

## 基于 PCC-LSTM 模型的全国碳排放权交易价格预测研究

朱亮亮<sup>1,2</sup>, 肖楚鹏<sup>1,2</sup>, 余梦<sup>1,2</sup>, 王迎秋<sup>3</sup>, 李野<sup>3</sup>, 赵猛<sup>3</sup>

1. 国网电力科学研究院有限公司, 南京 210000;
2. 国网电力科学研究院武汉能效测评有限公司, 武汉 430074;
3. 国网天津市电力公司, 天津 300090)

**摘要:** 为了有效提升碳价的预测精度和保证全国碳交易市场的稳定运行, 在利用皮尔逊相关系数法 (PCC) 提取关键碳价影响要素的基础上, 分别运用 ARIMA 模型、多项式回归算法和 LSTM 模型, 构建了全国碳交易市场的碳价预测模型。对比结果显示: 相较于其它两类模型, PCC-LSTM 模型的预测精度最高, 平均精度达到 99.26%, 标准偏差最小, 仅为 0.263 0, 很好地显示了深度学习算法在样本数据充足条件下的优越性和实用性。相较于上半年的平均碳价, 预测得到的 11 月碳价渐趋平稳, 且有所提高, 一定程度反映了全国碳交易市场整体运行平稳和年底交易活跃的特性。PCC-LSTM 模型在碳价预测领域的成功应用, 有助于企业熟悉碳市场的运行机制和变化规律, 对扩大碳市场的覆盖范围和确保碳市场的稳定、健康发展具有重要的推动作用。

**关键词:** 碳价预测; 皮尔逊相关系数; ARIMA; 多项式回归; LSTM; 全国碳市场

中图分类号: X22

文献标志码: A

DOI: 10.16803/j.cnki.issn.1004-6216.2022090050

## Research on national carbon emission trading price prediction based on PCC-LSTM model

ZHU Liangliang<sup>1,2</sup>, XIAO Chupeng<sup>1,2</sup>, YU Meng<sup>1,2</sup>, WANG Yingqiu<sup>3</sup>, LI Ye<sup>3</sup>, ZHAO Meng<sup>3</sup>

1. NARI Group Corporation/State Grid Electric Power Research Institute, Nanjing 210000, China;
2. Wuhan Energy Efficiency Evaluation Corporation of State Grid Electric Power Research Institute, Wuhan 430074, China;
3. State Grid Tianjin Electric Power Company, Tianjin 300090, China)

**Abstract:** To improve the prediction accuracy of the carbon price and ensure the stable operation of the national carbon market, the Pearson correlation coefficient (PCC) method was first used to determine the key influence factors of carbon price; 3 carbon price prediction models in the national carbon market were established based on 3 types of prediction approaches, named as ARIMA model, polynomial regression algorithm and LSTM model, respectively. The comparison results among 3 models demonstrated that, compared with the other 2 types of models, the PCC-LSTM model had the highest prediction accuracy, with an average accuracy of 99.26%, and the smallest standard deviation of 0.2630. It reflected the superiority and practicability of the deep learning algorithm under the condition of sufficient sample data. Compared with the average carbon price for the first half of 2022, the predicted carbon price in November would gradually stabilize and increase, which reflected overall stable operation and active trading characteristics at the end of the year of the national carbon trading market. The successful application of the PCC-LSTM model in carbon price prediction was capable of helping enterprises to familiarize themselves with the operation mechanism and variation rule of the carbon market, expanding the coverage range of the carbon market and ensuring its stable and healthy development.

**Keywords:** carbon price prediction; pearson correlation coefficient; ARIMA; polynomial regression; LSTM; national carbon market

CLC number: X22

人类社会快速发展产生的大量温室气体是导致全球变暖的最主要原因。中国作为能源消费和

碳排放大国, 在第七十五届联合国大会上, 习近平总书记宣布了中国力争 2030 年实现碳达峰和 2060

收稿日期: 2022-09-28

录用日期: 2022-11-07

基金项目: 国家电网公司总部科技项目(5400-202112582A-0-5-SF)

作者简介: 朱亮亮(1992—), 男, 硕士、工程师。研究方向: 电力需求侧管理、电碳耦合、综合能源。E-mail: zll\_seu@163.com

引用格式: 朱亮亮, 肖楚鹏, 余梦, 等. 基于 PCC-LSTM 模型的全国碳排放权交易价格预测研究[J]. 环境保护科学, 2023, 49(5): 55-62.

年前实现碳中和的“双碳”目标,充分表明了我国政府解决温室气体排放问题,坚决走节能减排、低碳可持续发展道路的决心。碳排放权交易机制作为有效的减排合作机制<sup>[1]</sup>,在碳减排过程中发挥着重大作用。2021年7月16日,全国碳排放权交易市场正式开启,我国碳减排工作迈入新阶段。碳排放权交易价格(以下简称“碳价”)作为碳交易市场的核心要素,其波动变化将直接影响企业的生产决策和投资者的买卖行为;碳价的准确预测,可以帮助企业熟悉碳市场的运行机制和变动规律,对扩大碳市场的覆盖行业和确保碳市场的稳定、健康发展具有重要的推动作用<sup>[2]</sup>。

目前,常用的碳价预测方法包括以下两类:分别是基于碳价历史数据的时间序列分析模型和基于碳价影响因素的回归预测模型<sup>[3]</sup>。在基于历史数据的碳价预测方面,王娜<sup>[4]</sup>运用 Boosting-ARMA 时间序列算法对洲际交易所(International Exchange, ICE)的 EUA(European Union Allowances, EUA)期货价格进行预测,显著提高了碳价预测的准确性。LI et al<sup>[5]</sup>以我国 5 个碳交易试点为研究对象,将 EMD-GARCH 模型与 CGE 模型相结合,给出全国碳市场合理碳价区间为 30~50 元/tCO<sub>2</sub>。ZHU et al<sup>[6]</sup>通过构建 EMD-ARIMA-ADD 模型,实现对欧盟碳交易市场四种碳期货价格的准确预测。JIANG et al<sup>[7]</sup>在将碳价原始序列分为线性和非线性两部分的基础上,采用 ARIMA-RF 集成模型进行碳价预测,预测精度高于传统时间序列模型。在基于影响因素的碳价预测方面,LI et al<sup>[8]</sup>以湖北试点碳市场为研究对象,将三种能源价格作为碳价影响因素,利用长短期记忆(Long-Short Term Memory, LSTM)神经网络算法对碳价进行预测,平均绝对误差仅为 0.617。高长征等<sup>[9]</sup>选取能源价格、股票市场、汇率等六类碳价影响要素,在运用完全集合经验模态分解方法(Complete EEMD with Adaptive Noise, CEEMDAN)对湖北碳市场的碳价序列进行分解的基础上,利用 Transformers 智能算法完成碳价预测,平均绝对百分比误差低至 1.43%。ZHANG et al<sup>[10]</sup>采用 TCN-Seq2Seq 深度神经网络模型,将金融指标和能源价格作为主要的模型输入对 EUA 期货价格进行预测,平均精确度达 96.97%。

上述研究主要以欧盟碳市场和我国的试点碳市场为研究对象,取得的成果很好地展示了两类预测方法的实用性和可靠性。但是,由于缺乏针对全国碳市场的碳价预测研究,无法检验预测方法的实际应用效果。ARIMA 模型作为常用的时间序列模型,结合历史碳价的波动变化情况<sup>[11]</sup>,完成未来碳价的趋势预测,模型操作简单且预测精度较高;多项式回归算法作为传统的回归分析方法,通过构建影响因素与碳价的响应关系,可以很好地考察各类因素对碳价的影响程度和探明碳价的变化趋势。LSTM 模型是一种以循环神经网络(Recurrent Neural Network, RNN)为基础并加以改进的深度学习算法<sup>[12]</sup>,通过建立影响因素和碳价的映射关系,真正有效地利用时序信息,实现对碳价的准确预测。综上所述,上述 3 类预测方法原理不同且特点各异,如何结合碳价本身的变化情况和掌握的数据信息质量,选取合适的预测方法,对于提高碳价预测精度至关重要。因此,本研究以全国碳市场碳价预测为主要研究目标,在采用皮尔逊相关系数法筛选和确定碳价影响因素的基础上,选取 3 类具有代表性,且广泛使用的 ARIMA 模型、LSTM 模型和多项式回归算法分别构建碳价预测模型,通过比较和选定最佳预测模型来实现碳价预测精度的有效提升,为企业制定合理的碳市场参与策略,政府平抑碳价波动,乃至全国碳市场未来的平稳、健康发展提供技术支持。

## 1 研究方法

### 1.1 总体思路

本文遵循“影响因素识别-模型构建-结果对比分析”技术路线,首先采用皮尔逊相关系数法对包括宏观经济、能源价格、人民币汇率、前一交易日碳价在内的 4 类要素,6 个变量和碳价进行相关性分析,提取出对碳价影响较大的 2 个变量,即前一交易日碳价、WTI(West Texas Intermediate)原油价格;然后,分别构建基于 ARIMA 模型、LSTM 模型和多项式回归算法的碳价预测模型,计算得到模型的预测精度和相对误差;最后,对比分析三种模型的预测精度,选定最适宜预测模型。总体技术路线见图 1。

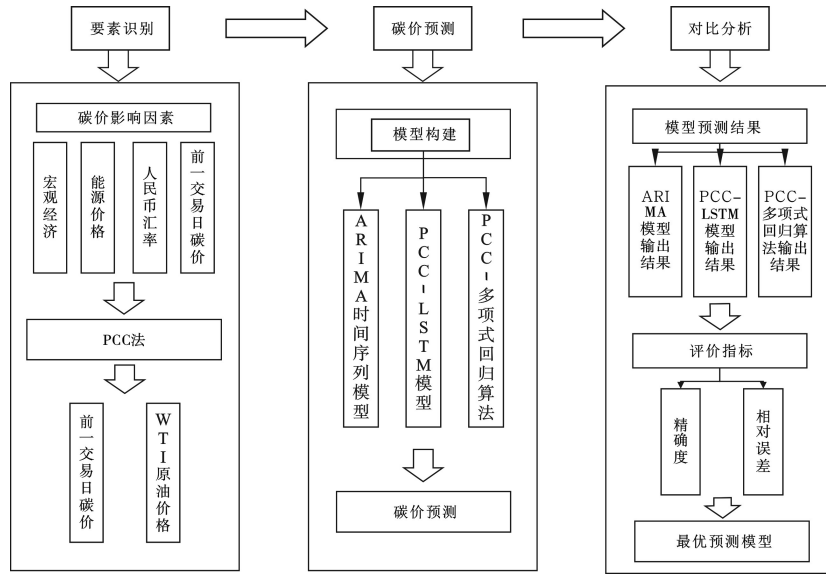


图 1 技术路线

Fig. 1 The technology roadmap

1.2 皮尔逊相关系数

皮尔逊相关系数(Pearson Correlation Coefficient, PCC)作为一种衡量变量间关联度的方法,常用于回归分析模型的输入变量选取<sup>[13]</sup>。运用该方法识别碳价的关键影响因素的核心计算,见式(1):

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{N}\right)\left(\sum Y^2 - \frac{(\sum Y)^2}{N}\right)}} \quad (1)$$

式中:  $r$  为碳价与影响因素的相关系数;  $Y$  为碳价;  $X$  为初始选定的碳价影响因素;  $N$  为  $X$ 、 $Y$  所包含的数据个数。

运用 PCC 法得到的相关系数  $r$  的取值范围为  $[-1,1]$ 。其中,  $r$  的绝对值可以表示变量间的相关性强弱程度。另外,  $r$  取值为正,代表两个变量之间的关系为正相关;相反,  $r$  取值为负,意味着两个变量之间的关系为负相关<sup>[14]</sup>。不同的相关系数  $r$  对应的相程度,见表 1。

1.3 ARIMA 时间序列模型

时间序列法是在分析预测对象的历史变化规律的基础上,对其未来变化趋势进行预测的方法。由于不需要考虑影响因素,模型建立过程简单;在变量波动有限、变化平稳的前提条件下,模型预测精度较高<sup>[15]</sup>。受制于全国碳市场外部因素的影响,碳价常呈现出非平稳特征,对于非平稳的数据序列,常用差分自回归移动平均模型(ARIMA),通过

差分变换将非平稳的碳价序列转为平稳序列,进而对碳价进行预测,见式(2):

$$\varphi(K)(1-K)^d x_t = \theta(K)\varepsilon_t \quad (2)$$

式中:  $\varphi(K)$  为自回归算子;  $(1-K)^d$  为差分算子,  $d$  为差分阶数;  $x_t$  为碳价时间序列;  $\theta(K)$  为移动平均算子;  $\varepsilon_t$  为残差序列。

表 1 相关系数  $r$  值与相程度对照表  
Table 1 The correlation coefficient  $r$  with the correlation degree

$r$ 值	相程度
$ r =0$	完全不相关
$0 <  r  < 0.3$	微弱相关
$0.3 <  r  < 0.5$	低度相关
$0.5 <  r  < 0.8$	显著相关
$0.8 <  r  < 1$	高度相关
$ r =1$	完全相关

1.4 多项式回归算法

多项式回归算法(Polynomial Regression)属于有监督的传统回归学习算法,在碳价预测中,以碳价为因变量,碳价影响因素为自变量,运用数据集训练得到二者之间的相关关系,进而完成碳价预测。

1.4.1 多项式回归模型的确定 常用的多项式模型一般不超过四阶,目前最常用的多项式回归模型是二阶多项式模型,见式(3):

$$y = \beta_0 + \sum_{i=1}^n \beta_i x_i + \sum_{i=1}^n \beta_{ij} x_i^2 + \sum_{i=1}^n \sum_{j=i+1}^n \beta_{ij} x_i x_j \quad (3)$$

式中:  $x_i$  为  $n$  维自变量;  $\beta_0, \beta_i, \beta_{ii}$ , 和  $\beta_{ij}$  为回归系数。二阶多项式回归模型中共有  $(n+1)(n+2)/2$  个函数项。

1.4.2 多项式回归模型的求解 将多项式回归方程转化为多元线性回归方程, 完成多项式回归系数的求解<sup>[16]</sup>。以一个包含两个变量的二阶多项式为例, 其求解步骤见式(4~5):

$$y = a_0 + a_1z_1 + a_2z_2 + a_3z_1^2 + a_4z_2^2 + a_5z_1z_2 \quad (4)$$

令  $x_1 = z_1, x_2 = z_2, x_3 = z_1^2, x_4 = z_2^2, x_5 = z_1z_2$ , 将上式转化为多元线性回归:

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 \quad (5)$$

利用最小二乘法, 即可求得多项式回归系数  $a_0, a_1, a_2, a_3, a_4$  和  $a_5$ 。

### 1.5 长短期记忆神经网络 (LSTM) 模型

LSTM 模型主要由记忆单元 (Cell)、输入门 (Input Gate)、输出门 (Output Gate) 和遗忘门 (Forget Gate) 等逻辑单元组成<sup>[17]</sup>。其中, 输入门和输出门主要用来接收、输出和修正参数, 分别记为  $l, \omega$ 。遗忘门表示是否保留当前隐藏层节点存储的历史信息<sup>[18]</sup>, 记为  $\varphi$ 。记忆单元表示神经元状态的记忆, 记为  $S_c^t$ 。3 个门以及独立的记忆单元设计, 使得 LSTM 单元具备保存、读取、重置和更新长距离历史信息的能力, 其实现步骤如下:

(1) 将  $t$  时刻数据代入到输入层, 经过激励函数输出结果。LSTM 输入门的输入由 3 部分组成, 包括  $t$  时刻输入层神经元的输入向量  $X_i^t$ , 前一个隐藏层的输出门的输出向量  $b_h^{t-1}$  和前一时刻记忆单元的保留信息  $b_c^{t-1}$ 。  $W_i$  为输入信息与各个门结构或输入单元状态的权重矩阵。  $t$  时刻输入门的输入向量  $a_i^t$  以及经过激励函数  $f$  得到的输出向量  $b_i^t$  的计算公式见式(6~7):

$$a_i^t = \sum_{i=1}^I W_{i1} X_i^t + \sum_{h=1}^H W_{h1} b_h^{t-1} + \sum_{c=1}^C W_{c1} b_c^{t-1} \quad (6)$$

$$b_i^t = f(a_i^t) \quad (7)$$

(2) 将输入层的输出结果信息输入 LSTM 结构的节点中, 并通过输入门、遗忘门和记忆单元的处理, 输出数据到下一隐藏层或输出层。与输入门一样, 遗忘门的输入  $a_\varphi^t$  也由相同 3 部分组成,  $t$  时刻的输出向量  $b_\varphi^t$  的计算, 见式(8~9):

$$a_\varphi^t = \sum_{i=1}^I W_{i\varphi} X_i^t + \sum_{h=1}^H W_{h\varphi} b_h^{t-1} + \sum_{c=1}^C W_{c\varphi} b_c^{t-1} \quad (8)$$

$$b_\varphi^t = f(a_\varphi^t) \quad (9)$$

记忆单元  $a_c^t$  的输入由两部分组成, 包括输入向量  $X_i^t$  和前一个隐含层的输出门的输出  $b_h^{t-1}$ , 见式(10):

$$a_c^t = \sum_{i=1}^I W_{ic} X_i^t + \sum_{h=1}^H W_{hc} b_h^{t-1} \quad (10)$$

记忆单元  $S_c^t$  依据遗忘门的判断决定是否保留过去的信息, 见式(11):

$$s_c^t = b_\varphi^t s_c^{t-1} + b_i^t g(a_c^t) \quad (11)$$

式中,  $g$  为激励函数。

(3) 将 LSTM 结构节点的输出传递到输出门神经元, 输出结果。将记忆单元的输出向量  $S_c^t$  作为神经网络输出门的输入向量, 并经过 Sigmoid 输出信息  $o^t$  和激励函数  $h$ , 得到最终结果  $b_h^t$ , 见式(12):

$$b_h^t = o^t h(s_c^t) \quad (12)$$

(4) 误差反向传播, 更新各个权值。对于 LSTM 的反向传播过程, 将误差记为  $L$ , 见式(13):

$$\delta_j^t = \frac{\partial L}{\partial a_j^t} \varepsilon_c^t = \frac{\partial L}{\partial b_c^t} \varepsilon_s^t = \frac{\partial L}{\partial s_c^t} \quad (13)$$

式中:  $\delta_s^t$  表示  $t$  时刻  $j$  节点的误差信号;  $\varepsilon_c^t$  表示  $t$  时刻记忆单元输出的误差信号;  $\varepsilon_s^t$  表示  $t$  时刻记忆单元中的保持值误差信号。

通过对记忆单元输出误差信号、输出门误差信号、记忆单元误差信号、遗忘门误差信号、输入门误差信号的求解, 对权值进行更新, 权值更新公式见式(14~15):

$$w_{ij} = w_{ij} - \eta \nabla L(w_{ij}) \quad (14)$$

$$\nabla L(w_{ij}) = \frac{\partial L}{\partial w_{ij}} = \frac{\partial L}{\partial a_j^t} \frac{\partial a_j^t}{\partial w_{ij}} = \delta_j^t b_j^t \quad (15)$$

基于上述 LSTM 神经网络训练步骤, 可以消除层数增加和计算梯度下降时出现的梯度消失和梯度爆炸问题, 有助于提高智能模型的预测精度。

### 1.6 模型预测性能指标

以模型的碳价预测结果作为输出参数, 对比全国碳交易市场的实际碳价, 选定精确度、相对误差和标准偏差作为主要指标, 量化评估 3 个模型的预



测性能,用于后续最适宜预测模型的确定,见式(16~18):

$$d = \frac{|p_{\text{实际}} - p_{\text{预测}}|}{p_{\text{实际}}} \times 100\% \quad (16)$$

$$A = 1 - d \quad (17)$$

$$S = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (P_i - \bar{P})^2} \quad (18)$$

式中: $d$ 为相对误差; $p_{\text{实际}}$ 为实际碳价; $p_{\text{预测}}$ 为碳价的预测值; $A$ 为模型的预测精度; $S$ 为标准偏差; $N$ 为样本交易日数量; $P_i$ 为交易日碳价; $\bar{P}$ 为交易日碳价的平均值。

## 2 案例分析

### 2.1 案例概述

本文选取 2021 年 7 月 16 日至 2022 年 6 月 2 日的全国碳市场碳价数据作为研究对象,选定宏观经济、能源价格、人民币汇率和前一交易日碳价 4 类因素作为候选的碳价影响要素,预测全国碳交易市场的碳价变化趋势和波动规律。碳价的主要影响因素,见表 2。全国碳市场碳价和上述影响因素的变化趋势,见图 2~图 4。

表 2 碳价主要影响因素  
Table 2 The main influence factors of carbon price

类别	变量含义	影响机理
宏观经济	中证工业指数	对实际碳排放量产生影响,导致市场需求发生变化,进而影响碳价
	沪深300	
能源价格	WTI原油	能源价格的变动影响化石能源的使用量,影响碳排放量,进而影响碳价
	液化天然气	
人民币汇率	美元	直接关系到商品的进出口贸易以及各类碳资产的交易
	欧元	

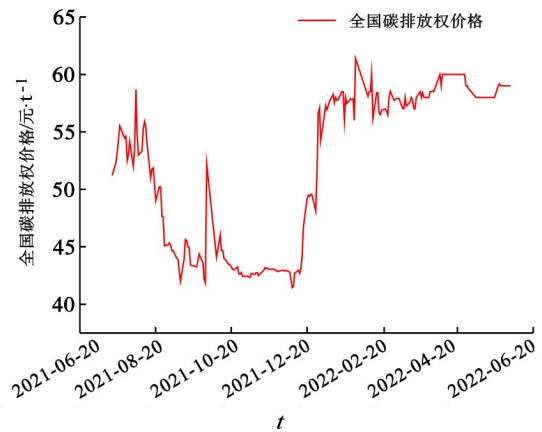


图 2 全国碳市场碳价走势  
Fig. 2 The variation in carbon price of national carbon market

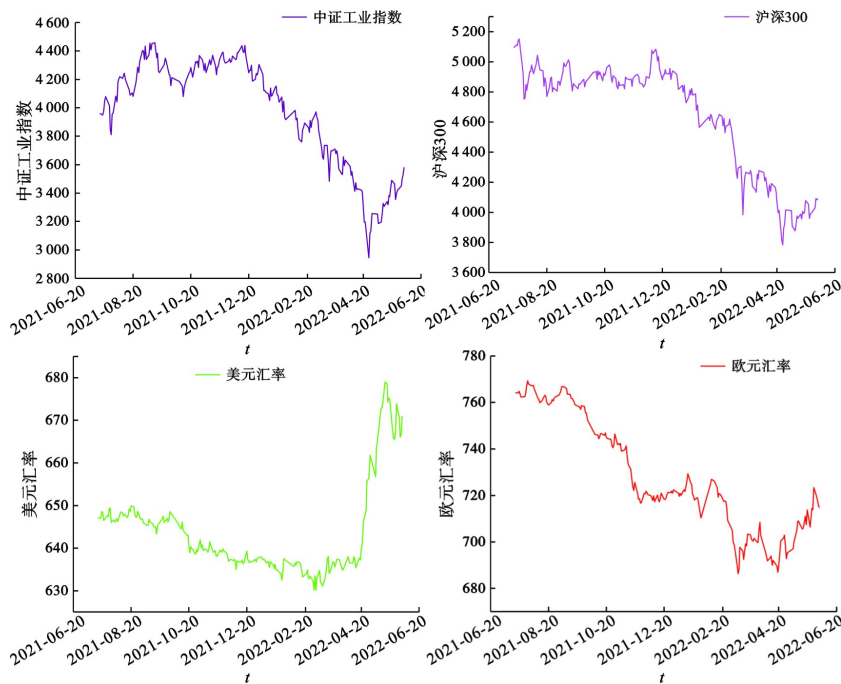


图 3 宏观经济及人民币汇率走势  
Fig. 3 The influence factors at economic aspect

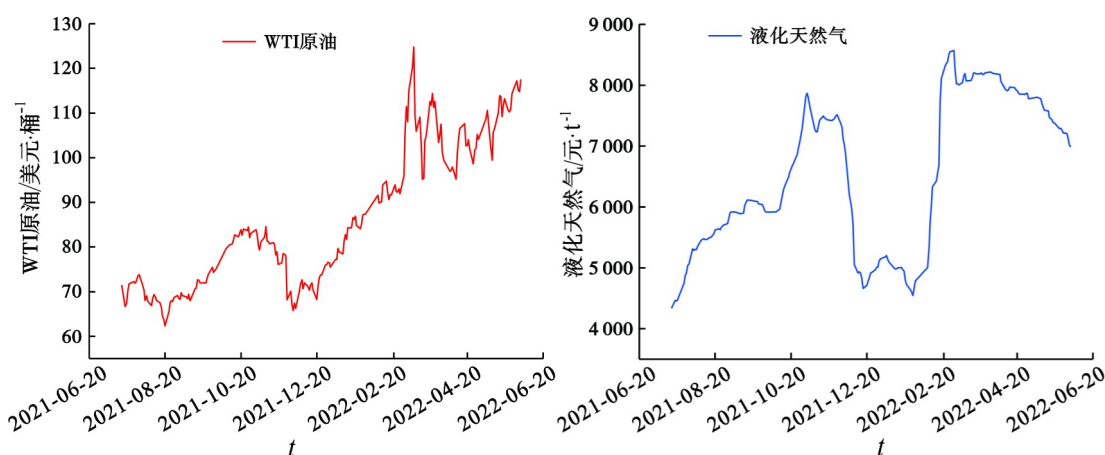


图 4 能源价格走势

Fig. 4 The variation tendency of energy price

## 2.2 变量相关性分析

为了准确识别碳价的关键影响因素和确保后续模型构建的合理性,采用 PCC 法对前述提到的 4 类、合计 6 个影响要素进行相关性分析,见表 3。

表 3 相关性分析结果汇总表

	相关系数值 $r$	相关程度
碳价-前一交易日碳价	0.973 3	高度正相关
碳价-中证工业指数	-0.779 8	显著负相关
碳价-沪深300	-0.662 7	显著负相关
碳价-液化天然气价格	0.140 9	微弱正相关
碳价-WTI原油价格	0.622 8	显著正相关
碳价-美元汇率	-0.250 8	微弱负相关
碳价-欧元汇率	-0.522 4	显著负相关

前一交易日碳价与碳价的相关程度最高, $r$  值达到 0.973 3,为高度正相关,表明全国碳市场的碳价走势呈现出一定的延续性;WTI 原油价格的相关程度为显著正相关, $r$  值为 0.622 8。相反,中证工业指数、沪深 300 和欧元汇率的相关程度为显著负相关,表明在近年来日益明显的股票经济与实体经济的相反关系的作用下,实体经济的下滑对碳价确实存在一定的抑制作用。至于液化天然气价格和美元汇率,相关性程度很低,可忽略不计。根据相关性分析结果,为保证模型拟合过程的稳定性和简易性,选定与碳价为正相关关系且  $r > 0.5$  的影响因素。因此,将前一交易日碳价和 WTI 原油价格作为 LSTM 和多项式回归算法的输入变量,构建基于影响因素的碳价预测模型。

## 2.3 结果分析

2.3.1 ARIMA 模型预测结果分析 以全国碳市场的碳价为研究对象,选取 2021 年 7 月 16 日至 2022 年 4 月 29 日的碳价数据作为训练样本用于构

建预测模型,2022 年 5 月 5 日—2022 年 6 月 2 日的碳价信息作为验证样本来考察 ARIMA 模型预测全国碳价的可靠性。ARIMA 模型的预测精度,见表 4。

表 4 ARIMA 模型的碳价预测情况

Table 4 The predicted results of ARIMA model

日期	预测碳价 /元·t <sup>-1</sup>	实际碳价 /元·t <sup>-1</sup>	预测精确 度/%	相对误差/%
2022-5-5	58.74	58	98.72	1.28
2022-5-6	59.06	58	98.17	1.83
2022-5-9	58.85	58	98.53	1.47
2022-5-10	59.19	58	97.94	2.06
2022-5-11	59.53	58	97.37	2.63
2022-5-12	59.90	58	96.72	3.28
2022-5-15	59.29	58	97.77	2.23
2022-5-16	59.79	58	96.91	3.09
2022-5-17	59.98	58	96.59	3.41
2022-5-18	59.72	58	97.03	2.97
2022-5-19	59.56	58	97.31	2.69
2022-5-20	59.59	58	97.27	2.73
2022-5-23	59.41	58.97	99.25	0.75
2022-5-24	59.74	59.18	99.06	0.94
2022-5-25	59.52	59	99.11	0.89
2022-5-26	59.87	59	98.53	1.47
2022-5-27	60.20	59	97.97	2.03
2022-5-30	60.57	59	97.33	2.67
2022-5-31	59.97	59	98.36	1.64
2022-6-1	60.46	59	97.52	2.48
2022-6-2	60.65	59	97.20	2.80
平均预测精确度			97.84%	
平均相对误差				2.16%
标准偏差				0.518 4

表 4 可知,总体预测精度较高,逐日的预测精度都在 96% 以上,平均精确度高达 97.84%。其中,2022 年 5 月 23 日的预测精度最高,为 99.25%;相反,2022 年 5 月 12 日、16 日和 17 日的预测精度普遍偏低,均低于 97%。这是因为 2022 年 5 月的前 10 个交易日(即 5 月 5 日至 5 月 20 日)的实际碳价

保持不变,背离了全年碳价先上升后下降的趋势。ARIMA 模型是基于碳价的整体变化情况进行预测,因此,该时段的碳价预测精度相对偏低。

2.3.2 多项式回归算法预测结果分析 二阶多项式模型的碳价预测结果,见表 5。该模型选定前一交易日碳价和 WTI 原油价格作为输入变量,预测得到 2022 年 5 月 5 日至 2022 年 6 月 2 日的碳价。表 5 可知,多项式回归算法的整体表现良好,逐日的预测精度均超过 96%,平均精度高于 ARIMA 模型,达到 98.17%。其中,2022 年 5 月 5 日的预测精度最高,高达 99.49%;相反,2022 年 5 月 23 日的预测精确度偏低,为 96.57%。主要原因在于,不同于前几个交易日的原油价格逐渐上涨,该交易日的 WTI 原油价格几乎不变,导致模型的预测结果出现一定偏差。另外,相较于 ARIMA 模型在前 10 个交易日的较差表现,多项式回归模型的表现并未受到碳价不变的影响,这与其更多依赖于输入变量和碳价的响应关系的合理性有关。

表 5 多项式回归算法的碳价预测情况  
Table 5 The predicted results of polynomial regression algorithm

日期	预测碳价 /元·t <sup>-1</sup>	实际碳价 /元·t <sup>-1</sup>	预测精确度/%	相对误差/%
2022-5-5	57.70	58	99.49	0.51
2022-5-6	56.95	58	98.19	1.81
2022-5-9	56.88	58	98.06	1.94
2022-5-10	56.85	58	98.02	1.98
2022-5-11	56.91	58	98.11	1.89
2022-5-12	56.91	58	98.13	1.87
2022-5-15	56.95	58	98.18	1.82
2022-5-16	56.98	58	98.24	1.76
2022-5-17	56.98	58	98.24	1.76
2022-5-18	56.94	58	98.17	1.83
2022-5-19	56.96	58	98.21	1.79
2022-5-20	56.97	58	98.23	1.77
2022-5-23	56.95	58.97	96.57	3.43
2022-5-24	57.88	59.18	97.81	2.19
2022-5-25	58.09	59	98.46	1.54
2022-5-26	57.95	59	98.22	1.78
2022-5-27	57.96	59	98.23	1.77
2022-5-30	57.97	59	98.26	1.74
2022-5-31	57.96	59	98.23	1.77
2022-6-1	57.95	59	98.22	1.78
2022-6-2	57.98	59	98.27	1.73
平均预测精确度			98.17%	
平均相对误差				1.83%
标准偏差				0.5132

2.3.3 LSTM 模型预测结果分析 与多项式回归算法相类似,LSTM 模型同样选定前一交易日碳价和 WTI 原油价格作为输入变量,构建二者与碳价的响应关系,完成碳价预测。LSTM 模型的碳价预测精度,见表 6。相较于前述两个模型,LSTM 模型的预测精度最高,逐日预测精度均超过了 98%,平均预测精度高达 99.26%。其中,2022 年 5 月 20 日的预测精度最高,为 99.69%;2022 年 5 月 23 日的预测精度最低,为 98.61%,与多项式回归模型的最差表现出现时间完全吻合。这也表明尽管二者的方法原理不同,但同样都是依赖于响应关系的合理性和输入变量的规律性变化。

表 6 LSTM 模型的碳价预测情况  
Table 6 The predicted results of LSTM model

日期	预测碳价 /元·t <sup>-1</sup>	实际碳价 /元·t <sup>-1</sup>	预测精确度/%	相对误差/%
2022-5-5	57.70	58	99.49	0.51
2022-5-6	58.67	58	98.84	1.16
2022-5-9	58.49	58	99.15	0.85
2022-5-10	58.50	58	99.14	0.86
2022-5-11	58.51	58	99.12	0.88
2022-5-12	58.48	58	99.17	0.83
2022-5-15	58.41	58	99.29	0.71
2022-5-16	58.31	58	99.46	0.54
2022-5-17	58.22	58	99.61	0.39
2022-5-18	58.21	58	99.63	0.37
2022-5-19	58.21	58	99.63	0.37
2022-5-20	58.18	58	99.69	0.31
2022-5-23	58.15	58.97	98.61	1.39
2022-5-24	58.39	59.18	98.67	1.33
2022-5-25	58.66	59	99.43	0.57
2022-5-26	58.76	59	99.60	0.40
2022-5-27	58.76	59	99.59	0.41
2022-5-30	58.69	59	99.48	0.52
2022-5-31	58.65	59	99.41	0.59
2022-6-1	58.65	59	99.40	0.60
2022-6-2	58.64	59	99.39	0.61
平均预测精确度			99.32%	
平均相对误差				0.68%
标准偏差				0.263 0

### 3 讨论

汇总了 ARIMA 模型、多项式回归算法和 LSTM 模型的全国碳价预测结果,见表 7。可知,3 种模型都展现了良好的预测性能,平均预测精度均超过了 97%。其中,LSTM 模型的预测精确度最高,为 99.32%,标准偏差最小,仅为 0.263 0;多项式回归算

法其次,预测精度为 98.17%,标准偏差为 0.513 2; ARIMA 模型相对较差,预测精度和标准偏差分别为 97.84% 和 0.518 4。这也表明在数据信息充足的前提下,深度学习算法具有更好的纠错能力,可以很好地捕捉到碳价的波动变化甚至静止不变的趋势。相反,当数据信息有限,特别是在影响因素的相关信息不够完备的条件下,ARIMA 模型将有更好的发挥空间。针对不同的样本信息条件,选择适宜的预测模型完成碳价预测,可为企业合理参与碳市场和政府制定相应措施平抑碳价波动指明方向,进而保障全国碳市场的健康、平稳发展。

表 7 模型预测精确度结果比较  
Table 7 The result comparison among three models

日期	ARIMA模型	多项式回归算法	LSTM模型	%
2022-5-5	98.72	99.49	99.49	
2022-5-6	98.17	98.19	98.84	
2022-5-9	98.53	98.06	99.15	
2022-5-10	97.94	98.02	99.14	
2022-5-11	97.37	98.11	99.12	
2022-5-12	96.72	98.13	99.17	
2022-5-15	97.77	98.18	99.29	
2022-5-16	96.91	98.24	99.46	
2022-5-17	96.59	98.24	99.61	
2022-5-18	97.03	98.17	99.63	
2022-5-19	97.31	98.21	99.63	
2022-5-20	97.27	98.23	99.69	
2022-5-23	99.25	96.57	98.61	
2022-5-24	99.06	97.81	98.67	
2022-5-25	99.11	98.46	99.43	
2022-5-26	98.53	98.22	99.60	
2022-5-27	97.97	98.23	99.59	
2022-5-30	97.33	98.26	99.48	
2022-5-31	98.36	98.23	99.41	
2022-6-1	97.52	98.22	99.40	
2022-6-2	97.20	98.27	99.39	
平均预测精确度	97.84%	98.17%	99.32%	
平均相对误差	2.16%	1.83%	0.68%	
标准偏差	0.5184	0.5132	0.2630	

为了更好地展示 3 类模型的预测功能,在运用 ARIMA 模型计算得到待预测日的前一交易日碳价和 WTI 原油价格的基础上,分别运用前述构建的 3 个预测模型完成全国碳市场 2022 年 11 月的碳交易日价格预测,见表 8。

表 8 全国碳市场的 11 月碳价预测结果  
Table 8 The predicted carbon price in November for the national carbon market

日期	预测碳价/元·t <sup>-1</sup>		
	ARIMA模型	多项式回归算法	LSTM模型
2022-11-1	62.65	62.54	60.79
2022-11-2	62.48	62.78	60.79
2022-11-3	62.72	62.62	60.78
2022-11-4	62.97	62.86	60.84
2022-11-7	63.34	63.10	60.95
2022-11-8	62.83	63.45	61.11
2022-11-9	63.21	62.97	61.12
2022-11-10	63.33	63.34	61.18
2022-11-11	63.13	63.46	61.23
2022-11-14	62.95	63.27	61.21
2022-11-15	62.99	63.11	61.15
2022-11-16	62.86	63.15	61.10
2022-11-17	63.12	63.02	61.03
2022-11-18	62.95	63.26	61.03
2022-11-21	63.19	63.10	61.02
2022-11-22	63.44	63.34	61.08
2022-11-23	63.81	63.58	61.19
2022-11-24	63.30	63.93	61.35
2022-11-25	63.68	63.45	61.36
2022-11-28	63.81	63.82	61.42
2022-11-29	63.60	63.94	61.47
2022-11-30	63.43	63.75	61.45
碳价平均值	63.17	63.27	61.12
标准偏差	0.3653	0.3929	0.2088

表 8 可知,3 个模型的 11 月碳价预测平均值分别为 63.17、63.27 和 61.12 元/t,相较于上半年的实际碳价平均值(58.28 元/t)有所提升。这是因为,年底临近履约窗口期,控排企业的碳配额和碳排放量已经基本确定,企业可以根据自身情况进行多余碳配额的卖出和不足碳配额的买入,在碳市场表现活跃,导致碳价提高。此外,对比 5 月的预测结果,3 个模型预测结果的标准偏差均有所下降,表明随着全国碳市场的日益规范化和管理机制的不断健全,碳价波动变小,碳价趋于稳定。

#### 4 结论

本文以全国碳市场的碳排放权交易价格为主要研究对象,在采用 PCC 法提取碳价关键影响因素的基础上,以应用广泛、理论成熟以及操作简便为选取依据,分别运用 ARIMA 时间序列、LSTM 和多项式回归算法 3 类方法,构建了预测精度较高的碳价预测模型,显示了 3 类模型的可靠性和实用性。

(下转第 129 页)



- area[J]. *advances in Atmospheric Sciences*, 2006, 23: 605 – 614.
- [9] FENG T, BEI N, HUANG R, et al. Summertime ozone formation in Xi'an and surrounding areas, China[J]. *Atmospheric Chemistry and Physics Discuss*, 2015, 15: 30563 – 30608.
- [10] 吕欢, 安俊琳, 苏筱倩, 等. 南京北郊地区黑碳气溶胶对光解系数的影响[J]. *中国环境科学*, 2020, 40(4): 1421 – 1428.
- [11] 杨豪. 气溶胶-辐射相互作用对中国东部臭氧浓度影响的数值模拟研究[D]. 南京: 南京信息工程大学, 2022.
- [12] JONSON J, KYLLING A, BERNTSEN T, et al. Chemical effects of UV fluctuations inferred from total ozone and tropospheric aerosol variations[J]. *Journal of Geophysical Research:Atmospheres*, 2000, 105(D11): 14561 – 14574.
- [13] REAL E and SARTELET K. Modeling of photolysis rates over Europe: impact on chemical gaseous species and aerosols[J]. *Atmospheric Chemistry and Physics*, 2011, 11(4): 1711 – 1727.
- [14] 赵帝, 卞思思, 王帅, 等. 沈阳市分季节环境空气 PM<sub>2.5</sub> 组分特征分析与来源解析[J]. *环境保护科学*, 2021, 47(2): 128 – 135.
- [15] 李陵. 北京市近年大气颗粒物和近地面臭氧的相互影响研究[D]. 北京: 中国环境科学研究院, 2020.
- [16] ZHANG B, WANG T, LIU J, et al. The surface aerosol optical properties in the urban area of Nanjing, West Yangtze River Delta, China[J]. *Atmospheric Chemistry and Physics*, 2017, 17(2): 1143 – 1160.
- [17] 张婉, 杨凌霄, 张雄飞, 等. 山东农村和背景地区雾霾天与清洁天气气溶胶光学特性[J]. *山东大学学报(工学版)*, 2020, 50(2): 1 – 8.
- [18] YAN W, YANG L, CHEN J, et al. Aerosol optical properties at urban and coastal sites in Shandong province, Northern China[J]. *Atmospheric Research*, 2017, 188: 39 – 47.
- [19] 许天峰, 马焉, 郑军, 等. 2019 年秋季南京北郊气溶胶光学特性研究[J]. *环境科学学报*, 2021, 41(11): 4353 – 4365.
- [20] 吴兑, 毛节泰, 邓雪娇, 等. 珠江三角洲黑碳气溶胶及其辐射特性的观测研究[J]. *中国科学:地球科学*, 2009, 11: 1542 – 1553.
- [21] 王贺, 曹念文, 王鹏, 等. 南京地区大气气溶胶综合观测与对比分析[J]. *遥感学报*, 2017, 21(1): 125 – 135.
- [22] 吴立新, 吕鑫, 秦凯, 等. 基于太阳光度计地基观测的徐州气溶胶光学特性变化分析[J]. *科学通报*, 2016, 61(20): 2287 – 2298.
- [23] 郑玉蓉, 王旭红, 张秀, 等. 基于 Landsat 数据的关中盆地腹地 AOD 时空格局及城市化对其影响[J]. *环境科学*, 2021, 42(6): 2699 – 2712.
- [24] 咎雅媛, 马晓燕, 田蓉. 基于 TUV 模式的对流层光解速率影响因子的研究[J]. *中国环境科学*, 2019, 39(9): 3707 – 3718.

### (上接第 62 页)

三种模型的对比结果显示, PCC-LSTM 模型的预测精度最高, 为 99.26%, 很好地显示了深度学习算法在数据充足条件下的优越性和实用性。相较于上半年的实际碳价, 预测得到的全国碳市场 11 月平均碳价有所提高, 变化幅度降低, 可以归结为全国碳交易市场运行总体平稳有序和年底活跃度持续攀升。在后续研究过程中, 采用数据预处理方法剔除样本数据异常值的干扰、结合与碳价负相关的影响因素到预测模型中、将碳价序列分解为多个子序列, 以及运用 GARCH(Generalized Autoregressive Conditional Heteroskedast)模型识别碳价的波动变化趋势, 有助于进一步提升碳价预测模型的优越性和实用性, 确保其发挥更大的作用。

### 参考文献

- [1] 崔连标, 范英, 朱磊, 等. 碳排放交易对实现我国“十二五”减排目标的成本节约效应研究[J]. *中国管理科学*, 2013, 21(1): 37 – 46.
- [2] 赵领娣, 王海霞. 碳交易价格预测研究-以深圳市为例[J]. *价格理论与实践*, 2019(2): 76 – 79.
- [3] 魏宇, 张佳豪, 陈晓丹. 基于 DMS 和 DMA 的我国碳排放权交易价格预测方法-来自湖北碳市场的经验证据[J]. *系统工程*, 2022, 40(4): 1 – 16.
- [4] 王娜. 基于 Boosting-ARMA 的碳价预测[J]. *统计与信息论坛*, 2017, 32(3): 28 – 34.
- [5] LI W, LU C. The research on setting a unified interval of carbon price benchmark in the national carbon trading market of China[J]. *Applied energy*, 2015, 155: 728 – 739.
- [6] ZHU B Z, HAN D, WANG P, et al. Forecasting carbon price using empirical mode decomposition and evolutionary least squares support vector regression[J]. *Applied energy*, 2017, 191: 521 – 530.
- [7] JIANG L B, WU P. International Carbon Market Price Forecasting Research Based on ARIMA-RF Model[C]. *International Conference on Strategic Management*, 2016, 1081 – 1084.
- [8] LI H J, HUANG X Y, ZHOU D H, et al. Forecasting Carbon Price in China: A Multimodel Comparison[J]. *International Journal of Environmental Research and Public Health*, 2022, 19(10): 6217.
- [9] 高长征, 李东伟, 王秀娜, 等. 利用智能机器学习方法对区域碳排放权交易价格预测研究-基于湖北碳市场数据的分析[J]. *价格理论与实践*, 2022, 4: 89 – 93.
- [10] ZHANG, F, WEN, N. Carbon price forecasting: a novel deep learning approach. *Environ Sci Pollut Res* 29, 54782–54795 (2022).
- [11] 欧阳迪, 崔佳. 碳中和战略下哈尔滨市空气污染物时空动态研究——基于 Holt-Winters 时间序列模型[J]. *环境保护科学*, 2023, 49(2): 112 – 119.
- [12] 杨丽, 吴雨茜, 王俊丽, 等. 循环神经网络研究综述[J]. *计算机应用*, 2018, 38(增 2): 1 – 6.
- [13] 赵海龙, 张丹丹, 黄松, 等. 基于皮尔逊相关系数的海南省地闪密度与雷击故障关系分析[J]. *高压电器*, 2019, 55(8): 186 – 192.
- [14] 孔宇, 王海起, 张浩然, 等. 基于集成学习算法的 PM<sub>2.5</sub> 浓度值预测[J]. *环境保护科学*, 2021, 47(4): 17 – 23.
- [15] 李文, 邓升, 段妍, 等. 时间序列预测与深度学习: 文献综述与应用实例[J]. *计算机应用与软件*, 2020, 37(10): 64 – 70.
- [16] 付凌晖, 王惠文. 多项式回归的建模方法比较研究[J]. *数理统计与管理*, 2004, 23(1): 48 – 52.
- [17] 杨丽, 吴雨茜, 王俊丽, 等. 循环神经网络研究综述[J]. *计算机应用*, 2018, 38(S2): 1 – 6.
- [18] 王秀. 基于 WD-ARIMA-LSTM 的短期风速预测[D]. 北京: 华北电力大学, 2021.