

基于 BP 神经网络的四川省碳排放预测



王长有^{*}, 温雅馨

成都信息工程大学应用数学学院, 四川成都 610225

摘要: 为应对全球气候变暖问题, 中国提出力争 2030 年前实现碳达峰。四川省是中国西部大省, 在中国西部地区具有重大示范作用。因此, 以四川省为例, 通过查询国家统计局网站和中国碳核算数据库, 收集四川省 1997 年至 2021 年的碳排放量及其影响因素的历史数据。分析四川省碳排放强度, 通过计算影响因素与碳排放量的皮尔逊相关系数, 最终选择 7 个重要影响因素并使用 EViews 软件预测其未来值。通过实验选择最佳节点数, 建立 BP 神经网络模型 (7-3-1 结构) 并预测 2022 年至 2035 年四川省碳排放量。结果表明该 BP 神经网络模型效果良好, 各样本相关系数值均高于 0.99, 均方根误差 RMSE 为 7.6078 (单位: 百万吨)。预测结果显示, 2022 年至 2035 年四川省碳排放量先增后降, 碳达峰时间为 2025 年, 峰值 281.2988 百万吨。

关键词: BP 神经网络; 四川省碳排放; 碳排放强度; 碳达峰; 相关系数; 预测

DOI: [10.57237/j.se.2024.01.001](https://doi.org/10.57237/j.se.2024.01.001)

Carbon Emission Prediction in Sichuan Province Based on BP Neural Network

Wang Changyou^{*}, Wen Yixin

College of Applied Mathematics, Chengdu University of Information Technology, Chengdu 610225, China

Abstract: To address the issue of global climate change, China proposed to strive to achieve the goal of peaking carbon emissions before 2030. Sichuan Province is a major province in western China and has a significant demonstration role in the western region. Therefore, taking Sichuan Province as an example, historical data on carbon emissions and their influencing factors from 1997 to 2021 were collected by querying the website of the National Bureau of Statistics and the CEADs. The paper analyzed the carbon emission intensity in Sichuan Province, selected 7 important influencing factors by calculating the Pearson correlation coefficient between the influencing factors and carbon emissions and predicted their future values by EViews. Then, selected the optimal number of nodes through experiments, established a BP neural network model (7-3-1 structure) and predicted the carbon emissions of Sichuan Province from 2022 to 2035. The results showed that the BP neural network model had good performance, with correlation coefficient values of all samples higher than 0.99 and root mean square error RMSE of 7.6078 million tons. The prediction results showed that from 2022 to 2035, the carbon emissions in Sichuan Province would first increase and then decrease, with a peak of 281.2988 million tons in 2025.

Keywords: BP Neural Network; Carbon Emissions in Sichuan Province; Carbon Emission Intensity; Carbon Peaking; Correlation Coefficient; Prediction

^{*}通信作者: 王长有, wangcy@cuit.edu.cn

1 引言

伴随着经济的飞速发展,人们面临的环境问题日益严峻,全球变暖俨然成为世界各国关注的重点问题。温室气体排放导致海平面上升、土地沙漠化、高温天气等,在所有温室气体中,二氧化碳(carbon dioxide, CO₂)排放量占比超过90% [1],因此,减少CO₂排放对减缓全球变暖有重要作用。中国作为世界上最大的发展中国家,在全球碳达峰、碳中和战略中扮演着十分重要的角色。2020年9月22日,中国政府在第七十五届联合国大会上承诺“二氧化碳排放力争于2030年前达到峰值”。四川省是中国西部大省,对中国西部地区的经济、社会、绿色发展均有重大影响。对四川省碳排放进行分析和预测,越来越具有重大的现实意义。

近年来,一些学者对四川省的工业、农业、交通运输业等领域进行了碳排放预测。如,方晓丽和骆勇对四川省公路水路货运能耗和碳排放趋势做出预测,为低碳交通的发展奠定理论基础[2];张丹等运用回归模型和灰色预测模型预估四川省工业碳排放量并提出合理建议[3];廖祖君等通过STIRPAT扩展模型定量分析影响四川省人均碳排放的因素,并在四种情景下预测四川省碳排放量的峰值及碳达峰时间[4];李乔楚等结合系统动力学(SD)的动态仿真优势与Visual Studio(VS)的智能分析优势,搭建SD-VS集成模型对四川省碳排放趋势进行预测[5, 6];卢奕亨等基于三大类碳源测算2005—2020年四川省农业碳排放量并构建面板Tobit模型探讨碳排放量影响因素[7]。

BP(back propagation)神经网络算法以其预测的精确度高被广大学者青睐。赵金元建立BP神经网络和多元线性回归模型对钢铁企业碳排放进行预测分析,BP神经网络模型具有很强的非映射能力,碳排放的变化趋势呈现非线性,结果表明BP神经网络模型对碳排放的预测优于多元线性回归模型[8]。赵成柏和毛春梅结合ARIMA模型和BP神经网络模型的优势对中国碳排放强度进行预测,取得良好效果[9]。胡振等利用BP神经网络模型对西安市家庭消费碳排放进行分析和预测[10]。

基于以上论述,本文采用BP神经网络模型,对四川省碳排放量进行建模分析和预测。考虑到数据的可得性,本文选取1997年至2021年的数据进行建模,预测区间为2022年至2035年。预测结果显示,四川省碳达峰时间为2025年,峰值281.2988百万吨。

2 方法与数据

2.1 BP 神经网络

1986年,以Rumelhart和McClelland为首的科学家提出BP(back propagation)神经网络的概念,BP神经网络是一种通过误差反向传播算法进行误差校正的多层前馈神经网络,是应用最广泛的神经网络模型之一,绝大多数人工神经网络模型都采用了BP算法或它的变形[11]。

BP神经网络具有强大的非线性映射能力、自适应能力和泛化能力,通常由一个输入层、一个输出层和一个或多个隐含层构成,每层由若干个节点组成,每个节点代表一个神经元。典型的BP神经网络是含一个隐藏层的三层结构网络。本文以四川省人均地区生产总值(元/人)、第二产业增加值占比、第三产业增加值占比、煤品燃料消费量(万吨)、油品燃料消费量(万吨)、天然气消费量(亿立方米)和电力消费量(亿千瓦小时)作为BP神经网络模型输入层,以四川省二氧化碳表观总排放量(百万吨)为输出层,建立3层BP神经网络结构对四川省碳排放量进行预测,如图1所示。

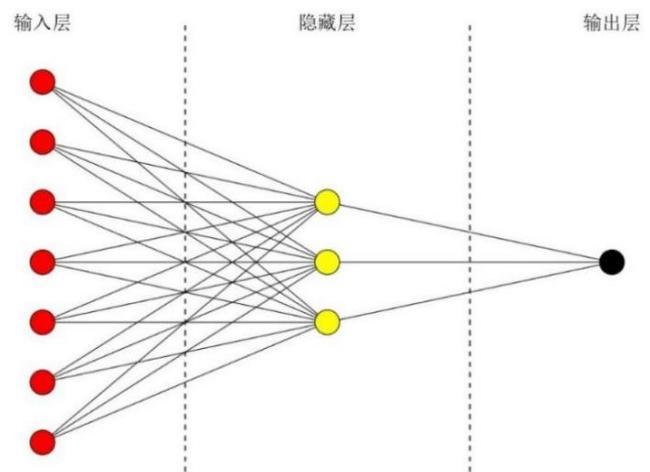


图1 BP 神经网络结构

2.2 软件

使用EViews 12软件,根据因变量的变动趋势建立线性回归模型或ARMA模型预测其2022年至2035年的变化趋势。使用Matlab R2023b训练BP神经网络模型并预测2022年至2035年的碳排放量。

2.3 数据来源

本文建模采用的数据区间为 1997 年至 2021 年。其中, 四川省年末常住人口、地区生产总值、各类能源消费量的原始数据是从国家统计局的网站查询所得; 四川省第二、三产业增加值占比分别由第二、三产业增加值和地区生产总值计算得到; 四川省表观二氧化碳排放量数据来源于中国碳核算数据库 (CEADS) [12-15]。

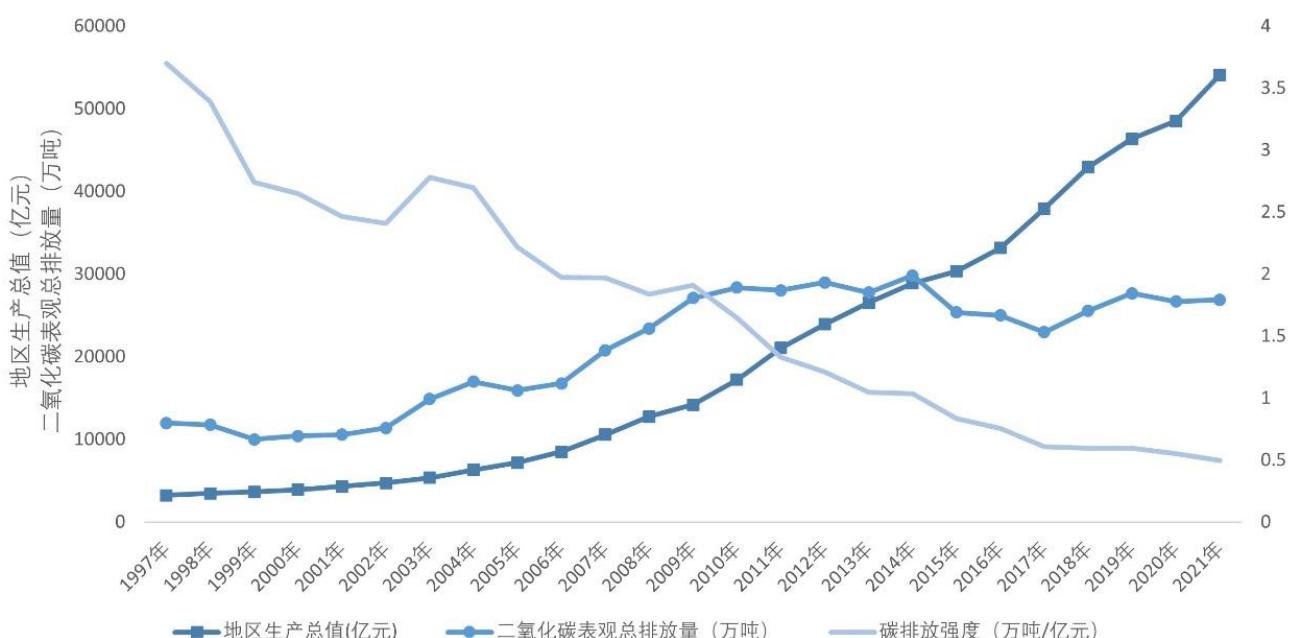


图 2 四川省 1997-2021 年地区生产总值、二氧化碳表观总排放量及碳排放强度趋势

由图可知, 四川省 GDP 呈持续增长的趋势, 2021 年四川省 GDP 是 1997 年的 16.69 倍。二氧化碳表观总排放量也呈现上升趋势, 但在 2014 年至 2017 年短暂下降, 随后又上升。其中, 2002 年至 2009 年增长迅速。碳排放强度整体呈下降态势, 由 1997 年的 3.697 万吨/亿元下降至 2021 年的 0.497 万吨/亿元。1997 年至 2009 年, GDP 和碳排放量均正向增长, 但经济发展速度高于碳排放量增速, 因此碳排放强度下降。2010 年及以后, 四川省经济仍然高质量发展, 然而, 受一系列绿色发展措施的影响, 碳排放量增长变缓甚至出现负增长, 使得碳排放强度快速下降。

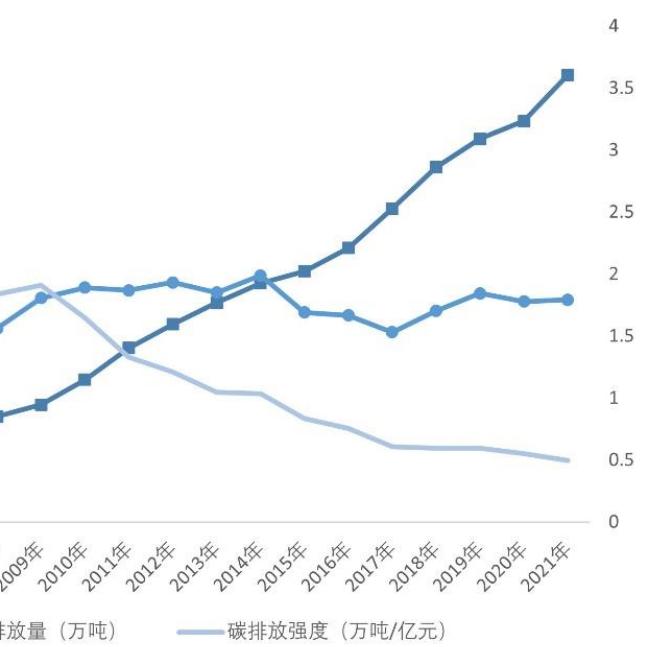
3.2 影响因素选择

除 GDP 以外, 还有许多其他因素对碳排放量产生影响, 如人口规模、产业结构[17]、煤炭能源和清洁能

3 结果与分析

3.1 四川省碳排放现状

有研究表明, 四川省 GDP 增速对碳排放规模影响显著[16], 因此, 了解近年来四川省的碳排放强度是进行碳排放预测的基础。其中, 碳排放强度=二氧化碳表观总排放量/地区生产总值。四川省 1997 年至 2021 年地区生产总值、二氧化碳表观总排放量及碳排放强度趋势如图 2 所示。



源消费量等。初步选取四川省年末常住人口 (万人)、人均地区生产总值(元/人)、第二产业增加值占比(%)、第三产业增加值占比 (%)、煤品燃料消费量 (万吨)、油品燃料消费量 (万吨)、天然气消费量 (亿立方米) 和电力消费量 (亿千瓦小时) 8 个影响因素。

使用皮尔逊相关系数 r 计算这 8 个影响因素与二氧化碳表观总排放量的相关关系, 并生成相关关系热力图, 如图 3 所示。其中, $r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$ 。

相关系数 r 的值介于 -1 与 1 之间, 可以根据相关系数的大小判断两个变量间的相关关系, r 的绝对值越接近 1, 表示相关程度越大; 反之, 相关程度越小。一般情况下, 我们认为如果相关系数的绝对值小于 0.2, 则两个变量极弱相关或不相关。

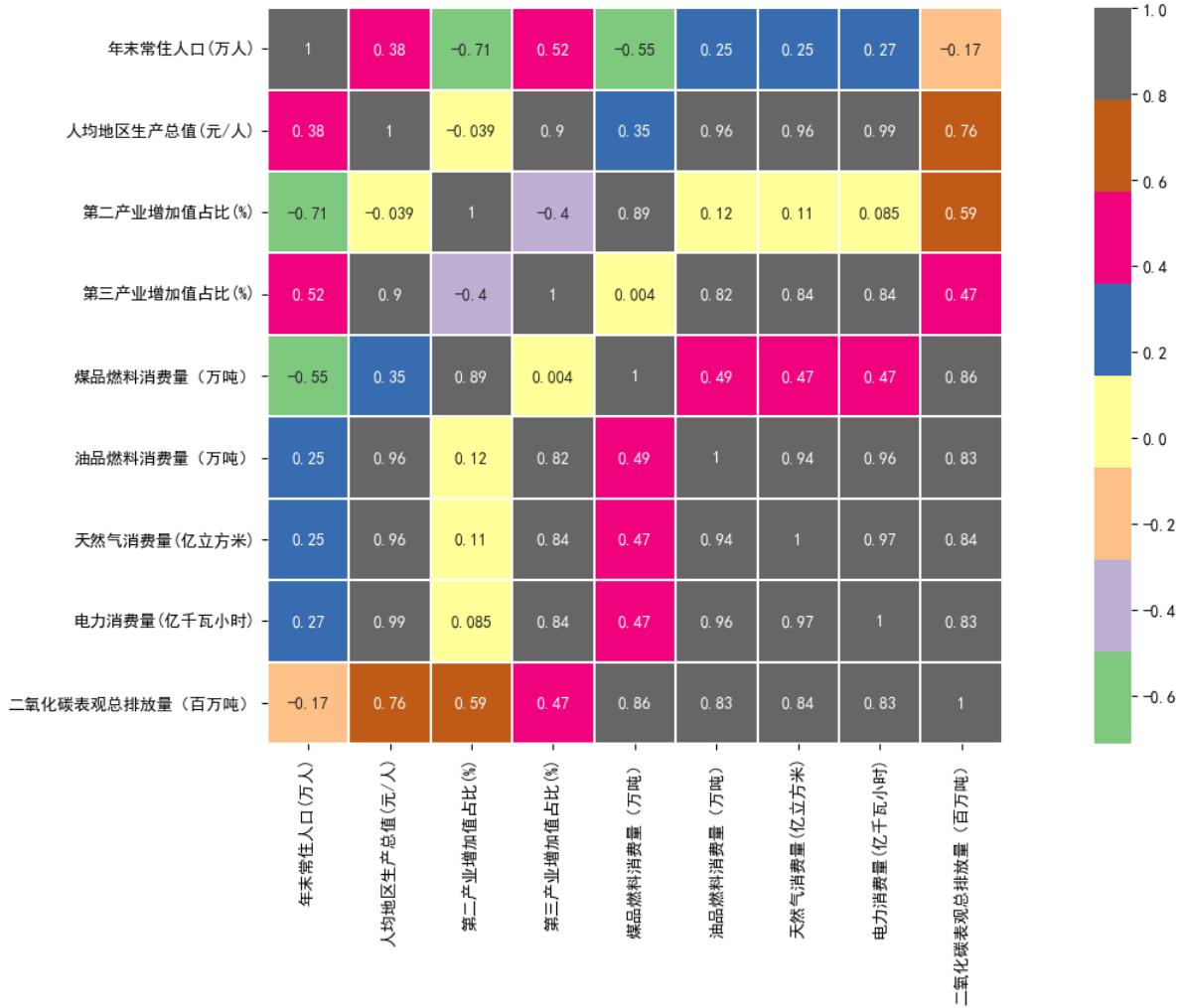


图 3 相关系数热力图

根据热力图显示, 年末常住人口与二氧化碳表观总排放量的相关系数仅为-0.17, 几乎不具有相关关系。其余指标均与二氧化碳表观总排放量有较强相关关系。因此, 选取除年末常住人口外的 7 个影响因素建立它们与碳排放量关系的模型。

3.3 建立 BP 神经网络模型

将上文选取的 7 个指标的历史数据作为输入数据, 二氧化碳表观总排放量的历史数据作为输出数据, 导入 Matlab 软件。样本数据区间为 1997 年至 2021 年, 共 25 年。使用 `randperm` 函数生成随机序列, 打乱数据集, 可以使模型精度更高。设置 80% 为训练集, 20% 为测试集, 选用 `newff` 函数构建 BP 神经网络结构。隐含层和输出层神经元传递函数选择 `tansig` 和 `purelin` 函数,

数, 选择结合了梯度下降法和高斯-牛顿法 (Levenberg-Marquardt) 的 `trainlm` 函数作为 BP 神经网络的训练函数[10]。设置训练次数为 1000 次, 学习速率为 0.01, 训练目标最小误差为 0.00001。

为寻找最优的神经网络结构, 对隐含层节点数量做逐点测试。结合经验方法, 将隐含层节点数区间设置为 [3, 12]。对每个不同的隐含层节点数分别训练 20 次, 选取训练 20 次得到的均方根误差 RMSE 的平均值为评价指标[18]。RMSE 值越小, 表明网络性能越好。

测试结果如表 1 所示, 当隐含层节点数为 3 时, 均方根误差 RMSE 的均值最小, 为 14.8954。因此, 设置隐含层节点数为 3, 本研究最优 BP 神经网络结构为“7-3-1”结构, 即输入层含 7 个节点, 隐含层含 3 个节点, 输出层含 1 个节点。

表 1 隐含层节点数逐点测试结果

均方根误差 RMSE (单位: 百万吨)	隐含层节点数									
	3	4	5	6	7	8	9	10	11	12
20 次实验最小值	3.6382	6.5224	7.4610	9.6332	5.7204	9.7896	6.7194	6.7368	5.1869	6.0388
20 次实验最大值	24.5868	38.6907	33.1715	26.9905	39.2815	47.0377	41.3318	30.8231	30.8806	38.0640
20 次实验均值	14.8954	18.8750	15.4698	17.3433	16.9762	20.3195	18.1494	15.5902	16.7742	21.6397

将仿真输出对真实值进行回归, 相关系数 R 越接近 1, 说明模型拟合效果越好。由图 4 可得, 训练样本、验证样本、测试样本和总体样本的仿真输出与期望输出的相关系数 R 分别为 0.99482、0.99961、0.9979、0.99531, 各样本相关系数值均高于 0.99, 表明模型拟合效果良好。

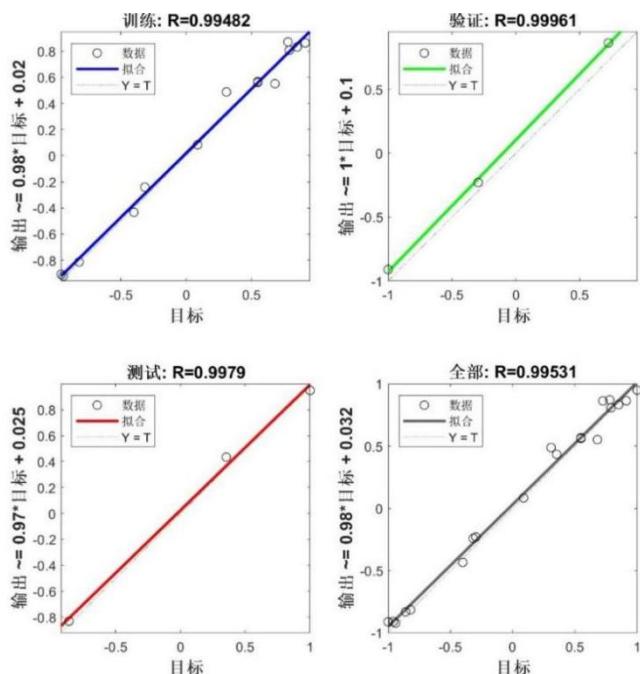


图 4 训练、验证、测试及全样本回归曲线

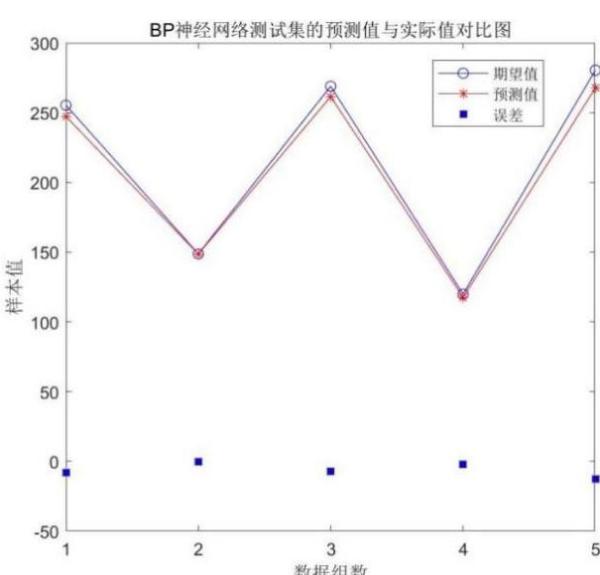


图 5 预测值与实际值对比

图 5 展示了该 BP 神经网络模型在测试集上的预测值和真实值的对比结果及误差, 由此图可直观感受到该模型性能良好, 能较好地预测未来四川省碳排放量。该模型的平均绝对误差 MAE 为 6.1384(单位: 百万吨), 均方误差 MSE 为 57.8784 (单位: 百万吨²), 均方根误差 RMSE 为 7.6078 (单位: 百万吨)。

3.4 预测

本研究旨在预测 2022 年至 2035 年四川省碳排放量, 得到四川省碳达峰时间和碳排放量的峰值。

首先需要预测 2022 年至 2035 年四川省人均地区生产总值、第二产业增加值占比、第三产业增加值占比、煤品燃料消费量、油品燃料消费量、天然气消费量和电力消费量的值。使用 EViews 软件, 根据原始数据特征, 分别对四川省人均地区生产总值、第二产业

增加值、第三产业增加值占比、煤品燃料消费量、电力消费量建立 ARMA 模型, 对四川省地区生产总值、油品燃料消费量、天然气消费量建立线性趋势模型, 预测其未来趋势。其中, 2022 年至 2035 年第二产业增加值占比由第二产业增加值除以地区生产总值得到。

预测结果如图 6 所示, 人均 GDP 持续增长; 第二产业增加值占地区生产总值比例先增后降, 第三产业增加值占地区生产总值比例持续下降, 但相对波动不大, 预示四川省产业结构的细微调整; 煤品燃料消费

量在 2022 年至 2028 年小范围波动, 随后逐渐趋于平稳; 其他燃料消费量呈上升趋势, 尤其电力消费量呈指数上升态势, 四川省清洁能源消费占比将越来越大。

将以上指标的预测值导入 Matlab 软件, 利用建立的 BP 神经网络模型预测 2022 年至 2035 年四川省二氧化碳表观总排放量, 预测结果如图 7 所示。由图可得, 2022 年以后, 四川省碳排放量将持续上升, 直至 2025 年达到峰值, 峰值为 281.2988 百万吨; 随后碳排放量不断下降, 预计 2035 年下降至 237.3688 百万吨。

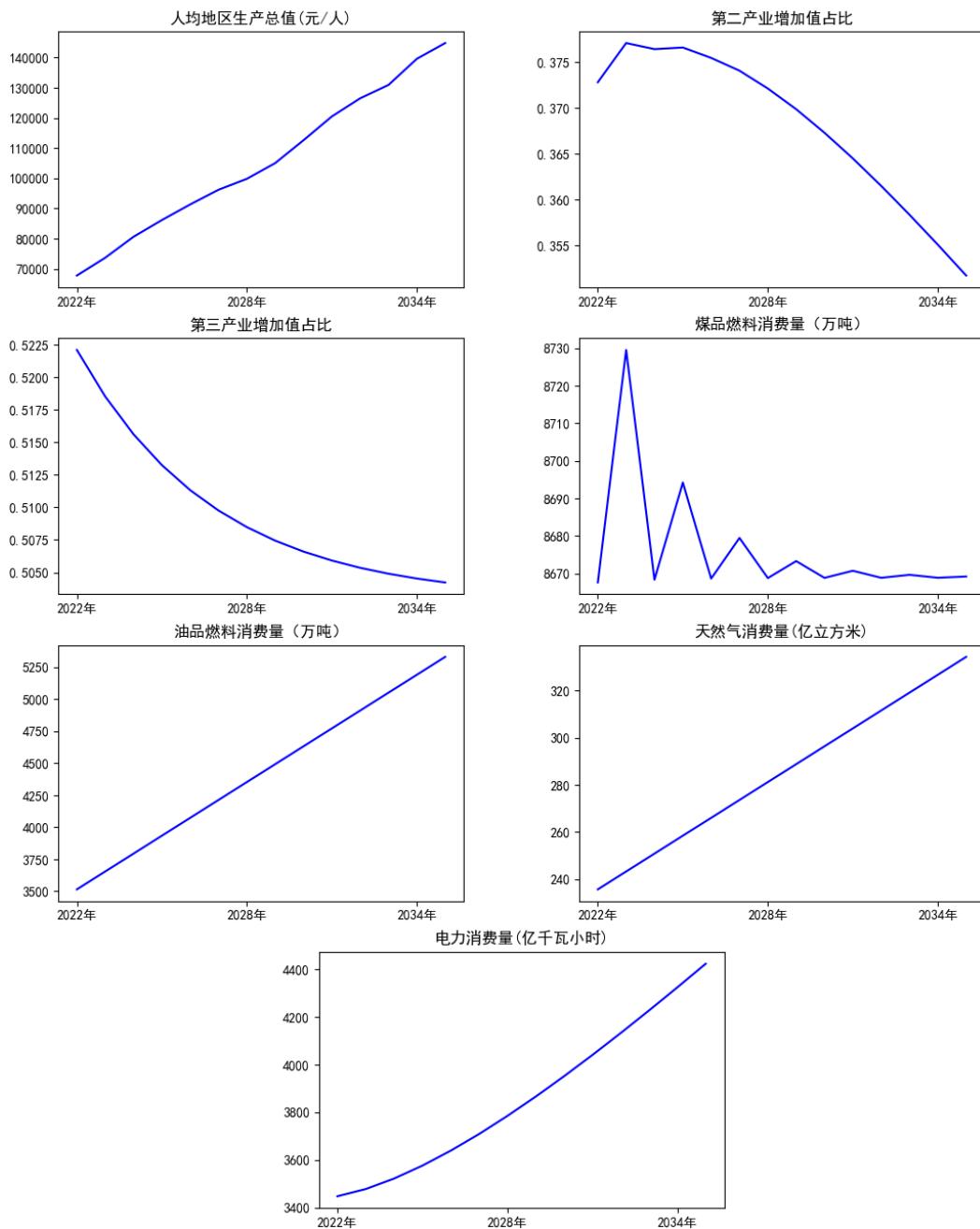


图 6 影响因素预测结果

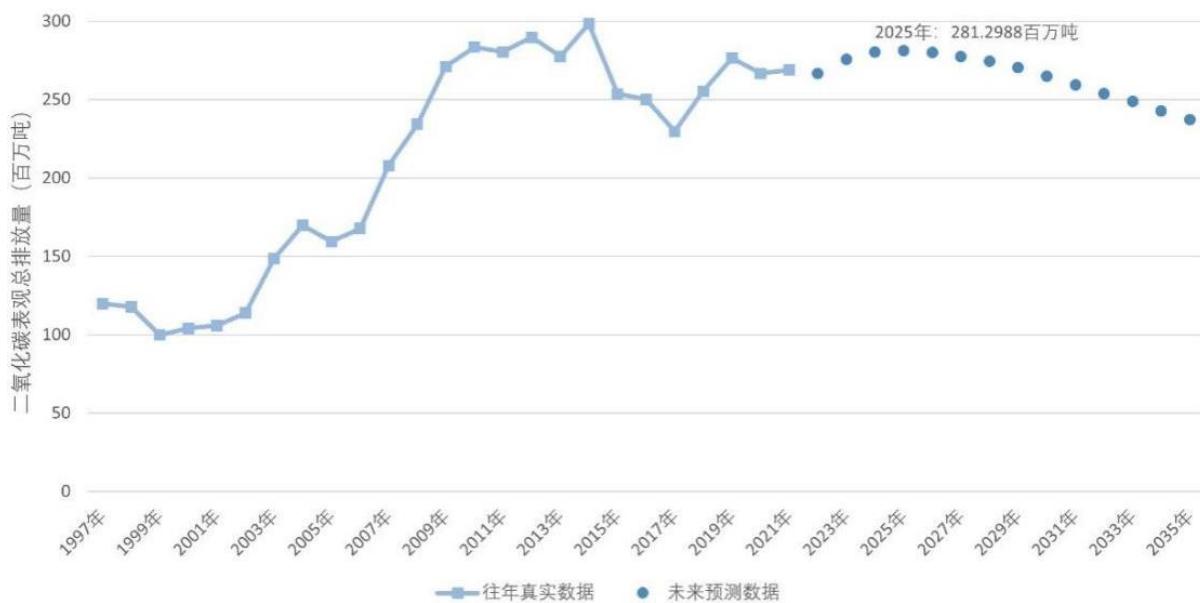


图 7 BP 神经网络模型预测结果

计算 2022 年至 2035 年四川省碳排放强度，如图 8 所示。随着四川省 GDP 一路高涨，碳排放量达峰后持续下降，碳排放强度也不断降低。预计 2035 年，碳排放强度将下降至 0.199 万吨/亿元。

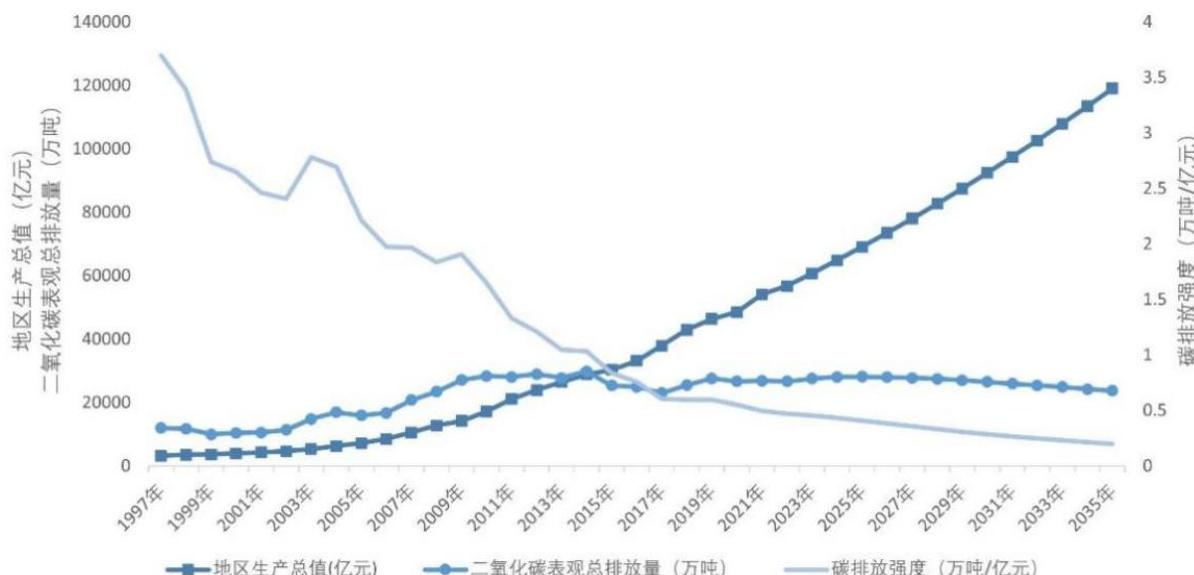


图 8 四川省 1997-2035 年地区生产总值、二氧化碳表观总排放量及碳排放强度趋势

4 讨论

本研究表明，四川省二氧化碳表观总排放量受四川省人均地区生产总值、产业结构、能源消费结构的影响。通过对往期数据建模并预测，可以预见 2022 年至 2035 年四川省经济水平不断上涨，产业结构有细微调整，天然气、电力等清洁能源将逐渐取代煤油燃料。

加强碳排放管理，持续优化产业结构，早日实现碳中和将会成为四川省面临的新的挑战。

BP 神经网络功能强大，除了能够用来预测碳排放趋势，还可以应用于许多其他领域，如经济学、生物学、化学等。除了回归预测，BP 神经网络还可以用来解决分类问题、控制工业过程、处理自然语言问题等。总而言之，BP 神经网络是一个通用模型，可以应用到各种各样的问题上，且大多效果良好。

由于碳排放量受很多因素影响, 因此对碳排放量的预测往往很难做到绝对准确, 预测结果还需要时间的检验。另外, BP 神经网络参数的设置主要靠经验和实验确定, 因此, 对 BP 神经网络的理论研究还有待进一步深入。

5 结论

本文通过皮尔逊相关系数法选择影响四川省碳排放量的重要因素, 建立 BP 神经网络模型并预测四川省碳达峰时间和峰值, 主要结论如下:

- (1) 四川省人均地区生产总值、第二产业增加值占比、第三产业增加值占比、煤品燃料消费量、油品燃料消费量、天然气消费量和电力消费量均与二氧化碳表观总排放量有较强相关关系。
- (2) 本研究建立的 BP 神经网络模型(7-3-1 结构)效果良好, 各样本相关系数值均高于 0.99, 均方根误差 RMSE 为 7.6078 (单位: 百万吨)。
- (3) 预测结果显示, 四川省碳排放量目前仍处于缓慢上升趋势, 预计将于 2025 年实现碳达峰目标, 峰值为 281.2988 百万吨。

参考文献

- [1] Li D Z, Huang G Y, Zhang G M, et al. Driving factors of total carbon emissions from the construction industry in Jiangsu Province, China [J]. Journal of Cleaner Production, 2020, 276: 123179.
- [2] 方晓丽, 骆勇. 四川省公路水路货运能耗及碳排放趋势预测方法研究 [J]. 交通节能与环保, 2019, 15(3): 23-26.
- [3] 张丹, 刘建文, 高一茹. 四川省工业碳排放量影响因素与预测 [J]. 重庆交通大学学报: 社会科学版, 2019, 19(4): 90-96.
- [4] 廖祖君, 张剑宇, 陈诗薇. 碳排放影响因素及达峰路径研究——基于四川省的分析 [J]. 软科学, 2023, 37(9): 95-101.
- [5] 李乔楚, 陈军华, 何京. 基于系统动力学-Visual Studio 集成模型的区域碳排放中长期变化趋势预测——以四川省为例 [J]. 环境污染与防治, 2022, 44(12): 1669-1675.
- [6] 李乔楚, 陈军华. 集成系统动力学和清单算法的能源结构调整对碳排放影响研究——以四川省为例 [J]. 安全与环境学报, 2023, 23(10): 100-106.
- [7] 卢奕亨, 田云, 周丽丽. 四川省农业碳排放时空演变特征及其影响因素研究 [J]. 中国农业资源与区划, 2023, 网络首发.
<https://doi.org/10.13637/j.issn.1009-6094.2023.0603>
- [8] 赵金元, 马振, 唐海亮. BP 神经网络和多元线性回归模型对碳排放预测的比较 [J]. 科技和产业, 2020, 20(11): 172-176.
- [9] 赵成柏, 毛春梅. 基于 ARIMA 和 BP 神经网络组合模型的我国碳排放强度预测 [J]. 长江流域资源与环境, 2012, 21(6): 665-671.
- [10] 胡振, 龚薛, 刘华. 基于 BP 模型的西部城市家庭消费碳排放预测研究——以西安市为例 [J]. 干旱区资源与环境, 2020, 34(7): 82-89.
- [11] Wang W X, Tang R C, Li C, et al. A BP neural network model optimized by Mind Evolutionary Algorithm for predicting the ocean wave heights [J]. Ocean Engineering, 2018, 162: 98-107.
- [12] Shan Y, Guan D, Zheng H, et al. China CO₂ emission accounts 1997-2015 [J]. Scientific Data, 2018, 5: 1-14.
- [13] Shan Y, Huang Q, Guan D, et al. China CO₂ emission accounts 2016-2017 [J]. Scientific Data, 2020, 7(1): 1-9.
- [14] Guan Y, Shan Y, Huang Q, et al. Assessment to China's Recent Emission Pattern Shifts [J]. Earth's Future, 2021, 9(11): 1-13.
- [15] Shan Y, Liu J, Liu Z, et al. New provincial CO₂ emission inventories in China based on apparent energy consumption data and updated emission factors [J]. Applied Energy, 2016, 184: 742-750.
- [16] 曾鸿, 耿源.“双碳”战略下四川省碳排放规模影响因素研究 [J]. 科技和产业, 2022, 22(11): 280-283.
- [17] 杨俊祺, 范晓军, 赵跃华等. 基于 PSO-BP 神经网络的山西省碳排放预测 [J]. 环境工程技术学报, 2023, 13(6): 2016-2014.
- [18] 燕振刚, 李薇, Yan T 等. BP 神经网络算法在河西绿洲玉米生产碳排放评估中的应用及算法有效性研究 [J]. 中国生态农业学报, 2018, 26(8): 1100-1106.