

DOI: 10.19783/j.cnki.pspc.241291

基于分行业用电特性与多因素影响的区域级 短期用电负荷曲线预测

郭耀扬¹, 张利¹, 郝颖², 赵波¹, 周颖³, 马笑天⁴, 李闯¹

(1. 北京信息科技大学自动化学院, 北京 100192; 2. 北京理工大学唐山研究院, 河北 唐山 063000; 3. 需求侧多能互补优化与供需互动技术北京市重点实验室(中国电力科学研究院有限公司), 北京 100192; 4. 国网河北营销中心, 河北 石家庄 050035)

摘要: 针对“双碳”目标推进下, 区域级全社会用电负荷曲线预测精度低的问题, 提出了一种基于分行业用电特性与多因素影响的全社会用电负荷曲线预测框架。首先, 结合行业负荷曲线聚类分析结果、综合用电评价指标构建方法, 从定性和定量两方面对不同行业用电特性进行分析, 验证分行业精细化用电特性挖掘的必要性。然后, 采用非线性相关系数、小提琴图可视化分析方法, 量化不同季节气温、日类型对不同行业用电特性的影响程度, 为后续分行业精细化预测奠定数据基础。最后, 利用卷积神经网络(convolutional neural network, CNN)提取行业用电特性, 结合双向长短期记忆网络(bi-directional long short-term memory, BiLSTM)与注意力机制(attention mechanism, Attention)构建集成预测模型, 精细化预测各行业负荷曲线, 采用间接预测方式实现全社会用电负荷曲线预测。结合华东某区域十一个行业以及居民用电负荷数据, 设置10组实验进行对比分析, 结果表明所提预测框架较传统预测方法能够显著降低预测误差。

关键词: 行业用电特性; 综合用电评价指标; 小提琴图; 负荷曲线预测

Regional short-term electricity load curve forecasting based on industry electricity consumption characteristics and multi-factor effects

GUO Yaoyang¹, ZHANG Li¹, HAO Ying², ZHAO Bo¹, ZHOU Ying³, MA Xiaotian⁴, LI Chuang¹

(1. School of Automation, Beijing Information Science and Technology University, Beijing 100192, China; 2. Tangshan Research Institute, Beijing Institute of Technology, Tangshan 063000, China; 3. Beijing Key Laboratory of Demand Side Multi-energy Carriers Optimization and Interaction Technique (China Electric Power Research Institute), Beijing 100192, China; 4. State Grid Hebei Marketing Service Center, Shijiazhuang 050035, China)

Abstract: Aiming at the problem of low accuracy of regional level total electricity load curve forecasting under the “dual carbon” goals, a forecasting framework that incorporates industry-specific electricity consumption characteristics and multi-factor influences is proposed. First, industry load curve clustering analysis and the construction of comprehensive electricity consumption evaluation indicators are used to qualitatively and quantitatively analyze the electricity consumption characteristics of different industries, demonstrating the necessity of refined, industry-level load profiling. Furthermore, the impact of various external factors, such as seasonal temperatures and day types, on industry-specific electricity usage patterns is quantified using nonlinear correlation coefficients and violin plot visualizations, laying a data foundation for fine-grained forecasting. Finally, a hybrid forecasting model is constructed by extracting electricity usage features via convolutional neural network (CNN), integrating them with the bi-directional long short-term memory (BiLSTM) network and the attention mechanism. This ensemble model is used to predict load curves for each industry, and an indirect forecasting approach is applied to reconstruct the total electricity load curve of the entire society. Using load data from eleven industries and residential electricity users in a region of East China, ten groups of comparative experiments are conducted. Results show that the proposed forecasting framework significantly reduces the prediction errors compared to traditional methods.

This work is supported by the Science and Technology Project of the Headquarters of State Grid Corporation of China (No. 5108-202218280A-2-379-XG).

Key words: industry electricity consumption characteristics; comprehensive electricity consumption evaluation index; violin diagram; load curve forecasting

基金项目: 国家电网公司总部科技项目资助(5108-202218280A-2-379-XG)

0 引言

电力负荷预测作为电力系统规划和运行管理的重要组成部分, 精准的预测结果对保障电力系统的稳定运行至关重要。近年来, 伴随着新型电力系统的持续建设以及电力市场改革的持续深化, 电网运行方式由“源随荷动”的控制模式逐渐转变为源-网-荷储协同互动模式, 电力负荷的波动性和随机性逐渐增大, 极大提升了负荷预测的难度。随着“双碳”目标的纵深推进、减碳政策的陆续出台, 给能源、工业、建筑和交通等重点行业带来深刻影响, 从而对行业用电特性产生影响。

目前, 伴随着人工智能技术的快速发展, 以数据驱动为基础的人工智能方法逐渐成为电力负荷预测的主要研究方法, 其中机器学习、深度学习由于具备良好的提取复杂抽象特征的能力, 逐渐成为主流负荷预测方法^[1-3]。文献[4]提出了一种时频尺度下的时间卷积网络与多元线性回归相融合的组合预测模型, 解决了随机因素造成的电力负荷成分愈发复杂、预测难度增大的问题; 文献[5]利用特征提取技术, 结合多种机器学习方法, 提出了一种新的混合预测模型 EWT-CNN-S-RNN-LSTM 模型, 提高了短期电力负荷预测的准确性; 文献[6]针对单一模型预测稳定性差的问题, 提出了一种基于改进 Q 学习算法的组合预测模型, 提升了在负荷波动场景下超短期负荷预测的精度; 文献[7]提出了一种基于空间自相关和卷积长短期记忆的新算法, 为城市配电网规划提供了参考; 文献[8]提出了一种基于神经基扩展分析与辅助编码器的短期电力负荷预测模型, 提取出与电力负荷存在潜在相关性的辅助信息特征, 从而提升了预测精度; 文献[9]通过分析数据间相互影响的内在联系和规律, 提高了短期负荷预测的精度; 文献[10]通过对负荷数据进行分解, 提取波动和趋势特征, 从而提升预测精度; 文献[11]提出了一种 CEEMDAN-SE-TR 模型, 通过对不同负荷分量进行学习, 提升了预测精度。总体来看, 当前负荷预测研究重点主要集中于对原始负荷数据进行特征提取并构建组合模型, 但对逐渐变化的行业用电特性研究较少。如: 文献[12]采用灰色关联度分析法定量分析多种影响因素对不同行业的影响程度, 分别预测不同行业的用电量; 文献[13]基于产业链间能源消耗关联性分析结果, 构建了用电量预测模型; 文献[14]从短期扰动和历史存续两方面分析不同行业间月用电量的关联关系, 提出了一种新型的中长期电量预测方法; 文献[15]结合收敛交叉映射算法, 辨识各行业用电需求变动的传导关系, 提出了内嵌

因果传导的行业电量预测方法。可以看出, 目前关于行业用电特性的研究基本都集中于用电量方面, 且多为中长期预测, 对于短期负荷预测研究较少。

因此, 本文聚焦于行业用电特性, 提出一种区域级用电负荷精细化预测框架。首先, 从定性和定量两个方面分析行业自身用电特性。通过聚类分析法构造聚类曲线, 对行业用电特性进行定性分析; 采用熵权法结合多个行业用电指标, 构建综合用电评价指标, 对行业用电特性进行定量分析。接着, 从影响因素方面分析不同因素对行业用电特性的影响: 采用斯皮尔曼相关系数(Spearman's rank correlation coefficient)分析气温与不同行业用电的相关性, 用以表征气温对不同行业用电影响的强弱; 引入小提琴图可视化日类型对不同行业的影响, 提出差异度指标, 量化日类型对不同行业用电特性的影响程度。进一步地, 为提取不同行业用电特性, 采用卷积神经网络(convolutional neural network, CNN)对不同行业进行精细化提取, 同时引入注意力机制, 构建 CNN-BiLSTM-Attention 预测模型, 分别预测各行业日负荷曲线, 采用间接预测方式实现对全社会用电日负荷曲线进行预测。最后, 分别设置 9 组对照实验与本文方法进行对比, 实验结果证明了所提方法的有效性。

1 行业用电特性分析

1.1 定性分析

行业用电需求一方面受到政策、产业链供给波动等外部因素影响, 另一方面受到生产工艺流程、技术创新等内部因素影响。因此, 不同行业同一天的用电负荷曲线可能会呈现出显著差异。以工业、住宿和餐饮业为例进行说明, 首先利用肘部法则确定聚类中心数量, 使用 K 均值聚类法分别对两个行业进行聚类分析, 提取典型日负荷曲线。其中, 肘部法则通过寻找误差平方和(sum of squared errors, SSE)下降的拐点来判断最佳聚类数。两个行业的肘部法则结果图分别见图 1 和图 2。

根据图 1 和图 2 可以确定工业、住宿和餐饮业的最佳聚类中心数均为 3。因此, 两个行业均取聚类中心数量为 3, K 均值聚类结果见图 3。

从图 3 可以看出, 工业、住宿和餐饮业日用电规律存在显著差异。工业用电在 01:00—06:00 用电较为平稳, 06:00 后用电负荷逐渐攀升, 09:00 左右达到首个峰值; 之后开始回落, 于 11:00 左右降至日间首个谷点; 紧接着开始回升, 于 13:00 左右增速放缓, 并于 15:00 达到全天峰值; 17:00 开始大幅下降, 但在 18:00 后下降放缓, 21:00 后出

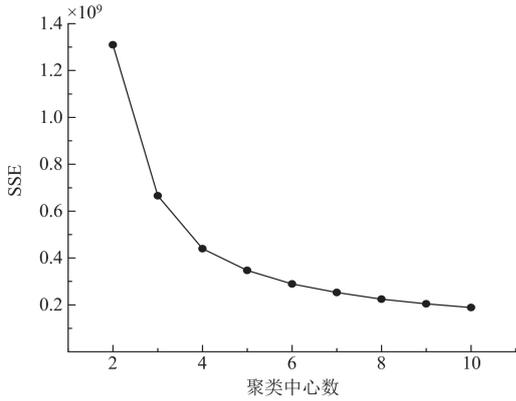


图1 工业负荷数据肘部法则结果图

Fig. 1 Graph of the results of elbow rule for industrial load data

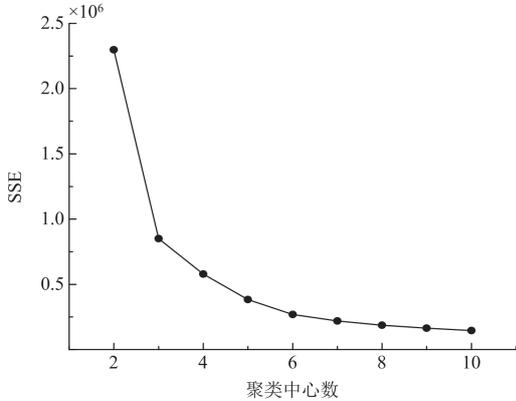


图2 住宿和餐饮业负荷数据肘部法则结果图

Fig. 2 Graph of the results of elbow rule for accommodation and food sector load data

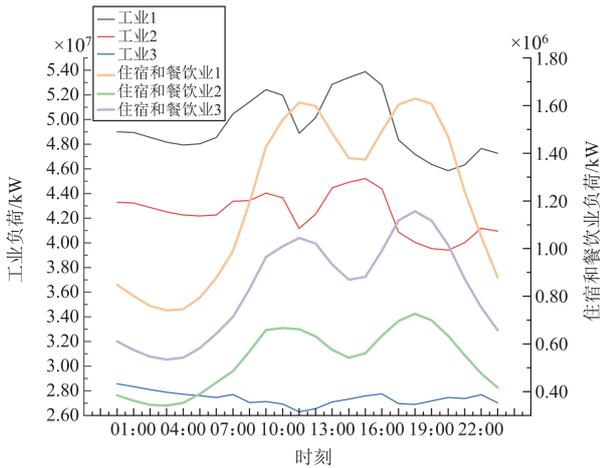


图3 工业和住宿和餐饮业聚类结果图

Fig. 3 Clustering results for industry and accommodation and food sector

现小幅上涨，整体呈现“双峰一谷”变化趋势。相较于工业，住宿和餐饮业用电规律较为简单，

01:00—04:00 呈现小幅度下降趋势，05:00 后大幅上升，10:00 增长放缓，并于 11:00 左右达到首个峰值；之后开始缓慢下降，于 14:00 达到日间谷点；15:00 再次回升，于 18:00 达到全天峰值，之后呈现下降趋势，整体同样呈现“双峰一谷”变化趋势。通过分析可以看出，虽然工业和住宿餐饮业日负荷曲线整体都呈现“双峰一谷”变化趋势，但由于各行业的特性以及外部因素的影响，二者在某些时段有着截然不同的用电变化规律。因此，为量化分析这种变化规律，提出了行业综合用电评价指标来定量分析不同行业的用电特性。

1.2 定量分析

本节参考文献[16]提出的方法，采用熵权法对波动度、冲击度、日峰谷差率和日负荷率等指标进行综合分析，构建行业用电特性的量化评价体系。波动度用于表征行业用电负荷的离散程度，其大小与负荷波动程度成正相关，即波动度越小说明负荷波动程度越小，负荷用电越平稳。波动度表示为

$$S_1 = \sqrt{\frac{\sum_{t=1}^N (P'_t - P'_{\text{mean}})^2}{N-1}} \quad (1)$$

式中： S_1 表示波动度； P' 表示归一化后的负荷数据； P'_{mean} 表示归一化平均值； N 表示采样点数。

冲击负荷是指变化幅度较大且具有周期性或非周期性的负荷，其数值可达到平均负荷的数倍。此类负荷通常持续时间很短，但峰值很高。冲击度表示为

$$S_2 = \sqrt{\frac{\sum_{t=1}^N (P_t - P_t^{\text{smooth}})^2}{N}} \quad (2)$$

式中： S_2 表示冲击度； P_t 表示负荷曲线数据； P_t^{smooth} 表示平滑后的负荷曲线数据。为保证各行业数据范围一致，可对 S_2 取归一化后进行对比。

日峰谷差率是指在一天之内电力系统日峰谷差与最大负荷的比值，其从极值的角度衡量了电力系统在一天内负荷波动的大小。日峰谷差率表示为

$$S_3 = \frac{P_t^{\text{max}} - P_t^{\text{min}}}{P_t^{\text{max}}} \quad (3)$$

式中： S_3 表示日峰谷差率； P_t^{max} 表示第 t 天日负荷曲线最大值； P_t^{min} 表示第 t 天日负荷曲线最小值。

日负荷率指一天之内平均负荷与最大负荷的比值，是一个重要的电力系统指标。日负荷率表示为

$$S_4 = \frac{P_t^{\text{mean}}}{P_t^{\text{max}}} \quad (4)$$

式中: S_4 表示日负荷率; P_t^{mean} 表示第 t 天日负荷曲线的平均值。

以上 4 个指标中, 波动度、冲击度、日峰谷差率越大, 表示行业用电负荷曲线越不平稳; 日负荷率越大则表示行业用电负荷曲线越平稳。为综合评价各指标对行业用电负荷曲线的表征能力, 引入熵权法计算各指标权重, 得到综合用电评价指标。

熵权法^[17]是一种客观赋权方法, 用于确定多指标综合评价系统中各指标的权重。其基于信息熵的概念, 通过计算每个指标的信息熵来衡量其在评价体系中的相对重要性。信息熵越小, 表明该指标的变异性越大, 所提供的信息量越多, 因此在综合评价中会赋予更大的权重。熵权法的基本流程如下。

1) 数据标准化

为消除各指标量纲和数值范围可能存在不同的情况, 需要对各指标进行标准化处理, 从而确保各指标量级一致。

2) 计算指标比重

各指标的比重计算公式为

$$r_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}} \quad (5)$$

式中: r_{ij} 表示第 i 个评价对象在第 j 个指标上的比重; x_{ij} 表示第 i 个评价对象在第 j 个指标上的标准化数据; m 表示评价对象总数。

3) 计算指标熵值

利用比重计算每个指标的熵值, 计算公式为

$$e_j = -\frac{1}{\ln m} \sum_{j=1}^m r_{ij} \ln r_{ij} \quad (6)$$

式中: e_j 表示第 j 个指标的熵值, 若 $r_{ij} = 0$, 则 $r_{ij} \ln r_{ij} = 0$ 。

4) 计算指标的差异系数

利用差异系数表示指标的有效性, 计算公式为

$$g_j = 1 - e_j \quad (7)$$

式中: g_j 表示第 j 个指标的差异系数。

5) 确定指标的权重

根据差异系数计算每个指标的权重, 计算公式为

$$w_j = \frac{g_j}{\sum_{j=1}^h g_j} \quad (8)$$

式中: w_j 表示第 j 个指标权重; h 表示指标的个数。

综合用电评价指标 S 为波动度、冲击度和日峰谷差率的加权平均值, 计算公式为

$$S = \sum_{j=1}^h w_j S_j \quad (9)$$

2 影响因素分析

常见的负荷影响因素有气温、日类型、突发事件等。其中, 气温主要通过温度变化改变用电行为, 从而对负荷产生影响; 日类型主要由于日常工作变化对负荷产生影响; 突发事件通常充满高随机性、偶然性, 如地震、洪水等, 对负荷有着很大的影响, 但通常影响时间较短。对于行业负荷而言, 气温变化通过影响设备环境温度及人体舒适度, 控制降温或取暖设备启停, 进而形成温度敏感型负荷; 日类型影响体现在是否需要工作, 主要影响基础负荷。研究时通常需要考虑这两种影响因素。而突发事件由于发生频次少、数据少, 一般只在特定情形下考虑。因此, 本文只考虑气温和日类型的影响。

2.1 气温影响

气温作为影响负荷的重要因素^[18-19], 其对负荷的影响主要体现在夏季的空调降温负荷和冬季的电采暖负荷等。目前, 主要通过皮尔逊相关系数大小来分析气温对负荷影响程度的强弱, 有关气温与负荷之间的非线性关系研究较少^[20]。另一方面, 皮尔逊相关系数要求数据服从正态分布, 而气温与负荷数据往往难以满足该假设。相比之下, Spearman 相关系数无需满足正态分布条件, 就能够有效分析非线性相关关系, 因此选择 Spearman 相关系数分析气温与负荷曲线数据之间的相关性。Spearman 相关系数 r_s 计算公式为^[21]

$$r_s = 1 - \frac{6 \sum_{k=1}^n d_k^2}{n(n^2 - 1)} \quad (10)$$

式中: n 为样本数量; d_k 为第 k 对观测值的秩次差。

通常情况下, 当 r_s 在 0~0.2 之间时, 表示两序列间相关性极弱或不相关; 在 0.2~0.4 之间时, 表示相关性较弱; 在 0.4~0.6 之间时, 表示相关性中等; 在 0.6~0.8 之间时, 表示相关性较强; 在 0.8~1 之间时, 表示相关性极强^[22]。

2.2 日类型影响

日类型对负荷的影响通常体现在工作日和休息日之间的差异。通常情况下, 工作日由于日常生产、办公等工作的开展, 导致其负荷较休息日更高。由于日类型数据为离散型数据, 负荷数据为时序数据, 而相关系数主要用于分析两个连续序列间的相关性, 故使用相关系数分析日类型和负荷之间相关性时, 可能会出现相关系数很低的情况, 因此考虑采用可视化工具分析不同日类型下负荷的模式和差异。

小提琴图是一种结合了箱线图和密度图优点的数据可视化方法，能够在有限的空间内展示大量数据分布特征，非常适合用于比较和分析不同类别的数据分布。小提琴图通常由箱线图部分和密度曲线两种结构组成，箱线图位于小提琴图中心部分，显示了数据的中位数、上下四分位数，以及可能的异常值(通常用点表示)；密度曲线位于箱线图外沿，曲线的宽度代表了数据点的密度估计。在小提琴图的“颈部”(箱线图的区域)，由于数据点较少，密度曲线较窄。而在“身体”部分，数据点密度较高，密度曲线变宽。小提琴图示意图见图4。

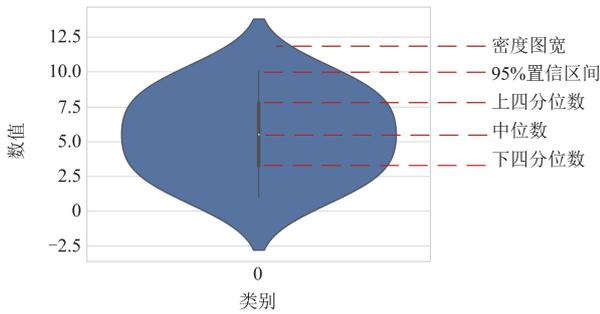


图4 小提琴图示意图

Fig. 4 Schematic diagram of violin diagram

通过对比工作日和休息日行业负荷小提琴图中的中位数和密度图宽，即可分析两种不同日类型下行业负荷的差异，从而体现日类型对负荷的影响。由于各行业在数据量级上相差较大，因此定义差异度对中位数和密度图宽进行对比分析，差异度公式为

$$D = \frac{|P_{mid,0} - P_{mid,1}|}{(P_{mid,0} + P_{mid,1})/2} \quad (11)$$

式中： D 表示差异度； $P_{mid,0}$ 表示工作日负荷曲线中位数或密度图宽； $P_{mid,1}$ 表示休息日负荷曲线中位数或密度图宽。

3 预测框架

3.1 CNN

CNN是一种广泛使用的深度学习模型，包括卷积层、池化层以及全连接层。其中，卷积层用于提取特征；池化层用于特征信息约简，以防止过拟合；全连接层用于进行权重加总，并输出结果^[23]。这种层次化的结构使其能够捕捉到负荷数据中的局部模式和时间依赖性，因此选择CNN提取行业特征，为后续负荷预测奠定基础。

3.2 双向长短期记忆网络

双向长短期记忆网络(bi-directional long short-term memory, BiLSTM)是长短期记忆网络(long short-term memory, LSTM)的一种优化网络，由于传统的LSTM具有单向性，不能全面反映序列的关联关系，因此BiLSTM通过引入一个前向LSTM层和一个后向LSTM层，使得网络能够充分考虑过去和未来的信息，有利于进一步提高模型预测的精度^[24]，BiLSTM的结构见图5。

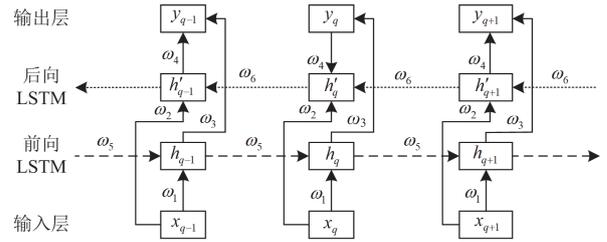


图5 BiLSTM网络结构图

Fig. 5 Structure of BiLSTM network

BiLSTM网络在 q 时刻的输出 y_q 由3部分决定，分别是 q 时刻的输入、前向LSTM网络前一刻的输出状态 h_{q-1} 以及后向LSTM网络后一刻的输出状态 h'_{q+1} 。具体公式为^[25]

$$h_q = z_1(\omega_1 x_q + \omega_5 h_{q-1}) \quad (12)$$

$$h'_q = z_2(\omega_2 x_q + \omega_6 h'_{q+1}) \quad (13)$$

$$y_q = z_3(\omega_3 h_q + \omega_4 h'_q) \quad (14)$$

式中： ω_p 为各层对应的权重， $p=1,2,\dots,6$ ； z_1 、 z_2 、 z_3 为不同层的激活函数。

3.3 注意力机制

由于时间序列输入较长时，容易出现信息丢失的情况，因此考虑使用注意力机制(attention mechanism, Attention)对不同特征进行动态权重调整，提升模型性能。具体如下：通过将BiLSTM层的输出输入至Attention层，采用加性注意力计算各个时间步的重要性评分，并通过softmax函数对隐层单元进行注意力打分，赋予各历史状态对应的注意力权重，最后采用加权平均得到注意力结果。

3.4 整体预测框架

本文构建的全社会用电负荷曲线预测框架的实现流程如下：首先，对各行业负荷进行定性与定量分析，判断行业用电特性是否存在差异；接着，分析各行业负荷与气温、日类型之间的关联关系，确定各行业负荷影响因素；然后构建预测模型分别对各行业进行预测，根据间接预测思想，将各行业预测结果累加得到全社会用电负荷预测值。具体流程见图6。

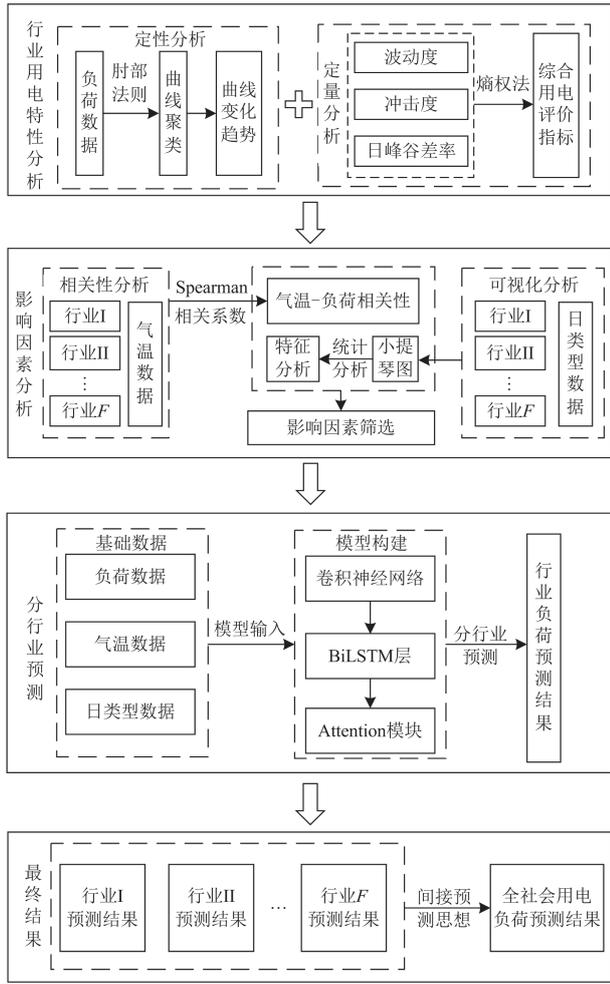


图6 全社会用电负荷预测框架

Fig. 6 Framework for forecasting power consumption load of the entire society

3.5 评价指标

电力负荷预测中常见的误差评价指标包括均方根误差(root mean squared error, RMSE)、平均绝对误差(mean absolute error, MAE)、平均绝对百分比误差(mean absolute percentage error, MAPE)等, 其中MAPE 由于其直观性和易于理解的特点, 使用较为广泛。本文选用这 3 种指标对预测结果进行评价, 3 种指标计算公式为

$$E_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{a=1}^n (Y_a - \hat{Y}_a)^2} \quad (15)$$

$$E_{\text{MAE}} = \frac{1}{n} \sum_{a=1}^n |Y_a - \hat{Y}_a| \quad (16)$$

$$E_{\text{MAPE}} = \frac{1}{n} \sum_{a=1}^n \left| \frac{Y_a - \hat{Y}_a}{Y_a} \right| \times 100\% \quad (17)$$

式中: Y_a 表示第 a 个样本的真实值; \hat{Y}_a 表示第 a 个样本的预测值。

4 算例分析

4.1 数据说明及预处理

算例使用负荷数据为华东某区域 2022 年全社会用电负荷和下属十一个行业以及居民用电负荷数据, 为方便表述, 后文统称为十二大行业。数据时间尺度为 15 min, 即每 15 min 采集一个数据点; 气温数据为该区域同时段下每日 24 h 对应的气温数据。负荷数据具体行业名称见表 1。使用线性插值法对数据集中缺失数据进行填补, 并计算每 1 h 的平均负荷, 每个类别共形成 8760 条负荷数据。

表 1 数据集类别名称情况

Table 1 Status of dataset category names

代号	类别名称
0	全社会用电
1	农/林/牧/渔业
2	工业
3	建筑业
4	交通运输/仓储和邮政业
5	信息传输/软件和信息技术服务业
6	批发和零售业
7	住宿和餐饮业
8	金融业
9	房地产业
10	租赁和商务服务业
11	公共服务及管理组织
12	城乡居民生活用电

另外, 为加速模型收敛并提高其泛化能力, 对原始负荷数据和气温数据进行归一化处理。采用的归一化方法为最大-最小归一化, 计算公式为

$$P' = \frac{P - P_{\min}}{P_{\max} - P_{\min}} \quad (18)$$

式中: P 为原始负荷数据; P_{\min} 为原始负荷数据最小值; P_{\max} 为原始负荷数据最大值。

4.2 行业用电特性分析

4.2.1 定性分析结果

根据 1.1 节分别对全社会及十二大行业用电负荷数据采用肘部法则进行分析, 确定各行业最佳聚类中心数, 结果见表 2。

根据肘部法则结果, 采用 K-means 聚类法对各行业用电负荷曲线进行聚类。全社会用电负荷曲线整体在 04:00 左右逐渐攀升, 10:00 左右达到首个峰值, 之后小幅度回落, 16:00 左右达到第二个峰值, 之后呈下降态势, 但在 19:00 会出现一次回升。

表 2 全社会及十二大行业肘部法则结果

Table 2 Results of elbow law for the entire society and the twelve major industries

行业名称	肘部法则	最佳聚类
	结果	中心数
全社会用电	4	4
农/林/牧/渔业	3	3
工业	3	3
建筑业	4	4
交通运输/仓储和邮政业	4	4
信息传输/软件和信息技术服务业	3	3
批发和零售业	3	3
住宿和餐饮业	3	3
金融业	3	3
房地产业	3	3
租赁和商务服务业	3	3
公共服务及管理组织	3	3
城乡居民生活用电	3	3

农/林/牧/渔业用电负荷曲线多在 06:00 达到全天峰值, 18:00 降至日间谷值; 建筑业与工业变化态势相似, 不再赘述; 交通运输/仓储和邮政业、信息传输/软件和信息技术服务业、批发和零售业、房

地产业、租赁和商务服务业变化态势相似, 在 10:00 左右达到峰值, 之后会保持较长一段时间平稳, 其中在 14:00 左右会出现小幅下降, 在 18:00 左右进一步下降; 住宿和餐饮业前文已介绍, 不再赘述; 金融业、公共服务及管理组织变化态势相似, 在 10:00 左右迎来峰值后开始逐步下降, 中间伴随着多段小幅度回升; 城乡居民生活用电在 00:00 后呈下降趋势, 05:00 左右逐渐回升, 通常在 06:00 迎来早高峰, 之后小幅下降, 在 10:00 再次迎来一个峰值, 15:00 后出现大幅上升, 在 20:00 达到全天峰值, 之后开始大幅下降。可以看出, 十二大行业用电负荷曲线变化各异, 与全社会用电负荷曲线在某些时段存在较大差异。因此, 在进行全社会用电负荷曲线预测时, 需考虑行业用电特性。

4.2.2 定量分析结果

根据 1.2 节分别计算全社会及十二大行业日用电负荷数据的波动度、冲击度、日峰谷差率、日负荷率, 并采用熵权法确定各子指标在综合用电评价指标中的权重。波动度、冲击度、日峰谷差率、日负荷率的平均值以及权重结果见表 3。

表 3 全社会及十二大行业各指标平均值和权重

Table 3 Average values and weights of indicators for the entire society and the twelve major industries

行业名称	平均值				权重			
	波动度	冲击度	日峰谷差率	日负荷率	波动度	冲击度	日峰谷差率	日负荷率
全社会用电	0.335	0.428	0.208	0.905	0.110	0.519	0.276	0.096
农/林/牧/渔业	0.310	0.300	0.229	0.890	0.252	0.512	0.157	0.079
工业	0.310	0.636	0.159	0.916	0.210	0.228	0.485	0.078
建筑业	0.346	0.624	0.532	0.731	0.295	0.179	0.133	0.393
交通运输/仓储和邮政业	0.371	0.498	0.492	0.797	0.106	0.482	0.216	0.196
信息传输/软件和信息技术服务业	0.365	0.348	0.162	0.925	0.040	0.433	0.324	0.204
批发和零售业	0.403	0.458	0.669	0.694	0.077	0.552	0.035	0.336
住宿和餐饮业	0.344	0.371	0.532	0.733	0.243	0.535	0.057	0.165
金融业	0.375	0.346	0.541	0.699	0.027	0.464	0.161	0.348
房地产业	0.377	0.387	0.515	0.767	0.178	0.558	0.060	0.204
租赁和商务服务业	0.385	0.343	0.596	0.710	0.115	0.558	0.058	0.269
公共服务及管理组织	0.331	0.328	0.415	0.784	0.111	0.504	0.175	0.210
城乡居民生活用电	0.304	0.451	0.524	0.692	0.064	0.467	0.116	0.354

根据表 3 可以得到全社会及十二大行业的综合用电评价指标平均值, 如图 7 所示, 其中行业代号见表 1。

由图 7 可以看出, 全社会用电的综合用电评价指标为 0.476, 而多数行业的综合用电评价指标在 0.4 左右, 只有工业、建筑业以及城乡居民生活用电的综合用电评价指标接近 0.476, 交通运输/仓储和邮政业甚至超过了 0.5, 故不同行业间的指标结果差

异较大, 说明各行业用电特性存在较大差异。因此在进行全社会用电负荷曲线预测时, 需考虑行业用电特性。

4.3 影响因素分析结果

4.3.1 气温影响分析

气温对用电负荷的影响具有显著的季节性特征。因此, 通过计算每个季节的相关系数能够更加准确地分析气温对用电负荷的影响。不同季节包含

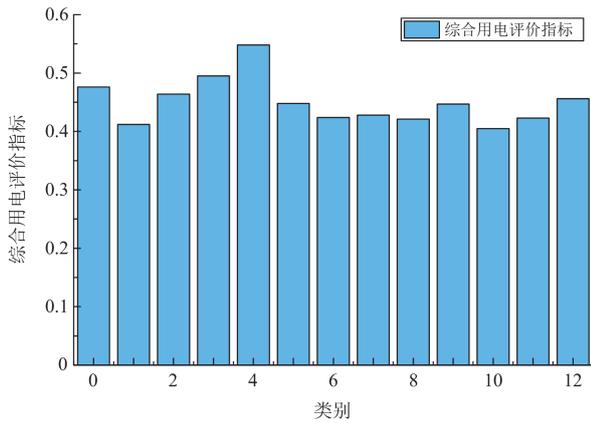


图 7 全社会及十二大行业综合用电评价指标情况

Fig. 7 Comprehensive electricity consumption evaluation indicators for the entire society and twelve major industries

的月份如下：春季为 3—5 月；夏季为 6—8 月；秋季为 9—11 月；冬季为 1—2 月与 12 月。

分别计算不同季节全社会及十二大行业日用电负荷曲线与气温的 Spearman 相关系数见图 8。

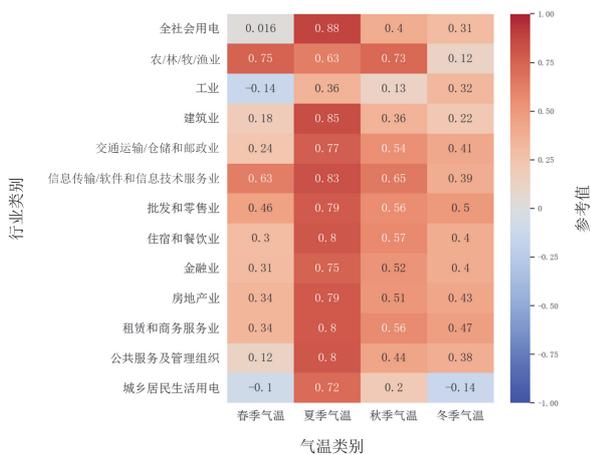


图 8 全社会及十二大行业负荷与气温相关性结果

Fig. 8 Results of the correlation between load and temperature for the entire society and the twelve major industries

由图 8 可以看出，农/林/牧/渔业在春季、夏季、秋季均与气温有着较强的相关性；夏季除工业外，其余行业均与气温有着强相关性；交通运输/仓储和邮政业、信息传输/软件和信息技术服务业、批发和零售业、住宿和餐饮业、金融业、房地产业、租赁和商务服务业等行业在秋季与气温也均有着较强的相关性；冬季各行业与气温的相关性均较低，主要原因为所选地区冬季气温相对较高，采暖需求低。

4.3.2 日类型影响分析

考虑到各行业数据量级相差较大，对负荷数据

进行归一化处理，绘制全社会及十二大行业用电负荷小提琴图见图 9。

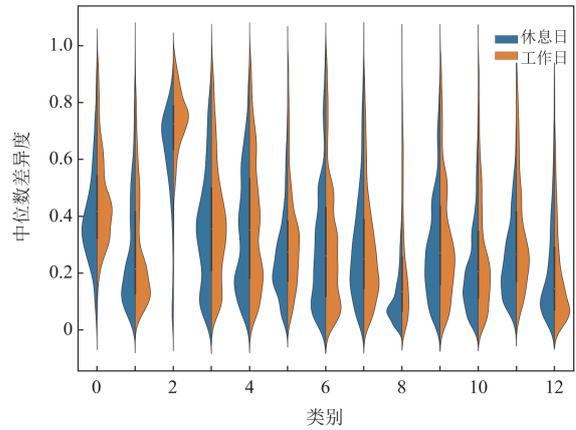


图 9 全社会及十二大行业负荷与日类型小提琴图

Fig. 9 Load and day type violin for the entire society and twelve major industries

由图 9 和图 10 可以看出，农/林/牧/渔业与批发和零售业，即类别 1 和类别 6，在工作日和休息日的负荷分布差异相对较小，其余行业在不同日类型负荷分布差异较大，存在显著区别。因此，在对全社会用电负荷进行预测时，需考虑各行业在不同日类型下的用电特性。

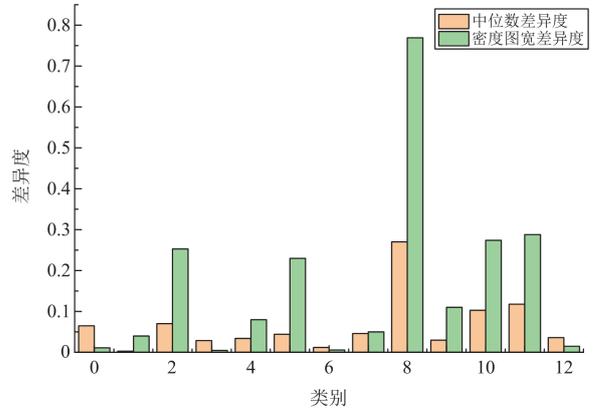


图 10 全社会及十二大行业差异度情况

Fig. 10 Differentials in the entire society and in the twelve major industries

根据以上分析结果，十二大行业需考虑的影响因素情况见表 4，其中“是”代表考虑该影响因素，“否”代表不考虑该影响因素。由于测试集位于秋冬两季，因此气温仅考虑秋冬两季的影响。

4.4 预测结果分析

本节共设置 10 组实验对预测框架进行验证，每组实验的内容见表 5。

表 4 十二大行业影响因素情况

Table 4 Factors affecting the twelve major industries

行业名称	气温(秋季)	气温(冬季)	日类型
农/林/牧/渔业	是	否	是
工业	否	否	是
建筑业	否	否	是
交通运输/仓储和邮政业	是	否	是
信息传输/软件和信息	是	否	是
技术服务业	是	否	是
批发和零售业	是	是	是
住宿和餐饮业	是	是	是
金融业	是	是	是
房地产业	是	是	是
租赁和商务服务业	是	是	是
公共服务及管理组织	是	否	是
城乡居民生活用电	否	否	是

表 5 实验内容清单

Table 5 List of experimental contents

实验序号	使用数据	是否考虑影响因素分析结果	预测方式
1	全社会用电负荷数据	否	直接预测
2	气温、全社会用电负荷数据	否	直接预测
3	日类型、全社会用电负荷数据	否	直接预测
4	气温、日类型、全社会用电负荷数据	否	直接预测
5	行业负荷数据	否	间接预测
6	气温、行业负荷数据	否	间接预测
7	气温、行业负荷数据	是	间接预测
8	日类型、行业负荷数据	否	间接预测
9	气温、日类型、行业负荷数据	否	间接预测
10	气温、日类型、行业负荷数据	是	间接预测

实验 1—实验 4 均为直接预测全社会用电负荷，即直接预测；实验 5—实验 9 均为先预测十二大行业，再将结果累加得到社会用电负荷预测值，即间接预测，两种预测方式具体思路如下。

直接预测指直接利用历史负荷数据和其他影响因素进行负荷预测；间接预测指先预测影响负荷的中间变量，再利用这些变量来推算待预测负荷。全社会用电负荷与十二大行业用电负荷存在如式(19)所示的数学关系。

$$P_0 = \sum_{b=1}^{12} P_b \quad (19)$$

式中： P_0 表示全社会用电负荷； P_b 表示十二大行业用电负荷； b 表示行业序号。

因此，全社会用电负荷满足间接预测条件，即可对十二大行业预测结果进行累加得到全社会用电负荷预测值。

实验 1—实验 4 作为对照组；实验 5—实验 9 需

要与实验 1—实验 4 的结果进行对比，分析行业用电特性是否被有效提取、利用；实验 10 采用本文预测框架，通过与前面 9 组实验进行对比验证本文方法的有效性。每组实验结果见图 11。

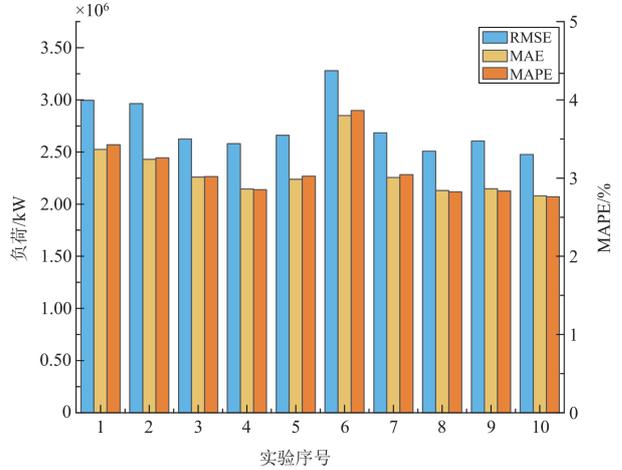


图 11 各组实验结果情况(RMSE、MAE 为左侧轴，MAPE 为右侧轴)

Fig. 11 Experimental results of each group (RMSE, MAE are left axis, MAPE is right axis)

由图 11 可以看出，采用间接预测的实验误差普遍低于直接预测，表明本文提出的提取行业用电特性的方法合理。其中，实验 6 仅考虑引入气温数据进行预测，而未考虑行业用电负荷气温相关性，导致预测误差大；实验 7 则先考虑了气温与行业负荷之间的相关性，再引入气温数据进行预测，明显降低了预测误差，验证了通过 Spearman 相关系数分析气温与行业负荷相关性的有效性。实验 10 采用本文提出的框架进行预测，预测误差均低于其他 9 组实验，验证了本文预测框架的有效性。其中，本文方法较传统方法直接对全社会历史负荷进行预测，即实验 1 的预测结果，RMSE 降低了 17.32%，MAE 降低了 17.63%，MAPE 降低了 19.39%。图 12 为测试集中一天各组的实验预测曲线对比情况。

5 结论

随着“双碳”目标的推进，伴随着一系列相关政策的出台，行业用电特性发生深刻变化，进而对全社会用电负荷造成影响。因此，本文聚焦于行业用电特性，从定性和定量两个角度分别分析行业用电特性；进一步采用 Spearman 相关系数分析各行业与气温间相关性，引入小提琴图可视化分析日类型对各行业的影响，并构建差异度指标量化影响程度；最后，构建 CNN-BiLSTM-Attention 组合预测模型，

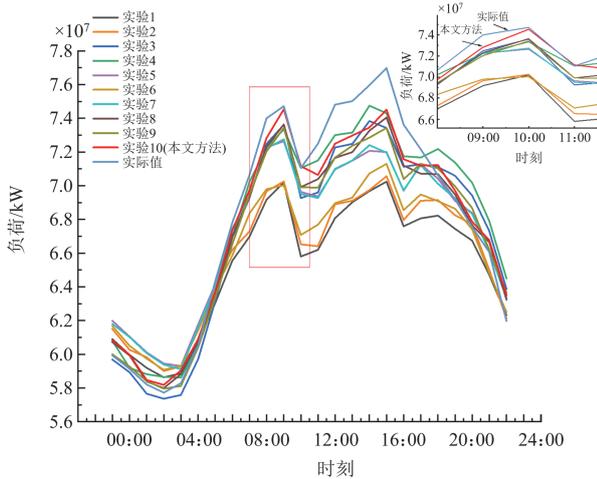


图 12 各组实验预测曲线对比情况

Fig. 12 Prediction curves comparison for experimental group

对各行业分别进行预测, 采用间接预测方式实现对全社会用电负荷进行预测。主要结论如下:

1) 各行业用电特性在定性和定量两方面均有差异, 其中交通运输/仓储和邮政业的综合用电评价指标差异显著, 因此需考虑各行业用电特性对全社会用电负荷预测的影响。

2) 不同行业在不同季节与气温的相关性存在显著差异, 并非所有行业均需考虑气温影响, 如工业、建筑业均与气温相关性较低。

3) 各行业用电量在不同日类型下差异显著, 其中金融业最为突出, 因此, 各行业均需考虑日类型的影响。

4) 相较传统方法直接对全社会历史负荷进行预测, 本文提出的预测框架预测误差显著降低。其中, RMSE 降低了 17.32%, MAE 降低了 17.63%, MAPE 降低了 19.39%。

由于本文在预测模型方面创新较少, 在个别时段的预测误差大于其他方法。因此下一步研究内容为对预测模型进行改进, 以提升行业用电特性提取能力, 进一步降低预测误差。

参考文献

[1] 韩富佳, 王晓辉, 乔骥, 等. 基于人工智能技术的新型电力系统负荷预测研究综述[J]. 中国电机工程学报, 2023, 43(22): 8569-8592.
HAN Fujia, WANG Xiaohui, QIAO Ji, et al. Review on artificial intelligence based load forecasting research for the new-type power system[J]. Proceedings of the CSEE, 2023, 43(22): 8569-8592.

[2] 陈敬德, 李雅晴, 杨欢红, 等. 基于改进 Shapley 特征优选及 ResNet50 深度学习模型的继电保护状态评估

研究[J]. 供用电, 2023, 40(12): 72-78, 106.

CHEN Jingde, LI Yaqing, YANG Huanhong, et al. Condition evaluation of relay protection system based on the modified Shapley feature optimization model and the ResNet50 deep learning algorithm[J]. Distribution & Utilization, 2023, 40(12): 72-78, 106.

[3] 杨铮宇. 基于大数据和机器学习的用电异常行为分析系统[J]. 电测与仪表, 2023, 60(5): 167-173.
YANG Zhengyu. Analysis system of abnormal behavior of electricity consumption based on big data and machine learning[J]. Electrical Measurement & Instrumentation, 2023, 60(5): 167-173.

[4] 李楠, 姜涛, 隋想, 等. 一种时频尺度下的多元短期电力负荷组合预测方法[J]. 电力系统保护与控制, 2024, 52(13): 47-58.
LI Nan, JIANG Tao, SUI Xiang, et al. A multi-component short-term power load combination forecasting method on a time-frequency scale[J]. Power System Protection and Control, 2024, 52(13): 47-58.

[5] FAN G F, HAN Y Y, LI J W, et al. A hybrid model for deep learning short-term power load forecasting based on feature extraction statistics techniques[J]. Expert Systems with Applications, 2024, 238.

[6] 张丽, 李世情, 艾恒涛, 等. 基于改进 Q 学习算法和组合模型的超短期电力负荷预测[J]. 电力系统保护与控制, 2024, 52(9): 143-153.
ZHANG Li, LI Shiqing, AI Hengtao, et al. Ultra-short-term power load forecasting based on an improved Q-learning algorithm and combination model[J]. Power System Protection and Control, 2024, 52(9): 143-153.

[7] JALALIFAR R, DELAVAR M R, GHADERI S F. SAC-ConvLSTM: a novel spatio-temporal deep learning-based approach for a short term power load forecasting[J]. Expert Systems with Applications, 2024, 237.

[8] 范茜茜, 王国强, 罗贺, 等. 基于 N-BEATS 与辅助编码器的短期电力负荷预测[J]. 电网技术, 2024, 48(4): 1612-1621.
FAN Xixi, WANG Guoqiang, LUO He, et al. Short-term electricity load forecasting based on N-BEATS with auxiliary encoders[J]. Power System Technology, 2024, 48(4): 1612-1621.

[9] 秦浩. 基于改进时间卷积网络的短期电力负荷预测研究[D]. 南昌: 南昌大学, 2023.

[10] LI K, HUANG W, HU G, et al. Ultra-short term power load forecasting based on CEEMDAN-SE and LSTM neural network[J]. Energy and Buildings, 2023, 279.

[11] RAN P, DONG K, LIU X, et al. Short-term load forecasting based on CEEMDAN and Transformer[J]. Electric Power Systems Research, 2023, 214.

[12] 张籍, 薛儒涛, 刘慧, 等. 基于深度信念网络的不同行

- 业中长期负荷预测[J]. 电力系统及其自动化学报, 2019, 31(9): 12-19, 27.
- ZHANG Ji, XUE Rutao, LIU Hui, et al. Medium-and long-term load forecasting for different industries based on deep belief network[J]. Proceedings of the CSU-EPSSA, 2019, 31(9): 12-19, 27.
- [13] 蒋贲. 用电市场景气分析研究[D]. 北京: 华北电力大学, 2009.
- [14] 王鹏, 陈启鑫, 夏清, 等. 应用向量误差修正模型的行业电力需求关联分析与负荷预测方法[J]. 中国电机工程学报, 2012, 32(4): 100-107, 12.
- WANG Peng, CHEN Qixin, XIA Qing, et al. Correlation analysis and forecasting method on industrial electricity demand based on vector error correction model[J]. Proceedings of the CSEE, 2012, 32(4): 100-107, 12.
- [15] 李龙飞, 钟海旺, 张广伦, 等. 基于因果分析的行业月度电量预测方法[J]. 电网技术, 2023, 47(9): 3629-3638.
- LI Longfei, ZHONG Haiwang, ZHANG Guanglun, et al. Monthly sectoral electricity demand prediction based on causal analysis[J]. Power System Technology, 2023, 47(9): 3629-3638.
- [16] 陆晓, 徐春雷, 冷钊莹, 等. 基于数据驱动方法的疫情阶段电力用户负荷特性画像模型[J]. 电力建设, 2021, 42(2): 93-106.
- LU Xiao, XU Chunlei, LENG Zhaoying, et al. Load characteristic portrait model of power users in epidemic stage applying data-driven method[J]. Electric Power Construction, 2021, 42(2): 93-106.
- [17] 邓燕国, 王冰, 曹智杰, 等. 基于熵权法与 GRA-ELM 的配电网空间负荷预测[J]. 电力工程技术, 2021, 40(4): 136-141.
- DENG Yanguo, WANG Bing, CAO Zhijie, et al. Spatial load forecasting of distribution network based on entropy weight method and GRA-ELM[J]. Electric Power Engineering Technology, 2021, 40(4): 136-141.
- [18] 付桂琴, 尤凤春, 曹欣, 等. 积温效应在电力日峰谷负荷中的应用及检验[J]. 应用气象学报, 2015, 26(4): 492-499.
- FU Guiqin, YOU Fengchun, CAO Xin, et al. Application and verification of accumulated temperature effects on daily peak load and daily valley load of power[J]. Journal of Applied Meteorological Science, 2015, 26(4): 492-499.
- [19] SATISH B, SWARUP K S, SRINIVAS S, et al. Effect of temperature on short term load forecasting using an integrated ANN[J]. Electric Power Systems Research, 2004, 72(1): 95-101.
- [20] 董家富, 薛一波. 气象因子与电力负荷的时空关联特性研究[J]. 电力信息与通信技术, 2022, 20(6): 11-23.
- DONG Jiafu, XUE Yibo. Research of spatio-temporal correlation characteristics between meteorological factors and power load[J]. Electric Power Information and Communication Technology, 2022, 20(6): 11-23.
- [21] 吴振龙, 莫艺鹏, 王荣花, 等. 基于 LSTM 和粒子群算法的多机组风电功率预测[J/OL]. 郑州大学学报(工学版): 1-8[2024-07-26]. <https://doi.org/10.13705/j.issn.1671-6833.2024.06.005>.
- WU Zhenlong, MO Yipeng, WANG Ronghua, et al. Multi-unit wind power prediction based on long short-term memory and particle swarm optimization[J/OL]. Journal of Zhengzhou University (Engineering Science): 1-8[2024-07-26]. <https://doi.org/10.13705/j.issn.1671-6833.2024.06.005>.
- [22] 吴玥珺, 赵二刚, 郭增良, 等. 基于 Spearman 系数和 TCN 的光伏出力超短期多步预测[J]. 太阳能学报, 2023, 44(9): 180-186.
- WU Junyue, ZHAO Ergang, GUO Zengliang, et al. Ultra-short-term photovoltaic power multi-step prediction based on Spearman coefficient and TCN[J]. Acta Energetica Solaris Sinica, 2023, 44(9): 180-186.
- [23] 季玉琦, 严亚帮, 和萍, 等. 基于 K-Medoids 聚类与栅格法提取负荷曲线特征的 CNN-LSTM 短期负荷预测[J]. 电力系统保护与控制, 2023, 51(18): 81-93.
- JI Yuqi, YAN Yabang, HE Ping, et al. CNN-LSTM short-term load forecasting based on the K-Medoids clustering and grid method to extract load curve features[J]. Power System Protection and Control, 2023, 51(18): 81-93.
- [24] 任建吉, 位慧慧, 邹卓霖, 等. 基于 CNN-BiLSTM-Attention 的超短期电力负荷预测[J]. 电力系统保护与控制, 2022, 50(8): 108-116.
- REN Jianji, WEI Huihui, ZOU Zhuolin, et al. Ultra-short-term power load forecasting based on CNN-BiLSTM-Attention[J]. Power System Protection and Control, 2022, 50(8): 108-116.
- [25] 吴小涛, 袁晓辉, 毛玉鑫, 等. 基于鹈鹕优化 CNN-BiLSTM 的电力负荷预测[J]. 水电能源科学, 2024, 42(8): 209-212, 172.
- WU Xiaotao, YUAN Xiaohui, MAO Yuxin, et al. Power load prediction based on pelican optimized CNN-BiLSTM[J]. Water Resources and Power, 2024, 42(8): 209-212, 172.

收稿日期: 2024-09-24; 修回日期: 2025-01-20

作者简介:

郭耀扬(2000—), 男, 硕士, 研究方向为电力系统负荷分析预测; E-mail: guo_090229@163.com

张利(1979—), 女, 通信作者, 博士, 副教授, 研究方向为微电网运行分析与控制、配电网故障检测技术。E-mail: ZZRZL@163.com

(编辑 许威)