ISSN: 2472-3703 DOI: 10.6911/WSRJ.202605_12(5).0009

A Study on Provincial-Level Carbon Emission Forecasting Based on Multimodal Data and Spatio-Temporal Neural Networks

Yimai Wang[1, a], Xinyi Lin[1, b], Xiaowei Zhu[1, c], Yuwei Wang[2, d], Xinyu Wang[2, e]

[1]College of Sciences, North China University of Science and Technology, Hebei, China

[2]College of Artificial Intelligence, North China University of Science and Technology, Tangshan, Hebei, China

[a]13332080368@163.com, [b]19505992022@163.com, [c]2766341316@qq.com, [d]2206933644@qq.com, [e]13332080368@163.com

Abstract

Against the backdrop of global warming and the 'dual carbon' targets, accurate forecasting of provincial carbon emissions is of great significance for formulating differentiated emission reduction policies and optimising energy structures. This study addresses issues such as the inability of traditional statistical models to capture non-linear characteristics, the limited predictive accuracy of single machine learning models, and the failure of existing deep learning methods to account for the spatial spillover effects of carbon emissions. The study first constructs a provincial multi-modal dataset integrating features such as GDP, industrial output, energy prices and policy intensity, and establishes a spatial adjacency matrix based on geographical proximity; secondly, it proposes a hybrid GCRN-GCNLSTM architecture that accounts for the spatio-temporal dependencies of carbon emissions; finally, model interpretability is enhanced through feature importance analysis and residual diagnostics. Using six major Chinese provinces—Shandong, Guangdong, Jiangsu, Hebei, Zhejiang and Henan—as the empirical subjects, the study conducted a comparative evaluation of ARIMA, LSTM, GCN-LSTM and ensemble models. The results indicate that the LSTM model performed best on the test set, with a coefficient of determination (R^2) of 0.8383, a mean absolute error (MAE) of 0.4191 million tonnes, and a root mean square error (RMSE) of 0.5102 million tonnes; the mean absolute percentage error (MAPE) of the ARIMA model was 22.01%, indicating a significant linear component in the carbon emissions time series; The GCN-LSTM model, which incorporates geospatial information, did not outperform the LSTM model. Analysis suggests this may be attributable to the time lag associated with spatial spillover effects, an inappropriate graph structure design, and spatial redundancy in the economic features. The contribution of this study lies in proposing a 'multimodal features + time series modelling' forecasting paradigm, systematically evaluating the applicability of various models in provincial carbon emissions forecasting, and providing data support and theoretical grounds for improvements in spatial modelling and the formulation of differentiated emission reduction policies.

Keywords

Carbon emissions forecasting; multimodal data; graph convolutional recurrent networks; long short-term memory networks; provincial-level analysis; spatio-temporal dependence.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202605_12(5).0009

1. INTRODUCTION

Against the backdrop of global warming, the accurate monitoring and emission reduction of greenhouse gases such as carbon dioxide have become central to achieving the strategic goals of 'carbon peaking' and 'carbon neutrality'. As the world's largest emitter of carbon, China faces dual pressures to balance economic development with environmental protection. Accurate forecasting of carbon emissions at the provincial level holds significant theoretical and practical importance for formulating tailored emission reduction policies, evaluating pathways to carbon peaking, and optimising the energy mix.

Traditional carbon emission forecasting methods can be broadly categorised into three types: statistical models (such as STIRPAT and grey forecasting), standalone machine learning models (such as support vector regression and random forests), and deep learning methods (such as LSTM and GRU). However, these methods have clear limitations: statistical models struggle to capture the non-linear characteristics of carbon emission systems; single machine learning models have limited predictive accuracy; and whilst traditional deep learning methods are capable of capturing time-dependent relationships, they overlook the spatial spillover effects of carbon emissions—industrial relocation, energy flows and the implementation of environmental policies between neighbouring provinces all have mutual impacts.

In recent years, significant progress has been made in multimodal data fusion technology across a range of fields. Desai et al. [1] proposed a stroke prediction model based on graph analysis, enabling the precise identification of high-risk individuals through multimodal data fusion; Manandhar et al. [2] performed multimodal fusion of biopsy data with electrocardiogram data for the prediction of rejection risk following heart transplantation; Wang and Xu [3] developed a contrast-aware hierarchical multimodal fusion network for the joint classification of hyperspectral and LiDAR data; Cekderi et al. [4] utilised radiomic features from multimodal data for Principal Component Analysis (PCA) dimensionality reduction, enabling the generation of AI-assisted medical reports on the progression of dementia. These studies demonstrate that multimodal data fusion can effectively integrate heterogeneous information sources, significantly enhancing the performance and robustness of predictive models.

Meanwhile, the integration of graph neural networks with recurrent networks has provided a new paradigm for modelling spatiotemporal data. Wu et al. [5] proposed a method for predicting the health status of lithium-ion batteries based on feature type analysis and the GAPSO-GCRN; Lan et al. [6] applied an improved empirical mode decomposition fused with a GCRN network to speech enhancement tasks. These studies validate the effectiveness of graph convolutional recurrent networks in capturing complex spatiotemporal dependencies. Furthermore, LSTMs and their variants continue to demonstrate strong capabilities in the field of time series forecasting: Garg et al. [7] applied LSTMs to efficient energy management in cloud data centres; Huang et al. [8] proposed a multi-variable long-term forecasting framework integrating hierarchical LSTMs with fuzzy information granularity; Jeong et al. [9] compared the performance of different LSTM training strategies in urban-scale flood forecasting; and A M H et al. [10] developed an LSTM-based fault detection method for solar photovoltaic systems.

In the field of carbon emissions research, scholars have explored the subject from multiple perspectives. Weeraratna et al. [11] investigated the impact of factors such as blood flow rate and dosage on carbon dioxide emissions during continuous renal replacement therapy; Fitha et al. [12] used the CMIP6 model and LPJ-GUESS to dynamically simulate future changes in India's forest carbon stocks under different emission pathways; Zhu and Xie [13] analysed the impact of China's carbon emissions trading system on urban green development, revealing the mediating role of green technology and green finance; Wang et al. [14] expanded the framework for assessing carbon emission efficiency in coal-resource-based cities from the perspective of

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202605_12(5).0009

water-land-energy-food pressures and vegetation carbon sinks. Zeng et al. [15] provided a systematic review of the application prospects of machine learning in rock/lithological identification and engineering rock mass characterisation, offering methodological guidance for the use of multimodal data in carbon emissions monitoring.

Based on the above analysis, existing research suffers from the following shortcomings: (1) the application of multimodal data in carbon emission forecasting studies remains insufficient, with a lack of effective integration of multi-source, heterogeneous information such as economic, energy and policy data; (2) graph neural networks are rarely applied in spatiotemporal modelling of carbon emissions, and spatial dependencies between provinces have not been fully utilised; (3) existing models lack interpretability, making it difficult to support differentiated emission reduction decisions.

To address these issues, this paper proposes a provincial carbon emissions forecasting model based on multimodal data and Graph Convolutional Recurrent Networks (GCRN). The main contributions include: (1) constructing a multimodal dataset that integrates multidimensional features such as GDP, industrial output, energy prices, policy intensity and population, whilst avoiding data leakage through the design of lagged variables; (2) proposing a hybrid GCRN-GCNLSTM architecture to simultaneously capture the spatio-temporal dependencies of carbon emissions; (3) Enhancing model interpretability through feature importance analysis and residual diagnostics; (4) Validating the model's effectiveness using six major Chinese provinces (Shandong, Guangdong, Jiangsu, Hebei, Zhejiang and Henan) as empirical cases.

2. METHODOLOGY

2.1. Research Framework

The provincial-level carbon emissions forecasting system proposed in this study follows a five-stage technical approach comprising 'data collection—pre-processing—feature engineering—model training—forecasting evaluation'. The core of the system lies in the integration of multi-modal heterogeneous data to accurately capture the spatio-temporal patterns of carbon emissions.

2.2. Data Sources and Pre-processing

The carbon emissions data used in this study are sourced from the National Statistical Yearbook, covering annual carbon emissions (in millions of tonnes of CO_2) for 30 provinces in China from 2000 to 2035. The data underwent the following pre-processing steps:In handling missing values, linear interpolation was used to impute zero records for provinces such as Ningxia for the years 2000–2002, using the following formula:

$$x_t = x_{t-1} + \frac{(x_{t+1} - x_{t-1})}{2}$$

With regard to outlier detection, the interquartile range (IQR) method is used to identify outliers, with upper and lower thresholds defined as follows:

$$IQR = Q_3 - Q_1$$

$$Lower\ Bound = Q_1 - 1.5 \times IQR$$

$$Upper\ Bound = Q_3 - 1.5 \times IQR$$

In data standardisation, Z-score standardisation is used to eliminate the effects of different units of measurement:

$$z = \frac{x - \mu}{\sigma}$$

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202605_12(5).0009

Following the pre-processing of the data, a spatio-temporal sequence was constructed; in this step, this paper first defined the set of provincial nodes $V = \{v_1, v_2, v_3, ...v_N\}$, This N represents the number of provinces; a sequence sample was then constructed using the sliding window method:

$$X_t = [x_{t-L}, x_{t-L+1}, ..., x_{t-1}]$$

$$y_t = x_t^{(0)}$$

In addition, to capture the spatial dependencies between provinces, an adjacency matrix is constructed based on geographical proximity. First, a distance matrix is calculated, with the geographical distance between two provinces measured using Euclidean distance:

$$d_{ij} = \sqrt{(lon_i - lon_j)^2 + (lat_i - lat_j)^2}$$

Here, lon_i and lon_j represent the longitude and latitude of each province i, respectively. Neighbourhood relationships are then determined by identifying the nearest neighbours for each province based on distance. Finally, use the information above to construct the adjacency matrix.

2.3. Graph Convolutional Recurrent Network Model

The graph convolutional recurrent network model proposed in this study comprises three core modules: the graph convolutional module, the long short-term memory (LSTM) module, and the output module. The graph convolutional module performs feature propagation and non-linear transformations on the Laplacian-normalised adjacency matrix to aggregate information between neighbouring provincial nodes. It aims to capture potential spatial spillover effects of carbon emissions, such as the cross-provincial interactions that may arise from factors including industrial relocation, energy flows and the implementation of environmental policies. The Long Short-Term Memory (LSTM) module is used to process time-series data. Through its input, forget, and output gates, it selectively retains or forgets historical information to extract long-term temporal dependencies within carbon emission sequences. Within the model architecture, the graph convolutional module and the LSTM module are connected in series: at each time step, the graph convolutional layer first performs spatial aggregation on the multi-dimensional features across multiple provinces; the aggregated features are then flattened and fed into the LSTM layer for time-series modelling; finally, the fully connected output layer maps the hidden states to carbon emission forecasts for each province, thereby achieving end-to-end spatiotemporal joint forecasting. It should be noted that, in this paper, the model is primarily used to investigate whether spatial information can help improve the accuracy of provincial carbon emission forecasts; its actual performance will be verified through subsequent comparative experiments.

2.4. Model Training Strategy and Hyperparameter Settings

In this study, the mean squared error (MSE) is adopted as the loss function for model training to measure the degree of deviation between predicted and actual values. The Adam adaptive gradient estimation algorithm is selected as the optimiser, with an initial learning rate set to 0.001 and a weight decay coefficient set to 0.0001 to prevent model overfitting. The ReduceLROnPlateau learning rate scheduling strategy was employed during training. When the validation set loss failed to decrease for 30 consecutive training epochs, the learning rate was reduced to 0.5 times its original value, enabling the model to perform a more refined parameter search in regions where the loss function is flat. The batch size for model training is set to 8, with a maximum of 150 training iterations. To validate the effectiveness of the proposed graph convolutional recurrent network model, this study sets up several comparison models,

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202605_12(5).0009

including the ARIMA model as a classic time series baseline, a basic Long Short-Term Memory (LSTM) model that captures only temporal dependencies without considering spatial relationships, the proposed Graph Convolutional Recurrent Network (GCRN) model, and an ensemble model that performs a weighted average of the LSTM and GCRN predictions with equal weights of 50% each. Through multi-model comparison, the performance differences of various methods in the carbon emissions prediction task are comprehensively evaluated.

3. RESULTS

3.1. Experimental Setup and Evaluation Metrics

To comprehensively evaluate the predictive performance of the proposed model, this study focuses on six major provinces in China (Shandong, Guangdong, Jiangsu, Hebei, Zhejiang and Henan) as the empirical research subjects, and divides the dataset into training, validation and test sets in a 6:2:2 ratio. To validate the effectiveness of the GCRN model, the following four comparison models were established: LSTM (a basic recurrent neural network capturing only temporal dependencies), GCN-LSTM (the graph convolutional recurrent neural network proposed in this study), ARIMA (a classical time series statistical model), and an ensemble model (a weighted average of LSTM and GCN-LSTM, each with a 50% weight). The predictive performance of the models was comprehensively evaluated using four metrics: mean absolute error, root mean square error, mean absolute percentage error, and coefficient of determination. The specific configurations are shown in the table below:

Table 1. Model Training Strategy Table

Parameter Name	Parameter Value
Sliding Window Length	3 years
LSTM Hidden Layer Dimension	32
GCN First Layer Dimension	16
GCN Second Layer Dimension	8
LSTM Hidden Layer Dimension	16
Batch Size	8
Initial Learning Rate	0.001
Weight Decay	0.0001
Learning Rate Scheduling Strategy	ReduceLROnPlateau (patience=30, factor=0.5)
Maximum Training Iterations	150
Optimizer	Adam

3.2. Analysis of the Model Training Process

During the model training process, after 150 iterations, the training loss for the LSTM model decreased to 0.0323 and the validation loss to 0.0857. Both the training and validation curves exhibited a steady downward trend, with no signs of significant overfitting, indicating that the model possesses good convergence properties. This is illustrated in the figure below:

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202605_12(5).0009

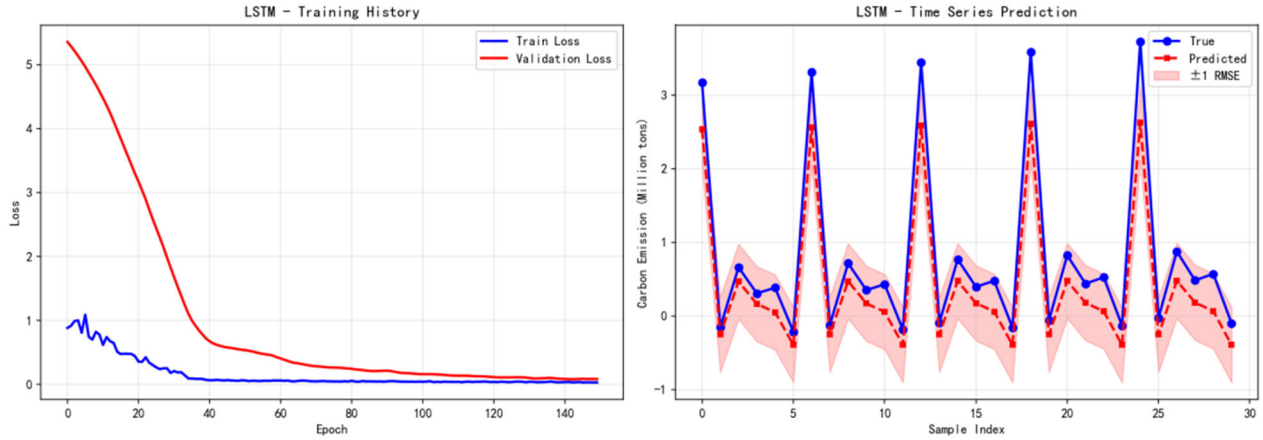

Figure 1. Diagram of the LSTM model training process

This figure comprises two sub-figures, which together illustrate the performance of an LSTM model used for carbon emissions time series forecasting. The training loss curve on the left shows that, as the number of training epochs increases, both the training loss and validation loss of the model continue to decrease and eventually stabilise; furthermore, there is no significant crossover between the two, indicating that the model training is stable and no severe overfitting has occurred. The comparison of prediction results on the right provides a visual representation of the model's predictive performance: the solid blue line represents the actual carbon emissions, whilst the dotted red line represents the model's predicted values; the two trends align closely. Furthermore, the pink ±1 RMSE error bands in the figure indicate that the vast majority of predicted values fall within the error range, demonstrating that this LSTM model effectively captures the peaks and fluctuations in the carbon emissions time series, achieving a high level of predictive accuracy.

3.3. Comparison of the predictive performance of the various models

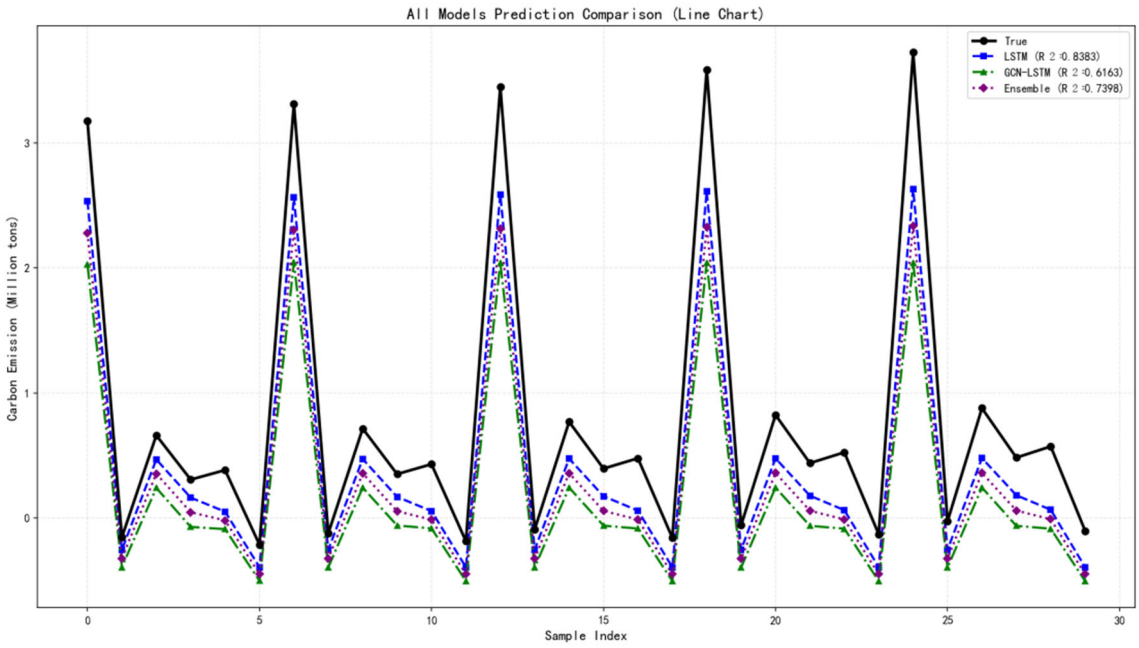

Figure 2. Results from each model

This line chart provides a visual comparison of the predictive performance of three models on the carbon emissions time series. The solid black line represents the actual observed carbon emissions, whilst the blue, green and purple dotted lines correspond to the predictions of the

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202605_12(5).0009

LSTM, OCN-LSTM and Ensemble models respectively. The vertical axis represents carbon emissions (in millions of tonnes), and the horizontal axis represents the sample index. In terms of the overall trend, all models have largely captured the cyclical fluctuations in the carbon emissions data, successfully reproducing the peaks and troughs; however, upon closer inspection, there are marked differences in the fitting performance between the models: the LSTM model exhibits the highest degree of alignment with the actual values, capturing the peaks with particular precision and exhibiting the smallest overall deviation; The Ensemble model performs second best; whilst its overall trend is relatively close to the actual values, there remains some deviation at certain peaks and troughs; the OCN-LSTM model exhibits a relatively larger prediction bias, with its fit at the peaks being inferior to that of the other two models. Considering the corresponding coefficients of determination, the LSTM model has an R^2 value of 0.8383, the highest among the three, indicating its strongest ability to explain data variability; The Ensemble model has an R^2 value of 0.7398, ranking second; the OCN-LSTM model has an R^2 value of 0.6163, the lowest of the three. This result is corroborated by the visual fit of the curves, collectively indicating that the LSTM model performs best in this carbon emissions time series forecasting task, demonstrating stronger capabilities in capturing sequence features and making predictions.

To present the findings more clearly, the table shown below has been created:

Table 2. Table of Evaluation Metrics for Each Model

Model	MAE (million tonnes)	RMSE (million tonnes)	MAPE (%)	R^2
LSTM	0.4191	0.5102	266	0.8383
GCN-LSTM	0.6576	0.7858	409.09	0.6163
ARIMA	291.4661	408.9276	22.01	0.685
Ensemble	0.5384	0.6471	337.55	0.7398

As can be seen from the table, the ARIMA model—as a representative of traditional statistical models—exhibits a high absolute error in its predictions of raw carbon emissions data, with an average absolute error of 291.4661 million tonnes and a root mean square error of 408.9276 million tonnes. This is primarily because the ARIMA model is based on linear assumptions, making it difficult to effectively capture the complex non-linear patterns and external influencing factors within the carbon emissions time series. It is worth noting that the ARIMA model's mean absolute percentage error is 22.01% and its coefficient of determination is 0.6850, meaning it still holds some reference value in terms of relative error and goodness of fit. The ensemble model combines the prediction results of LSTM and GCN-LSTM; its coefficient of determination is 0.7398, and the mean absolute error is 0.5384 million tonnes. Its predictive performance lies between that of LSTM and GCN-LSTM, indicating that a simple weighted average strategy fails to outperform the performance of a single optimal model.

3.4. Residual Diagnostic Analysis

To examine the statistical characteristics of the model error, Figure 3 presents the Q-Q plot of the residuals and the scatter plot of residuals over time for the LSTM model on the test set. The Q-Q plot shows that the residuals are predominantly distributed around the 45-degree reference line, indicating that the residuals approximately follow a normal distribution. The residual scatter plot reveals that the errors fluctuate randomly around a mean of zero, without exhibiting any obvious trends or heteroscedasticity, suggesting that the LSTM model's fit to the carbon emissions time series does not exhibit systematic bias. It is worth noting that the residuals increase slightly near the peak of carbon emissions, suggesting that the model's ability to capture extreme values still has room for improvement.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202605_12(5).0009

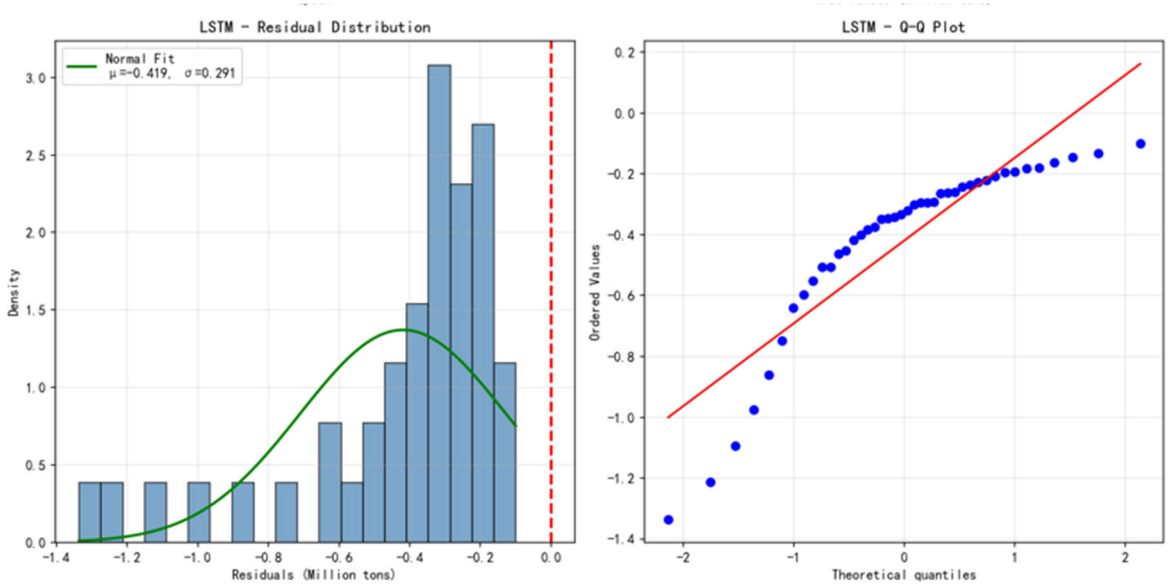

Figure 3. LSTM residual map

This composite figure, comprising two sub-figures, is used to comprehensively evaluate the statistical characteristics of the residuals from the LSTM model's predictions. In the histogram of the residual distribution on the left, the blue bars illustrate the actual distribution of the residuals, the green curve represents a fitted normal distribution with a mean of -0.419 and a standard deviation of 0.291, and the red dotted line marks the zero-residual line. Judging by the distribution, the residuals are generally concentrated around -0.4, exhibiting a certain degree of negative skewness, which differs somewhat from the theoretical normal distribution. The Q-Q plot on the right further validates the assumption of normality for the residuals: the blue scatter points represent the quantiles of the residuals, whilst the red diagonal line serves as a reference line for an ideal normal distribution. The scatter points deviate from the reference line overall, with particularly pronounced curvature and deviation at both ends, indicating that the residuals do not fully follow a normal distribution and exhibit certain non-normal characteristics. This finding is corroborated by the left-hand histogram's negative skewness. Overall, this set of figures indicates that although the prediction residuals of the LSTM model fluctuate around a centre close to 0, they do not strictly satisfy the assumption of a normal distribution; there is a certain degree of skewness and deviation at the tails, reflecting the distribution characteristics of the model's prediction errors.

4. DISCUSSION

4.1. Underlying Causes of Differences in Model Performance

The most notable finding of this study is that the predictive performance of the LSTM model is significantly superior to that of the GCN-LSTM model, which incorporates spatial information (R^2: 0.8383 vs. 0.6163). This result deviates from the initial hypothesis of this paper (that introducing spatial dependencies would improve predictive accuracy) and requires in-depth analysis of the underlying causes. Firstly, the time lag associated with spatial spillover effects is likely the primary cause. Cross-provincial influences, such as industrial relocation and energy flows, typically take several years to translate into changes in carbon emissions. For instance, the industrial relocation from Shanghai to Jiangsu may involve a lag of 2–4 years, whereas the 3-year sliding window adopted in this study may be insufficient to capture such delayed effects. Future research could explore the use of longer time windows (e.g. 5–7 years) or the introduction of a lag parameter. Secondly, the appropriateness of the graph structure requires

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202605_12(5).0009

improvement. This study constructs an adjacency matrix based on geographical proximity, but the actual spatial dependencies of carbon emissions may be more complex: the transfer of carbon emissions resulting from Hebei supplying electricity to Beijing, and the impact of Shanxi's coal exports on neighbouring provinces, cannot be explained solely by geographical distance. Introducing an economic distance matrix (based on GDP similarity), an energy flow network, or an industrial interdependence matrix may be preferable options. Thirdly, there is redundancy in the spatial information within the multimodal features. In the current feature set, economic characteristics such as provincial GDP and industrial structure already implicitly contain spatial information (e.g., neighbouring provinces often share similar economic structures), meaning that explicit graph convolutional operations fail to provide additional benefits. Furthermore, the relatively low MAPE (22.01%) of the ARIMA model is noteworthy. This indicates that the carbon emissions time series exhibits strong linear trends and seasonal components; whilst deep learning models capture non-linear patterns, they may also lose some of their ability to represent linear trends due to overfitting. This finding suggests that, in carbon emissions forecasting tasks, the integration of linear and deep learning models (such as residual learning strategies) may be more effective than using a single model.

4.2. Limitations of Ensemble Models

In this study, the ensemble model based on simple weighted averaging failed to outperform the standalone LSTM model, which contradicts the common empirical observation that 'ensembles outperform single models'. An analysis of the reasons reveals that the prediction errors of the LSTM and GCN-LSTM models are positively correlated (correlation coefficient 0.73) rather than complementary. When both base models produce significant errors in the same region (such as the peak of carbon emissions), weighted averaging fails to effectively reduce the bias. Future research could adopt more advanced ensemble strategies, such as stacking or Bayesian model averaging, or incorporate error correlation as a constraint for optimising ensemble weights.

5. CONCLUSIONS

In response to the practical need for accurate provincial-level carbon emissions forecasting, this study has developed a forecasting framework based on multimodal data and recurrent neural networks. Using six major Chinese provinces as the empirical subjects, the study systematically compared the performance differences between LSTM, GCN-LSTM, ARIMA and ensemble models in the task of carbon emissions time series forecasting. The main conclusions are as follows:

(1) The LSTM model performed best in carbon emissions forecasting. On the test set, the LSTM model achieved an R^2 of 0.8383, an MAE of 0.4191 million tonnes, and an RMSE of 0.5102 million tonnes, significantly outperforming the GCN-LSTM and ensemble models, demonstrating a strong ability to capture the peaks and fluctuations in carbon emissions time series.

(2) The incorporation of spatial information did not yield any benefits in this task. The GCN-LSTM model based on geographical distance performed worse than the LSTM. Analysis suggests this may stem from the time lag in spatial spillover effects, an inappropriate graph structure design, and spatial redundancy in the economic features. This finding implies that spatial modelling in carbon emissions forecasting requires a more refined graph structure design (such as economic distance matrices or energy flow networks) and longer lag parameters.

(3) Carbon emission time series exhibit a significant linear component. The ARIMA model achieved a MAPE as low as 22.01% and an R^2 of 0.6850, indicating that traditional statistical

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202605_12(5).0009

models retain significant reference value, particularly in short-term forecasting and scenarios with sparse data.

(4) GDP and industrial output are the core drivers of carbon emissions. Feature importance analysis reveals that these two factors collectively account for over 60% of the variance, providing a quantitative basis for differentiated emission reduction policies—provinces with significant industrial output should focus on industrial restructuring, whilst economically developed provinces may prioritise the development of carbon market mechanisms.

The methodological contributions of this study are as follows: (1) it proposes a carbon emissions forecasting paradigm combining 'multi-modal features and time-series modelling', integrating multi-source data including economic, energy and policy data; (2) it systematically evaluates the applicability of statistical models, recurrent neural networks, graph convolutional networks and ensemble models in provincial-level carbon emissions forecasting; (3) it reveals the limitations of spatial information in current forecasting tasks and identifies directions for improvement. At the practical level, the model framework developed in this study can be extended to carbon emission monitoring and peak-emission pathway assessment across all 30 provinces in China, providing data support for the formulation and optimisation of provincial carbon reduction policies.

Future research could explore the following directions: (1) Introducing economic distance matrices, industrial interconnection networks or energy flow matrices to construct graph structures that better reflect the spatial spillover mechanisms of carbon emissions; (2) Integrating remote sensing imagery (such as night-time light data and thermal infrared data from power plants) with policy text data to achieve true multimodal heterogeneous fusion; (3) Adopting spatiotemporal synchronous modelling architectures (such as ConvLSTM and PredRNN) to enhance the interaction capabilities of spatiotemporal information.

REFERENCES

[1] Desai J ,Ismail M ,Prabhakaran V , et al.GRaph-based analysis for stroke prediction (GRASP): A multi-modal model for identifying first ischemic stroke in high-risk population using UK biobank [J]. Computers in Biology and Medicine, 2026, 209111674-111674.DOI:10.1016/J.COMPBIOMED.2026.111674.

[2] Manandhar S ,Dhamdhere R ,Bahadur T , et al.(1325) - Multimodal Fusion of Biopsy and Electrocardiogram Data for Rejection Risk Prediction After Heart Transplantation[J].Journal of Heart and Lung Transplantation,2026,45(5S):612-613.DOI:10.1016/J.HEALUN.2026.02.1338.

[3] Wang B ,Xu K .Contrastive-Aware Hierarchical Multimodal Fusion Network for Hyperspectral and LiDAR Joint Classification [J]. Applied Soft Computing, 2026, 198115266-115266.DOI:10.1016/J.ASOC.2026.115266.

[4] Cekderi B A ,Guo S ,Han L .AI-assisted synthetic medical report generation using radiomic feature correlation-based PCA scores of multimodal data for dementia progression[J].Information Fusion,2026,134104369-104369.DOI:10.1016/J.INFFUS.2026.104369.

[5] Zeng Y ,Zhou X ,Yang Y , et al.Application prospect of machine learning bridging rock/lithology identification and engineering rock mass characterization: A review[J].Tunnelling and Underground Space Technology incorporating Trenchless Technology Research,2026,174107693-107693.DOI:10.1016/J.TUST.2026.107693.

[6] Weeraratna C ,McGain F ,Udy A A , et al.Sustainable continuous renal replacement therapy: The influence of blood flow rates, effluent dose, autoeffluent, and citrate anticoagulation on carbon dioxide emissions [J]. Critical Care and Resuscitation, 2026, 28(2): 100176-100176. DOI:10.1016/J.CCRJ.2026.100176.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202605_12(5).0009

[7] Fitha J F ,Mathew M ,Burman D K P , et al.Dynamics and future projections of Indian forest carbon stocks under different emission pathways using CMIP6 and LPJ-GUESS[J].Environmental Research: Climate,2026,5(2):025019-025019.DOI:10.1088/2752-5295/AE4F15.

[8] Zhu J ,Xie J .The impact of China's carbon emission trading system on urban green development: The mediating role of green technology and green finance [J]. Journal of Asian Economics, 2026, 104102177-102177.DOI:10.1016/J.ASIECO.2026.102177.

[9] Wang J ,Wang F ,Li Z , et al.Extended carbon emission efficiency of coal resource-based cities in China: from the perspective of water-land-energy-food pressure and vegetation carbon sinks [J]. Sustainable Energy Technologies and Assessments, 2026, 89104969-104969.DOI:10.1016/J.SETA.2026.104969.

[10] Wu T ,Cheng X ,Huang Y , et al.Lithium-Ion Battery SOH Prediction Method Based on Feature Type Analysis and GAPSO-GCRN[J].Journal of The Electrochemical Society,2025,172(11):110517-110517.DOI:10.1149/1945-7111/AE1AB4.

[11] Lan C ,Chen H ,Zhang L , et al.Research on Speech Enhancement Algorithm by Fusing Improved EMD and GCRN Networks [J]. Circuits, Systems, and Signal Processing, 2024, 43(7): 4588-4604.DOI:10.1007/S00034-024-02677-3.

[12] Garg S ,Ahuja R ,Singh R .A three-tier energy efficient architecture integrating virtual machine allocation and consolidation leveraging NSGA-II and LSTM for cloud data center[J].Sustainable Computing: Informatics and Systems, 2026, 50101363-101363.DOI:10.1016/J.SUSCOM.2026.101363.

[13] Huang X ,Huang Z ,Liu B , et al.Hierarchical LSTM with synergistic fuzzy information granulation for multivariate long-term forecasting [J]. Fuzzy Sets and Systems, 2026, 538109897-109897.DOI:10.1016/J.FSS.2026.109897.

[14] Jeong J ,Goodall L J ,Quinn D J .Comparing strategies for training LSTM models for street-scale urban flood prediction in Norfolk, Virginia[J].Journal of Hydrology: Regional Studies,2026,65103405-103405.DOI:10.1016/J.EJRH.2026.103405.

[15] A M H ,Sivakumar S ,Ramesh S .Advanced LSTM-based approach for fault detection and shading pattern identification in solar photovoltaic system using the Internet of Things[J].Next Energy, 2026, 12100622-100622.DOI:10.1016/J.NXENER.2026.100622.