

# 基于 ARIMA-BP 神经网络的碳排放预测模型研究

刘云<sup>1</sup> 李琼<sup>2</sup>

1 合肥职业技术学院, 安徽合肥, 230001;

2 徽商职业学院, 安徽合肥, 230001;

**摘要:** 随着全球气候变化的加剧, 碳排放预测成为各国政府和科研机构关注的重点。本文旨在利用 python 语言实现一种基于 ARIMA-BP 神经网络的碳排放预测, 以提高碳排放预测的准确性和稳定性。通过结合 ARIMA 模型的线性预测能力和 BP 神经网络的非线性映射能力, 本文提出的组合应用能够更全面地捕捉碳排放时间序列数据的特征。实证结果表明, 该组合使用在碳排放预测方面表现优异, 具有较高的预测精度和稳定性。本文的研究为碳排放预测提供了一种新的思路和方法, 对于制定有效的减排政策、推动绿色低碳发展具有重要意义。

**关键词:** 碳排放预测; ARIMA 模型; BP 神经网络; 组合预测

**DOI:** 10.64216/3080-1486.25.12.040

## 引言

在全球气候变化的大背景下, 我国提出了“双碳”目标, 即力争 2030 年前实现碳达峰, 2060 年前实现碳中和。然而, 碳排放预测涉及众多因素, 包括能源结构、产业结构、经济发展水平等, 这些因素使得碳排放时间序列数据呈现出复杂的非线性特性。传统的预测方法往往难以准确捕捉这些非线性特征, 导致预测结果不准确。因此, 探索新的预测方法, 提高碳排放预测的准确性和稳定性, 具有重要的研究意义。

长江经济带对我国生态安全具有重要的战略支撑作用, 是支撑我国经济社会可持续发展的重要基础。长江中下游地区作为我国经济发展的重点区域, 其碳排放情况对全国双碳目标的实现具有重要影响。因此, 准确预测该地区的碳排放趋势, 对于制定针对性的减排措施具有重要意义。

## 1 长江中下游地区碳排放现状

长江经济带碳排放量在 15 年间整体呈持续上升的态势, 由 2005 年的 2365.31 百万吨提升至 2020 年的 4230.67 百万吨, 总体增长 78.86%。长江中下游地区碳排放的主要来源有能源消费(煤炭、石油、天然气)、工业生产(重工业、制造业)、交通运输(公路运输、水运)、建筑业和农业活动以及土地利用变化等多方面。2005-2010 年, 长江经济带城镇化与工业化快速推进, 碳排放增速高达 46.44%。2010-2015 年间, 区域能源经济结构开始逐步调整, 碳排放增速下降至 15.22%<sup>[1-2]</sup>。

据相关研究, 长江经济带下游城市的总体碳排放量远远高于其他两个区域, 占长江经济带所有城市碳排放的 52% 及以上。2020 年下游地区城市平均碳排放为 52.48 百万吨, 远高于上游地区的 23.19 百万吨<sup>[1]</sup>。而长江中游城市群的碳排放情况也不容忽视, 其碳排放效率与产业结构优化之间存在一定的时空耦合关系。

## 2 ARIMA-BP 神经网络模型构建

考虑到碳排放数据的时间序列特性和非线性数据特征, 本文选择构建 ARIMA-BP 神经网络组合进行预测。其中, ARIMA 模型侧重处理时间序列数据的平稳性和趋势性, BP 神经网络主要处理数据中的非线性复杂关系。

### 2.1 ARIMA 模型

ARIMA 模型被广泛运用于各类时间序列数据分析和建模。碳排放总量是典型的时间序列数据, 本文采用 ARIMA 模型进行初步预测。ARIMA 模型的本质是将非平稳的时间序列转化为平稳时间序列, 建立一个因变量对滞后值以及随机误差项当前值和滞后值进行回归的模型<sup>[4]</sup>。具体公式如下:

$$y_t = \mu + \sum_{i=1}^p r_i y_{t-i} + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (3.1)$$

其中:

$y_t$ : 当期值  $y_{t-i}$ : 前一期值  $\mu$ : 常数项

$p$ : 阶数  $r_i$ : 自相关系数  $\varepsilon_t$ : 误差项  $q$ : 当前误差与前几项误差的关系  $\theta_i$ : 系数

ARIMA 模型的基本思想是通过对历史数据的分析来确定时间序列的趋势和周期性, 从而实现对未来值的预

测。具体步骤如下：

(1) 平稳性检验：利用单位根检验判断时间序列数据是否平稳。如果数据不平稳，则进行差分处理直至平稳。差分阶数记为  $d$  (如  $d=1$  为一阶差分)。若原始数据平稳，则  $d=0$ 。

(2) 模型定阶：根据自相关函数和偏自相关函数的分析结果，确定 ARIMA 模型的阶数  $(p, d, q)$ 。

(3) 模型参数估计：利用最大似然估计或者最小二乘法等方法估计模型的参数，并验证参数的显著性。

(4) 模型诊断与调整：对模型进行诊断检验，确保模型的拟合程度和残差的随机性。如果残差未通过检验，调整参数，重复步骤 (2) – (4)。

(5) 模型预测：使用训练好的模型进行对未来时间点值的预测。

## 2.2 BP 神经网络

BP 神经网络是一种使用反向传播的多层前馈神经网络<sup>[5]</sup>，其主要特点是输出的计算是前向传播的，而误差是反向传播的，也就是会依据输出值误差反向调整权重。

BP 神经网络具体流程如下：

信号的前向传播：从输入层经过隐含层，最后到达输出层。训练样本从输入层输入，经过隐层处理后传向输出层，如果输出层的预测结果与期望值不同，则转入下个过程，否则输出。

误差的反向传播：从输出层到隐含层，最后到输入层，依次调节隐含层到输出层的权重和偏置，输入层到隐含层的权重和偏置。如果输出层的预测值没有达到期望值，或者说没有达到标准时，输出将以某种形式反向传递，在这个传递过程中将误差均摊到每一层神经元，从而调整各层的权重和偏置。

## 2.3 组合预测构建

为了克服单一模型的局限性，提高碳排放预测的准确性和稳定性，本文将 ARIMA 模型与 BP 神经网络相结合，构建组合预测。具体思路如下：

表 1 2005–2020 年长三角省份年总碳排放量

年份	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
碳排放量	1393.4	1534.31	1611.71	1711.42	1815.56	2011.88	2233.91	2261.16	2276.4	2233.47	2243.96	2266.86	2322	2275.19	2324.79	2287.67

### (2) BP 神经网络预测

把“年份”当输入、把“排放量”当输出，利用 BP 神经网络实现一次性映射，再预测 2025–2031 结果即可。

1、数据预处理：对原始碳排放数据进行清洗和整理，去除异常值和缺失值；然后进行标准化处理，消除量纲影响。

2、ARIMA 模型预测：首先利用 ARIMA 模型对数据进行预测，得到预测误差序列。主要是线性数据预测，获取到的误差序列输入到构建的 BP 神经网络中，以此来修正模型。

3、BP 神经网络构建：建立一个包含 1 个神经元和 3 个神经元的两层隐含层 BP 神经网络。将误差数据按时间顺序输入到神经网络中，输入节点设置为 4，输出节点设置为 1。通过滚动窗口的方式，将上一期的误差继续传入神经网络，作为输入的一部分，不断修正并预测模型。输入层有设计 4 个输入节点，代表当前期的 4 个误差数据。第一隐含层包括 1 个神经元，第二隐含层包括 3 个神经元，最后输出层设计 1 个输出节点，代表预测的误差。其中使用滚动窗口机制，每次输入 4 个误差数据到神经网络中，输出预测的误差，将预测的误差加入到输入数据中，形成新的 4 个输入数据(即滚动窗口)，继续输入到神经网络中进行下一次预测。将 ARIMA 模型的预测误差作为 BP 神经网络的输入，通过 BP 神经网络的非线性映射能力对误差进行修正。

4、组合预测：将 ARIMA 模型的预测结果与 BP 神经网络的修正结果相结合，得到最终的预测值。

## 3 模型预测与分析

### 3.1 碳排放预测

本文样本数据主要采用的是中国碳核算数据库 (CEADs) 提供的 2005–2020 年长三角省份碳排放数据。

#### (1) ARIMA 模型预测

首先根据 CEADs 提供的数据获取 2005–2020 年的每年总碳排放数量，利用 ADF 检验数据是否平稳，是否需要差分。然后结合表 1 的数据利用 python 代码计算获取差分阶数  $d$  的值，通过运算得到  $d=0$ 。

然后再进行自动网格输入，确定 ARIMA 阶数，从而得出 ARIMA 阶数值  $p=2, q=1$ 。

#### (3) ARIMA–BP 组合预测

首先根据读取表 1 的数据，利用 ADF 检验确定  $d$ ；然后利用 ARIMA 模型实现 2021–2035 的预测；接着利用 ARIMA 残差训练 BP 网络；最后利用把 2021–2031 的 ARI

MA 预测值 + BP 残差修正值, 得到最终 2025-2031 结果。核心代码实现如图 1 所示。最后获得如图 2 所示的预测结果。

### 3.2 结果分析

预测结果表明, 利用构建的 ARIMA-BP 神经网络组合对长江中下游地区 2025-2035 年的碳排放量进行预测。结果显示, 该地区碳排放量在未来十年内将呈现先上升后下降的趋势。随着经济发展和能源消费的增长, 碳排放量在短期内仍将上升, 但是在 2025 年到达峰值。但随着产业结构调整、能源转型以及减排政策的实施, 碳排放量有望在 2025 年后逐步下降, 最终实现碳达峰目标。

```
look_back = 2 # ② 构造监督学习样本: 用前 2 期线差预测当期残差
X, y_bp = [], []
for i in range(len(arima_resid) - look_back):
    X.append(arima_resid[i:i+look_back].flatten())
    y_bp.append(arima_resid[i+look_back, 0])
X, y_bp = np.array(X), np.array(y_bp)
from sklearn.neural_network import MLPRegressor # ③ 训练 BP 网络
bp = MLPRegressor(hidden_layer_sizes=(8, 4),
                   activation='tanh',
                   solver='lbfgs',
                   max_iter=1000,
                   random_state=42)
bp.fit(X, y_bp)
steps_2031 = 2031 - 2020 # ④ 生成 2021-2031 的 ARIMA 预测 11 步
arima_pred = model.forecast(steps=steps_2031)
arima_pred.index = range(2021, 2032)
last_resids = arima_resid[-look_back: ].flatten() # ⑤ 预测残差
future_resid = []
for k in range(steps_2031):
    feat = last_resids.reshape(1, -1)
    pred_res = bp.predict(feat)[0]
    future_resid.append(pred_res)
    last_resids = np.append(last_resids[1:], pred_res)
future_resid = np.array(future_resid)
final_pred = arima_pred + future_resid # ⑥ 组合结果
result_2025_2031 = final_pred.loc[2025:2031]
```

图 1 ARIMA-BP 组合预测核心代码

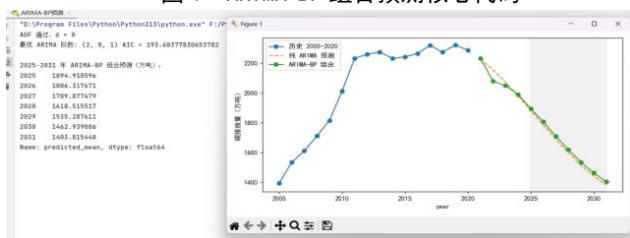


图 2 2025-2031 碳排放 ARIMA-BP 预测结果

### 4 结论

本文基于 ARIMA-BP 神经网络模型对长三角地区的碳排放进行了预测研究。通过结合 ARIMA 模型的线性预测能力和 BP 神经网络的非线性映射能力, 构建了组合

预测。实证结果表明, 该预测方法在碳排放预测方面表现优异, 具有较高的预测精度和稳定性。但是在 BP 神经网络训练因子较少, 未来可以考虑更多的影响因素, 进一步提高预测精准度和稳定性。本文的研究为碳排放预测提供了一种新的思路和方法, 对于制定有效的减排政策、推动绿色低碳发展具有重要意义。

### 参考文献

- [1] 冯新惠, 李艳, 等. “双碳”目标下城市形态对碳排放的影响: 以长江经济带为例[J]. 环境科学, 2024, 6(45): 3389-3401
- [2] 马明媚, 殷文琦, 吴亚齐, 等. 碳中和目标下黄河流域全要素碳生产率时空演化研究[J]. 人民黄河, 2024, 46(1): 12-18.
- [3] 钱志权, 韩佳银. 长三角城市群新型城镇化与低碳发展耦合过程与机理[J]. 长江流域资源与环境, 2023, 32(11): 2285-2297.
- [4] 肖丹. 基于 ARIMA 模型的四川省 GDP 分析与预测[J]. 生产力研究, 2023(10): 62-66
- [5] 王永利, 李颐雯, 王欢, 等. 基于改进 BP 神经网络的河北省碳排放预测[J]. 生态经济, 2024, 6(40): 30-37
- [6] 谭川江, 王超, 常昊, 等. 基于人工神经网络模型的碳排放预测研究进展[J]. 天然气与石油, 2024, 2(42): 124-132
- [7] 杨绍华, 张宇泉, 耿涌. 基于 LMDI 的长江经济带交通碳排放变化分析[J]. 中国环境科学, 2022, 42(10): 4817-4826.
- [8] 《安徽省统计年鉴》<https://tjj.ah.gov.cn/ssah/qwfbjd/tjnj/index.html>

作者简介: 刘云 (1988—), 女, 安徽省郎溪县人, 讲师, 硕士研究生, 主要研究方向为计算机应用, 大数据技术, 人工智能等。

李琼 (1983—), 女, 安徽省泾县人, 副教授, 硕士研究生, 主要研究方向为大数据技术, 人工智能等。

基金项目: 安徽省高校自然科学研究重点项目 (2023AH053108) 双碳目标下碳排放预测模型研究。