Proceedings of CONF-APMM 2025 Symposium: Simulation and Theory of Differential-Integral Equation in Applied Physics
DOI: 10.54254/2753-8818/2025.DL27564

Carbon Emission Prediction: Methods and Applications Across Fields

Xuantong Chen

Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China
xtchen886@connect.hkust-gz.edu.cn

Abstract. The environmental problems hold a close connection to humans' daily life, among which the carbon emissions serve an important role. To promote sustainable development, contributions measures need to be conducted to deal with the environmental problems effectively. Currently, abundant studies have discussed the methods to predict carbon emissions so as to help the relevant regulations legislated reasonably and practically. This document focused on the currrent 15 studies regarding carbon emissions prediction methods. The methods examined include traditional statistical models (e.g., grey models, ARIMA), machine learning techniques (e.g., Random Forest, XGBoost), deep learning approaches (e.g., LSTM, CNN), hybrid models (e.g., ARIMA-LSTM, VMD-SSA-LSSVM), and innovative domain-specific methods (e.g., G-code analysis for 3D printing). The results indicate a rather mature studying progress in certain fields, as the individual models, hybrid models and innovative models analyzed all showed great performance and progress to some extent in respective fields compared to specific previous studies.Generally, in spite of the fact that it lacks of a systematic prediction method that is suitable for all diverse industries commonly, various fields are riding on a rising way of improving the methods effectiveness to support the regulation works.

Keywords: Carbon emissions, Prediction method, Machine learning, Deep learning, Hybrid model

1. Introduction

As the environment problems getting more and more serious nowadays,humans are now aware of the necessity of environmental protection to a huge extent. Carbon emissions refer to the amount of greenhouse gas emissions generated during the production, transportation, use, and recycling of a product, which is a crucial constitution of greenhouse gas emissions. Dealing with the environmental problems can significantly help improve people's living statement and life experience, ameliorate city appearance, promote the economic growth and foster a virtuous cycle. In this case, the suppression of carbon emissions appears to be a crutial mission for every geographical areas and industries. By calculating the total amount of carbon emissions, researchers can take a further step to make a progress to decide the overall structure of carbon emissions. Hence, the prediction of carbon emissions can be a helpful tool for the decisions of organizations. Nowadays, countries around the world are making contributions to encourage low carbon lifestyle and

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2025.DL27564

regulations are legislated to restrict everyone's behaviours. This systematic review analyzes fifteen landmark studies aiming to provide the researchers an overall aspect about the general current development and potential direction for future research in the fiels of carbon emissions prediction. This document categorizes the 15 research into five distinct approaches by the main innovative methodologies employed in the studies: Traditional Statistics and Time Series Models, which provide foundational analytical frameworks; Machine Learning Methods, which offer powerful pattern recognition capabilities; Deep Learning Techniques, which excel in handling complex temporal and spatial relationships; Combination and Hybrid Models, which integrate multiple approaches to enhance predictive performance; and Innovative Domain-Specific Methods, which address particular industrial applications through tailored solutions. The application and innovative integration of these five categories of prediction methods play a crucial role in advancing carbon emission forecasting capabilities. This integration enables more comprehensive analysis that captures both linear trends and complex nonlinear relationships within emission data. The integration between different methodologies allows for the development of robust hybrid systems that outperform individual models in accuracy, reliability, and adaptability across diverse scenarios. By searching and analyzing the relative 15 studies, this review ultimately accelerated progress toward effective carbon management solutions.

2. Prediction methods

A huge amount of research regarding the topic of carbon emissions is conducted by many scholars. This document classify these premier 15 methods introduced in the 15 dissertations into mainly 5 categories: Traditional statistics and time series models, Machine learning methods, Deep learning methods, Combination and Hybrid Models and Other innovative methods.

2.1. Traditional statistics and time series models

Traditional statistical and time series models form the foundation of carbon emission prediction. These methods rely on historical data patterns and statistical relationships. They assume that data trends can be captured through mathematical formulations linearly. The main advantage of these approaches is their interpretability and relatively low computational requirements.

The improved multivariate grey model (IPSO_OGM(1,N)) marks a significant advancement in grey system theory [1]. This model addresses three major limitations of traditional grey models [1]. It uses particle swarm optimization for parameter tuning, generalized inverse matrix solutions for handling multicollinearity, and Markov chains for residual correction [1]. The study focused on Hunan Province's carbon emissions, analyzing factors including economic development level, industrial structure, and energy intensity [1]. Compared to traditional grey models, support vector machines, and autoregressive moving average models, this improved multivariate grey model achieved superior accuracy [1]. The research predicted that under low-carbon development scenarios, Hunan could achieve peak carbon emissions by 2027 with 332 million tons [1]. This study is the only case that focused on Hunan province in China [1].

The K-means cluster based Logistic model approach offers another traditional methodology [2]. This study used energy consumption data from all Chinese provinces, applying K-means clustering to group provinces with similar energy characteristics before applying Logistic regression [2]. The method processed data including coke, coal oil, gasoline, diesel, fuel oil, crude oil, natural gas, and power consumption [2]. This two-stage approach took the difference of energy consumption pattern

DOI: 10.54254/2753-8818/2025.DL27564

between China's diverse provinces into account, allowing for more accurate and flexible predictions [2].

The Python-based grey prediction model study demonstrates the application of traditional methods with modern coding tools [3]. This research applied grey system theory to predict carbon emissions at the city level, using data from 2017 to 2021 [3]. In the aspect of short-term forecasting, the model showed satisfactory performance [3]. Moreover, the research focus on a particular period of time and a scope of a city, hence this method is especially valuable for certain cities with limited historical data [3]. The study highlighted Python's effectiveness in implementing grey models and generating accurate calculation results for emission prediction tasks [3].

These traditional methods share several common characteristics. They all work well with limited data points, require relatively simple computational resources, and provide easily interpretable results. However, they differ in their specific approaches: the improved grey model focuses on algorithmic enhancements, the Logistic model incorporates clustering for regional differentiation, while the Python implementation emphasizes computational accessibility and reproducibility.

2.2. Machine learning methods

Generally, many scenarios of real-life cases can not be explained effectively by linear models. Machine learning methods learn from data, and with the strong base of the collected environmental data, these approaches can construct a model to deal with diverse complex situations. Their characteristic of learning nonlinear relationships through training data makes machine learning methods suitable for high-dimensional or complex data.

The machine learning combination model for regional agricultural carbon emissions represents a comprehensive approach [4]. This study developed a triple framework combining feature engineering, optimization algorithms, and machine learning [4]. It employed weight value algorithms for feature selection, dimensionality reduction techniques, and multiple machine learning regression algorithms [4]. The research covered three agricultural scenarios: regional cultivation, animal husbandry, and planting-grazing-mixed farming systems [4]. For crop cultivation, factors included fertilizer use, pesticides, agricultural film, diesel fuel, tillage, and irrigation. For animal husbandry, it considered emissions from intestinal fermentation and manure management of various livestock. The model achieved remarkable accuracy rates: 99.65% in Heilongjiang and 98.87% in Hunan for crop emissions, 98.06% in Inner Mongolia and 99.37% in Xinjiang for livestock emissions, and 99.00% in Liaoning and 99.19% in Sichuan for mixed systems [4]. The utilization of weight value algorithm indicates the innovation point of this research, with the transfer of regression methods into agricultural carbon emissions prediction field [4]. In addition, the triple form combination prediction model of feature engineering optimization algorithm machine learning highlighted the study's unique advancement [4].

The differential evolution optimized hybrid XGBoost (DEOH-XGBoost) study presents an advanced machine learning approach [5]. This method employed differential evolution to optimize XGBoost parameters automatically [5]. And the three main components of the innovative model in this study are feature engineering, model construction, and model integration [5]. The research collected data from seven global regions including Africa, Asia-Pacific, CIS, Europe, Middle East, North America, and South America [5]. Thirty-four factors were considered, such as natural gas consumption, coal consumption, and renewable energy generation [5]. The optimization process enhanced the model's ability to handle feature interactions and avoid overfitting [5]. The results demonstrated superior accuracy compared to other advanced techniques across all regions, showing

DOI: 10.54254/2753-8818/2025.DL27564

particular strength in capturing complex, nonlinear relationships between multiple predictors and carbon emissions.

The improved neural network methods study (PSO-BP) focused on optimizing neural network architectures using particle swarm optimization [6]. This approach addressed the challenge of parameter selection in neural networks by automatically determining optimal network configurations [6]. The particle swarm optimization component efficiently searched the parameter space without requiring gradient information, making it suitable for complex emission prediction tasks where traditional optimization methods might struggle. Focusing on a certain position, this new method presented a particular effect exclusively [6].

2.3. Deep learning methods

Deep learning methods have emerged as powerful tools for carbon emission prediction, particularly effective in handling temporal patterns and complex spatial relationships. These approaches use multiple processing layers to learn representations of data with multiple levels of abstraction, making them suitable for capturing intricate patterns in emission data.

The LSTM-GBRT model for provincial-level emissions prediction combines two powerful deep learning architectures [7]. This research utilized electricity data as a substitution standard for carbon emissions, analyzing the correlation between power load and emissions [7]. The study employed K-means clustering to group provinces with similar energy consumption patterns before applying the prediction model [7]. The hybrid approach leveraged LSTM's strength in capturing temporal dependencies and GBRT's effectiveness in handling complex feature relationships [7]. Results showed superior performance compared to single digits models like LSTM, CNN, and GRU models in the aspect of mean absolute percentage error, mean absolute error, and root mean square error [7].

The CNN-based approach for residential building emissions presents a novel application of computer vision techniques [8]. This research is the only one that focused on Peking's architecture industry in China within the 15 studies analyzed [8]. Two pain points were considered: Simulation-based methods and regression models were hard to adhere the workflow in early designing stage, while input variables of regression-based models neglected many details of the building space [8]. Hence, floor plan images were used as input to predict operational carbon emissions during early design stages [8]. The CNN architecture directly encoded spatial features from building layouts, capturing details that traditional parameter-based methods might miss [8]. The research created a dataset of approximately 2000 residential units in Beijing, training models to predict emission ranges rather than point estimates [8]. This approach provided architects with real-time feedback during design processes, achieving 89.5% accuracy with 200ms inference times, making it practical for design-stage decision support [8]. Constructing a corresponding relationship between architechtural shape and carbon emissions highlighted the originality of this research [8].

The LSTM-based regional emission trends research demonstrates the effectiveness of recurrent networks for temporal prediction [9]. This study trained LSTM networks on historical emission data to forecast future trends, showing high accuracy and stability [9]. The method proved particularly valuable for analyzing regional emission patterns and providing references for carbon reduction strategies [9]. The attention mechanisms incorporated in the model helped identify important time steps and features, enhancing both performance and interpretability [9].

The secondary decomposition BAS-LSTM study for intelligent buildings introduces a sophisticated preprocessing approach [10]. This method combined variational mode decomposition (VMD) and ensemble empirical mode decomposition (EEMD) before LSTM processing [10]. The beetle antennae search algorithm optimized LSTM parameters efficiently [10]. This comprehensive

DOI: 10.54254/2753-8818/2025.DL27564

approach handled the non-stationary and nonlinear characteristics of building emission data effectively, showing improved prediction accuracy for intelligent building emissions management [10].

The ANN approach for developing countries' emissions focused on practical implementation considerations [11]. ANN approach is a machine learning algorithm and prediction model created from the basic principle of human neuron [11]. This research predicted emissions for 20 developing nations using variables including green technology innovations, energy consumption, trade openness, economic growth, and population [11]. The study employed Levenberg-Marquardt optimization for network training, taking advantage of its faster convergence compared to gradient descent methods [11]. The approach demonstrated ANN's capability to handle diverse economic and environmental factors while providing accurate emission predictions for policy planning [11].

2.4. Combination and hybrid models

Compared to previous categories, hybrid approaches combine multiple methodologies to leverage their respective strengths while mitigating individual limitations. These models typically show improved accuracy and robustness by integrating different theoretical foundations and technical approaches.

The ARIMA-LSTM model for Chinese regional emissions represents a classic hybrid approach [12]. This research divided China into eastern, central, and western regions based on economic development and energy consumption patterns [12]. The study identified 14 influencing factors across population, economy, transportation, and technology dimensions [12]. For the preprocessing stage of the 27 years' data, linear interpolation method were applied to supplement the null values, Min-Max standardization method were employed for normalization [12]. Finally, random forest algorithm selected 10 most important factors [12]. The hybrid model used ARIMA to capture linear trends and LSTM to handle nonlinear residuals, along with the carbon emissions prediction model built from the definition and model combination of factors [12]. Overall, the study built the ARIMA-LSTM carbon emissions prediction model by introducing the non- linear prediction methods based on the reciprocal error combination method for balancing model prediction error [12]. And results showed this combination outperformed individual models (LR, BPNN, ARIMA, LSTM) in accuracy and effectiveness for both short-term and long-term predictions [12].

The VMD-SSA-LSSVM (variational mode decomposition-sparrow search algorithm-least square support vector machine) model for short-term emissions forecasting introduces a sophisticated signal processing approach [13]. This method used variational mode decomposition to separate emission data into four intrinsic mode functions with different center frequencies plus a residual component [13]. The sparrow search algorithm optimized least squares support vector machine parameters for each component [13]. And the final prediction results were calculated from the combination results generated by the fusion of every reconstructed sequences [13]. This approach significantly improved prediction accuracy for short-term emissions characterized by randomness and volatility, outperforming conventional methods in handling irregular data patterns and sudden changes [13].

The OLS-CNN fusion model addresses the interpretability challenge in deep learning approaches [14]. This innovative method first applied ordinary least squares regression to identify statistically significant emission drivers, then used these selected features as input to convolutional neural networks [14]. The OLS component provided econometric rigor and interpretability, while the CNN handled complex nonlinear relationships [14]. This study is innovative as the model simultaneously predicted carbon emissions and explanatory variables, maintaining both accuracy and

DOI: 10.54254/2753-8818/2025.DL27564

interpretability [14]. This approach effectively bridged the gap between traditional econometrics and modern deep learning, offering a balanced solution for policy-relevant emission forecasting [14].

2.5. Innovative approaches

Innovative approaches address specific application scenarios with tailored methodologies, often introducing novel data sources or prediction paradigms that expand the boundaries of conventional emission forecasting, which is rather unique for a particular field.

The G-code based method for 3D printing emissions presents a unique approach to manufacturing emission prediction [15]. This research analyzed the G-code instructions from fused deposition modeling (FDM) 3D printers to predict emissions before physical production [15]. The study extracted 27 printing parameters including speed, layer height, and fill density that directly affect energy consumption and emissions [15]. Twelve experimental sets validated the method's accuracy and reliability [15]. This approach enabled emission optimization during the design phase rather than post-production assessment, supporting sustainable development of additive manufacturing technologies [15].

This innovative method differs fundamentally from other approaches by using manufacturing instructions rather than historical emission data as primary input. It demonstrates how domain-specific knowledge can lead to novel prediction paradigms that address particular industry needs. The approach also shows potential for application in other computer-numerical-control manufacturing processes.

3. Conclusion

This review has systematically analyzed 15 studies covering a wide spectrum of carbon emission prediction methods, from traditional time-series models to advanced AI hybrids.

The comprehensive review of these fifteen studies reveals a diverse and evolving landscape of carbon emission prediction methodologies. Traditional statistical methods continue to provide value, particularly in data-scarce environments, while machine learning and deep learning approaches offer increasingly sophisticated pattern recognition capabilities. Hybrid models demonstrate how combining methodologies can achieve superior performance, and innovative specialized approaches address particular application needs. As emission prediction becomes increasingly important for climate policy and sustainable development, this methodological diversity provides multiple pathways for addressing the complex challenge of accurate carbon emission forecasting. Future research should focus on enhancing model interpretability, reducing computational costs, and improving adaptability to real-time data.

References

[1] Zhang, T. (2023). Research on carbon emission prediction method based on improved multivariate grey model [Master's thesis, Central South University]. (2023).

[2] Ma, L., Lin, K., Guan, M., & Lin, M. (2017). The prediction of carbon emission in all provinces of China with the K-means cluster based logistic model. In 2017 International Conference on Service Systems and Service Management (pp. 1–6). Dalian, China: IEEE.

[3] Peng, T., et al. (2023). Application of grey prediction model based on Python in carbon emission prediction and low-carbon economic development analysis. In 2023 International Conference on Network, Multimedia and Information Technology (NMITCON) (pp. 1–6). Bengaluru, India: IEEE.

[4] Qi, Y. W. (2024). Research on regional agricultural carbon emission prediction based on machine learning combined model [Master's thesis, Xidian University].

DOI: 10.54254/2753-8818/2025.DL27564

[5] Zhang, L., Lu, G., Yan, X., Xia, P., Chen, Z., & Wu, D. (2025). A differential evolution optimized hybrid XGBoost for accurate carbon emission prediction. Environmental Modelling & Software, 193, 106627.
[6] Zou, M., Ma, J., & Dai, J. (2024). A study of improved neural network methods applied to carbon emission prediction. In 2024 5th International Conference on Computer Engineering and Application (ICCEA) (pp. 1633–1637). Hangzhou, China: IEEE.
[7] Lin, C., Lyu, J. B., Zhang, Y., et al. (2025). Provincial-level carbon emission prediction method based on LSTM-GBRT model. Southern Power System Technology, 1–10.
[8] Yan, S., Zhang, Y., Sun, H., & Wang, A. (2023). A real-time operational carbon emission prediction method for the early design stage of residential units based on a convolutional neural network: A case study in Beijing, China. Journal of Building Engineering, 75, 106994.
[9] Jiang, X. (2025). Prediction method of carbon emissions of intelligent buildings based on secondary decomposition BAS-LSTM. Clean Technologies and Environmental Policy, 27, 1903–1913.
[10] Sakilu, O. B., & Chen, H. (2025). Carbon dioxide emissions prediction of selected developing countries using artificial neural network. Journal of the Knowledge Economy.
[11] Weiheng, G., Fulin, Z., Xiao, L., & Zhengfeng, W. (2023). Research on regional carbon emission trends based on LSTM. In 2023 6th Asia Conference on Energy and Electrical Engineering (ACEEE) (pp. 562–566). Chengdu, China: IEEE.
[12] Liu, Y. Z. (2024). Research on regional carbon emission prediction in China based on ARIMA-LSTM [Master's thesis, Liaoning Technical University].
[13] Xu, Z. L., Cheng, Z. Y., Zhang, S., et al. (2025). Short-term carbon emission prediction based on sequence-reconstructed VMD-SSA-LSSVM combined model. Journal of Anhui University (Natural Science Edition), 49(4), 28–37.
[14] Yao, J. (2024). A fusion method integrated econometrics and deep learning to improve the interpretability of prediction: Evidence from Chinese carbon emissions forecast based on OLS-CNN model. Computational Economics.
[15] Shujun Yu, Heng Liu, Gang Zhao, Hua Zhang, Feng Hou, Kuankuan Xu.A code-based method for carbon emission prediction of 3D printing: A case study on the fused deposition modeling (FDM) 3D printing and comparison with conventional approach. (n.d.).