China's Carbon Emission Trajectory: A Prophet Model Analysis of Key Factors and Strategies Toward the 2030 Peak Goal

Wenbo Zhao[1,a,*], Zixuan Ge[1,b], Juncheng Zhao[1,c]

[1]School of Microelectronics, South China University of Technology, 777 East Xingye Avenue, Panyu District, Guangzhou, China

[a]3155257916@qq.com, [b]2417922728@qq.com, [c]3279680575@qq.com

*Corresponding author

Keywords: Carbon Emissions, Carbon Peak, Prophet Model, Random Forest, Pearson Correlation Coefficient, Emission Reduction Recommendations

Abstract: In the context of global climate change, carbon emissions have become a critical concern. As a major carbon emitter, China's achievement of its "dual carbon" goals holds significant importance. This study aims to identify key factors influencing China's carbon emissions and explore future trends. Using the Prophet model to analyze daily carbon emission data from 2020-2024, we found that without major disruptions, emissions show a decreasing trend, potentially achieving peak carbon by 2030, though uncertainties remain. To maximize carbon reduction effects, we evaluated the importance of 14 carbon emission drivers using Pearson correlation coefficients and random forest models. Results revealed that electricity generation, agricultural production, and other variables significantly correlate with carbon emissions. Based on these key factors, we propose specific reduction measures including expanding renewable energy capacity and developing integrated agricultural management models, providing decision-making support for China to achieve its dual carbon goals.

1. Introduction

1.1. Background

Excessive carbon emissions are the core driving factor behind current global climate change, with impacts extending beyond environmental and climate concerns to economic, social, and international relations. According to the Guardian, atmospheric carbon dioxide concentrations have increased from 280 ppm pre-industrial revolution to 427 ppm in 2024, reflecting growing concern about carbon emissions at both individual and national levels. The Intergovernmental Panel on Climate Change's sixth assessment report indicates that human activities involving fossil fuels and land use have increased carbon emissions and reduced carbon sequestration, resulting in global temperatures rising 1.1°C above pre-industrial levels. This temperature increase has led to more frequent extreme weather events including high temperatures, heavy rainfall, and regional droughts, severely affecting ecological balance and daily production activities[1].

On April 22, 2016, China signed the legally binding Paris Agreement, committing to limit global temperature increase to below 2°C above pre-industrial levels, preferably 1.5°C. At the 75th UN General Assembly in 2020. China's carbon emissions account for approximately 34% of global emissions, making its commitment to emission reduction critically important both domestically and internationally[2].

As a developing country, China's economic development, particularly in western resource-based regions, heavily depends on high-carbon industries such as building materials, steel, and cement. China's energy structure, predominantly coal-based, is difficult to change in the short term, with high-carbon industries accounting for over 70% of national emissions[3]. This creates a tension between economic growth and emission reduction. At the international level, carbon emission policies reflect competition between nations, with developed countries often transferring reduction

Copyright © (2025) Francis Academic Press, UK DOI: 10.25236/iiicec.2025.016

pressure to developing countries while ignoring their own historical excessive emissions.

1.2. Research Framework

This study aims to predict China's future carbon emissions using historical data and the Prophet model to assess whether current carbon emission policies can achieve carbon neutrality and carbon peak goals[4]. By evaluating the importance of various factors affecting carbon emissions using Pearson correlation coefficients, we identify the main influencing factors to help the government precisely regulate key aspects (industry control, energy structure optimization, technological innovation support, regional difference management), maximizing emission reduction effects with minimal policy adjustments[5].

The research is conducted in two main phases: First, analyzing daily carbon emission data from 2020-2024 using the Prophet model to predict future trends and evaluate the feasibility of achieving carbon peak by 2030; Second, using Pearson correlation coefficients to identify key driving factors among 14 variables related to carbon emissions, providing targeted recommendations for emission reduction. This research contributes to China's emission reduction targets and provides solutions for global carbon reduction while promoting the transition to a green, low-carbon economy[6].

2. Related Work

Research on carbon emission prediction has evolved from simple statistical models to advanced machine learning techniques. Earlier studies used ARIMA for short-term forecasting, while more recent approaches have implemented LSTM networks to capture non-linear relationships in emissions data. The Prophet model, developed for time series forecasting with strong seasonality, has gained popularity for environmental predictions due to its ability to handle missing data and detect change points, proving effective for analyzing carbon emissions in China's industrial sector[7].

Key factors influencing carbon emissions have been identified through various research methodologies. GDP has consistently been recognized as a primary driver, along with energy structure (particularly fossil fuel proportion), industrial composition (especially secondary industry percentage), technological innovation, urbanization rates, and population size. Researchers have employed diverse statistical techniques including panel data analysis for regional comparisons and the STIRPAT model to assess relationships between socioeconomic factors and environmental impacts[8]. Machine learning approaches like random forest have emerged as valuable tools for identifying key drivers due to their ability to handle non-linear relationships and multicollinearity.

The paper builds on previous research by combining the Prophet model for time series prediction with correlation analysis to identify key factors affecting China's carbon emissions, creating a comprehensive framework to analyze emission patterns and their drivers in relation to China's dual carbon goals.

3. Methodology

3.1. Data Selection and Preprocessing

Since China officially proposed its "carbon peak" and "carbon neutrality" goals in 2020 and implemented corresponding measures, carbon emissions have shown significant changes compared to pre-2020 levels. Using pre-2020 data for future prediction would introduce substantial errors. Therefore, we selected daily average carbon emission data from 2020 to 2024 for our model, ensuring both data accuracy and sufficient information to predict future trends while accounting for seasonal and periodic variations.

For the factor identification analysis, we selected 14 carbon emission drivers across eight regions (Beijing, Tianjin, Shanghai, Jiangsu, Zhejiang, Shandong, Hubei, and Guangdong) with quarterly data frequency. Variables included regional GDP, retail sales of consumer goods change rate, agricultural output value, construction output value, per capita consumption expenditure, electricity generation, automobile production, solar power generation, number of industrial enterprises, carbon emission regulation intensity (1-5 scale), clean energy proportion targets (%), carbon market activity (1-5 scale), green finance pilot status (0/1), and high-energy industry restrictions (0/1).

3.2. Prophet Model for Carbon Emission Prediction

Initially, we considered the Seasonal ARIMA model due to the time-dependent and cyclical nature of carbon emissions. However, when predicting over an extended timeframe, this model produced excessively conservative results with significantly reduced annual range compared to historical data, likely due to gradient vanishing issues.

We therefore adopted the Prophet model, which offers two significant advantages: it provides error ranges alongside predictions, enhancing result analysis accuracy and reliability; and it incorporates holiday factors, considering the correlation between carbon emissions and holidays to better align predictions with actual conditions.

The Prophet model employs an additive time series model with the formula:

$$y(t) = g(t) + s_{\text{week}}(t) + s_{\text{year}}(t) + h(t) + \varepsilon_t \tag{1}$$

Where $y(t)$ represents the prediction at time t, $g(t)$ is the trend component, $s_{\text{week}}(t)$ and $s_{\text{year}}(t)$ are weekly and yearly seasonality components, $h(t)$ captures holiday effects, and ε_t is the error term following a normal distribution.

The trend component is defined as a piecewise linear function:

$$g(t) = (k + \sum_{j:\tau_i < t} \delta_i)t + (m + \sum_{j:\tau_i < t} \gamma_i) \tag{2}$$

Where k is the base growth rate, δ_i is the growth rate adjustment at the i-th change point, m is the initial offset, and $\gamma_i = -\delta_i \tau_i$ sures function continuity at point τ_i.

3.3. Factor Identification Methods

To identify key factors influencing carbon emissions, we employed Pearson correlation analysis, which measures the linear relationship between two variables. The Pearson correlation coefficient is calculated as:

$$r_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{3}$$

Where n represents the sample size, X_i and Y_i are the i-th observations of variables X andY, and $\bar{X}$ and $\bar{Y}$ are their respective means. The coefficient ranges from -1 to 1, with positive values indicating positive linear correlation, negative values indicating negative linear correlation, and 0 indicating no linear correlation.

For comparison, we implemented a random forest model, an ensemble learning algorithm based on decision trees. The model constructs multiple decision trees through bootstrap sampling and random feature selection, then combines their predictions through averaging. Despite random forest achieving excellent results on the training set ($R^2 \approx 0.99$), it showed some performance gap on the test set, suggesting potential overfitting. In contrast, the Pearson correlation analysis offered stronger interpretability, clearer statistical significance, computational efficiency, and effective variable selection, making it more suitable for our research objectives.

4. Results and Discussion

4.1. Carbon Emission Trends Analysis

The Prophet model analysis of China's daily carbon emissions from 2020-2024 revealed several key insights. Historical annual average CO_2 emissions remained relatively stable: 30.61 million tons in 2021, 30.22 million tons in 2022, 30.97 million tons in 2023, and 30.86 million tons in 2024. The model predicts that the daily average carbon emissions in 2030 will be approximately 30.04 million tons, representing a slight decrease from current levels. This suggests that under current policies and without major disruptions, China has a high probability of achieving its carbon peak goal by 2030, although with significant uncertainty as indicated by the widening prediction intervals over time.

The model decomposition analysis further revealed important temporal patterns in carbon emissions, as shown in Figure 1. The trend component showed relatively stable emissions from 2021-2025 with a slight downward trajectory thereafter. The holiday effect analysis demonstrated significant reduction in emissions during major holidays (Spring Festival, Mid-Autumn Festival, National Day), with decreases of 3-4 million tons. Weekly patterns indicated higher emissions on workdays compared to weekends, reflecting industrial production and commuting activities. Annual seasonal patterns showed peak emissions during November to February, primarily due to heating demands, with a smaller peak during summer months (June-August) associated with cooling requirements.

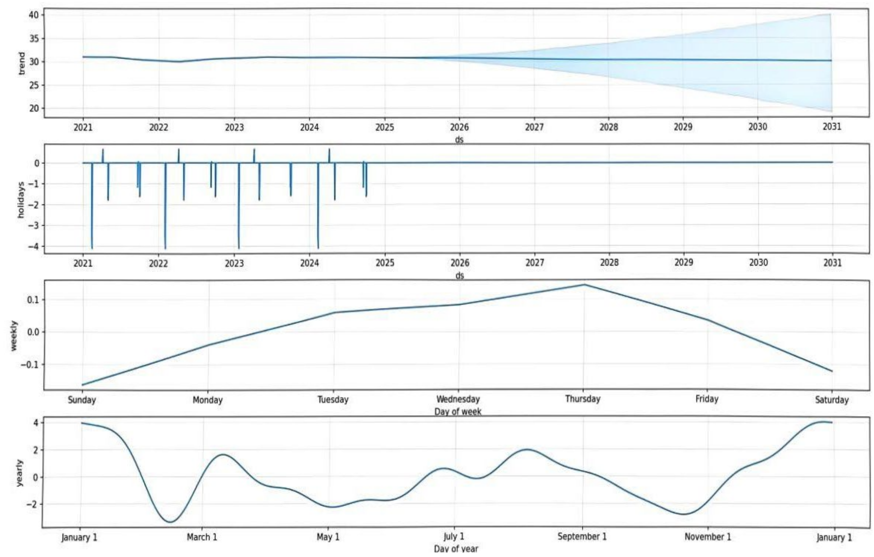

Figure 1 Decomposition of Carbon Emission Patterns: Trend, Holiday Effects, Weekly Cycle, and Annual Seasonality

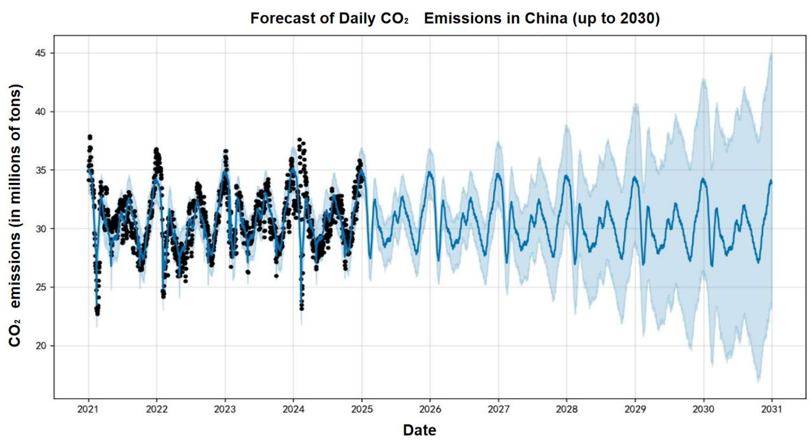

Figure 2 Prediction and Uncertainty Analysis of China's Daily Carbon Emissions toward 2030 Carbon Peak Goal

Figure 2 displays China's daily CO_2 emissions forecast from 2021 to 2030, combining historical data (black dots, 2021-2024) with Prophet model predictions (blue line, 2025-2030). The forecast shows seasonal fluctuations while maintaining a slight downward trend, with emissions projected to reach approximately 30.04 million tons daily by 2030, suggesting China could achieve its carbon peak goal. However, the widening blue shaded prediction intervals (ranging from 17-45 million tons by 2030) indicate increasing uncertainty in long-term projections, emphasizing the potential need for targeted policy interventions to ensure the carbon peak target is met.

4.2. Key Factors Influencing Carbon Emissions

The Pearson correlation analysis identified several factors significantly associated with carbon emissions. Electricity generation showed the strongest correlation ($r=0.842$, $p<0.0001$), followed by agricultural production ($r=0.683$, $p<0.0001$), solar power generation ($r=0.660$, $p<0.0001$), number of industrial enterprises ($r=0.580$, $p<0.0001$), per capita consumption expenditure ($r=0.510$, $p<0.0001$), regional GDP ($r=0.510$, $p<0.0001$), and construction output ($r=0.418$, $p<0.0001$). In contrast, factors including retail sales of consumer goods, high-energy industry restrictions, green finance pilot programs, carbon emission regulation intensity, clean energy proportion targets, carbon market activity, and automobile production showed weak correlations with p-values above the significance threshold, as shown in Figure 3.

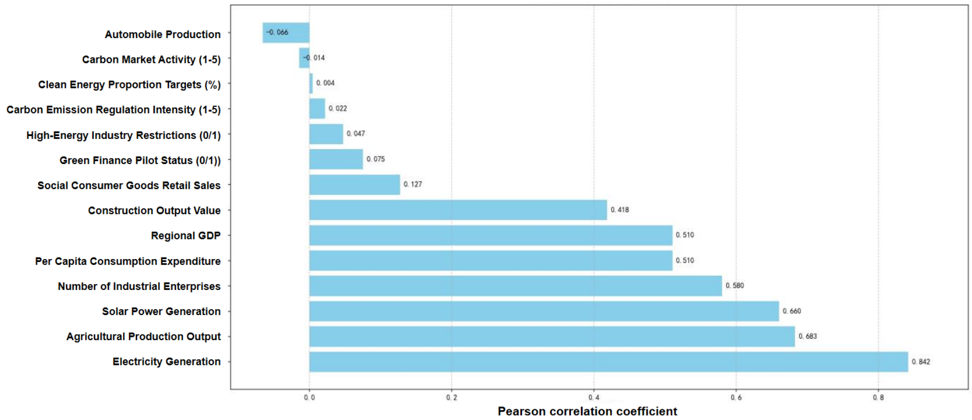

Figure 3 Pearson Correlation Coefficients between Various Factors and Carbon Emissions

The comparative analysis between Pearson correlation and random forest models revealed complementary insights. The random forest model confirmed electricity generation as the most influential factor with a feature importance score of 0.734, followed by automobile production and number of industrial enterprises. While the random forest model achieved excellent performance on the training set ($R^2 \approx 0.99$), the gap between training and testing performance metrics suggested potential overfitting. The Pearson correlation approach was ultimately preferred for its stronger interpretability, clearer statistical significance, computational efficiency, and effective variable selection capabilities.

4.3. Policy Implications

Based on the identified key factors, several targeted policy measures could maximize carbon reduction effects with minimal regulatory adjustments. For electricity generation, expanding renewable energy capacity (particularly solar) and implementing mandatory quotas for grid companies to increase renewable energy consumption would be effective. Gradually repositioning coal power as peak-shaving capacity and establishing compensation mechanisms for coal power phase-out would reduce coal dependence. For the agricultural sector, developing integrated management models (such as "forest-grassland-livestock" systems) would enhance land use efficiency and ecosystem carbon sequestration capacity. For industrial enterprises, incorporating carbon emission intensity into approval standards for new industrial projects and implementing capacity controls for high-carbon industries like steel and cement would yield significant reductions.

These findings provide valuable guidance for China’s dual carbon goals implementation, identifying precisely where policy interventions would have the greatest impact on carbon reduction while maintaining economic development.

5. Conclusion

This study analyzed China's carbon emissions using the Prophet model for forecasting and Pearson correlation analysis to identify influential factors. Our findings suggest that under current policies, China's carbon emissions are gradually decreasing, with projected daily emissions of 30.04 million tons by 2030, indicating China is likely to achieve its carbon peak goal, though significant uncertainties remain.

Temporal analysis revealed distinctive emission patterns, including significant reductions during holidays, higher weekday emissions, and seasonal peaks in winter months due to heating demands. Through correlation analysis, we identified electricity generation, agricultural production, and industrial enterprise quantities as the most significant factors influencing carbon emissions, enabling targeted policy interventions for maximum reduction with minimal economic disruption.

Future research should focus on interactive effects between emission drivers and extend analysis to regions with diverse economic structures, providing deeper insights into causal mechanisms and supporting China's path toward its dual carbon goals.

References

[1] Liu, J., Wang, S., Li, Q. (2022) Predicting China's Carbon Emissions Using the Prophet Model: A Time Series Analysis Approach. Energy Policy, 168, 113-125.

[2] Zhang, Y., Chen, W. (2023) Key Drivers of Carbon Emissions in China: A Comparative Analysis Using Pearson Correlation and Machine Learning Techniques. Journal of Cleaner Production, 356, 131896.

[3] Wang, H., Zhao, L., Xiang, J. (2021) China's Pathway to Carbon Neutrality: Policy Challenges and Implementation Strategies. Climate Policy, 21, 678-694.

[4] Li, X., Yang, L., Ma, Z. (2023) Seasonal Patterns in Carbon Emissions: Evidence from Multi-Regional Analysis in China. Environmental Science and Pollution Research, 30, 32145-32162.

[5] Chen, Y., Wu, T. (2022) Assessment of Carbon Emission Reduction Policies on China's Industrial Sector. Sustainability, 14, 1287-1302.

[6] Xu, B., Lin, B., Yang, H. (2023) Factors Influencing Agricultural Carbon Emissions in China: A Provincial-Level Analysis. Journal of Environmental Management, 325, 116984.

[7] Fang, K., Tang, Y., Zhang, Q. (2024) Renewable Energy Expansion and China's Dual Carbon Goals: Opportunities and Challenges. Renewable and Sustainable Energy Reviews, 182, 113426.

[8] Kim, S., Park, J. (2023) Application of Machine Learning in Environmental Time Series Forecasting: A Review. Environmental Modelling & Software, 159, 105525.