

基于用户用电行为聚类与经济指标关联的工业 sector 中长期电力需求预测方法

冯胜涛, 马宏图, 杨娜, 董宇, 王少敏
国网邯郸供电公司经研所, 河北 邯郸 056001
DOI:10.61369/EPTSM.2025060018

摘 要 : 本文聚焦工业 sector 中长期电力需求预测方法, 提出一种融合用户用电行为聚类与经济指标关联的预测方法。在数据层面, 整合多源异构的工业用户用电数据、生产数据及外部环境数据, 通过数据清洗、标准化与特征工程, 构建了反映用户用电特性的高维特征体系。在用户画像层面, 综合比较 K-Means、层次聚类、DBSCAN 等算法, 选取最优聚类模型对工业用户进行细分, 识别并提炼出稳定型、波动型、高耗低效型等典型用电模式, 为差异化预测奠定基础。在经济关联层面, 筛选出工业增加值、产业结构、单位产值电耗等关键经济指标, 运用相关性分析、回归模型及时间序列分析等方法, 量化各指标与不同用电模式群体电力需求的内在关联机制与动态影响。在模型构建层面, 设计“分群-关联-预测”的融合框架, 针对不同用电模式群体建立差异化预测子模型, 并通过加权集成形成总需求预测。研究表明, 该方法通过精细化用户分群与深度经济关联分析, 能够有效捕捉工业用电的复杂性与规律性, 相比传统单一模型, 显著提升了预测的准确性和稳定性, 可为电力系统规划、调度及能源政策制定提供更科学、可靠的决策支持。

关 键 词 : 电力需求预测; 中长期预测; 用户用电行为聚类; 经济指标关联

A Medium-to-long-term Electricity Demand Forecasting Method for the Industrial Sector Based on the Correlation between User Electricity Consumption Behavior Clustering and Economic Indicators

Feng Shengtao, Ma Hongtu, Yang Na, Dong Yu, Wang Shaomin
Economic Research Institute of State Grid Handan Power Supply Company, Handan, Hebei 056001

Abstract : This paper focuses on medium- to long-term electricity demand forecasting methods for the industrial sector and proposes a forecasting approach that integrates user electricity consumption behavior clustering with economic indicator correlations. At the data level, it integrates multi-source heterogeneous industrial user electricity consumption data, production data, and external environmental data. Through data cleaning, standardization, and feature engineering, a high-dimensional feature system reflecting user electricity consumption characteristics is constructed. At the user profiling level, it comprehensively compares algorithms such as K-Means, hierarchical clustering, and DBSCAN, selecting the optimal clustering model to segment industrial users. It identifies and refines typical electricity consumption patterns, such as stable, fluctuating, and high-consumption-low-efficiency types, laying the foundation for differentiated forecasting. At the economic correlation level, key economic indicators including industrial added value, industrial structure, and electricity consumption per unit of output value are screened. Methods such as correlation analysis, regression models, and time series analysis are employed to quantify the internal correlation mechanisms and dynamic impacts of these indicators on the electricity demand of different electricity consumption pattern groups. At the model construction level, a "clustering-correlation-forecasting" integrated framework is designed. Differentiated forecasting sub-models are established for different electricity consumption pattern groups, and a total demand forecast is formed through weighted integration. Research indicates that this method, through refined user clustering and in-depth economic correlation analysis, can effectively capture the complexity and regularity of industrial electricity consumption. Compared to traditional single models, it significantly enhances the accuracy and stability of forecasts, providing more scientific and reliable decision support for power system planning, scheduling, and energy policy formulation.

Keywords : electricity demand forecasting; medium-to-long-term forecasting; clustering of user electricity consumption behavior; correlation with economic indicators

引言

中长期电力需求预测不仅需反映工业生产的季节性、周期性规律，更需适配产业结构调整、技术升级及宏观经济波动带来的动态变化，是保障电力系统安全稳定运行、推动电力企业高质量发展的关键前提。本文提出一种融合用户用电行为聚类与经济指标关联的工业领域中长期电力需求预测方法，该方法通过多源用电数据的清洗与特征提取，结合 K-Means、DBSCAN 等聚类算法对工业用户进行分群，识别稳定型、波动型、高耗低效型等差异化用电模式；筛选工业增加值、产业结构占比、单位工业增加值电耗等关键经济指标，运用灰色关联分析、VAR 模型等方法量化其与各用电群体电力需求的动态关联机制；构建“分群—关联—预测”的一体化模型，实现从微观用电行为到宏观经济趋势的多维度耦合预测。研究为工业领域中长期电力规划、电力市场交易及能源政策制定提供科学、精准的决策依据，助力新型电力系统在工业领域的落地应用。

一、用户用电行为聚类剖析

（一）用电行为数据收集与整理

随着智能配电网信息化、自动化、互动化程度的不断提高，各种先进的配用电自动化和管理系统得到广泛的应用，由此产生并积累了海量多源异构数据。对这些数据进行有效挖掘和科学合理地利用，可以有效提升智能配电网的运行管理水平，同时也是大数据背景下电力企业发展的必然要求^[1]。在用电行为数据收集与整理环节，需明确数据收集范围，按制造业、重工业、轻工业等细分用户群体进行区分，避免数据混杂导致特征失真，同时确定时间维度，中长期预测需覆盖3-5年历史数据以反映季节性与周期性规律，短期数据则用于捕捉突发波动^[2]。数据来源应整合多维度信息，包括智能电表采集的实时负荷数据、用户生产调度系统中的生产计划与产能数据、电力营销系统中的用户基本信息及缴费记录，以及气温、湿度、节假日等外部环境数据，并明确各数据源的采集频率、格式与传输方式，保障数据获取的及时性与稳定性^[3]。数据预处理是关键步骤，先进行数据清洗，通过箱线图等可视化方法识别异常值，结合生产日志排查原因后采用插值或删除处理；针对缺失值，缺失比例低时用均值、中位数或相邻数据填充，比例高时则基于历史用电模式构建预测模型填充。接着进行数据标准化或归一化处理，采用 Z-score 或 Min-Max 方法消除量纲差异，对时间序列数据进行平稳性检验，非平稳数据通过差分、对数变换转为平稳序列^[4]。最后进行数据特征提取，构建包含时间维度、生产关联和用户属性的特征体系，并通过主成分分析、因子分析等方法对高维特征降维，保留核心特征以降低计算复杂度。

（二）聚类算法原理与选择

主流聚类算法各有特点，选择时需结合数据特征、规模与业务需求。划分式聚类以 K-Means 为代表，通过距离度量将数据划分为 K 个簇，计算效率高，适合大规模用电数据，但对初始中心敏感且难以处理非球形分布^[5]。层次式聚类如层次凝聚聚类，通过构建树状结构展示数据层级关系，无需预设簇数，但计算复杂度高，不适合大规模数据。密度聚类如 DBSCAN，基于数据点密度识别任意形状簇，对异常值不敏感，但对参数设置依赖强，在密度不均匀数据中表现不佳。算法选择需综合考虑数据分布特征、规模大小及簇数确定方法，如肘部法则、轮廓系数或业务经验，确保聚类结果具备实际业务意义。

（三）聚类结果分析与用电模式识别

聚类有效性评估需从内部与外部两方面开展，内部评估通过簇内紧凑性与簇间分离度衡量聚类质量，常用 Davies-Bouldin 指数和 Calinski-Harabasz 指数等指标^[6]。外部评估则结合业务知识，将聚类结果与用户行业、规模、工艺等已知信息匹配，验证同一簇内用户业务属性是否相似，不同簇间是否存在显著差异，若出现行业混杂则需回溯数据或算法环节。用电模式特征提取从时间序列与经济关联两方面入手，时间序列分析通过绘制日、周、季节负荷曲线，识别稳定型、波动型和突发型等模式；经济关联分析结合产值、能耗等数据，提炼高效型、高耗低效型和规模驱动型等模式^[7]。用电模式分类与标签化需基于聚类结果与特征分析，构建涵盖时间、经济、行业等维度的标签体系，并为每个标签设定明确阈值标准，确保新增用户可快速匹配模式标签，为中长期电力需求预测提供依据。

二、经济指标关联研究

（一）关键经济指标选取

关键经济指标的选取需紧密贴合工业电力需求的影响逻辑，确保指标能精准反映工业生产规模、结构变化及发展趋势，同时兼顾数据的可获取性、时效性与连续性^[8]。工业增加值是核心指标，直接体现产出总量并与电力消耗显著正相关，辅以工业总产值验证关联强度；产业结构方面需关注高耗能产业产值占比和战略性新兴产业占比，分析结构调整对电力需求的差异化影响，其中高耗能产业拉高需求，新兴产业可能提升效率；投入与效率维度则以单位工业增加值电耗为核心，反映用电集约化程度，其下降会抑制需求增长，同时考虑工业固定资产投资，特别是设备与产能投资，预示未来电力需求变化；宏观环境方面引入地区 GDP 及增速作为背景指标，反映经济活跃度对工业的带动作用，并纳入大宗商品价格指数，其波动影响企业生产意愿与盈利水平，间接作用于电力需求。指标筛选过程中需剔除关联性弱或数据质量差的指标，通过相关性检验优化，形成层次清晰、逻辑连贯的指标体系，为后续关联分析奠定基础。

（二）经济指标与电力需求的关联分析方法

经济指标与电力需求的关联分析需综合运用多种方法，以全面揭示线性与非线性、静态与动态的内在联系^[9]。相关性分析中，Pearson 系数适用于线性关系判断，Spearman 等级相关系数

则可捕捉单调非线性关联；灰色关联分析在样本量不足时能有效评估各指标对电力需求的影响权重。回归分析进一步量化影响程度，线性回归模型可明确各指标单位变化对电力需求的具体作用，非线性回归或引入交互项则可提升复杂关系的拟合度。时间序列分析聚焦动态关联，通过向量自回归模型（VAR）和脉冲响应函数研究经济指标波动对电力需求的滞后影响，协整检验与误差修正模型（ECM）则揭示长期均衡与短期调整机制，为动态预测提供支持。

（三）经济指标的预测与趋势分析

经济指标的预测与趋势分析是衔接经济关联研究与电力需求预测的关键环节，需根据指标特性选择适宜方法，并考虑宏观环境的不确定性影响^[10]。时间序列法适用于数据连续、趋势稳定的指标，移动平均法可平滑短期波动，Holt-Winters指数平滑法可处理趋势性与季节性波动；回归预测法适合多变量驱动指标，可引入政策虚拟变量量化政策影响；情景分析与德尔菲法则适用于数据少或受突发因素影响的指标，通过设定乐观、基准、悲观情景，结合专家意见提供多维度参考。在趋势分析中，需提炼指标的长期规律与短期波动特征，例如通过工业增加值增速预判工业生产整体趋势，奠定电力需求基础增速；通过产业结构比例变化评估对电力需求的抑制或拉动效应；通过单位工业增加值电耗趋势测算用电效率提升带来的节电量。同时需关注指标间联动关系，避免孤立解读，确保趋势分析全面支撑电力需求的中长期预测逻辑。

三、中长期电力需求预测模型构建

（一）融合用电行为聚类与经济指标的预测模型原理

融合用电行为聚类与经济指标的中长期电力需求预测模型，通过“分群—关联—预测”框架，打破单一维度局限，实现预测精度与解释性的双重提升。模型以用电模式分层为基础，将聚类得到的稳定型、波动型、高耗低效型等群体作为分组依据，分别提取用电行为特征与经济关联特征，构建多维特征体系。针对不同群体采用差异化建模策略，稳定型用户适用线性回归或时间序列模型；波动型用户引入非线性模型并考虑经济指标滞后影响。通过加权集成等方法融合各群体预测结果，形成整体需求预测。模型具备动态适配能力，可在用户模式或经济结构变化时自动调整特征与参数，确保预测结果既贴合实际用电行为，又反映经济发展趋势。

（二）模型参数确定与优化

模型参数的确定与优化是保障预测精度的关键环节，需遵循

“初步设定—迭代优化—验证筛选”的流程，确保参数既符合模型数学逻辑，又适配工业电力需求实际变化规律。在初步设定阶段，根据模型类型与特征属性设定参数初始范围，如线性回归模型需设定正则化参数，非线性模型如SVR需确定核函数参数与惩罚系数，时间序列模型如ARIMA则通过ACF与PACF图判断自回归、差分与移动平均阶数，集成学习权重可按用电规模占比分配。在迭代优化过程中，采用网格搜索、随机搜索或贝叶斯优化等方法，结合交叉验证策略评估参数性能，避免过拟合，同时需平衡精度与可解释性，确保经济指标相关参数符合逻辑。最终形成参数文档，记录参数含义、优化过程、最优值及性能指标，为模型复现与更新提供依据。

（三）模型验证与评估

模型验证与评估需从准确性、稳定性和适用性三个维度展开，通过多组验证实验与量化指标全面检验模型性能，并结合业务场景判断其是否满足中长期电力需求预测的实际需求。在准确性验证中，采用历史数据回测结合MAE、RMSE、MAPE等指标评估预测误差，确保误差分布合理且精度达标；在稳定性验证中，通过数据扰动、时间分段和场景迁移测试，检验模型在不同条件下的性能一致性；在适用性评估中，关注模型的可解释性、预测周期适配性及决策支持能力，确保预测结果能为电力规划提供有效依据。最终形成验证评估报告，若模型满足要求则投入使用，否则需迭代优化直至达标。

四、结束语

本研究围绕工业领域中长期电力需求预测的精准性与实用性需求，构建了一套融合用户用电行为聚类与经济指标关联的预测方法，系统解决了传统预测中维度单一、忽视用户异质性及宏观经济关联不足的核心问题。从应用价值来看，本研究提出的方法不仅有效提升了工业领域电力需求预测的精度，更通过用电模式标签化与经济关联机制分析，为电力企业提供了兼具数据支撑与逻辑解释的决策工具。研究构建的预测方法为工业领域中长期电力需求分析提供了新的思路与范式，后续可通过更多实证数据验证与多维度优化，持续完善方法体系，使其在电力系统规划、能源政策制定等领域发挥更大作用，为工业能源转型与电力行业高质量发展提供更坚实的技术支撑。

参考文献

- [1] 陈俊艺. 计及用户用电行为聚类的不确定性需求响应概率模型研究 [D]. 湖北：武汉大学，2019.
- [2] 黄薇，温蜜，张照贝. 考虑用户用电行为聚类的电力负荷预测方法 [J]. 计算机仿真，2022, 39(12): 148–153. DOI: 10.3969/j.issn.1006-9348.2022.12.027.
- [3] 汪颖，杨维，肖先勇，等. 基于去噪自编码器网络特征降维与改进小批优化K均值算法的海量用户用电行为聚类及分析 [J]. 电力自动化设备，2022, 42(6): 146–153. DOI: 10.16081/j.epae.202203017.
- [4] 冯志颖，唐文虎，吴青华，等. 考虑负荷纵向随机性的用户用电行为聚类方法 [J]. 电力自动化设备，2018, 38(9): 39–44, 53. DOI: 10.16081/j.issn.1006-6047.2018.09.007.
- [5] 于文龙. 基于用电特征选择的用户用电行为聚类分析及应用 [D]. 湖南：湖南大学，2020.
- [6] 龚钢军，陈志敏，陆俊，等. 智能用电用户行为分析的聚类优选策略 [J]. 电力系统自动化，2018, 42(2): 58–63. DOI: 10.7500/AEPS20170726009.
- [7] 李春燕，蔡文悦，赵落生，等. 基于优化SAX和带权负荷特性指标的AP聚类用户用电行为分析 [J]. 电工技术学报，2019, 34(z1): 368–377. DOI: 10.19595/j.cnki.1000-6753.tces.L80097.
- [8] 国网上海市电力公司，上海电力大学. 一种基于用户用电行为聚类的电力负荷概率预测方法：CN202110635884.8[P]. 2021-09-07.
- [9] 朱炎平. 基于聚类的用户用电行为分析研究 [D]. 华北电力大学，华北电力大学（北京），2017. DOI: 10.7666/d.Y3264136.
- [10] 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. DOI: 10.3969/j.issn.1009-1831.2019.03.012.