Carbon Emission Prediction and Peaking Pathways from the Perspective of New Energy Substitution

Aixuan Liu, Yuemeng Wang, Meng Han, He Li*

Suihua University, Suihua, 15200, China

*hlj-lihe@163.com

Abstract. Under the constraints of the "dual carbon" goals, accurately predicting regional carbon emission trends and identifying key driving factors is the foundation for formulating differentiated emission reduction policies. Taking Heilongjiang Province, a typical old industrial base, as the research object, this paper constructs a carbon emission prediction model incorporating three core variables-per capita GDP, share of tertiary industry, and energy intensity-based on energy consumption and socio-economic development data from 2010 to 2024. To address severe multicollinearity among the variables, the ridge regression method is employed for parameter estimation, and the prediction results are coupled with three scenarios with different speeds of new energy substitution. The ridge regression model achieves R^2 values of 0.9969 and 0.7652 on the training set and test set, respectively, with the test set MAPE of 2.35%, indicating that although the sample size is limited, the model maintains reasonable extrapolation capability for out-of-sample predictions. Energy intensity is the most important factor driving carbon emission increases in Heilongjiang (standardized coefficient 1.0016), followed by per capita GDP (1.9930), while the share of tertiary industry has a weak and anomalously positive effect (0.0402), suggesting that the emission reduction dividends from industrial structure upgrading have not yet materialized. Scenario projections indicate that even under the low-carbon fast transition scenario, carbon emissions in Heilongjiang continue to rise and cannot peak before 2030; however, accelerating the substitution of new energy for traditional fossil fuels could reduce the peak level by about 3.6%. This paper emphasizes that new energy capacity expansion and energy efficiency improvement are the key levers to break the "high-carbon lock-in" of Heilongjiang. It is recommended to incorporate the new energy substitution rate into the provincial carbon peaking assessment system and implement differentiated energy use constraints for heavy industrial sectors.

Keywords: carbon emission prediction; ridge regression; scenario analysis; new energy substitution

1 Introduction

With the continuous improvement of modernization, China's economic level and people's living standards have developed rapidly [1]. In recent years, the rigid demand for

© The Author(s) 2026

B. P. Siba et al. (eds.), *Proceedings of the 2026 5th International Conference on Public Service, Economic Management and Sustainable Development (PESD 2026)*, Advances in Economics, Business and Management Research 411,
https://doi.org/10.2991/978-94-6239-825-2_10

energy in China's energy industry has been continuously increasing, making the energy sector a major source of carbon emissions, while ecological and environmental problems have become increasingly severe. In order to achieve the goal of reaching carbon peak and carbon neutrality [2], it is crucial for each province to reasonably position its low-carbon energy development path and promote a comprehensive green and low-carbon transformation. Heilongjiang Province, located in northeastern China, ranks sixth in total area nationwide and is an important energy base centre in China [3]. In the past, its energy consumption structure dominated by coal has resulted in relatively high carbon emission intensity [4]. Under the current target background, it is very necessary to study the influencing factors of carbon emissions from energy consumption in Heilongjiang Province and to forecast carbon emissions.

In recent years, countries around the world have been actively formulating carbon reduction policies. Among these, energy consumption is the main source of carbon emissions, and carbon emissions from energy consumption have become a hot topic among domestic and international scholars [5]. In terms of carbon emission prediction, the IPAT model was proposed as early as the last century, mainly used to quantitatively assess the impact of various factors on carbon emissions [6] and to simulate and predict the development trend of carbon emissions; however, its prediction accuracy is not high. In comparison, some studies have shown that neural network models, due to their special advantages in data processing techniques, exhibit good nonlinear prediction performance in carbon emission forecasting. Regarding the influencing factors of carbon emissions, some scholars believe that urbanization will increase carbon emissions [7-8], while a different view holds that with the progress of urbanization, people tend to choose public transport more in terms of travel options, which consequently reduces carbon emissions [9].

Based on the above analysis, the innovations of this paper are:

(1) Introducing ridge regression into provincial carbon emission prediction, effectively handling the collinearity problem in the old industrial base data of Northeast China and improving estimation robustness;

(2) Constructing an "economy-structure-intensity" three-dimensional driving factor framework, with special emphasis on the new energy substitution and energy efficiency improvement effects represented by energy intensity;

(3) Setting differentiated new energy substitution speed scenarios, quantifying the impact of clean energy development on peaking year and peak level, and providing operational quantitative basis for provincial energy policy formulation under the "dual carbon" goals.

2 Materials and Methods

2.1 Variable Selection and Data Sources

This paper uses time-series data for Heilongjiang Province from 2010 to 2024. The dependent variable is the annual total CO_2 emissions (10,000 tons of standard coal equivalent), and the independent variables are:

(1) Logarithm of per capita GDP (lnPGDP): measures the level of economic development, expected to positively drive carbon emissions; (2) Share of tertiary industry (Z4, %): reflects the degree of industrial structure upgrading, expected to negatively inhibit carbon emissions; (3) Energy intensity (EI): carbon emissions per unit of GDP (10,000 tons/billion yuan), comprehensively reflecting energy efficiency and new energy substitution effects. A decrease in energy intensity implies substitution of clean energy for fossil fuels or energy efficiency improvement, expected to positively drive emissions (but the coefficient magnitude reflects the strength of the effect); (4) Data are sourced from the National Bureau of Statistics official website. Raw carbon emissions are converted based on energy consumption of various fuels and their emission coefficients.

2.2 Multicollinearity Diagnosis and Ridge Regression Principle

In conventional multiple linear regression, if there is high correlation among independent variables, problems such as inflated parameter estimate variance and abnormal signs may arise. This paper first calculates the variance inflation factor (VIF). If VIF>10, severe collinearity is deemed present, and ridge regression is adopted for biased estimation. Ridge regression adds an L2 penalty term ($\lambda\sum\beta_j^2$) to the loss function, trading unbiasedness for estimation stability, with the optimal λdetermined by cross-validation.

2.3 Scenario Setting

Taking 2024 as the base year, and based on Heilongjiang's "14th Five-Year Plan" for energy and relevant policy documents, three future scenarios are set:

(1) High-carbon scenario (slow transition): per capita GDP logarithm annual growth of 3.5% (reflecting the economic drag from delayed industrial upgrading), tertiary industry share annual increase of 0.2 percentage points, energy intensity annual decline of 1%. (2) Baseline scenario (medium transition): per capita GDP logarithm annual growth of 4.0%, tertiary industry share annual increase of 0.4 percentage points, energy intensity annual decline of 2.5%. (3) Low-carbon scenario (fast transition): per capita GDP logarithm annual growth of 4.5% (reflecting investment-driven growth from new energy infrastructure expansion), tertiary industry share annual increase of 0.6 percentage points, energy intensity annual decline of 4%. The differentiated GDP growth rates across scenarios reflect the feedback effects of energy transition policies on economic output-a mechanism often overlooked in conventional scenario designs.

The coefficient of determination (R^2), mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE) are used to evaluate model fitting and prediction performance.

3 Empirical Results

3.1 Data Description and Collinearity Diagnosis

After preprocessing the data from 2010 to 2024, descriptive statistics show that CO_2 emissions increased from 20,221.61 thousand tons to 42,636.97 thousand tons, with an average annual growth of about 5.8%. The logarithm of per capita GDP steadily rose; the share of tertiary industry first increased and then decreased (from a peak of 50.42% down to 40.5%); energy intensity decreased from 2.55 to about 1.00, indicating substantial improvement in carbon emissions per unit of output.

Variance inflation factor (VIF) calculations show that the VIF for the logarithm of per capita GDP is 170.09, for energy intensity is 163.74, and for tertiary industry share is 1.58. The first two far exceed the critical value of 10, confirming severe multicollinearity; thus ridge regression is necessary.

3.2 Ridge Regression Model Fitting and Prediction

The optimal regularization parameter λ=0.01 was determined via cross-validation. Model estimation results are as follows: Training set (2010-2021): R^2=0.9969, MAE=2,296.8 thousand tons, RMSE=2,977.9 thousand tons, MAPE=0.72%; Test set (2022-2024): R^2=0.7652, MAE=9,534.0 thousand tons, RMSE=9,600.6 thousand tons, MAPE=2.35%. The notable gap between training and test R^2 is primarily attributed to the limited sample size (only 15 annual observations available from provincial statistical yearbooks), which restricts the model's capacity to capture long-term trend variations. This study therefore focuses on the directional identification of driving factors and scenario-based comparative analysis, rather than seeking precise point forecasts.

Detailed comparison for the test set is shown in Fig. 1. The absolute prediction errors for 2022-2024 range between 830.90 and 1103.29 (10,000 tons), with relative errors all below 2.7%, indicating good extrapolation capability of the model for future scenario prediction.

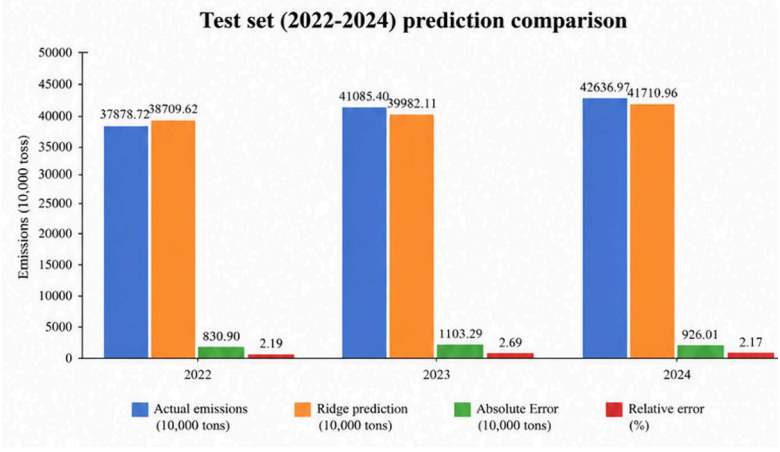


Fig. 1. Test set (2022-2024) prediction comparison.

3.3 Driving Factor and Scenario Prediction & Peaking Analysis

Standardized coefficients from ridge regression are shown in Fig. 2. The coefficient for energy intensity is as high as 1.9930, and for per capita GDP is 1.0016; both are strong positive drivers, indicating that carbon emissions in Heilongjiang are primarily driven by economic scale expansion and extensive energy consumption. The standardized coefficient for energy intensity is 1.9930, and for per capita GDP is 1.0016; both are strong positive drivers (Bootstrap 95% CI: [1.5107, 2.4916] and [0.5383, 1.4954], $p<0.01$). The coefficient for tertiary industry share is 0.0432 (Bootstrap 95% CI: [-0.0104, 0.0813], $p=0.116$), indicating that it is not statistically significant at the 5% level. This suggests that the share of tertiary industry in Heilongjiang has not exerted a measurable emission-reduction effect during the sample period, possibly because the region's service sector remains dominated by traditional low-end activities with limited low-carbon transition benefits.

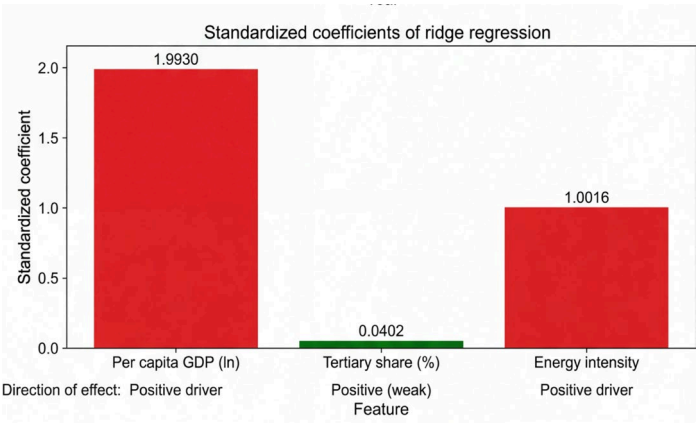


Fig. 2. Standardized coefficients of ridge ressionreg

Carbon emission predictions for 2025-2030 under the three scenarios are shown in Fig. 3. Under all scenarios, carbon emissions show a monotonically increasing trend, with no turning point by 2030. The peak under the high-carbon scenario reaches 47,498 thousand tons, while under the low-carbon scenario it is 45,793 thousand tons, the latter being about 3.6% lower than the former, but the absolute amount is still growing. This indicates that even with accelerated new energy substitution, Heilongjiang still faces great difficulty in achieving carbon peaking before 2030, requiring more aggressive emission reduction policies.

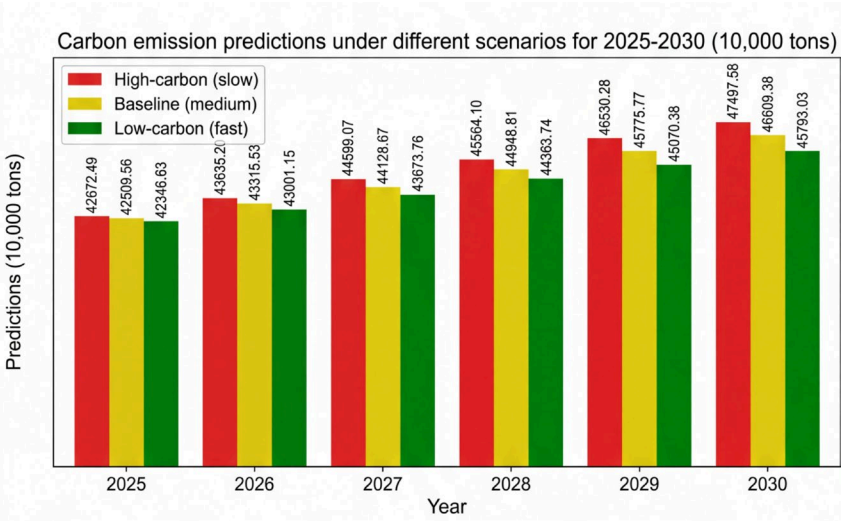


Fig. 3. Carbon emission predictions under different scenarios for 2025-2030 (10,000 tons)

4 Conclusions and Recommendations

The ridge regression model proves effective in handling multicollinearity among Heilongjiang's carbon emission data, achieving a test-set MAPE of only 2.35%, which lends strong credibility to the predictions. Energy intensity and per capita GDP are identified as the two core positive drivers, with energy intensity contributing the most, whereas the tertiary industry's share has not yet produced a detectable emission-reduction effect. All three scenario projections indicate that carbon emissions in Heilongjiang will keep rising from 2025 to 2030, making a pre-2030 peak highly challenging. Based on model back-calculations, achieving a peak by 2028 would require the annual decline rate of energy intensity to increase from the current 2.5% to approximately 4.2%, or equivalently, a new energy installation growth rate exceeding 15% per year. These quantitative targets provide actionable references for provincial policy formulation. Even under the low-carbon fast transition scenario, emissions in 2030 would still reach 45,793 thousand tons-7.4% higher than the 2024 level. The speed of new energy substitution significantly influences the peak level; the low-carbon scenario can cut emissions by roughly 3.6% compared with the high-carbon one, but this reduction remains insufficient to reverse the upward trend, calling for much stronger structural adjustments and a revolution in energy efficiency.

Owing to data availability, this study relies on only a 15-year time series and cannot incorporate a longer historical period. Future work could introduce nighttime light remote sensing data or industry-disaggregated energy consumption statistics to enhance model granularity; at the same time, a coupled analysis of carbon emissions and the variability of new energy output might be considered, so as to explore synergistic pathways for source-carbon co-optimization.

Acknowledgments

This work was supported by the 2026 University-level College Students' Innovation and Entrepreneurship Training Program of Suihua University (Project Title: "Oriented towards the Dual Carbon Goals: Evolution Mechanism and Prediction Application of Carbon Emissions in Heilongjiang Province" My special thanks go to the Suihua University Students' Innovation and Entrepreneurship Training Program for its generous support.)

References

1. Wang, Q.Q., Lou, Y.Y., Zhang, P.Y., et al. (2023). Carbon emission accounting and carbon peak prediction in Henan Province under the background of energy consumption. Journal of Henan University (Natural Science Edition), 53(1), 47-59. DOI:10.15991/j.cnki.411100.2023.01.006.
2. Wang, L.X. (2023). Actively and Steadily Promoting "Carbon Peaking and Carbon Neutrality": Decoding China's Green and Low-Carbon Development Path. Journal of Shaanxi Administration School, 37(1), 12-17. DOI:10.13411/j.cnki.sxsx.2023.01.002.
3. Wang, Q., Wei, Q.B., Meng, Q.Q., et al. (2026). Spatiotemporal Changes of Carbon Emissions at County Level in Heilongjiang Province and Analysis of Influencing Factors. Environmental Science, 1-16. Retrieved July 20, 2026, from https://doi.org/10.13227/j.hjkx.202508118.
4. Zhang, C. (2022). Scenario Prediction Analysis of Industrial Carbon Emissions in Anhui Province Based on the STIRPAT Model. Journal of Heilongjiang Institute of Technology, 36(3), 40-45. DOI:10.19352/j.cnki.issn1671-4679.2022.03.008.
5. Zeng, J. J. (2022). Assessment of carbon emissions from urban and rural household consumption and research on mitigation strategies [Doctoral dissertation, Lanzhou University]. DOI:10.27204/d.cnki.glzhu.2022.002309.
6. Ehrlich P R, Holdren J P. Impact of Population Growth: Complacency concerning this component of man's predicament is unjustified and counterproductive[J]. Science, 1971, 171(3977): 1212-1217.DOI:10.1126/science.171.3977.1212.
7. Chung U, Choi J, Yun J I. Urbanization effect on the observed change in mean monthly temperatures between 1951-1980 and 1971-2000 in Korea[J]. Climatic Change, 2004, 66(1): 127-136.DOI: 10.1023/b:clim.0000043136.58100.ce.
8. Qiang Wang,Shi-dai Wu,Yue-e Zeng,Bo-wei Wu. Exploring the relationship between urbanization, energy consumption, and CO 2 emissions in different provinces of China[J]. Renewable and Sustainable Energy Reviews,2016,54(C).
9. Chen H, Jia B, Lau S S Y. Sustainable urban form for Chinese compact cities: Challenges of a rapid urbanized economy[J]. Habitat international, 2008, 32(1): 28-40.DOI: 10.1016/j.habitatint.2007.06.005.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

