

基于电力大数据分析的电力消费指数研究与应用

郑 斌¹, 裘炜浩¹, 侯素颖¹, 杨世旺², 戚胡佳¹

(1. 国网浙江省电力有限公司, 浙江 杭州 310000; 2. 国网浙江省电力有限公司丽水供电公司, 浙江 丽水 323000)

摘要: 用电量作为国民经济的“温度计”和“晴雨表”, 与经济运行息息相关, 用电量消费的变化折射出经济运行的活跃度, 深刻反映了地方经济和社会运转状况。为了深入剖析电力消费的动态变化程度与变化趋势, 基于某省近三年电力相关数据, 通过引入电力消费现状因子和期望因子, 提出并构建了电力消费指数模型, 实现了对各区域、产业和行业的电力消费水平精准分析。另外利用深度学习长短记忆神经网络(LSTM)对电力消费指数进行预测, 分析社会经济发展趋势, 为政府制定政策提供重要参考依据。

关键词: 用电量; 电力大数据; 电力消费指数; LSTM

Research and application of power consumption index based on power big data analysis

ZHENG Bin¹, QIU Weihao¹, HOU Suying¹, YANG Shiwang², QI Hujia¹

(1. State Grid Zhejiang Electric Power Co., Ltd., Hangzhou 310000, China; 2. State Grid Zhejiang Electric Power Co., Ltd., Lishui power supply company, Lishui 323000, China)

Abstract: As the “thermometer” and “barometer” of the national economy, electricity consumption is closely related to the economic operation. The change of electricity consumption reflects the activity of economic operation and profoundly reflects the local economic and social operation. In order to deeply analyze the dynamic change degree and trend of power consumption, based on the power related data of Zhejiang Province in recent three years, this paper proposes and constructs the power consumption index model by introducing the power consumption status factor and expectation factor, which realizes the accurate analysis of the power consumption level of various regions, industries and industries. In addition, the deep learning long short term memory (LSTM) neural network is used to analyze the power consumption level of various regions, industries and industries. This paper forecasts the power consumption index, analyzes the social and economic development trend, and provides important reference for the government to formulate policies.

Key words: electricity consumption; power big data; power consumption index; LSTM

0 引言

2020年初受疫情影响, 各行各业均受到了不同程度的冲击, 宏观经济出现较大波动。

国网总部专项基金 (6211XT20002J)

当前某省民营企业数量众多, 涉及行业领域广泛, 民营企业受疫情影响的程度关乎社会经济平稳运行发展, 为了及时掌握疫情期间企业生产经营状况, 亟需从宏观层面对各地区、各产业和各行业运行状况开展动态监测和实时分

析。为了积极响应中央“六稳六保”加强宏观经济运行监测要求，充分利用电力数据“晴雨表”“风向标”的特性，发挥电力数据覆盖面广、实时性好和真实性高的优势，开展电力大数据分析，提出并构建了电力消费指数，通过电力客户用电数据、业扩数据等海量电力运行数据进行特征挖掘，真实客观反映各地区、各产业和各行业电力消费情况。电力需求变化能够在一定程度上反映出经济发展的真实情况。

目前电力景气指数算法普遍采用回归分析法，缺点是没有考虑数据的非线性问题，以及没有考虑电力数据客观性。马哈茂德等学者提出了时间序列的方法，比如回归分析法^[1]、指数平滑模型法等，该方法是从数据的时序性出发，忽略了数据的非线性问题，因此在实际应用中存在局限性。为了提高预测模型的准确度以及使用的广泛性，本文利用神经网络（Neural Networks）技术建立城市经济走势预测模型，并对经济发展趋势进行预测，为经济政策提供重要参考依据。最后研究结果表明，本文所提出的经济走势预测模型（LSTM预测模型）具有较高的实用价值。该模型兼顾了数据的时序性和非线性关系，能够反映时间序列中的历史过程，相比传统的预测方法具有更高的准确率和实用性。

1 研究思路

传统电力消费（景气）指数构建方法一般会将会电力数据与宏观经济统计指标进行相关性分析，选取用电量、业扩报装容量和用户数等电力数据作为模型输入指标，而电力数据本身为全量数据，客观性强于宏观经济统计指标。通过建立电力负荷与宏观经济之间的联动关系来协助政府合理地规划电网建设，优化电力资源的配置^[2]。通过研究经济发展方式与电力经济的关系，提供预判分析^[3]。另外利用神经网络（Neural Networks）技术，从电力角度出发对城市发展的因素进行分析，并对经济发展形式

进行预测^[4]，能够在一定程度上反映出当今经济发展的形势。

用户数等相关性较低的指标与用电量等核心指标本身存在强关联，引入过多冗余的变量反而会降低模型构建的易读性和可追溯性。因此，选取与电力消费水平现状相关性最强的用电量指标，与电力消费水平预期相关性最强的业扩容量指标作为模型的输入数据^[5]。基于用电量指标和业扩容量指标，引入电力消费现状因子和电力消费期望因子，构建电力消费指数模型，计算公式为

$$EPI = \frac{Q_i + \Delta C_i \alpha T}{Q_{i-12} + \Delta C_{i-12} \alpha T} \times 100\% \quad (1)$$

式中， Q_i 为本期用电量，反映了电力消费的现状水平，称为电力消费现状因子； ΔC_i 为本期业扩净增容量； α 为单位容量的每小时电量转化率； T 为转化时间； $\Delta C_i \alpha T$ 体现了业扩净增容量带来的期望电力消费水平，称为本期电力消费期望因子； $Q_i + \Delta C_i \alpha T$ 为当期电力消费综合因子；EPI指数100为荣枯线分界。

2 单位容量的每小时电量转化率计算

为了计算单位容量的每小时电量转化率，需要搜集用户的用电数据信息，具体包括2018年、2019年全省高压非居用户的月用电量数据和月运行容量数据。不同行业的用户，其用电特性不同，对应的单位容量的每小时电量转化率也不同。

单位容量每小时的电量转化率 = 2018~2019年未发生过业扩变更的用户总用电量 / 2018~2019年未发生过业扩变更的用户总运行容量 / 总小时数。

3 实证研究

3.1 数据获取

本文所需业务数据来源于一体化云平台，主要获取2018年1月~2020年7月某省全省用户的单位代码、单位名称、月份、行业序号、行业、用电量、运行容量、增容用户合同容量、

新装用户合同容量、减容用户合同容量和销户用户合同容量数据。

3.2 数据预处理

数据预处理主要包含以下步骤：

1) 确定用户范围。根据实际需要，确定用户范围为全省高压用户，包括光伏发电用户，并剔除趸售用户。另外，防疫、外贸和园区等专题指数的用户群体由政府等相关单位提供，单独处理。

2) 加工月电量和业扩净增容量等基础数据。高压用户月电量=用户月度电量电费台账+自备电厂自发自用部分电量，光伏用户月电量=发电量-上网电量；业扩报装数据来自当期归档的高压用户业扩新装、增容、永久减容和销户流程，净增容量=增容容量+新装容量-减容容量-销户容量。

3) 计算行政区域。用户的结构化地址数据包含省、市、县和街道四个层级，由于历史数据遗留问题，存在地址信息不完整、不准确和不能归属到90个行政区县等情况。利用用户的用电地址、供电单位等信息，通过近邻算法构建匹配模型，绝大部分用户都能得到行政区县归属，剩下部分用户通过电话确认、基层单位实地走访等方式得到确认。

4) 制作数据宽表。根据电力消费指数模型输入指标的计算逻辑，从2018年1月~2020年7月，整合高压用户的行政区域、行业等维度数据以及月电量和月业扩净增容量数据，形成累计30个月的用户明细数据宽表，适用于电力消费指数模型开展当期同比、累计同比计算，提升后续模型的计算效率。

5) 指数计算。基于数据宽表，按照指数公式进行计算。计算结果包含的区域维度为全省、11个地市和90个区县公司，行业维度包含全行业、三大产业和11大行业，另外包含了防疫、外贸和园区等专题指数。

3.3 模型结果

某省EPI指数趋势图如图1所示。从指数来

看，全省经济形势稳中向好、稳中有进。2020年7月，全省EPI指数持续反弹，达114.38，较6月上升9.20个点。2020年1~7月，某省累计EPI指数为98.21，较1~6月累计上升3.27个点。

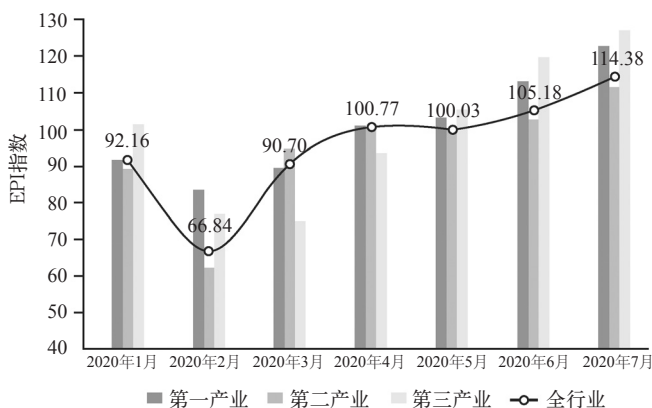


图1 某省全行业及各产业单月EPI指数趋势图

Fig. 1 Trend chart of EPI index of the whole industry and each industry in a province

(1) 电力看5G发展

5G基站运行保持快速增长态势如图2所示。2020年7月EPI指数为123.75，仍保持较高水平。1~7月累计EPI指数为119.08，保持高于全行业和第二产业面上的水平。从电力消费看，1~7月累计用电量为6.72亿kW·h，同比增长21.08%，5G基站的运营保持着快速的增长。

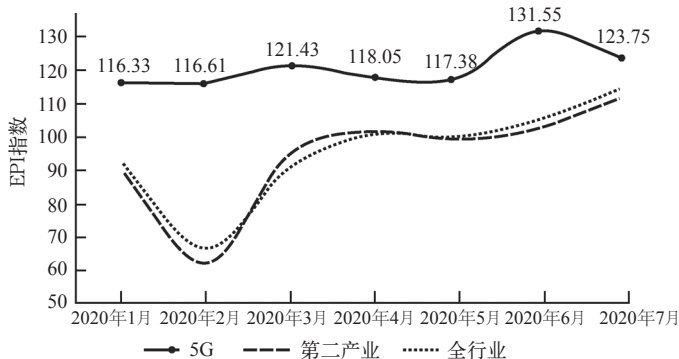


图2 某省2020年1~7月5G基站建设和运营EPI指数情况

Fig. 2 EPI index of 5G base station construction and operation in a province from January to July 2020

(2) 电力看充电桩发展

充换电服务业继续高位增长如图3所示。2020年以来，某省电力公司不断加快充电桩建设，优化充换电服务，并陆续推出充电服务优惠举措。在各类政策推动下，充换电服务保持

高速增长。7月充换电服务业EPI指数为170.86，但仍明显高于荣枯线100以上，较2019年同期仍保持高速增长。1~7月累计EPI指数为128.52，保持较快的增长态势。

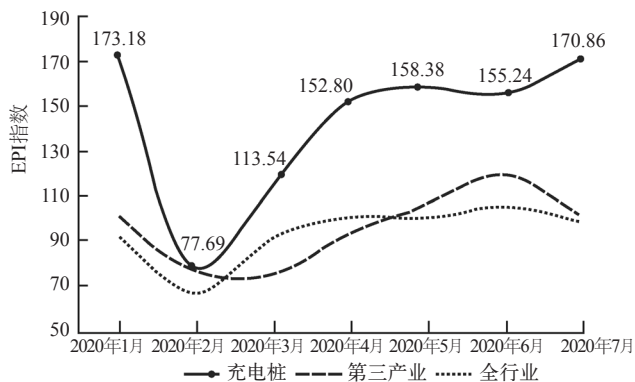


图3 2020年1~7月充电桩服务业EPI指数情况

Fig. 3 EPI index of charging pile service industry in a province from January to July 2020

(3) 电力看高新园区发展

高新园区加速回升，如图4所示。2020年以来，全省各地抢抓国家大力支持新基建项目建设政策机遇，精准发力加速推进，高新园区作为新经济增长点，保持着较快的增长速度。浙江省43家高新园区7月EPI指数为112.82，较6月提升8.52个点，呈现较快的反弹速度。1~7月累计EPI指数为96.72，稍落后于同期水平。

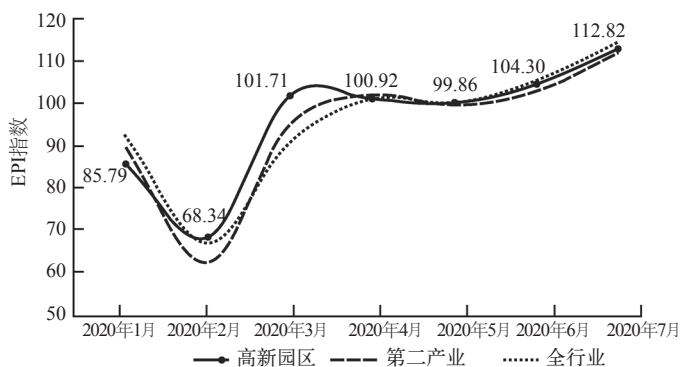


图4 2020年1~7月高新园区EPI指数情况

Fig. 4 EPI index of high tech parks from January to July 2020

3.4 电力消费指数预测

为了更好地对各地区、各产业和各行业运行状况开展动态监测，有必要对电力消费指数进行提前预测，以便为各级政府预测经济发展

趋势、制定经济政策提供重要的参考依据，助力政府服务由事后向事前转型。从电力消费指数的时序散点图可以明显地看出，随着时间的推移，电力消费指数呈明显的趋势，长短记忆神经网络(LSTM网络)算法^[17-21]在处理具有趋势的时间序列中相比传统的机器学习算法更有优势。LSTM网络在使用反向传播时间训练并克服消失梯度和梯度爆炸的问题，适合长周期的电力消费指数数据学习，找出其隐藏的映射关系并进行预测。LSTM网络的基本单元组成如图5所示。

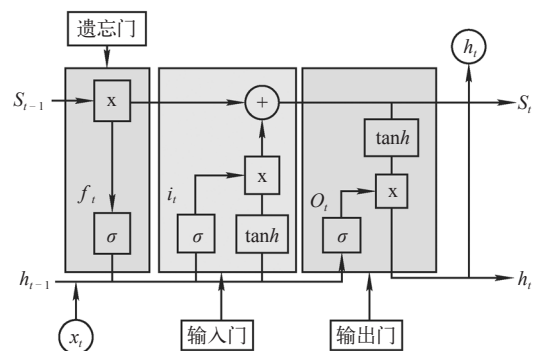


图5 LSTM网络的基本单元

Fig. 5 basic unit of LSTM network

LSTM网络的基本单元包括遗忘门、输入门和输出门。遗忘门中输入由状态记忆单元和中间输出单元共同决定状态记忆单元遗忘门。输入门中的变量分别经过Sigmoid函数和tanh函数变化后共同决定状态记忆单元中保留向量。计算公式如下

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \quad (2)$$

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \quad (3)$$

$$g_t = \phi(W_{gx}x_t + W_{gh}h_{t-1} + b_g) \quad (4)$$

$$O_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \quad (5)$$

$$S_t = g_t \otimes i_t + S_{t-1} \otimes f_t \quad (6)$$

$$h_t = \phi(S_t) \otimes O_t \quad (7)$$

式中， f_t 、 i_t 、 g_t 、 O_t 、 h_t 和 S_t 分别为遗忘门、输入门、输入节点、输出门、中间输出和状态单元的状态； W 为分别相对应的门与输入 x_t 和中间输出 h_{t-1} 相乘的矩阵权重； b 为相对应门的偏执项； $\otimes$ 表示向量元素按位相乘； σ 为Sigmoid函数

变化； ϕ 表示 $\tanh$ 函数变化。

本文的电力消费指数预测是利用深度学习LSTM网络的数学模型得到的，分别对某省全行业、第一产业、第二产业和第三产业的2020年1月1日~2020年7月13日的日累计电力消费指数中70%用于训练，30%用于测试。经过交叉验证和网格搜索得到10天最佳滑动时间窗口，利用近10天的指数预测下一天指数情况，预测效果如图6~图9所示，红色为真实值，蓝色为预测值。

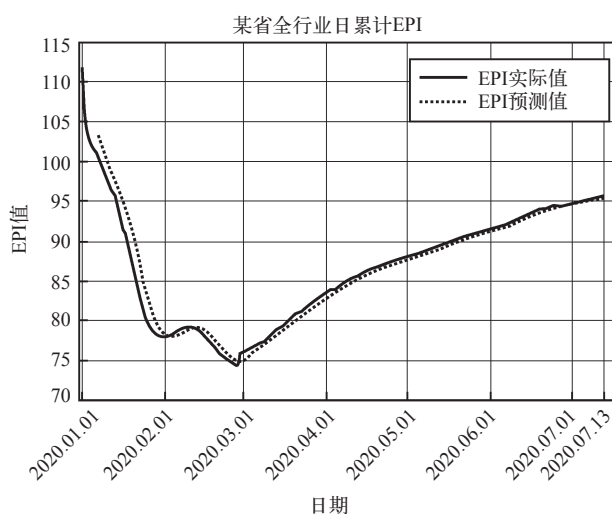


图6 某省全行业日累计EPI预测曲线

Fig. 6 Prediction curve of daily cumulative EPI of the whole industry in Zhejiang Province

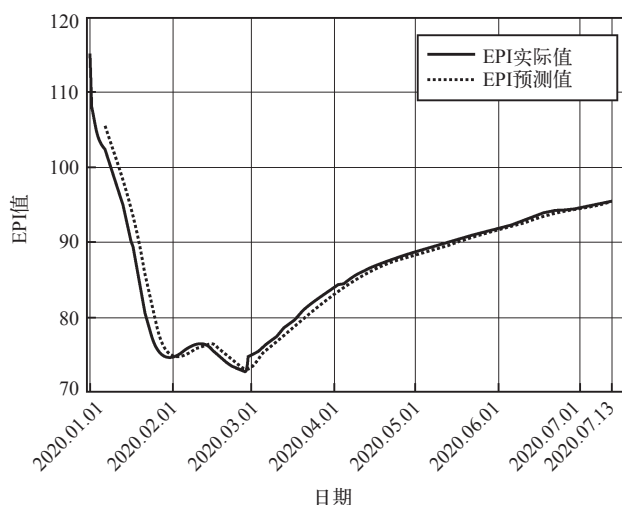


图7 某省第一产业日累计EPI预测曲线

Fig. 7 Prediction curve of daily cumulative EPI of primary industry in Zhejiang Province

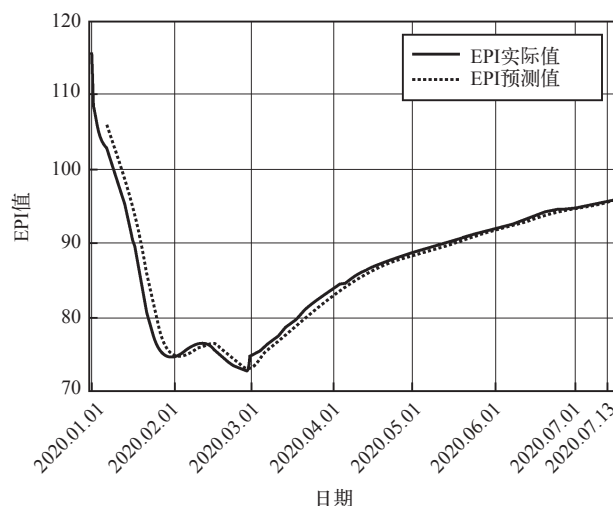


图8 某省第二产业日累计EPI预测曲线

Fig. 8 Prediction curve of daily cumulative EPI of secondary industry in Zhejiang Province

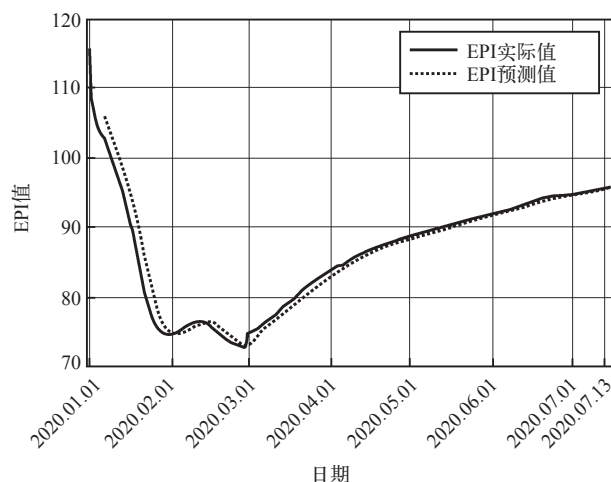


图9 某省第三产业日累计EPI预测曲线

Fig. 9 Prediction curve of daily cumulative EPI of tertiary industry in Zhejiang province

某省全行业、第一产业、第二产业和第三产业的预测均方根误差分别为1.239 1、0.513 1、1.499 0和0.955 5。

模型验证结论如下：利用跨时间窗口的数据对训练好的模型进行验证，预测值和真实值的均方误差为0.854 32，认为新模型的预测效果较好，具有实用价值。

4 结束语

本文为及时掌握全省经济运行状况及趋势提出了电力消费指数，基于电力消费指数的计

算方法,对某省的电力消费水平进行计算,并利用深度学习LSTM网络的数学模型实现了对电力消费指数的预测,为各级政府评估区域、行业(产业)经济发展趋势及制定经济政策提供重要参考依据,助力政府服务由事后向事前转型,提升政府数字化管理水平。

最后与实际宏观经济运行状况进行对比分析,验证本计算方法的有效性和实用性。通过实际应用,本文提出的电力消费指数可以准确全面反映当前经济运行水平和未来经济景气预期,实现“电力看产业”“电力看区域”和“电力看园区”等,在实际应用中每天、每月定时输出指数结果。

参考文献

- [1] MAHMUD M A. Isolated area load forecasting using linear regression analysis: practical approach[J]. Energy and Power Engineering, 2011, 3(4): 547-550.
- [2] 刘玉娇,宋坤煌,王向.基于电力大数据的经济景气数分析[C].2019电力行业信息化年会.
- [3] 朱家明,关珊,蒋萌萌.宏观经济与短期电力负荷的联动关系分析[D].蚌埠:安徽财经大学,2017.
- [4] 田传波,尹玉,金鹏,等.基于电力数据影响城市宏观经济发展分析与预测[J].电气应用,2015(增刊1):232-233.
- [5] 史静,李琥,李冰洁,等.基于长短记忆神经网络的配电网负荷预测方法研究[J].供用电,2019,36(7):71-52.
- [6] 李冬辉,尹海燕,郑博文,等.改进的LSTM方法在冷水机组传感器故障检测中的应用[J].电工技术学报,2019,34(11):2324-2332.
- [7] 陈亮,王震,王刚.深度学习框架下LSTM网络在短期电力负荷中的应用[J].电力信息与通信技术,2017(5):8-11.
- [8] 陈卓,孙龙祥.基于深度学习LST网络的短期电力负荷预测方法[J].电子技术,2018,47(1):39-41.
- [9] SUNDERMEYER M, NEY HSCHLUTER R. From feedforward to recurrent LSTM neural networks for modeling[J]. IEEE/ACM Transactions on Audio, Speech and Language Processing, 2015, 23(3): 517-529.
- [10] 梁妍,祖向荣.基于LSTM网络的入侵检测模型研究[J].数字化用户,2018,24(35):12-13.
- [11] 张冲冲.基于LSTM深度网络和改进核极限学习机的短期电力负荷预测[D].保定:华北电力大学,2019.
- [12] 陈亮,王震,王刚.深度学习框架下LSTM网络在短期电力负荷预测中的应用[J].电力信息与通信技术,2017,15(5):8-11.
- [13] 陈先昌.基于卷积神经网络的深度学习算法与应用研究[D].杭州:浙江工商大学,2014.
- [14] 代林.基于神经网络的电力负荷预测方法研究实现[D].成都:电子科技大学,2012.
- [15] 高志强,黄剑,李永,等.深度学习:从入门到实战[M].北京:中国铁路出版社,2018.
- [16] 张铃.M-P神经元模型的几何意义及其应用[J].软件学报,1998(5):15-19.
- [17] 田娟,李英祥,李彤岩.激活函数在神经网络中的对比研究[J].计算机系统应用,2018,27(7):43-49.
- [18] 姚朝,辛平安,施卜今,等.基于LSTM时间地柜神经网络的短期电力负荷预测[J].云南水利发电,2019,35(3):163-165.
- [19] 包子阳.神经网络与深度学习,基于Tensorflow框架和Python技术实现[M].北京:电子工业出版社,2019.
- [20] 朱安.应用时间序列的自适应模型预测短期电力负荷[J].电力系统自动化,1984(2):12-19.

作者简介



郑 斌

郑斌(1974—),男,硕士,高级工程师,研究方向为电力营销管理。

(收稿日期:2020-08-17)EA