

1-16-2023

Research on forecasting electricity consumption of high-energy-consuming industries based on Granger causality and ARDL model

Yu SHEN

Economic and Technological Research Institute, State Grid Fujian Electric Power Co., Ltd., Fuzhou 350012, China

Xianan HUANG

Economic and Technological Research Institute, State Grid Fujian Electric Power Co., Ltd., Fuzhou 350012, China

Lin LIU

Economic and Technological Research Institute, State Grid Fujian Electric Power Co., Ltd., Fuzhou 350012, China

Zhenda HU

Economic and Technological Research Institute, State Grid Fujian Electric Power Co., Ltd., Fuzhou 350012, China

Jiu GU

School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

and additional works at: <https://jepst.researchcommons.org/journal>

Recommended Citation

See next page for additional authors

SHEN, Yu; HUANG, Xianan; LIU, Lin; HU, Zhenda; GU, Jiu; and JIANG, Chuanwen (2023) "Research on forecasting electricity consumption of high-energy-consuming industries based on Granger causality and ARDL model," *Journal of Electric Power Science and Technology*. Vol. 37: Iss. 6, Article 19.

DOI: 10.19781/j.issn.1673-9140.2022.06.019

Available at: <https://jepst.researchcommons.org/journal/vol37/iss6/19>

This Article is brought to you for free and open access by Journal of Electric Power Science and Technology. It has been accepted for inclusion in Journal of Electric Power Science and Technology by an authorized editor of Journal of Electric Power Science and Technology.

Research on forecasting electricity consumption of high-energy-consuming industries based on Granger causality and ARDL model

Authors

Yu SHEN, Xianan HUANG, Lin LIU, Zhenda HU, Jiu GU, and Chuanwen JIANG

基于格兰杰因果与 ARDL 模型的高能耗 产业用电量预测

沈 豫¹, 黄夏楠¹, 刘 林¹, 胡臻达¹, 顾 玫², 蒋传文²

(1. 国网福建省电力有限公司经济技术研究院, 福建 福州 350012; 2. 上海交通大学电子信息与电气工程学院, 上海 200240)

摘 要:为挖掘产业经济发展与电力大数据之间的耦合关系, 构建高耗能产业用电量与多元经济指标数据之间的向量自回归 (VAR) 模型。通过使用格兰杰 (Granger) 因果关系检验的方法, 提取对用电量预测有显著影响的产业经济指标数据, 在此基础上建立考虑经济因素影响的高耗能产业用电量自回归分布滞后 (ARDL) 模型。以某地区 2016—2020 年产业用电和经济数据进行算例分析, 结果表明: Granger 因果关系检验可以有效地挖掘与细分产业用电量相关联的经济指标; 将这些经济因素考虑到产业用电预测模型中, 可以有效地提高模型的预测精度。

关 键 词: Granger 因果关系; 关联分析; ARDL 模型; 产业经济指标; 电量预测

DOI: 10.19781/j.issn.1673-9140.2022.06.019 中图分类号: TM621 文章编号: 1673-9140(2022)06-0165-08

Research on forecasting electricity consumption of high-energy-consuming industries based on Granger causality and ARDL model

SHEN Yu¹, HUANG Xianan¹, LIU Lin¹, HU Zhenda¹, GU Jiu², JIANG Chuanwen²

(1. Economic and Technological Research Institute, State Grid Fujian Electric Power Co., Ltd., Fuzhou 350012, China;
2. School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: In order to explore the coupling relationship between the industrial economic development and power big data, a vector auto regressive (VAR) model is constructed between the electricity consumption of high-energy-consuming industries and the data of multiple economic indicators. By using the Granger causality test method, industrial economic index data are extracted that has a significant impact on electricity consumption forecasting. Based on this, an Auto Regressive Distributed Lag (ARDL) model of electricity consumption is established in high-energy-consuming industries that takes economic factors into account. An example analysis of industrial electricity consumption and economic data in a certain region from 2016 to 2020 shows that the Granger causality test can effectively dig out the economic indicators related to the electricity consumption of subdivided industries. Considering these economic factors in the regression model, the prediction accuracy of the model is effectively improved.

Key words: Granger causality; correlation analysis; ARDL model; industrial economic indicators; electricity forecast

收稿日期: 2021-09-04; 修回日期: 2021-11-19

基金项目: 福建省科学技术协会科技创新智库课题研究项目 (Fjkk-B2009); 国家重点研发计划 (2018YFB0905200)

通信作者: 蒋传文 (1966—), 男, 博士, 教授, 主要从事电力系统优化运行、电力市场的研究; E-mail: jiangcw@sjtu.edu.cn

电力行业是国民经济的基础行业,产业经济的发展与电力消费有着密切的联系,挖掘两者之间的关联关系,有利于合理规划产业集群周边的电力发展、提高产业用能经济性,从而助力产业经济健康发展^[1-2]。随着“双碳”目标的提出,产业节能减排得到社会的高度关注,高能耗产业作为地区用电量的风向标,更是电力部门关注的重点。对高能耗产业开展用电量预测研究,可以及时发现高能耗产业在生产过程中存在的问题^[3-5]。

电力消费与经济发展的关系作为电力经济研究的关键问题,传统的做法是以计量经济学模型作为基础,对电力经济数据进行分析,典型的方法是基于时间序列的计量模型^[6-7]。文献[8]以安徽省作为研究对象,采用了多源回归模型对经济增长与用电量偏移的原因进行分析,得到了影响电力消费的多类经济指标以及各指标的影响强度。而文献[9]以北京作为研究对象,以多部门的经济与电力数据建立 Granger 因果关系分析模型,结果显示:总体上存在经济对能源消费的单向因果关系,不同产业的关系各不相同。文献[10]基于新型分位数方法对1995—2015年的统计数据分析,结果表明工业产值与电力消费存在正相关关系;文献[11-12]采用 HP 滤波技术对实际 GDP 的趋势成分、周期成分的处理,研究结果表明能源消费与 GDP 之间存在协整关系;文献[13]则分析了电力消费与固定资产投资、人均可支配收入等经济指标之间的协整关系。上述分析结果可以看出,对于时间序列数据的研究,不同的时间区间、变量类别、计量分析模型都对分析结果有着极大的影响。另一方面,目前大多数的计量研究主要是分析能源总量与经济总量之间的关系,并没有考虑到具体地区的细分产业发展的特点,使得模型的构建不够精细化,误差难以缩小。

负荷预测问题同样是长期以来电力行业的研究热点,大体上可以分为基于数理统计的预测方法和基于人工智能的预测方法。数理统计的方法主要以时间序列分析法^[14]、回归分析模型以及灰色模型^[15]为主。应用于电力负荷预测的人工智能算法包括长短期记忆网络模型(long short-term memory network, LSTM)^[16]、支持向量机(support vector machines, SVM)^[17]、神经网络模型^[18]等。数

理统计的方法对于数据的解释性更强,可以挖掘负荷历史数据与其他外部变量之间的关联性,在处理数据量小的问题时适应性较好;而人工智能方法虽然对预测结果有着更高的精度,但是其对于模型的解释力却略显不足。

基于上述分析,本文以高能耗产业作为研究对象,研究产业用电量与经济发展指标之间的关联与影响。首先根据产业划分原则,建立产业用电量与经济指标的对应关系;基于协整—格兰杰因果关系检验方法,分析高耗能行业电力消费与经济指标数据之间的相互影响关系,并构建多元经济指标与细分产业用电量的自回归分布滞后(autoregressive distributed lag, ARDL)模型,对地区高能耗产业的用电量进行预测,以实际用电数据对算法精度进行验证,从经济性的角度解释预测误差的来源。

1 产业经济与电力大数据的关联指标选择

1.1 产业电力消费数据细分

在中国《统计年鉴》中,电力平衡表是反映国家或地区的电力供应和消费的重要指标,其中,电力消费量统计量则反映着地区的电力负荷的情况。将地区的用电量比作整体,当对地区电力负荷进行预测时,仅仅考虑对地区总体的电力负荷预测,而不对负荷各分量的变化进行量化分析,则很有可能造成预测的误差较大^[19]。而以具体的产业作为研究对象,其用能模型的影响因素较少,自变量与因变量之间的关联性更强。

1.2 产业经济指标细分

采用分行业的方式对某省的经济发展状况进行再划分,可实现更细化的研究。本文列出以行业划分的时序经济指标,如表1所示,可以看出,对于每一个用于刻画某地区经济状况的经济指标都可以按

表 1 产业经济指标分类

Table 1 Industrial economic index classification

行业	主要统计指标
农业	农林牧渔业总产值
工业	规上工业分行业增加值 规上工业主要产品产量
服务业	规上服务业分行业门类营业收入

照分行业的方式进一步细分,这与电力消费统计量的划分原则一致,为后续研究提供基础。

1.3 基于协整—格兰杰因果关系的指标关系分析

本文运用计量经济学中时间序列分析方法,包括平稳性分析、协整检验以及 Granger 因果关系检验等,其流程如图 1 所示。

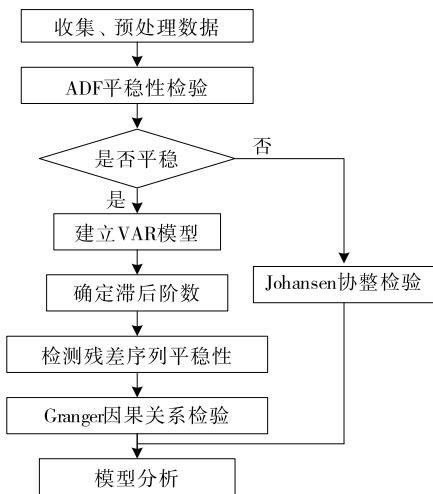


图 1 基于协整及 Granger 因果关系分析

Figure 1 Analysis flowchart based on the cointegration and Granger causality analysis

在对时序数据收集后,首先需要对数据进行预处理,去除异常数据。然后采用 ADF 检验对数据做平稳性分析,若平稳,则可直接建立 VAR 多元平稳时序模型以及 Granger 因果关系检验对数据进行回归分析;若数据非平稳,则需要使用协整分析,当数据间存在协整关系时,则可构建 Granger 因果关系检验,若变量间不存在协整关系,则为“伪回归”,Granger 因果检验失去意义。

1.3.1 单位根检验

一阶自回归模型(autoregressive model, AR)为

$$y_t = c + \alpha y_{t-1} + u_t \quad (1)$$

式中 c 为截距; u_t 为期望为 0 的误差; α 为自回归系数; y_{t-1} 、 y_t 分别为 $t-1$ 、 t 时刻的观测值。

单位根检验: AR(1)模型中 y_{t-1} 的参数是否为 1,目的是对时间序列数据的平稳性检验。若 y_t 为单位根过程,则 y_t 为非平稳序列。DF 检验原假设与备择假设为

$$\begin{cases} H_0: \alpha = 1 \\ H_1: \alpha < 1 \end{cases} \quad (2)$$

ADF 检验是 DF 检验的推广,即考虑 AR(p)模

型的平稳性,其原理与 DF 检验类似,不再赘述。

1.3.2 向量自回归(VAR)模型

多变量的因果关系检验可以通过构建 VAR 模型进行测试,VAR 模型是将单变量自回归模型推广到多变量组成的向量自回归模型,典型的 p 阶滞后的 VAR(p)模型为

$$Y_t = C + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_p Y_{t-p} + U_t \quad (3)$$

式中 $Y_t = [y_{1,t-i}, y_{2,t-i}, \cdots, y_{n,t-i}]^T$ ($i = 1, 2, \cdots, p$) 为滞后第 i 期的 n 维变量 $y_{1,t-i}, y_{2,t-i}, \cdots, y_{n,t-i}$ 组成的向量; $C = [c_1, c_2, \cdots, c_n]^T$ 为 $n \times 1$ 维常数向量; $U_t = [u_1, u_2, \cdots, u_n]^T$ 为 $n \times 1$ 维随机误差列向量; Φ_i ($i = 1, 2, \cdots, p$) 为 $n \times n$ 维自回归系数矩阵,即

$$\Phi_i = \begin{bmatrix} \varphi_{11}^{(i)} & \varphi_{12}^{(i)} & \cdots & \varphi_{1n}^{(i)} \\ \varphi_{21}^{(i)} & \varphi_{22}^{(i)} & \cdots & \varphi_{2n}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{n1}^{(i)} & \varphi_{n2}^{(i)} & \cdots & \varphi_{nn}^{(i)} \end{bmatrix}$$

1.3.3 Granger 因果关系检验

Granger 因果关系检验是一种假设检验的统计方法,检验一组时间序列 x 是否为另一组时间序列 y 的原因,利用 x 的历史数据提高对 y 变量的预测能力。

令 x, y 为广义平稳序列,先建立 y 的 p 阶自回归模型,再引入 x 的滞后期建立增广回归模型,即

$$\begin{cases} y_t = c_1 + \sum_{i=1}^p \alpha_{11}^{(i)} y_{t-i} + \sum_{i=1}^p \alpha_{12}^{(i)} x_{t-i} + u_{1t} \\ x_t = c_2 + \sum_{i=1}^p \alpha_{21}^{(i)} y_{t-i} + \sum_{i=1}^p \alpha_{22}^{(i)} x_{t-i} + u_{2t} \end{cases} \quad (4)$$

式中 c_1, c_2 为常数; $\alpha_{st}^{(i)}$ ($s, t = 1, 2$) 为第 i 期滞后自回归系数; u_{1t}, u_{2t} 为误差; x_{t-i}, y_{t-i} ($i = 1, 2, \cdots, p$) 分别为 x, y 第 i 期的观测值。

以 y_t 为例,若自回归系数满足 $\alpha_{12}^{(1)} = \alpha_{12}^{(2)} = \cdots = \alpha_{12}^{(p)} = 0$,则不存在 $x_t \rightarrow y_t$ 的 Granger 因果关系。通常利用 F 检验判断是否存在 Granger 因果关系。

1.3.4 协整关系分析

协整关系是指多个非平稳时间序列变量的线性组合形成的变量是平稳时间序列。协整检验的方法主要有 Engle-Granger(E-G 两步法)、Johansen 协整检验法,其中,前者协整适用单个协整关系的检验,而后者适用于多个协整关系的检验。

2 高能耗产业电量预测模型建模

2.1 预测模型的框架

产业电量预测模型的建模思路如图 2 所示,分为数据预处理、产业用电分类、细分产业用电量预测 3 个部分。

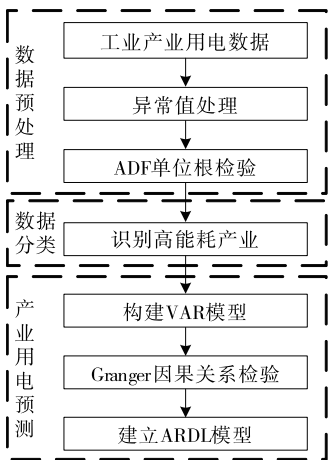


图 2 高能耗产业用电量预测流程

Figure 2 Flow chart of electricity consumption forecast for high energy consumption industries

1)数据预处理。首先对收集到的地区工业产业用电数据的异常值进行识别和处理,然后通过对比用电数据的同比增长率,确定该月度用电数据是否存在异常。

2)产业用电分类。结合地区电力部门统计的行业用电数据,包括产业月度用电量、用电装接容量、单位 GDP 电耗等数据,最终确定高能耗行业为非金属矿物制品业、黑色金属冶炼和压延加工业、纺织业、化学原料和化学制品制造业以及有色金属冶炼和压延加工业。

3)产业用电量预测。针对高能耗产业选取经济指标数据,包括主要产品产量、主营业务收入及利润总额。对细分产业构建 VAR 模型并进行 Granger 因果关系检验,基于此建立产业用电量预测的 ARDL 模型。

2.2 高能耗产业 ARDL 模型

ARDL 模型最早由 Charemza^[20] 等提出,在 Granger 因果关系检验结果的基础上,将高能耗产业中具有较强因果关系经济指标变量作为构建模型的首选变量,建立 ARDL 模型为

$$y_{\text{esum},t} = \alpha + \sum_{i=1}^p \varphi_i y_{\text{esum},t-i} + \sum_{j=1}^5 \sum_{k=1}^3 \sum_{l_{jk}=0}^{q_{jk}} \beta_{jkl_{jk}} x_{jk,t-l_{jk}} + v_t \quad (5)$$

式中 α 为常数; φ_i 、 $\beta_{jkl_{jk}}$ 为系数; v_t 为白噪声; p 、 q_{jk} 为最大滞后阶数; $y_{\text{esum},t}$ 、 $y_{\text{esum},t-i}$ 分别为 t 、 $t-i$ 时期地区高能耗产业的总用电量; $x_{jk,t-l_{jk}}$ 为第 j 个高能耗产业的第 k 个经济指标变量滞后 l_{jk} 个时期的值。

3 仿真算例

3.1 数据来源

本文选取 2016 年 1 月至 2021 年 6 月中国某地区月度用电数据作为产业用电数据集,月度经济数据来源于地区统计局。纺织业是该地区的重点和高能耗产业,其用电量与选取的经济指标如图 3 所示。

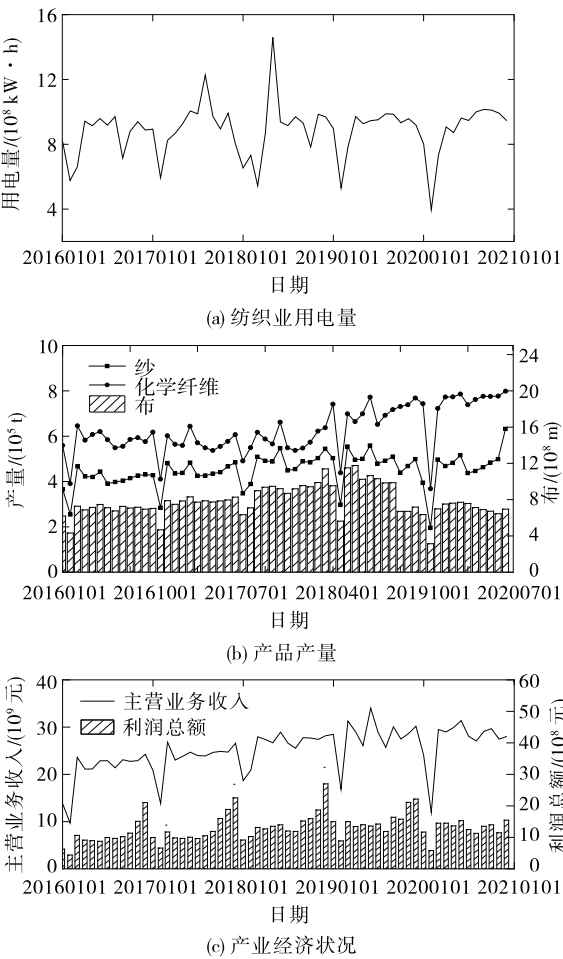


图 3 纺织业用电量与典型经济指标

Figure 3 Electricity consumption of textile industry and typical economic indicators

由图 3 可知,纺织业作为高能耗产业,其用电量和经济变量折线图的形状存在一定程度上的相似;同时,用电量和经济数据都存在明显的周期性循环波动,且受春节影响明显,在春节期间,产业用电和经济数据都呈现明显跌落。由于数据可能为非平稳时间序列,因此,在进行用电量合计与高耗能行业主要经济指标关系研究之前,需要对原始时间序列数据进行平稳性检验。

3.2 产业用电数据计量经济分析

3.2.1 单位根检验

根据图 1 对产业的分类,依次对工业中的高能耗产业的用电数据进行 ADF 单位根检验。 y_{esum} 、 y_{eh1} 、 y_{eh2} 、 y_{eh3} 、 y_{eh4} 、 y_{eh5} 分别表示高能耗产业、纺织业、化学原料和化学制品制造业、黑色金属冶炼和压延加工业、有色金属冶炼和压延加工业以及非金属矿物制品业的用电数据(采用 2016—2020 年产业用电量数据)。平稳性检验如表 2 所示, t 统计量即 T 检验,通过 t 分布理论来比较 2 个平均数的差异是否显著; P 值(P -value)为伴随概率,用于与原假设设定的显著性水平 α 进行比较。原假设为变量不平稳,若 $P \leq \alpha$,则拒绝原假设;反之,则接受原假设。

表 2 ADF 单位根

Table 2 ADF unit root				
变量	检验形式 (C, T, K)	t 统计量	P 值	原假设 结论
y_{esum}	($C, T, 0$)	-5.865 6	0.000	趋势平稳
y_{eh1}	($C, N, 0$)	-5.277 0	0.000	平稳 I(0)
y_{eh2}	($C, N, 0$)	-3.621 0	0.009	平稳 I(0)
y_{eh3}	($C, T, 0$)	-4.044 5	0.012	趋势平稳
y_{eh4}	($C, T, 0$)	-5.027 9	0.001	趋势平稳
y_{eh5}	($C, T, 9$)	-8.376 0	0.000	趋势平稳

由表 2 可知,高能耗产业的平稳性各不相同,部分为平稳时间序列;部分属于趋势平稳序列,其区别在于方程中是否含有时间趋势项。采用典型消除时间趋势项的差分方法,对趋势平稳的时间序列处理后再次检查该序列是否平稳,若平稳,则不需要进行 Johansen 协整检验。对表 2 中趋势平稳的变量进行差分处理后均为平稳序列。

3.2.2 Granger 因果关系分析

为了反映产业用电量与经济指标之间的关系,在满足平稳性的基础上,构建各细分高能耗产业的 VAR 模型并进行 Granger 因果关系检验。变量选

取:产业利润总额 x_{prf} 、产业主营业务收入 x_{inc} 、产业产量 x_{prd} 。分析各细分高耗能产业用电数据与经济指标的因果关系,考虑到样本容量限制,本文选择最大滞后阶数为 4,最终得到各个细分产业的因果关系,如表 3 所示。原假设指经济指标不能引起产业用电数据变化的 Granger 原因。

表 3 格兰杰因果校验结果

Table 3 Granger causality check result				
电力数据	经济指标	t 统计量	P 值	原假设结论
y_{eh1}	$x_{eh1, inc}$	24.610	0.00	拒绝
	$x_{eh1, prd}$	6.700	0.15	不拒绝
	$x_{eh1, prf}$	46.760	0.00	拒绝
y_{eh2}	$x_{eh2, inc}$	0.340	0.55	不拒绝
	$x_{eh2, prd}$	7.200	0.01	拒绝
	$x_{eh2, prf}$	0.660	0.41	不拒绝
y_{eh3}	$x_{eh3, inc}$	1.611	0.20	不拒绝
	$x_{eh3, prd}$	0.770	0.37	不拒绝
	$x_{eh3, prf}$	0.340	0.55	不拒绝
y_{eh4}	$x_{eh4, inc}$	8.140	0.00	拒绝
	$x_{eh4, prd}$	0.840	0.36	不拒绝
	$x_{eh4, prf}$	3.610	0.06	拒绝
y_{eh5}	$x_{eh5, inc}$	0.920	0.33	不拒绝
	$x_{eh5, prd}$	0.550	0.45	拒绝
	$x_{eh5, prf}$	11.560	0.00	不拒绝

由表 3 的检验结果表明:不同的高能耗产业的经济指标对于其用电量的 Granger 因果关系检验结果各不相同。化学原料和化学制品制造业、非金属矿物制品业都是由产业产量对用电量有 Granger 因果关系;而产业主营业务收入和利润总额是纺织业、有色金属冶炼和压延加工业用电量的 Granger 原因;黑色金属冶炼和压延加工业的经济指标数据对用电量都没有 Granger 因果关系。若特定的经济指标对具体产业的用电量具有 Granger 因果关系,则该经济指标的增长会引起该产业用电量的增长。

格兰杰因果关系检验的结论作为统计意义上的“格兰杰因果性”,不是真正意义上的因果关系,因此不能作为肯定或否定因果关系的根据。但是,其作为统计结论对于后续的用电量预测模型具有一定的参考价值。

3.2.3 高能耗产业月度用电量预测

在不同高能耗产业的 Granger 因果关系检验的

基础上,将这些对产业用电量增长具有 Granger 因果关系的经济变量添加入自回归模型之中,构建最大滞后阶数为 4 的 ARDL 模型,选取赤池信息量准则(akaike information criterion, AIC)作为模型的信息准则,剔除不显著的变量后得到的回归结果如表 4 所示,其中, $x_{eh1,prf}$ 、 $x_{eh4,inc}$ 、 $x_{eh4,prf}$ 、 $x_{eh5,prd}$ 分别表示纺织业的利润、有色金属工业的收入、有色金属行业的利润、非金属矿物制品业的产业产量。

表 4 产业用电量回归估计结果

Table 4 Regression estimation results of industrial electricity consumption			
回归变量	系数	标准差	P 值
$y_{eh}(-1)$	0.113 637	0.121 535	0.935 014
$y_{eh}(-2)$	-0.212 998	0.140 703	-1.513 818
$y_{eh}(-3)$	0.332 709	0.124 175	2.679 * * *
$x_{eh1,prf}$	-3 060.478 000	1 437.219 000	-2.129 * *
$x_{eh4,inc}$	12.060 980	119.257 800	0.101 1
$x_{eh4,inc}(-1)$	12.408 640	111.944 500	0.111 0
$x_{eh4,inc}(-2)$	317.934 300	111.242 400	2.858 * * *
$x_{eh4,prf}$	-372.759 400	1 275.338 000	-0.292 0
$x_{eh4,prf}(-1)$	-590.469 200	1 190.893 000	-0.496 0
$x_{eh4,prf}(-2)$	1 813.577 000	1 145.001 000	1.584 0
$x_{eh4,prf}(-3)$	3 003.167 000	1 183.386 000	2.538 * *
$x_{eh5,prd}$	338.099 800	41.603 970	8.127 * * *
$x_{eh5,prd}(-1)$	67.984 380	54.076 920	1.257 0
$x_{eh5,prd}(-2)$	-87.741 460	58.382 540	-1.503 0
$x_{eh5,prd}(-3)$	-82.201 160	57.921 380	-1.419 0
$x_{eh5,prd}(-4)$	-136.675 000	38.338 670	-3.565 * * *
C	239 328.400	45 075.940	5.309 * * *

注: * * *、* * 和 * 分别表示 1%、5% 以及 10% 水平上显著

由表 4 可见,该地区高能耗产业的用电量合计受到该地区的纺织业、有色金属、非金属矿物制品业的某些具体经济指标的显著影响,如:纺织业当期的利润总额、有色金属加工业的滞后 2 期的主营业务收入和利润;非金属矿物制品业的产物水泥当月与上一季度的产量。地区某个产业的经济指标的 增长,对于区域整体的高能耗产业的总用电量增长具有一定的影响作用。将经济指标与具体的产业现象相联系,进一步提升了模型的解释能力。

2016—2020 年高能耗产业用电量 ARDL 模型的拟合结果如图 4 所示,本文构建的 ARDL 模型和自回归模型的拟合效果如图 5 所示。

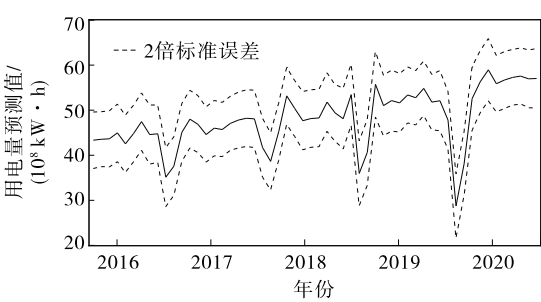


图 4 ARDL 模型拟合结果

Figure 4 ARDL model fitting results

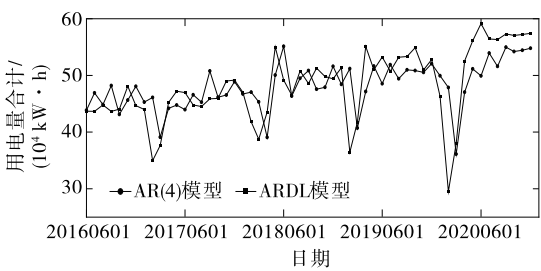


图 5 中长期电量预测模型拟合效果

Figure 5 Fitting diagram of mid-and long-term electricity forecasting model

由图 4 可知,与用电量实际值进行对比,模型的均方根误差(RMSE)为 $22\ 888.12\ 10^4\ \text{kW}\cdot\text{h}$,平均绝对值误差(MAE)为 $17\ 166\ 10^4\ \text{kW}\cdot\text{h}$,平均相对误差(MAPE)为 3.628 8%,Theil 不等系数(U)为 0.023 74,偏差比为 0,方差比为 0.023 74,协方差比为 0.976 25。从偏差比、方差比以及协方差比的结果看出,该模型对高能耗产业用电量的拟合效果较好。

由图 5 可得,ARDL 模型的残差平方和为 2.72×10^{10} ,回归标准差为 26 385.24,而 AR(4)模型的残差平方和为 1.59×10^{11} ,回归标准差为 55 850.97,均大于本文构建的 ARDL 模型,对比显示,ARDL 模型对数据的拟合效果更好。

通过最新收集的 2021 年 1—6 月的该地区高耗能行业主要经济指标与用电量数据,验证模型中选取的经济指标是否与用电量存在长期均衡关系,模型的预测效果如图 6 所示,可见除了春节期间的相对误差约为 10% 外,其他月份都约为 6% 左右,整体的平均相对误差为 6.84%,考虑新冠肺炎疫情对产业的冲击,该预测结果的精度处于可以接受的范围之内。

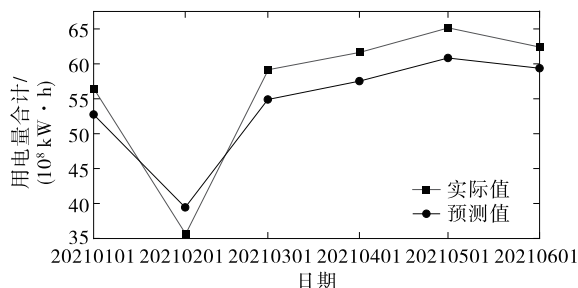


图 6 高耗能行业用电量模型预测效果

Figure 6 Prediction effect diagram of electricity consumption model for high energy-consuming industries

进一步分析预测结果可以发现,春节期间的预测值高于实际值;春节后的预测值都低于实际值。由于过往年份的预测结果中并没有显著的变化趋势,根据控制变量的原则,该现象与国内春节期间的高强度疫情防控策略相关;同时由于国内疫情得到有效控制,各个产业的用电需求在年后迅速反弹,导致年后实际值大于预测值。因此,对于预测模型的改进,还需结合产业政策、进出口贸易等多方面因素。

4 结语

本文研究细分高能耗产业的用电量与经济指标之间的关联分析,通过 VAR 模型和 Granger 因果关系挖掘影响细分产业用电量的经济变量,并在此基础上建立 ARDL 模型,对地区高能耗产业的用电量进行中长期用电预测。通过对地区产业的实际电力消费和经济数据进行方法验证,得出结论:

1) Granger 因果关系检验和协整理论可以有效挖掘产业经济与用电的影响关联关系,对于含多元变量的模型分析有着重要的作用;将数理统计的理论应用于数据挖掘可以提高模型的解释性;

2) Granger 因果关系检验表明,不同细分产业发电量与各经济指标之间的关联性不同,具有关联性的产业经济指标的增长会促进该产业电力消费的增长;同时,考虑产业经济指标对于用电量预测模型的精度提升具有明显作用;

3) 基于 ARDL 和 Granger 因果关系的高能耗产业用电量预测模型,得到的预测与实际数据的相对误差较小;同时,研究细分产业的用电量影响因素对高能耗产业的整体用电量预测起到了较好的效果。

参考文献:

- [1] 李泽文,邓拓夫,曾祥君,等. 智能电网能量流的时空多尺度大数据探讨[J]. 电力科学与技术学报, 2015, 30(1): 22-27.
LI Zewen, DENG Tuofu, ZENG Xiangjun, et al. Study on big data with time-space multi-scale for smart grid power flow[J]. Journal of Electric Power Science and Technology, 2015, 30(1): 22-27.
- [2] 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.
- [3] 韩肖清,李廷钧,张东霞,等. 双碳目标下的新型电力系统规划新问题及关键技术[J]. 高电压技术, 2021, 47(9): 3036-3046.
HAN Xiaoqing, LI Tingjun, ZHANG Dongxia, et al. New issues and key technologies of low-carbon power system planning under double carbon goals[J]. High Voltage Engineering, 2021, 47(9): 3036-3046.
- [4] 谭显东,刘俊,徐志成,等. “双碳”目标下“十四五”电力供需形势[J]. 中国电力, 2021, 54(5): 1-6.
TAN Xiandong, LIU Jun, XU Zhicheng, et al. Power supply and demand situation in the 14th Five-Year Plan under the "Double Carbon" target[J]. Electric Power, 2021, 54(5): 1-6.
- [5] 童光毅. 基于双碳目标的智慧能源体系构建[J]. 智慧电力, 2021, 49(5): 1-6.
TONG Guangyi. The construction of a smart energy system based on dual carbon goals[J]. Smart Power, 2021, 49(5): 1-6.
- [6] ZHANG C, ZHOU K L, YANG S L, et al. On electricity consumption and economic growth in China[J]. Renew Sustain Energy Reviews, 2017, 76: 353-368.
- [7] LIN B Q, OMOJU O E, OKONKWO J U. Factors influencing renewable electricity consumption in China[J]. Renew Sustain Energy Reviews, 2016, 55: 687-696.
- [8] GE F, YE B, XING S N, et al. The analysis of the underlying reasons of the inconsistent relationship between economic growth and the consumption of electricity in China—a case study of Anhui province[J]. Energy, 2017, 128: 601-608.

- [9] LIU D, RUAN L, LIU J C, et al. Electricity consumption and economic growth nexus in Beijing: a causal analysis of quarterly sectoral data[J]. Renewable Sustainable Energy Reviews, 2018, 82: 2498-2503.
- [10] LI K, YUAN W H. The nexus between industrial growth and electricity consumption in China-new evidence from a quantile-on-quantile approach[J]. Energy, 2021, 231: 120991.
- [11] 庞传军, 张波, 余建明, 等. 基于结构化负荷模型的电力负荷概率区间预测[J]. 中国电力, 2021, 54(9): 89-95.
PANG Chuanjun, ZHANG Bo, YU Jianming, et al. Probabilistic interval forecasting of power load based on structured load model[J]. Electric Power, 2021, 54(9): 89-95.
- [12] 袁家海, 丁伟, 胡兆光. 电力消费与中国经济发展的协整与波动分析[J]. 电网技术, 2006, 30(9): 10-15.
YUAN Jiahai, DING Wei, HU Zhaoguang. Analysis on cointegration and co-movement of electricity consumption and economic growth in China[J]. Power System Technology, 2006, 30(9): 10-15.
- [13] 李强, 赵健, 王磊, 等. 配电网馈线负荷预测及风险评估预警分析方法[J]. 供用电, 2020, 37(5): 50-55.
LI Qiang, ZHAO Jian, WANG Lei, et al. Distribution network feeder load forecasting and risk assessment warning analysis method[J]. Distribution & Utilization, 2020, 37(5): 50-55.
- [14] 王雁凌, 吴梦凯. 经济新常态下基于偏最小二乘回归的中长期负荷预测模型[J]. 电力自动化设备, 2018, 38(3): 133-139.
WANG Yanling, WU Mengkai. Medium and long-term load forecasting model based on partial least squares regression under the new economic normal[J]. Electric Power Automation Equipment, 2018, 38(3): 133-139.
- [15] 张瑶, 王傲寒, 张宏. 中国智能电网发展综述[J]. 电力系统保护与控制, 2021, 49(5): 180-187.
ZHANG Yao, WANG Aohan, ZHANG Hong. Overview of smart grid development in China[J]. Power System Protection and Control, 2021, 49(5): 180-187.
- [16] 刘江永, 刘文翰, 易灵芝. 多时序协同中期负荷预测模型[J]. 电力系统及其自动化学报, 2020, 32(2): 48-53.
LIU Jiangyong, LIU Wenhan, YI Lingzhi. Multi-sequence coordinated medium-term load forecasting model[J]. Proceedings of the CSU-EPSC, 2020, 32(2): 48-53.
- [17] 王飞, 李正辉, 李渝, 等. 基于数据序列分辨率压缩尺度优化的月度电量预测方法[J]. 电力系统保护与控制, 2020, 48(11): 62-68.
WANG Fei, LI Zhenghui, LI Yu, et al. Data series resolution compression scale optimization based monthly electricity consumption forecasting[J]. Power System Protection and Control, 2020, 48(11): 62-68.
- [18] 陈浩文, 刘文霞, 李月乔. 基于奇异谱分析与神经网络的中期负荷预测[J]. 电网技术, 2020, 44(4): 1333-1347.
CHEN Haowen, LIU Wenxia, LI Yueqiao. Medium-term load forecast based on singular spectrum analysis and neural network[J]. Power System Technology, 2020, 44(4): 1333-1347.
- [19] 王永伟, 李新龙, 田斐, 等. 基于人群搜索算法的电网短期用电负荷预测研究[J]. 电网与清洁能源, 2020, 36(12): 35-40.
WANG Yongwei, LI Xinlong, TIAN Fei, et al. Research on short-term electric load forecasting of power grid based on crowd search algorithm[J]. Power System and Clean Energy, 2020, 36(12): 35-40.
- [20] CHAREMZA W W, DEADMAN D F. New directions in econometric analysis[M]. Oxford: Oxford University Press, 1997: 150-153.