工复产相关政策制定过程中起到至关重要的作用。

参考文献

- [1] 陈万波, 李根东, 宋弘景, 等. 基于某电力企业特大事故整改复工历程的思考[J]. 能源研究与管理, 2021(1):6-11.
- [2] 温苑红, 林韵. 电力施工企业应对疫情影响的发展策略研究[J]. 农电管理, 2020(4):31-34.
- [3] 聂新伟, 史丹, 蔺通. 新型冠状病毒疫情对我国电力行业的影响分析[J]. 中国能源, 2020, 42(2):25-31, 14.