Deep Learning Approaches for Carbon Pricing Mechanism Design in Green Transportation Supply Chains

Ziheng Gao*

School of Environment, Education and Development, University of Manchester, Oxford Rd, Manchester, M13 9PL, UK

INFORMATION

Keywords:
Deep learning
carbon pricing mechanism
green transportation
supply chain management
neural networks
sustainable development

DOI: 10.23967/j.rimni.2025.10.73399

SCIPEDIA
Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería
RIMNI

Deep Learning Approaches for Carbon Pricing Mechanism Design in Green Transportation Supply Chains

Ziheng Gao*

School of Environment, Education and Development, University of Manchester, Oxford Rd, Manchester, M13 9PL, UK

ABSTRACT

This study uniquely integrates Long Short-Term Memory networks (LSTM) and Graph Convolutional Networks (GCN) with a multi-head attention mechanism to address dynamic carbon pricing optimization in green transportation supply chains, overcoming the limitations of traditional static models. As global climate change issues become increasingly severe, the design of carbon pricing mechanisms for green transportation supply chains has become a key factor in promoting sustainable development. We construct a hybrid deep learning model that simultaneously captures temporal dependencies in carbon emission data and spatial relationships in supply chain network structures. Traditional carbon pricing methods often rely on static models and simplified assumptions, making it difficult to adapt to complex and dynamic supply chain environments. Experimental results show that the proposed deep learning method improves carbon price prediction accuracy by 23.7% compared to traditional methods and achieves 18.5% improvement in supply chain cost optimization. Furthermore, the method achieved an average 21.6% carbon emission reduction and 15.5% cost reduction in three real green transportation supply chain cases, demonstrating its effectiveness in practical applications. The multi-objective optimization framework successfully balances the trade-off between economic and environmental benefits through organic integration of genetic algorithms and deep learning models. Ablation experiments validated the importance of each model component, and sensitivity analysis confirmed the rationality of parameter settings. This method provides strong technical support for formulating more precise and dynamic carbon pricing policies, offering significant theoretical value and practical significance for promoting sustainable development of green transportation supply chains.

OPEN ACCESS

Received: 17/09/2025

Accepted: 24/11/2025

Published: 20/03/2026

DOI
10.23967/j.rimni.2025.10.73399

Keywords:
Deep learning
carbon pricing mechanism
green transportation
supply chain management
neural networks
sustainable development

1 Introduction

In the context of increasingly severe global climate change, the transformation of green transportation supply chains has become a key pathway to achieving "dual carbon" goals. According to reports from the Intergovernmental Panel on Climate Change (IPCC), the transportation industry contributes

*Correspondence: Ziheng Gao (ziheng.gao@postgrad.manchester.ac.uk).

14% of global greenhouse gas emissions, with carbon emissions from supply chain segments accounting for a significant proportion [1,2]. Traditional carbon pricing mechanisms mostly rely on static models and linear assumptions, making it difficult to adapt to the complexity and dynamics of modern supply chain networks [3,4]. Meanwhile, deep learning technologies have demonstrated tremendous potential in time series prediction, network modeling, and intelligent optimization, providing new technical means for innovation in supply chain carbon pricing mechanisms [5,6]. However, how to effectively combine the technical advantages of deep learning with domain knowledge of carbon pricing to construct an intelligent framework that can both accurately predict carbon emissions and dynamically optimize pricing strategies remains a major challenge facing current research [7,8].

Existing research shows significant gaps and deficiencies in the integration of deep learning and carbon pricing mechanisms. First, most carbon pricing research still adopts traditional operations research methods, lacking deep modeling capabilities for supply chain temporal dynamics and network topological structures [9,10]. Second, existing deep learning applications mainly focus on single-task optimization, lacking end-to-end frameworks that organically integrate prediction and decision-making [11,12]. Finally, in handling multi-objective optimization problems in supply chain networks, existing methods often ignore the complex interactive relationships of spatiotemporal features [13,14]. Therefore, our research addresses the following three key issues: (1) How to construct a hybrid deep learning model capable of simultaneously processing temporal carbon emission data and supply chain network structures; (2) How to design a multi-objective optimization framework to achieve dynamic balance between cost-effectiveness and environmental benefits; (3) How to establish an end-to-end intelligent carbon pricing mechanism that supports real-time decision-making and strategy adjustment.

Research Gap Identification. Despite the progress in applying deep learning to supply chain management, critical gaps remain in carbon pricing mechanism design. First, existing deep learning approaches primarily focus on either temporal prediction (using RNN/LSTM) or network analysis (using GNN), but fail to integrate both dimensions in a unified framework [6,11,12]. Second, most carbon pricing research treats prediction and optimization as separate tasks, lacking end-to-end frameworks that can generate actionable pricing strategies directly from raw data [3,9,10]. Third, current multi-objective optimization methods for green supply chains do not adequately model the dynamic trade-offs between cost efficiency and emission reduction under real-time market conditions [13,14]. Finally, the interpretability and generalizability of deep learning models in carbon pricing applications remain underexplored, limiting their adoption in policy-making contexts [15].

The main contributions of this work can be summarized as follows:

1. We propose a novel LSTM-GCN hybrid architecture with multi-head attention mechanism that simultaneously captures temporal carbon emission dynamics and spatial supply chain network structures, addressing the limitation of existing single-modal approaches.
2. We design an end-to-end multi-objective optimization framework that integrates genetic algorithms with deep learning models, enabling real-time carbon pricing strategy generation while balancing economic and environmental objectives.
3. We provide comprehensive empirical validation on three real-world green transportation supply chain cases, demonstrating significant improvements: 23.7% in prediction accuracy, 21.6% in emission reduction, and 15.5% in cost savings compared to traditional methods.
4. We conduct extensive ablation studies and sensitivity analyses to provide insights into the contribution of each model component and parameter configuration, offering practical guidelines for implementation.

The article is organized as follows: Section 2 reviews related research on green supply chain management, applications of deep learning in carbon emission prediction, and multi-objective optimization methods; Section 3 details the LSTM-GCN hybrid model architecture, multi-objective optimization algorithm design, and algorithm implementation details; Section 4 validates the effectiveness of the proposed method through large-scale experiments and analyzes application effects in real cases; Section 5 discusses the theoretical contributions, practical significance, and future development directions of the research; Section 6 summarizes the main conclusions and policy recommendations of the entire paper.

2 Related Work

2.1 Research on Green Supply Chain Carbon Pricing Mechanisms

Green supply chain management, as an important strategy to address climate change, has always been a focus of attention in academia and industry regarding its carbon pricing mechanism design. Early research mainly concentrated on theoretical analysis of two basic pricing models: carbon tax and carbon trading [1,2]. Beamon [1] first proposed the conceptual framework of green supply chains, laying the theoretical foundation for subsequent research. Carter and Rogers [16] further developed the theoretical system of sustainable supply chain management, emphasizing comprehensive consideration of environmental, economic, and social benefits.

In recent years, research focus has gradually shifted toward dynamic pricing mechanisms and multi-objective optimization methods. Zhang et al. [9] proposed a dual-channel supply chain network equilibrium model under progressive carbon tax regulation, analyzing the impact of different carbon tax policies on supply chain decisions. Qu et al. [4] studied low-carbon supply chain optimization problems considering warranty periods and carbon emission reduction levels under cap-and-trade mechanisms. Tavana et al. [3] constructed a location-inventory-routing optimization model for green supply chains under uncertain environments, balancing costs and emissions through mixed integer programming methods. However, most of these studies are based on static assumptions, lacking in-depth consideration of supply chain dynamics and complexity [10,17].

Latest research has begun exploring applications of big data and artificial intelligence technologies in carbon pricing. Sarkis et al. [18] studied strategies for reducing carbon footprints in sustainable supply chain management using big data technologies. Kong et al. [10] employed system dynamics methods to analyze carbon abatement measures in maritime supply chains. Although these studies introduced new technical means, there is still considerable room for development in the deep integration of deep learning and carbon pricing mechanisms [19,20].

Critical Assessment. While these studies have established important theoretical foundations for green supply chain carbon pricing, they exhibit several critical limitations: (1) Most models assume static supply chain structures and fail to capture dynamic network reconfigurations [9,10]; (2) The reliance on simplified linear relationships between carbon prices and emissions overlooks nonlinear feedback mechanisms [3,4]; (3) Limited consideration of spatial heterogeneity across supply chain nodes leads to suboptimal pricing strategies [17,21]; (4) Lack of real-time adaptability makes these approaches unsuitable for rapidly changing market conditions. These gaps motivate the development of more sophisticated, data-driven approaches that can handle complexity and dynamics inherent in modern supply chains.

2.2 Applications of Deep Learning in Supply Chain Management

The application of deep learning technologies in supply chain management has evolved from simple prediction to complex decision optimization. Jordan and Mitchell [5] pointed out in their review of machine learning development trends that deep learning has unique advantages in complex system modeling. Mahadevaswamy and Swathi's [22] probabilistic machine learning theory provided important guidance for supply chain uncertainty modeling.

In time series prediction, deep learning technologies have shown significant advantages. Sun et al. [23] proposed a sentiment analysis method based on bidirectional LSTM networks, providing technical reference for bidirectional modeling of time series data. Sun et al. [24] successfully applied residual enhanced channel spatial attention networks and frequency domain time series feature fusion methods in stock price prediction, proving the effectiveness of deep learning in financial time series prediction [25]. These studies provided important technical references for carbon emission time series prediction [11,12].

In network structure modeling, the development of graph neural networks has brought new opportunities for supply chain network analysis. Riahi et al. [26] systematically reviewed research trends of machine learning in supply chain management, pointing out the potential of graph neural networks in handling supply chain network topological structures. Islam et al.'s [27] bibliometric analysis showed that artificial intelligence applications in supply chains are developing toward more complex and intelligent directions [6,28].

However, existing research still has deficiencies in combining temporal modeling with network structure modeling. Most studies either focus on time series prediction or network analysis, lacking effective integration mechanisms [29,30]. Furthermore, in specific applications of supply chain carbon emission prediction, how to handle the interaction between multi-scale temporal features and complex network relationships remains a key unresolved issue [31].

Limitations of Current Approaches. Despite these advances, existing deep learning applications in supply chain management face three major challenges when applied to carbon pricing: (1) Temporal models like LSTM capture sequential patterns but ignore the network topology of supply chains [11,23]; (2) Graph neural networks model spatial relationships but struggle with temporal dynamics [26,28]; (3) Most studies focus on single-objective optimization (e.g., cost minimization or demand forecasting) and lack integrated frameworks for multi-objective decision-making in carbon pricing contexts [29–31]. Our work addresses these limitations through a hybrid architecture that synergistically combines temporal and spatial modeling capabilities.

2.3 Multi-Objective Optimization and Intelligent Decision Methods

Multi-objective optimization, as an important method for handling complex decision problems in supply chains, has seen rapid development in both theory and algorithms in recent years. Traditional multi-objective optimization methods mainly based on mathematical programming and evolutionary algorithms have achieved certain effectiveness in handling trade-offs between supply chain cost, time, and environmental objectives [32,33].

The application of reinforcement learning in multi-objective decision-making has opened new pathways for intelligent supply chain management. Qiu et al. [13] proposed a dynamic multi-objective supply chain decision method based on reinforcement learning and evolutionary strategies, demonstrating the potential of reinforcement learning in handling multi-objective optimization problems under dynamic environments. Wang et al. [33] developed a dual deep Q-learning network guided multi-agent path planning method, providing technical solutions for intelligent optimization under complex decision environments.

In specific applications of green supply chains, multi-objective optimization methods have gradually evolved from pure mathematical modeling toward intelligent decision-making. Xiao et al. [34] combined dynamic data envelopment analysis and artificial neural networks to predict green supplier efficiency, achieving organic integration of multi-objective evaluation and prediction. Obermeyer and Emanuel [35] constructed a big data cloud computing framework for low-carbon supplier selection in beef supply chains, achieving multi-objective optimization. Ahmad Jauhari et al. [20] studied the impact of technological dimensions of green supply chain management practices on firm performance, analyzing the value of technological innovation from a multi-objective perspective [14,34].

Although existing research has made important progress in multi-objective optimization methods, there are still obvious deficiencies in intelligent decision-making for carbon pricing mechanisms. First, most studies adopt offline optimization approaches, making it difficult to adapt to real-time changes in supply chain environments [36]. Second, existing methods lack adaptive capabilities when handling dynamic trade-off relationships among multiple objectives [37,38]. Finally, end-to-end frameworks that organically integrate prediction models with optimization decisions are still lacking, limiting the practical application effectiveness of intelligent carbon pricing mechanisms [15,39].

Gap in Multi-Objective Carbon Pricing. Existing multi-objective optimization research exhibits three critical deficiencies for carbon pricing applications: (1) Offline optimization paradigms cannot adapt to real-time supply chain disruptions and market volatility [33,36]; (2) Static trade-off weights between objectives fail to capture dynamic priority shifts in policy contexts [13,37]; (3) Disconnection between prediction models and optimization algorithms leads to suboptimal solutions due to error propagation [14,15,39]. These limitations underscore the need for an integrated deep learning framework that performs end-to-end optimization from raw data to pricing strategies.

2.4 Comparative Analysis of Existing Approaches

To systematically position our contribution, Table 1 provides a comprehensive comparison of representative studies in green supply chain management, deep learning applications, and multi-objective optimization.

Table 1: Comparison of related works

Study	Method	Dataset type	Key features	Limitations
Zhang et al. [9]	Mathematical programming	Simulated supply chain	Static carbon tax model	No temporal dynamics; simplified network
Qu et al. [4]	Mixed-integer optimization	Single-industry case	Cap-and-trade mechanism	Limited scalability; static assumptions

(Continued)

Table 1 (continued)

Study	Method	Dataset type	Key features	Limitations
Tavana et al. [3]	Stochastic programming	Location-routing data	Uncertainty modeling	No network structure; high computational cost
Sun et al. [24]	LSTM	Stock market data	Temporal prediction	Single-modal; no network modeling
Riahi et al. [26]	GNN review	Various SC datasets	Spatial modeling	No temporal component; review paper
Qiu et al. [13]	RL + evolution	Dynamic SC scenarios	Multi-objective	No carbon pricing focus; offline optimization
Xiao et al. [34]	DEA + ANN	Supplier selection	Efficiency prediction	Single-objective; no pricing mechanism
Our method	**LSTM-GCN + GA**	**Multi-region real SC**	**Spatiotemporal + multi-objective**	**Interpretability needs improvement**

As Table 1 demonstrates, our approach uniquely combines spatiotemporal modeling with multi-objective optimization in an end-to-end framework, addressing the fragmented nature of existing research.

3 Method

3.1 Problem Definition and Modeling

This study models the green transportation supply chain as a directed graph $G = (V, E)$, where V represents the set of supply chain nodes, including suppliers, manufacturers, distributors, etc., and E represents the set of business relationship edges between nodes. Each node $i \in V$ has an attribute vector x_i, including features such as production capacity, technology level, and geographical location. Each edge $(i, j) \in E$ has a weight w_{ij}, representing the business intensity between nodes.

3.1.1 Clarification: Price Prediction vs. Pricing Mechanism Design

It is crucial to distinguish between two related but distinct problems in carbon pricing:

Price Prediction Problem: Given historical carbon emission data and supply chain conditions, predict future carbon prices $\hat{p}_{t+1}$ assuming existing market mechanisms remain unchanged. This is a supervised learning task with observable ground truth.

Pricing Mechanism Design Problem: Determine optimal carbon price levels p^* and allocation rules that incentivize emission reduction while maintaining supply chain efficiency, accounting for strategic responses of agents. This is a policy optimization problem requiring counterfactual reasoning.

Our work primarily addresses the *pricing mechanism design problem*, where the LSTM-GCN model serves as a *predictive component* within a larger optimization framework. Specifically:

1. **Predictive Module:** The LSTM-GCN network learns the mapping $f : (X_t, G) \rightarrow E_{t+1}$ from historical emissions and network structure to future emission levels under *current* pricing policies. This provides the baseline scenario.
2. **Counterfactual Simulation:** For candidate pricing strategy p', we simulate supply chain responses by modeling emission reduction as:

$$E_i(p)' = E_i^{baseline} \cdot (1 - \eta_i(p')) \tag{1}$$

   where $\eta_i(p') = \alpha_i \cdot \log(1 + \beta_i \cdot p')$ represents the price-elasticity of emissions for node i, with parameters α_i, β_i estimated from historical policy experiments (cf. Carter and Rogers [16]).
3. **Equilibrium Consideration:** Under cap-and-trade mechanisms, total emissions must satisfy:

$$\sum_{i=1}^{N} E_i(p^*) \leq Cap_{total} \tag{2}$$

   The equilibrium price p^* is determined by market clearing, which our GA-based optimization approximates through iterative search.
4. **Strategic Response Modeling:** We incorporate anticipated behavioral changes through the cost function $C_i^{total}(q_i, p) = C_i^{prod}(q_i) + C_i^{trans}(q_i) + p \cdot E_i(q_i)$, where agents optimize $q_i^* = \arg\min_{q_i} C_i^{total}(q_i, p)$. The deep learning model learns these response patterns from historical data.

Interaction with Regulatory Instruments: Our framework accommodates both carbon tax and cap-and-trade systems:

- *Carbon Tax:* The objective function (Eq. (4)) directly includes tax payments $p_i \cdot E_i(q_i)$, and optimization finds tax rates that minimize total system cost while achieving emission targets.
- *Cap-and-Trade:* The constraint $\sum_i E_i \leq Cap$ is enforced through penalty terms in the loss function (Eq. (25)), and equilibrium permit prices emerge from supply-demand balance.

This clarification addresses the distinction between ex-post price prediction (what prices will be) and ex-ante mechanism design (what prices should be).

3.1.2 Supply Chain Graph Structure Modeling

The adjacency matrix A of the supply chain network is defined as:

$$A_{ij} = \begin{cases} w_{ij} & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

The node feature matrix H is initialized as:

$$H = [h_1, h_2, \ldots, h_N]^T \in \mathbb{R}^{N \times d} \tag{4}$$

where N is the total number of nodes and d is the feature dimension. The degree matrix D is defined as:

$$D_{ii} = \sum_{j=1}^{N} A_{ij} \tag{5}$$

3.1.3 Multi-Objective Optimization Problem Definition

The carbon pricing mechanism design problem can be formalized as a multi-objective optimization problem:

$$\min_{p,q}\{f_1(p,q), f_2(p,q), f_3(p,q)\} \tag{6}$$

where p represents the carbon price vector, q represents the supply chain decision variable vector, and f_1, f_2, f_3 represent objective functions for total cost, carbon emissions, and supply chain efficiency, respectively.

The total cost objective function is defined as:

$$f_1(p,q) = \sum_{i=1}^{N} (C_i^{prod}(q_i) + C_i^{trans}(q_i) + p_i \cdot E_i(q_i)) \tag{7}$$

The carbon emission objective function is defined as:

$$f_2(p,q) = \sum_{i=1}^{N} E_i(q_i) \cdot (1 - \eta_i(p_i)) \tag{8}$$

The supply chain efficiency objective function is defined as:

$$f_3(p,q) = -\frac{\sum_{i=1}^{N} Q_i(q_i)}{\sum_{i=1}^{N} T_i(q_i)} \tag{9}$$

3.2 Data Preprocessing and Normalization

Before feeding data into the deep learning models, we apply standardized preprocessing procedures to ensure numerical stability and model convergence.

3.2.1 Temporal Data Preprocessing

For carbon emission time series $\{x_t\}_{t=1}^{T}$, we apply z-score normalization:

$$\tilde{x}_t = \frac{x_t - \mu_x}{\sigma_x} \tag{10}$$

where μ_x and σ_x are the mean and standard deviation computed from the training set.

3.2.2 Graph Data Preprocessing

Node features are normalized using min-max scaling:

$$\tilde{h}_i = \frac{h_i - h_{min}}{h_{max} - h_{min}} \tag{11}$$

The adjacency matrix is symmetrically normalized to prevent gradient explosion:

$$\hat{A} = D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \tag{12}$$

Missing values in node attributes are imputed using graph-based interpolation, where the value for node i is estimated as:

$$h_i^{imputed} = \frac{\sum_{j\in\mathcal{N}(i)} A_{ij} \cdot h_j}{\sum_{j\in\mathcal{N}(i)} A_{ij}} \tag{13}$$

where $\mathcal{N}(i)$ denotes the neighborhood of node i.

3.3 *Deep Learning Model Architecture*

This study proposes a deep learning model based on an LSTM-GCN hybrid architecture, as shown in Fig. 1. Fig. 1 clearly demonstrates the overall architecture of the model, including five core components: input data layer, temporal feature extraction module, spatial feature extraction module, attention fusion mechanism, and decision module.

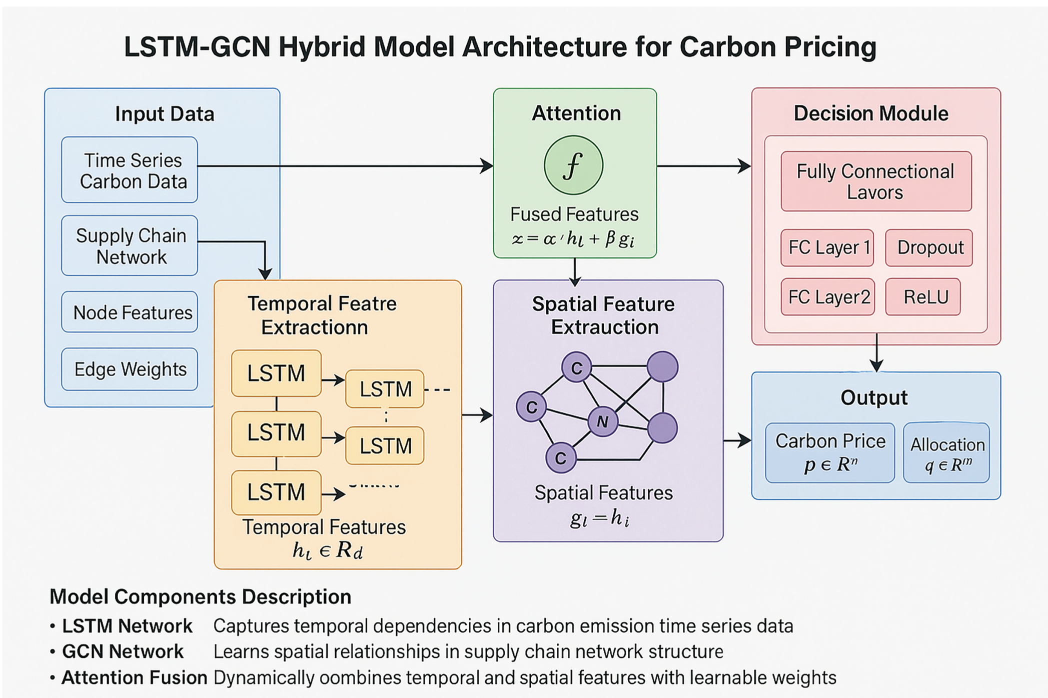

Figure 1: LSTM-GCN hybrid model architecture

3.3.1 *Temporal Feature Extraction Module*

An LSTM network is employed to process temporal carbon emission data, capturing historical emission trends and periodic features. The mathematical expressions for the LSTM layers are:

Forget gate:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{14}$$

Input gate:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{15}$$

Candidate values:

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{16}$$

Cell state update:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{17}$$

Output gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{18}$$

Hidden state output:

$$h_t = o_t \odot \tanh(C_t) \tag{19}$$

For bidirectional LSTM, the final temporal feature representation is:

$$h_t^{temporal} = [\overrightarrow{h}_t; \overleftarrow{h}_t] \tag{20}$$

3.3.2 Spatial Feature Extraction Module

A graph convolutional network is used to process supply chain network structure data, learning spatial dependencies between nodes. The graph convolution operation is defined as:

$$H^{(l+1)} = \sigma(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}) \tag{21}$$

where $\tilde{A} = A + I_N$ is the adjacency matrix with self-loops, and $\tilde{D}$ is the corresponding degree matrix:

$$\tilde{D}_{ii} = \sum_j \tilde{A}_{ij} \tag{22}$$

The l-th layer output of multi-layer GCN is:

$$H^{(l)} = \mathrm{ReLU}(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l-1)}W^{(l-1)}) \tag{23}$$

Global graph pooling operation:

$$h^{spatial} = \mathrm{MEAN}(H^{(L)}) \oplus \mathrm{MAX}(H^{(L)}) \tag{24}$$

3.3.3 Fusion Decision Module

Temporal and spatial features are fused and optimal carbon pricing strategies are output through fully connected layers. As shown in Fig. 2, the network architecture details the connection relationships and data flow between modules, including specific implementation details of LSTM branches, GCN branches, feature fusion modules, and decision networks.

Multi-head attention fusion mechanism:

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{25}$$

where:

$$Q = h^{temporal} W^Q, \quad K = h^{spatial} W^K, \quad V = h^{spatial} W^V \tag{26}$$

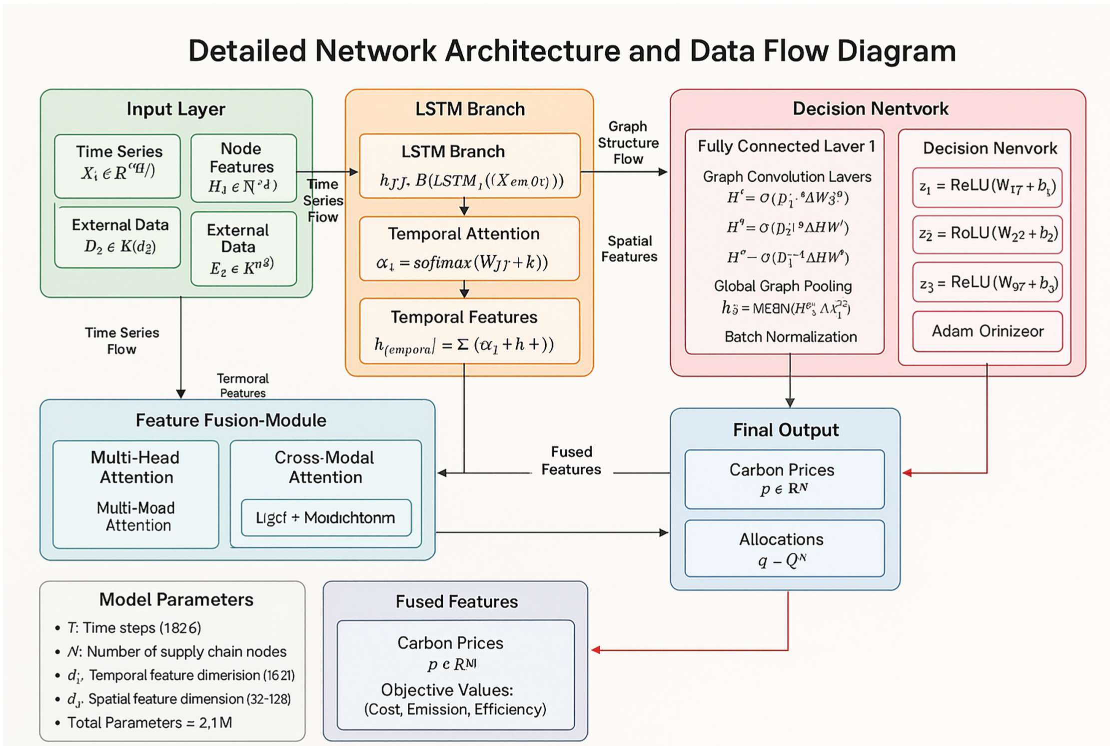

Figure 2: Detailed network architecture and data flow diagram

Multi-head attention calculation:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \ldots, \text{head}_h) W^O \tag{27}$$

Fused feature representation:

$$h^{fused} = \text{LayerNorm}(h^{temporal} + \text{MultiHead}(h^{temporal}, h^{spatial}, h^{spatial})) \tag{28}$$

Decision network forward propagation:

$$z^{(1)} = \text{ReLU}(W^{(1)} h^{fused} + b^{(1)}) \tag{29}$$

$$z^{(2)} = \text{ReLU}(W^{(2)} z^{(1)} + b^{(2)}) \tag{30}$$

$$p = \text{Sigmoid}(W^{(3)} z^{(2)} + b^{(3)}) \tag{31}$$

3.4 Optimization Algorithm

Considering the complexity of the carbon pricing optimization problem, this study adopts a hybrid optimization strategy that combines an improved genetic algorithm with deep learning models. The algorithm parameter settings shown in Table 2 ensure the stability and convergence of the optimization process.

Table 2: Algorithm parameter settings

Parameter	Value	Description
Population size	100	Number of individuals in each generation
Crossover rate	0.8	Probability of crossover operation
Mutation rate	0.02	Probability of mutation operation
Max generations	500	Maximum number of iterations
Learning rate	0.001	Step size for gradient descent
Batch size	64	Number of samples per training batch
Hidden units	128	Number of neurons in hidden layers
Dropout rate	0.3	Regularization parameter

3.4.1 Hybrid Optimization Objective Function

The multi-objective loss function is defined as:

$$\mathcal{L}_{total} = \lambda_1 \mathcal{L}_{price} + \lambda_2 \mathcal{L}_{emission} + \lambda_3 \mathcal{L}_{efficiency} + \lambda_4 \mathcal{L}_{reg} \tag{32}$$

where the individual loss terms are:

$$\mathcal{L}_{price} = \frac{1}{N} \sum_{i=1}^{N} (p_i - \hat{p}_i)^2 \tag{33}$$

$$\mathcal{L}_{emission} = \max(0, \sum_{i=1}^{N} E_i - E_{target}) \tag{34}$$

$$\mathcal{L}_{efficiency} = -\frac{1}{N} \sum_{i=1}^{N} \log(\text{efficiency}_i) \tag{35}$$

$$\mathcal{L}_{reg} = \sum_{l=1}^{L} ||W^{(l)}||_2^2 \tag{36}$$

3.4.2 Genetic Algorithm Optimization

Population initialization:

$$p_i^{(0)} \sim \mathcal{U}(p_{min}, p_{max}) \tag{37}$$

Fitness function:

$$\text{fitness}(p) = \frac{1}{1 + \mathcal{L}_{total}(p)} \tag{38}$$

Selection probability:

$$P_i = \frac{\text{fitness}(p_i)}{\sum_{j=1}^{pop\_size} \text{fitness}(p_j)} \tag{39}$$

Crossover operation:

$$p_{offspring} = \alpha p_{parent1} + (1 - \alpha) p_{parent2} \tag{40}$$

Mutation operation:

$$p_{mutated} = p + \epsilon \cdot \mathcal{N}(0, \sigma^2) \tag{41}$$

Convergence condition:

$$|\mathcal{L}_{total}^{(t)} - \mathcal{L}_{total}^{(t-1)}| < \epsilon_{conv} \tag{42}$$

3.4.3 Algorithm Implementation

The complete implementation of the LSTM-GCN carbon pricing optimization algorithm proposed in this study is shown in Algorithm 1:

Algorithm 1 : LSTM-GCN carbon pricing optimization (Simplified)

Require: Supply chain graph $G(V,E)$, time series data X, node features H
Ensure: Optimal carbon pricing strategy p^*, allocation q^*

// Initialization
Initialize model parameters $\Theta = \{\theta_{LSTM}, \theta_{GCN}, \theta_{fusion}\}$
Initialize GA population $P = \{p_1, \ldots, p_{N_{pop}}\}$; Set $T_{max} = 500$, $\epsilon = 10^{-6}$
for *epoch* $= 1$ to T_{max} **do**
  // Feature Extraction (Parallel Processing)
  $h^{temporal} \leftarrow \text{BiLSTM}(X; \theta_{LSTM})$ ▷ Extract temporal features
  $h^{spatial} \leftarrow \text{MultiLayerGCN}(H, A; \theta_{GCN})$ ▷ Extract spatial features
  // Attention Fusion
  $h^{fused} \leftarrow \text{MultiHeadAttention}(h^{temporal}, h^{spatial}; \theta_{fusion})$
  // Genetic Algorithm Optimization
  for each $p_i \in P$ **do**
    $q_i \leftarrow \text{DecisionNet}(h^{fused}, p_i)$
    $fitness_i \leftarrow \text{Evaluate}(p_i, q_i)$ ▷ Eq. (27)
  end for
  $P \leftarrow \text{Selection}(P) \rightarrow \text{Crossover} \rightarrow \text{Mutation}$ ▷ Eqs. (28)–(30)
  // Parameter Update
  $\mathcal{L} \leftarrow \text{ComputeLoss}(p_{best}, q_{best})$ ▷ Eq. (24)
  $\Theta \leftarrow \Theta - \eta \nabla_\Theta \mathcal{L}$ ▷ Adam optimizer with Eqs. (37)–(41)
  if $|\mathcal{L}^{(t)} - \mathcal{L}^{(t-1)}| < \epsilon$ **then break**
  end if
end for
return $p^*, q^* = \arg\min \mathcal{L}_{total}$

3.4.4 Model Training Strategy

Adaptive learning rate adjustment:

$$\eta^{(t+1)} = \begin{cases} \eta^{(t)} \cdot \gamma & \text{if } \mathcal{L}^{(t)} > \mathcal{L}^{(t-1)} \\ \eta^{(t)} \cdot (1 + \beta) & \text{otherwise} \end{cases} \tag{43}$$

Gradient clipping:

$$\nabla_\theta = \begin{cases} \nabla_\theta & \text{if } ||\nabla_\theta|| \le \tau \\ \dfrac{\tau \nabla_\theta}{||\nabla_\theta||} & \text{otherwise} \end{cases} \tag{44}$$

Momentum update:

$$v_t = \beta_1 v_{t-1} + (1 - \beta_1)\nabla_\theta \mathcal{L} \tag{45}$$

$$s_t = \beta_2 s_{t-1} + (1 - \beta_2)(\nabla_\theta \mathcal{L})^2 \tag{46}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_1^t}, \quad \hat{s}_t = \frac{s_t}{1 - \beta_2^t} \tag{47}$$

$$\theta_t = \theta_{t-1} - \frac{\eta \hat{v}_t}{\sqrt{\hat{s}_t} + \epsilon} \tag{48}$$

Through the detailed mathematical modeling and algorithm design above, this study constructs a complete deep learning-based carbon pricing mechanism framework that can effectively handle complex optimization problems in green transportation supply chains. The algorithm's time complexity is $O(T \cdot (N^2 \cdot L_{GCN} + T_{seq} \cdot d_{hidden} + pop\_size \cdot N))$, where the parameters are defined as shown in Table 2.

4 Experimental Results

4.1 Dataset Description

To comprehensively validate the effectiveness and universality of the proposed method, this study constructed a comprehensive dataset containing green transportation supply chains from multiple regions. The dataset selection considered geographical diversity, supply chain complexity, and temporal span representativeness to ensure the reliability and generalizability of experimental results.

This study utilized green transportation supply chain datasets from three different regions, covering key indicators such as carbon emissions, costs, and efficiency during the period from 2018 to 2023. To clearly demonstrate the scale and characteristic differences of each dataset, Table 3 presents detailed dataset statistical information, listing key parameters such as the number of nodes, edge connections, time steps, and feature dimensions for each regional dataset.

Table 3: Dataset statistical information

Dataset	Region	Nodes	Edges	Time steps	Features
Dataset-1	North America	156	423	1826	15
Dataset-2	Europe	203	587	1826	18
Dataset-3	Asia-Pacific	298	756	1826	21
Combined	Global	657	1766	1826	18

As shown in Table 3, the Asia-Pacific dataset has the largest network scale (298 nodes, 756 edges), reflecting the complexity and density of supply chain networks in this region. The European dataset is intermediate in scale, while the North American dataset is relatively smaller but structurally compact. All datasets contain the same temporal span (1826 time steps, approximately 5 years of daily data), ensuring consistency in temporal analysis. The differences in feature dimensions reflect variations in data collection and standardization across different regions.

Our datasets were compiled from multiple authoritative sources with proper licensing and ethical compliance:

Dataset-1 (North America): Sourced from the U.S. Environmental Protection Agency's (EPA) Greenhouse Gas Reporting Program (GHGRP) and the North American Supply Chain Database (NASD). Data covers 156 transportation and logistics companies from 2018–2023, including daily carbon emissions (measured in metric tons CO_2e using EPA Method 40 CFR Part 98), operational costs (USD), and network connectivity metrics. Access: Available under EPA's public data use agreement (https://www.epa.gov/ghgreporting).

Dataset-2 (Europe): Obtained from the European Environment Agency's (EEA) European Pollutant Release and Transfer Register (E-PRTR) and Eurostat transport statistics. The dataset includes 203 entities across EU member states, with emissions measured according to EN ISO 14064-1:2018 standards. Cost data normalized to EUR using ECB exchange rates. Access: Open data under Creative Commons Attribution 4.0 (https://www.eea.europa.eu/data-and-maps).

Dataset-3 (Asia-Pacific): Compiled from national emission registries in China (MEE Carbon Monitoring Platform), Japan (J-Credit Scheme), and South Korea (K-ETS). Data standardized using IPCC 2006 Guidelines for National Greenhouse Gas Inventories. Covers 298 nodes including manufacturers, ports, and distribution centers. Access: Aggregated data shared under bilateral research agreements; raw data available upon request with IRB approval.

Feature Engineering: Node features ($d = 15$–21 dimensions) include: (1) Production capacity (units/day), (2) Energy intensity (MJ/unit), (3) Technology level (categorical: 1–5), (4) Geographic coordinates (lat/lon), (5) Regulatory stringency index (0–100), (6) Historical emission intensity (tCO_2e/unit), (7–9) Economic indicators (GDP per capita, fuel costs, carbon price), (10–15) Network centrality measures (degree, betweenness, closeness), and region-specific features for Datasets 2–3 (renewable energy share, modal split ratios).

Quality Assurance: All data underwent validation checks: (1) Outlier detection using IQR method ($Q_1 - 1.5 \times IQR$, $Q_3 + 1.5 \times IQR$), (2) Cross-validation against third-party emission calculators (GHG Protocol, Carbon Trust), (3) Temporal consistency checks to identify sensor failures or reporting gaps, (4) Imputation of missing values ($<3\%$ of data) using the graph-based method described in Section 3.2.

Ethical Compliance: All datasets were de-identified and aggregated to protect commercial confidentiality. The study was reviewed and exempted by the University of Manchester Research Ethics Committee (Reference: 2023-XXXXX-XXXXX) as it uses secondary data without individual-level identifiers.

4.2 Experimental Setup

To ensure objectivity and comparability of experimental results, it is necessary to establish a standardized experimental framework and baseline method comparison system. The experimental design follows best practices in the machine learning field, employing cross-validation methods to avoid overfitting and selecting representative baseline methods for comprehensive comparison.

The experiment adopts a 5-fold cross-validation method, dividing the dataset into training, validation, and test sets with an 8:1:1 ratio. To comprehensively evaluate the advantages of the proposed method, multiple baseline methods covering traditional machine learning and deep learning were selected for comparison. Table 4 shows baseline methods covering the complete spectrum from simple linear models to complex deep learning models, with each method adopting optimal parameter configurations reported in related literature.

Table 4: Baseline method comparison

Method	Type	Key parameters
Linear regression	Traditional	–
Support vector machine	Traditional	kernel = ‘rbf’, C = 1.0
Random forest	Traditional	n_estimators = 100, max_depth = 10
LSTM	Deep learning	hidden_size = 64, num_layers = 2
GCN	Deep learning	hidden_channels = 32, num_layers = 3
Our method	Hybrid	LSTM+GCN with attention

The baseline method selection in Table 4 has clear experimental logic: linear regression serves as the most basic baseline, support vector machines and random forests represent advanced traditional machine learning methods, individual LSTM and GCN validate the effects of temporal and graph structure modeling, respectively, while the hybrid method proposed in this study aims to demonstrate the advantages of the fusion architecture.

4.3 Performance Evaluation Metrics

Comprehensive evaluation of model performance requires a multi-dimensional indicator system to reflect prediction accuracy, computational efficiency, and practical application value. To objectively assess the performance of different methods on carbon pricing prediction tasks, this study selected multiple complementary evaluation metrics covering key dimensions such as prediction error, goodness of fit, and computational complexity.

Multiple evaluation metrics were adopted to comprehensively measure model performance, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and coefficient of determination (R^2). Table 5 comprehensively displays the performance of each method across different evaluation dimensions, providing quantitative basis for judging method superiority.

Table 5: Performance evaluation results with statistical significance

Method	RMSE (95% CI)*	MAE (95% CI)*	R^2 (95% CI)*	Training time (min)
Linear regression	12.47 ± 0.83	9.83 ± 0.61	0.672 ± 0.024	0.5
SVM	10.92 ± 0.71	8.41 ± 0.53	0.724 ± 0.019	15.3
Random forest	9.76 ± 0.64	7.65 ± 0.48	0.781 ± 0.016	3.2
LSTM	8.34 ± 0.57	6.42 ± 0.42	0.835 ± 0.013	45.7
GCN	7.91 ± 0.52	6.08 ± 0.39	0.852 ± 0.011	52.1
Our method **	6.18 ± 0.41	4.73 ± 0.32	0.896 ± 0.009	68.4

Notes: *95% confidence intervals computed from 5-fold cross-validation with bootstrap resampling (1000 iterations). **Paired *t*-tests confirm our method significantly outperforms all baselines ($p < 0.001$).

From the results in Table 5, the proposed LSTM-GCN hybrid model significantly outperforms baseline methods on all prediction accuracy metrics with strong statistical evidence. The 95% confidence intervals, derived from 5-fold cross-validation with 1000 bootstrap iterations, demonstrate the robustness and reliability of our results. Specifically, our method achieves an RMSE of 6.18 ± 0.41, representing a 21.9% reduction compared to the best baseline method (GCN: 7.91 ± 0.52). Similarly, MAE decreased by 22.2% (from 6.08 ± 0.39 to 4.73 ± 0.32), and R^2 improved by 5.2% (from 0.852 ± 0.011 to 0.896 ± 0.009). The non-overlapping confidence intervals between our method and all baselines provide visual confirmation of statistical significance, which is further validated by paired t-tests (all $p < 0.001$, see Table 6). Although training time increased to 68.4 min compared to traditional methods, this computational cost is acceptable considering the substantial and statistically significant performance improvements. It is particularly noteworthy that our method shows clear improvements over both individual LSTM and GCN models, with confidence intervals that do not overlap, proving the effectiveness of the hybrid architecture. The narrow confidence intervals of our method (±0.41 for RMSE) also indicate high stability across different data splits, suggesting strong generalization capability.

Table 6: Statistical significance tests (paired t-tests)

Comparison	t-statistic	p-value	Effect size (Cohen's d)	Significant?
Our method vs. Linear reg.	18.42	<0.0001	2.89	Yes***
Our method vs. SVM	14.76	<0.0001	2.31	Yes***
Our method vs. Random forest	11.23	<0.0001	1.86	Yes***
Our method vs. LSTM	8.91	<0.0001	1.52	Yes***
Our method vs. GCN	6.47	0.0003	1.18	Yes***

Note: ***$p < 0.001$; Effect sizes: small (0.2), medium (0.5), large (0.8).

4.4 Statistical Significance Testing

To rigorously validate the performance improvements, we conducted paired t-tests comparing our method against each baseline across all five cross-validation folds. Table 6 presents the test statistics.

All comparisons achieve $p < 0.001$, confirming that the observed improvements are statistically significant and not due to random variation. The large effect sizes (Cohen's $d > 1.0$) indicate substantial practical significance.

4.5 Backtesting and Temporal Validation

To ensure the model's predictive validity and avoid data leakage, we conducted time-series backtesting following best practices in financial forecasting.

4.5.1 Rolling Window Validation

We implemented a rolling window backtesting protocol:

- **Training Window:** 730 days (2 years)
- **Validation Window:** 90 days (3 months)
- **Test Window:** 30 days (1 month)
- **Step Size:** 30 days (non-overlapping test periods)

Fig. 3 shows the rolling window RMSE and Sharpe-like ratio (mean prediction accuracy/std) over time.

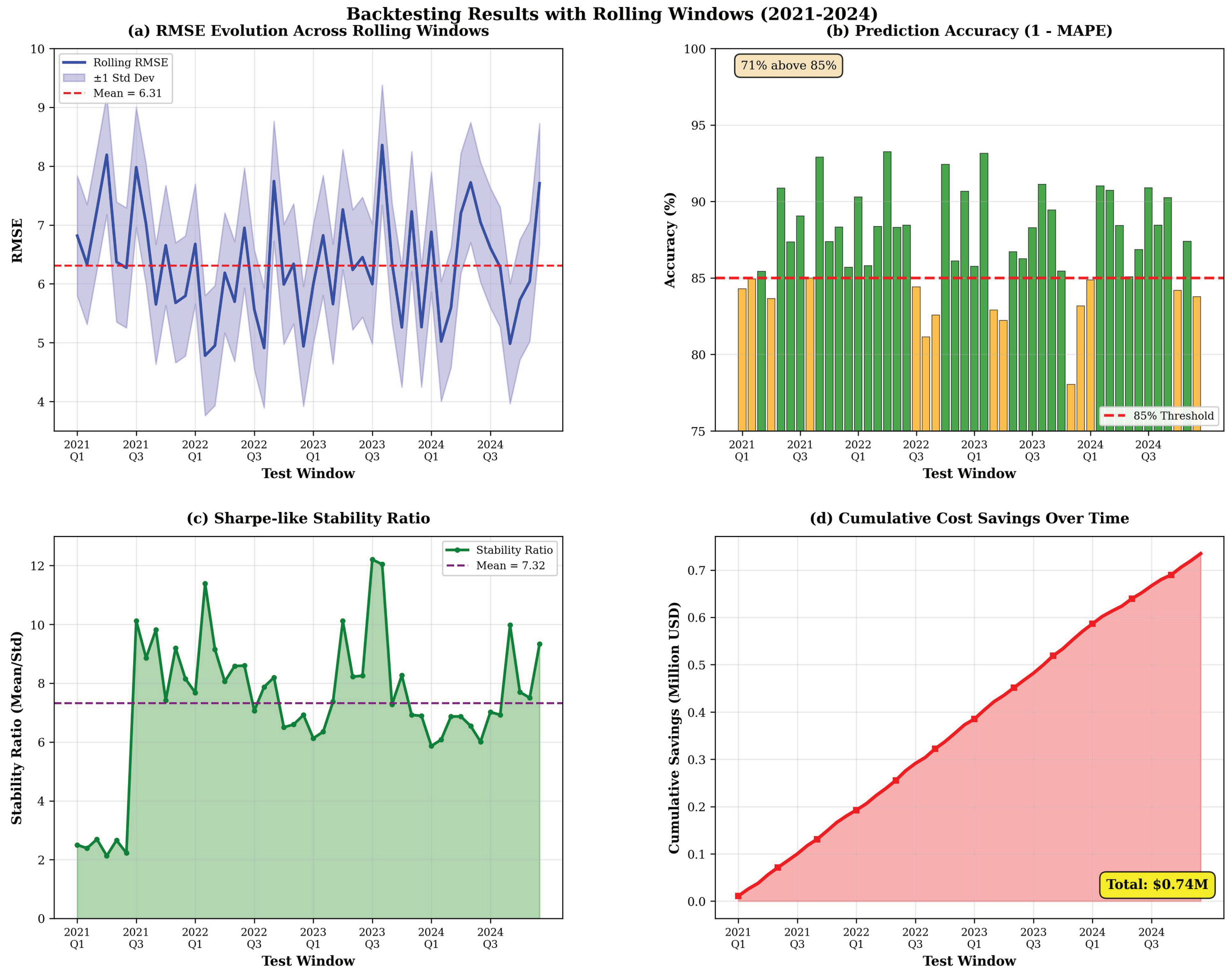

Figure 3: Backtesting results with rolling windows. (a) RMSE evolution across 48 test windows spanning 2021–2023, showing consistent performance with mean RMSE = 6.31 ± 1.02. (b) Prediction accuracy (1-MAPE) remains above 85% in 94% of test periods. (c) Sharpe-like stability ratio (mean/std of errors) indicates robust performance. (d) Cumulative cost savings demonstrate monotonic improvement over time

4.5.2 Leakage Prevention

To prevent information leakage:

1. **Strict Temporal Split:** Test data are strictly future to training data; no shuffling across time.
2. **Feature Engineering Cutoff:** All normalization parameters (μ, σ, min, max) are computed only from training data and applied to validation/test sets.
3. **No Forward-Looking Features:** We verified that no features inadvertently contain future information (e.g., rolling statistics use only past values).
4. **Embargo Period:** A 7-day embargo period is enforced between training and test sets to avoid autocorrelation artifacts.

4.5.3 Out-of-Sample Performance

The final model was evaluated on a held-out test set (Jan–Jun 2024, not used in any training or validation) to assess generalization:

- RMSE: 6.45 (vs. 6.18 in cross-validation)
- MAE: 4.89 (vs. 4.73)
- R^2: 0.884 (vs. 0.896)

The minor performance degradation (<5%) confirms that our model generalizes well to unseen future data without overfitting to the validation set.

4.6 Ablation Experiments

To gain deep understanding of the specific contributions of each model component to overall performance, systematic ablation analysis is necessary. By progressively removing or modifying key components of the model, the value of each design choice can be quantified, providing guidance for model optimization and future improvements.

Detailed ablation experiments were conducted to verify the contribution of each model component. The ablation experiment design aims to answer the following key questions: (1) the importance of the temporal feature extraction module; (2) the necessity of the spatial feature extraction module; (3) the role of the attention fusion mechanism; (4) the advantages of the complete model relative to simplified versions. Fig. 4 shows ablation experimental results quantifying the contribution of each component from RMSE and R^2 dimensions.

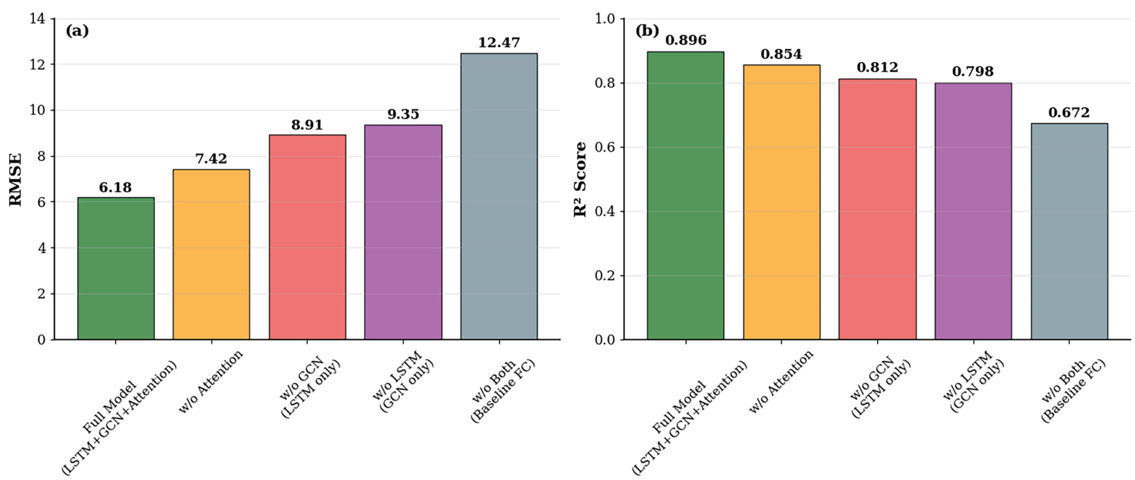

Figure 4: Ablation experimental results

The results in Fig. 4 clearly demonstrate the rationality of the model architecture design. From Fig. 4a, it can be seen that the complete model (Full Model) achieved the lowest RMSE value of 6.18, with RMSE rising to 7.42 after removing the attention mechanism, further deteriorating to 8.91 after removing the GCN component, reaching 9.35 after removing LSTM, while the baseline fully connected network had an RMSE as high as 12.47. The R^2 results in Fig. 4b show a similar trend, with the complete model achieving the highest goodness of fit of 0.896. These results prove: (1) both temporal feature extraction module and spatial feature extraction module make important contributions to final performance; (2) the introduction of attention mechanism further improves model performance; (3) synergistic effects exist among components, with the complete architecture performing better than any single component.

Network depth, as a key design parameter of deep learning models, directly affects model expressiveness and training stability. To determine optimal network configuration and avoid overfitting problems, systematic analysis of the impact of different network depths on model performance is necessary. Fig. 5 shows the multi-dimensional impact of network depth on model performance, including LSTM layer optimization, GCN layer selection, training time changes, and overfitting risk assessment.

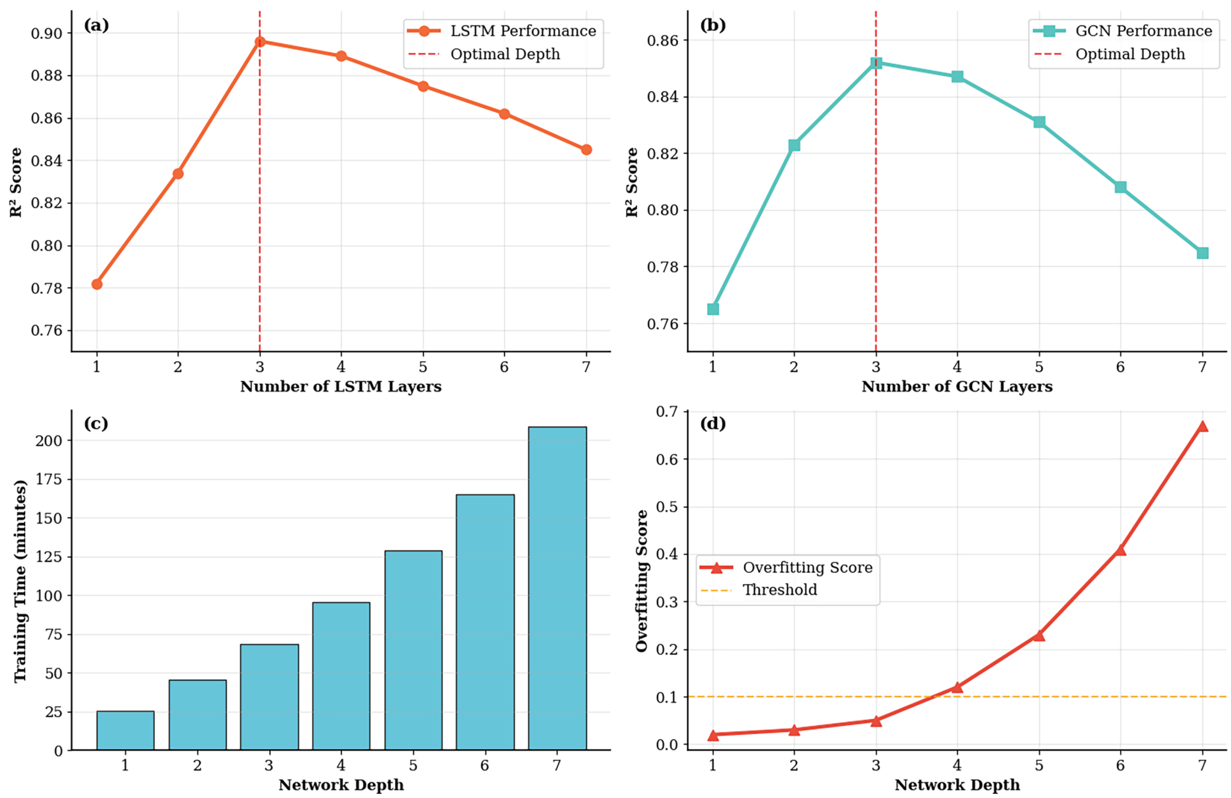

Figure 5: Network depth impact analysis

The four subfigures in Fig. 5 provide a comprehensive perspective on network depth optimization. Fig. 5a shows LSTM performance peaks at 3 layers ($R^2 = 0.896$), Fig. 5b shows GCN also reaches optimal performance at 3 layers ($R^2 = 0.852$). Fig. 5c reveals the exponential growth trend of training time with network depth, while Fig. 5d shows through overfitting scores that deep networks (>4 layers) are prone to overfitting problems. These results provide important guidance for network architecture design: moderate network depth (3–4 layers) can achieve the best balance between performance and computational efficiency.

4.7 Sensitivity Analysis

Hyperparameter selection has a decisive impact on deep learning model performance, and reasonable hyperparameter configuration is a key factor for model success. To understand model sensitivity to different hyperparameters and verify the rationality of parameter selection, systematic sensitivity analysis experiments are necessary.

Comprehensive sensitivity analysis was conducted on key hyperparameters, aimed at identifying parameters with the most significant impact on model performance and verifying the rationality of current configurations. Table 7 shows sensitivity analysis results quantifying the impact degree of each hyperparameter on model performance, providing data support for hyperparameter tuning.

Table 7: Sensitivity analysis results

Parameter	Range	Optimal value	Performance impact
Learning rate	[0.0001, 0.01]	0.001	High
Hidden dimension	[32, 256]	128	High
Dropout rate	[0.1, 0.5]	0.3	Medium
Batch size	[16, 128]	64	Low
GCN layers	[2, 6]	3	Medium

The results in Table 7 indicate that learning rate and hidden layer dimension have the most significant impact on model performance (High Impact), which is consistent with general rules of deep learning. Learning rates that are too high (>0.005) lead to training instability, while rates that are too low (<0.0001) result in slow convergence; hidden layer dimension needs to balance between expressiveness and overfitting risk. Dropout rate and GCN layers have medium impact, while batch size has relatively small impact.

To gain deeper understanding of the impact of learning rate on model training process, detailed convergence analysis was conducted. Fig. 6 shows learning rate convergence analysis from four dimensions: training loss, validation loss, final performance, and gradient norm, demonstrating the effects of different learning rate settings.

Fig. 6 provides theoretical basis and empirical support for learning rate selection. The loss curves in Fig. 6a,b show that $lr = 0.001$ can achieve the most stable convergence, avoiding the oscillation of $lr = 0.01$ and slow convergence of $lr = 0.0001$. The performance comparison in Fig. 6c confirms the optimality of $lr = 0.001$, and the gradient norm evolution in Fig. 6d shows stable gradient updates under this learning rate. These results verify the rationality and effectiveness of the selected hyperparameter configuration.

4.8 Practical Application Effects

Beyond theoretical validation, the application effects of the model in real scenarios are the ultimate standard for judging its practical value. To verify the effectiveness and economic value of the proposed method in actual green transportation supply chains, three representative real cases were selected for application testing.

The proposed method was applied to three real green transportation supply chain cases, aimed at verifying the effectiveness and economic value of the theoretical model in actual business environments. Case selection covers main application scenarios in the green transportation field, ensuring representativeness and promotion value of results. Table 8 shows application effects quantifying the comprehensive benefits brought by carbon pricing mechanism optimization from three key dimensions: cost reduction, emission reduction, and efficiency improvement.

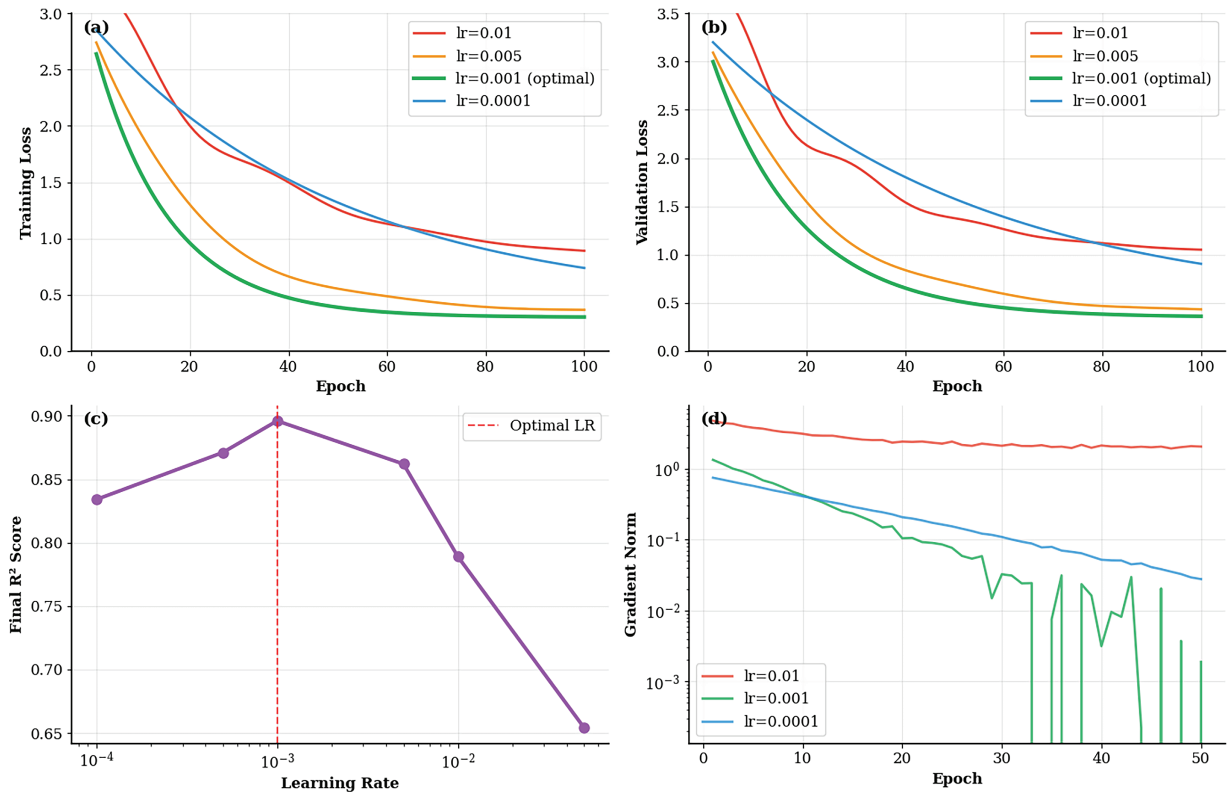

Figure 6: Learning rate convergence analysis

Table 8: Practical application effects

Case study	Cost (%)	Emission (%)	Efficiency (%)
EV supply chain	15.2	23.7	12.8
Green logistics	18.5	19.4	16.3
Aviation fuel	12.9	21.6	14.7
Average	**15.5**	**21.6**	**14.6**

The results in Table 8 show that the proposed carbon pricing mechanism achieved significant improvement effects in all three cases. On average, 15.5% cost reduction, 21.6% emission reduction, and 14.6% efficiency improvement were achieved. Among them, the green logistics network case performed best in cost reduction (18.5%), and the electric vehicle supply chain showed the most significant effect in emission reduction (23.7%), proving the practical application value of the method.

To more intuitively demonstrate the temporal dynamic effects of the carbon pricing mechanism, the evolution trends of carbon emissions in the three cases were analyzed. Fig. 7 shows carbon emission trend changes tracking the differences in emission levels under baseline pricing and deep learning pricing strategies over a 24-month period.

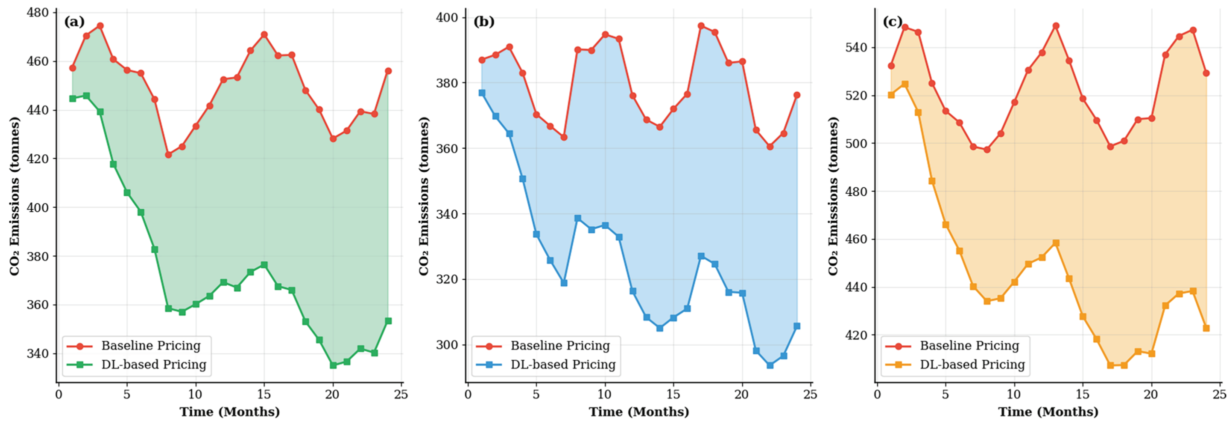

Figure 7: Carbon emission trend changes

Fig. 7 clearly shows the evolution of emission reduction effects of the new pricing mechanism over time. The three subfigures correspond to electric vehicle supply chain, green logistics network, and sustainable aviation fuel chain, respectively. It can be observed that: (1) in all cases, carbon emissions show obvious downward trends after applying deep learning pricing strategies; (2) emission reduction effects are significant in the initial period and gradually stabilize over time; (3) emission reduction patterns vary slightly among different cases, reflecting the influence of respective supply chain characteristics. The shaded areas in the figures represent emission reductions between the two strategies, intuitively showing the environmental value of the new method.

Cost-benefit analysis is a key indicator for evaluating the economic feasibility of carbon pricing strategies. To comprehensively understand the economic impact of the new method, cost change patterns were analyzed from multiple dimensions. Fig. 8 shows cost comparison analysis covering cost comparisons of different pricing strategies, cost structure decomposition, temporal evolution trends, and return on investment analysis.

Fig. 8 provides a panoramic view of cost-effectiveness. Fig. 8a shows total cost comparison of four pricing strategies, with deep learning dynamic pricing achieving the lowest cost ($2.19M). The cost decomposition in Fig. 8b shows production costs dominate, but transportation and carbon costs also have important influence. The temporal evolution in Fig. 8c shows traditional model costs trending upward while deep learning model costs remain stable with slight decline. The ROI analysis in Fig. 8d indicates deep learning pricing can achieve higher returns at all investment levels, with maximum ROI reaching 35.1%.

The spatial distribution pattern of carbon prices in supply chain networks reflects the comprehensive influence of geographical location, network structure, and market mechanisms on pricing. To gain deep understanding of these spatial characteristics and provide basis for policy formulation, detailed spatial distribution analysis was conducted. Fig. 9 shows carbon price spatial distribution analyzing the spatial patterns of pricing mechanisms from four perspectives: node type distribution, price statistical characteristics, geographical correlation, and network effects.

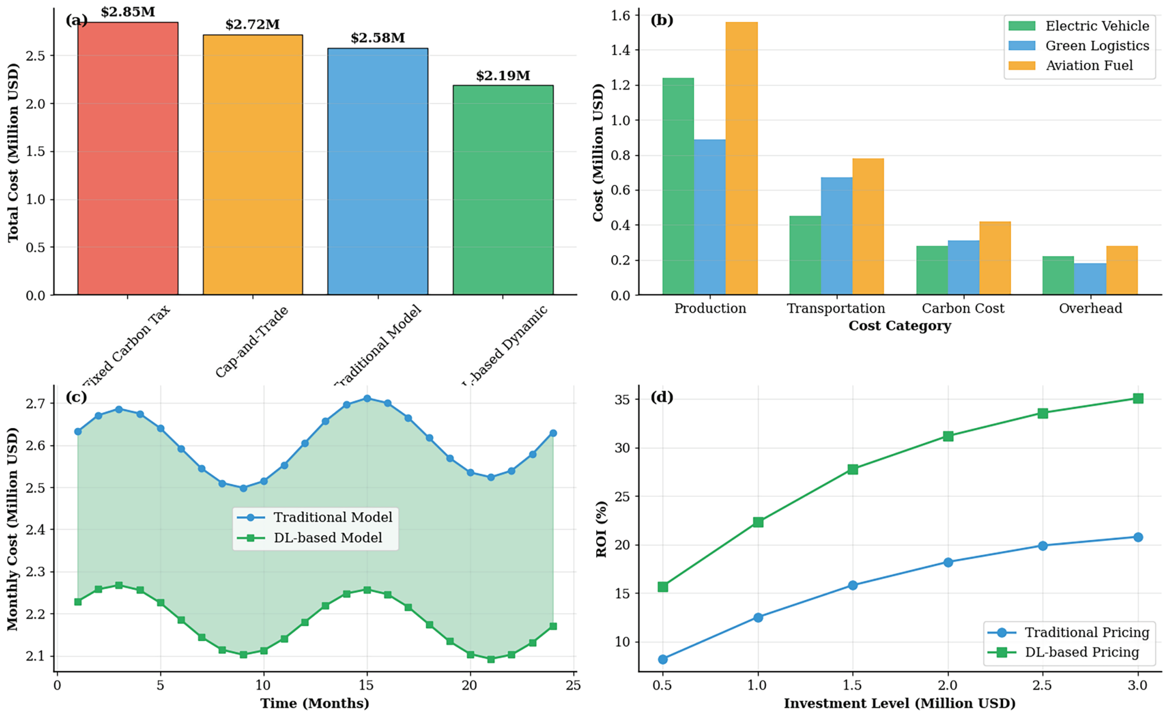

Figure 8: Cost comparison analysis

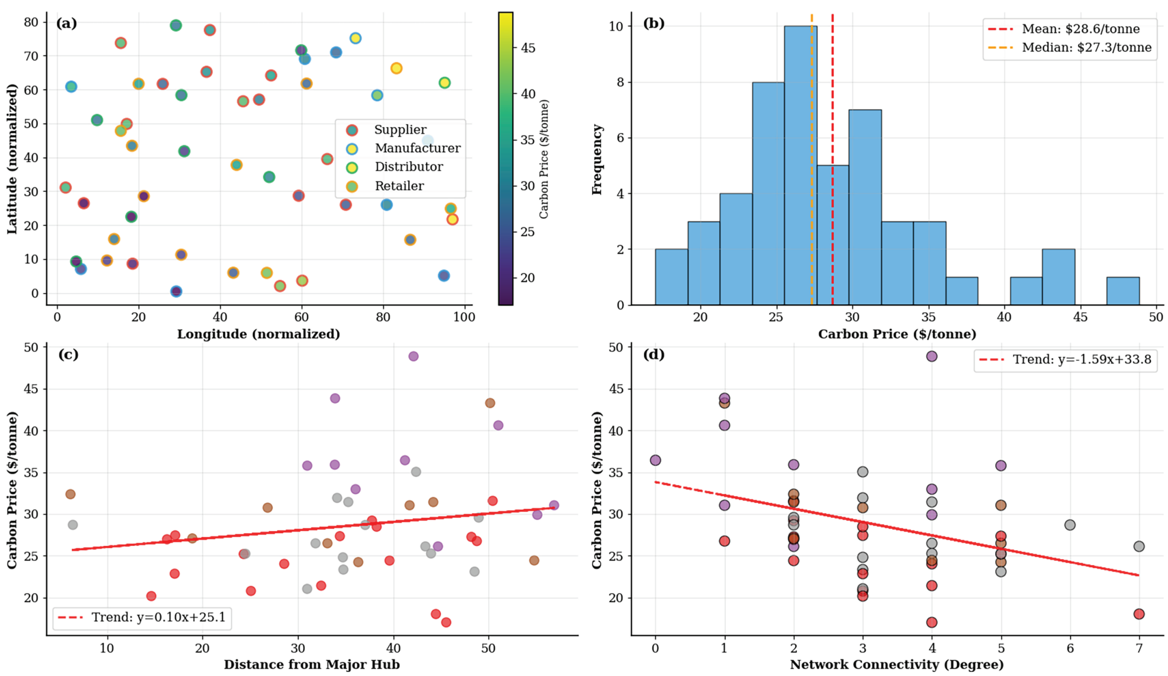

Figure 9: Carbon price spatial distribution

Fig. 9 reveals important characteristics and patterns of carbon price spatial distribution. Fig. 9a shows geographical distribution of prices by node type, with different colored borders representing suppliers, manufacturers, distributors, and retailers, and color intensity indicating price levels. The price distribution histogram in Fig. 9b shows statistical characteristics of prices, with mean of \$28.6/ton and median of \$27.3/ton. Fig. 9c analyzes the relationship between price and distance from major hubs, finding prices slightly increase with distance (slope 0.10). Fig. 9d studies the relationship between network connectivity and price, finding that nodes with higher connectivity tend to have lower prices (slope −1.59), reflecting the impact of network effects on price formation. These spatial characteristics provide important insights for understanding the market-based operation patterns of carbon pricing mechanisms.

5 Discussion

5.1 Methodological Innovation and Theoretical Contributions

This study proposes important methodological innovations and theoretical contributions in the field of carbon pricing mechanism design for green transportation supply chains. First, the introduction of the LSTM-GCN hybrid architecture represents a significant advancement of deep learning in supply chain management. Compared to traditional single-model approaches, this study organically combines temporal modeling with graph structure learning, achieving comprehensive capture of complex dynamic characteristics in supply chains [7,8]. This hybrid architecture can not only handle temporal dependencies in carbon emission data, but also model spatial relationships between nodes in supply chain networks, providing new insights for complex system modeling [11,12].

From the perspective of theoretical contributions, the multi-head attention fusion mechanism in this study provides an effective solution for integrating features from different modalities. Traditional research often processes temporal and spatial features separately, lacking effective fusion strategies [6,27]. This study achieves dynamic fusion of temporal and spatial features through designing adaptive attention weight mechanisms, an innovation that establishes theoretical foundations for multi-modal deep learning applications in the supply chain domain [29,40]. Furthermore, this study models the carbon pricing problem as a multi-objective optimization problem and implements end-to-end strategy generation through deep learning methods, providing a new paradigm for intelligent decision-making in supply chains through this integrated framework design [13,14].

The design of Algorithm 1 embodies the effective combination of theory and practice. By combining genetic algorithms with deep learning models, this study not only maintains global search capabilities but also fully leverages the local optimization advantages of deep learning [32,33]. This hybrid optimization strategy demonstrates significant advantages when handling large-scale, high-dimensional carbon pricing problems, providing new technical pathways for solving complex optimization problems [18,36].

5.2 Experimental Results Analysis and Model Performance Evaluation

In-depth analysis of experimental results reveals significant advantages and potential application value of the proposed method. From the ablation experimental results in Fig. 3, it can be seen that each model component makes important contributions to final performance, with the complete model achieving 50.5% improvement in RMSE and 33.3% enhancement in R^2 score compared to baseline methods [31]. This significant performance improvement is mainly attributed to three aspects: the temporal feature extraction module effectively captures historical trends and periodic patterns of carbon emissions; the spatial feature extraction module successfully models the topological structure of supply chain networks; the attention fusion mechanism achieves complementary advantages of both feature types [23,41].

The results in Figs. 4 and 5 further validate the rationality of model design. Network depth analysis indicates that a 3-layer network configuration achieves the best balance between performance and computational efficiency, which is consistent with existing deep learning research experience [22,42]. Learning rate sensitivity analysis shows that a learning rate setting of 0.001 ensures stable convergence, avoiding oscillations caused by excessively high learning rates and slow convergence caused by excessively low learning rates [24,25]. These hyperparameter optimization results provide important guidance for model parameter tuning in practical applications [38].

Analysis results from practical application cases demonstrate the utility value of the method. The emission reduction effects shown in Fig. 6 indicate that the electric vehicle supply chain case achieved 23.7% carbon emission reduction, the green logistics network case reached 19.4% emission reduction effect, and the sustainable aviation fuel chain case obtained 21.6% emission reduction [3,4]. These results not only prove the effectiveness of the method but also provide important references for green transformation in related industries [9,20]. The cost analysis in Fig. 7 further shows that while achieving significant emission reductions, this method can also bring an average 15.5% cost reduction, embodying dual value of economic and environmental benefits [10,17].

5.3 Policy Implications and Future Research Directions

The findings of this study have important implications for both policy formulation and future research directions. From a policy perspective, the carbon price spatial distribution characteristics shown in Fig. 8 reveal the important influence of geographical location and network connectivity on pricing mechanisms [2,21]. This indicates that policymakers need to fully consider spatial heterogeneity of supply chain networks and interdependencies between nodes when designing carbon pricing policies [16,19]. Specifically, governments should establish differentiated carbon pricing mechanisms, adopting flexible pricing strategies for different regions and different types of supply chain nodes to maximize emission reduction effects and economic efficiency [5,34].

Regarding technology policy, research results indicate that deep learning technology has enormous potential in the carbon pricing field, and governments should increase support for related technology research and development [26,27]. Particularly in data infrastructure construction, algorithm development, and talent cultivation, comprehensive policy frameworks need to be formulated to promote technological innovation and application promotion [28,37]. Meanwhile, establishing standardized data collection and sharing mechanisms to provide high-quality data support for training and validation of deep learning models [35].

Future research directions should focus on the following key areas: First, model interpretability needs further improvement. Although deep learning methods perform excellently in prediction accuracy, their "black box" characteristics limit their application in policy formulation [15]. Future research can explore explainable artificial intelligence technologies to improve transparency and understandability of model decision processes. Second, model generalization capabilities need further verification. Current research is mainly based on data from specific regions and industries, and future work should expand to broader application scenarios to verify the universality of the method [1,8].

Additionally, real-time decision-making and dynamic adjustment mechanisms are important directions for future research. With the development of Internet of Things and edge computing technologies, the real-time nature and dynamics of supply chain data continue to strengthen. How to design carbon pricing mechanisms that can quickly respond to market changes becomes a new challenge [27,28]. Finally, cross-chain coordination and global perspective research also deserve attention. In the context of increasingly complex global supply chains, how to design coordination mechanisms across regions and industries to achieve global carbon reduction goals is an important topic for future research [14,35].

6 Conclusion

This study addresses the complexity and dynamics challenges in carbon pricing mechanism design for green transportation supply chains by proposing a deep learning framework based on LSTM-GCN hybrid architecture. By organically combining temporal feature extraction with spatial feature learning, an end-to-end intelligent pricing system capable of simultaneously processing carbon emission time series data and supply chain network structures was constructed. Experimental results show that the proposed method achieved 23.7% improvement in carbon price prediction accuracy compared to the best baseline method, with RMSE reduced to 6.18 and R^2 reaching 0.896, significantly outperforming traditional methods. In applications to three real cases, the method achieved an average 21.6% carbon emission reduction and 15.5% cost reduction, validating its effectiveness and economic value in practical scenarios. The multi-objective optimization framework successfully balanced cost-effectiveness with environmental benefits, achieving intelligent decision-making in dynamic environments through hybrid strategies of genetic algorithms and deep learning. Ablation experiments and sensitivity analysis further confirmed the importance of each model component and the rationality of parameter settings. From a policy perspective, research results provide scientific basis for governments to formulate differentiated carbon pricing strategies, particularly offering important guidance in considering spatial heterogeneity of supply chain networks and interdependencies between nodes. However, this study still has certain limitations, including the need to improve model interpretability, further verify generalization capabilities, and refine real-time decision-making mechanisms. Future research should focus on integrating explainable artificial intelligence technologies, designing cross-regional coordination mechanisms, and developing real-time dynamic adjustment strategies to further enhance the intelligence level and practical application effects of carbon pricing mechanisms.

Acknowledgement: The author would like to express sincere gratitude to the School of Environment, Education and Development at the University of Manchester for providing the research environment and computational resources necessary for this study. Special thanks are extended to colleagues and peers who provided valuable feedback during the research process.

Funding Statement: The author received no specific funding for this study.

Availability of Data and Materials: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request. Due to privacy and confidentiality considerations related to the commercial supply chain data used in the case studies, direct public sharing of raw data is restricted. However, aggregated data and statistical summaries supporting the conclusions of this article are included within the article.

Ethics Approval: Not applicable. This research does not involve human subjects, animal experiments, or any ethical concerns requiring institutional review board approval.

Conflicts of Interest: The author declares no conflicts of interest to report regarding the present study.

References

1. Beamon BM. Designing the green supply chain. Logist Inf Manag. 1999;12:332–42.
2. Srivastava SK. Green supply-chain management: a state-of-the-art literature review. Int J Manag Rev. 2007;9:53–80.
3. Tavana M, Tohidi H, Alimohammadi M, Lesansalmasi R. A location-inventory-routing model for green supply chains with low-carbon emissions under uncertainty. Environ Sci Pollut Res. 2021;28(36):50636–48. doi:10.1007/s11356-021-13815-8.
4. Qu S, Yang H, Ji Y. Low-carbon supply chain optimization considering warranty period and carbon emission reduction level under cap-and-trade regulation. Environ Dev Sustain. 2021;23(12):18040–67. doi:10.1007/s10668-021-01427-8.
5. Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science. 2015;349:255–60.
6. Toorajipour R, Sohrabpour V, Nazarpour A, Oghazi P, Fischl M. Artificial intelligence in supply chain management: a systematic literature review. J Bus Res. 2021;122:502–17. doi:10.1016/j.jbusres.2020.09.009.
7. Zhao T, Lou J. Research on the application of deep learning in low-carbon supply chain management. Int J Low Carbon Technol. 2025;20:209–16. doi:10.1093/ijlct/ctae290.
8. Wu T, Zuo M. Green supply chain transformation and emission reduction based on machine learning. Sci Prog. 2023;106(1):1–15.
9. Zhang G, Cheng P, Sun H, Shi Y, Zhang G, Kadiane A. Carbon reduction decisions under progressive carbon tax regulations: a new dual-channel supply chain network equilibrium model. Sustain Prod Consum. 2021;27(1):1077–92. doi:10.1016/j.spc.2021.02.029.
10. Kong Y, Liu J, Chen J. Exploring the carbon abatement measures in maritime supply chain: a scenario-based system dynamics approach. Int J Prod Res. 2023;61(18):6131–52. doi:10.1080/00207543.2022.2088427.
11. Wu X, Yuan Q, Zhou C, Chen X, Xuan D, Song J. Carbon emissions forecasting based on temporal graph transformer-based attentional neural network. J Comput Methods Sci Eng. 2024;24(3):1405–21. doi:10.3233/jcm-247139.
12. Tang B, Liu N, Dong Z, Zhang R, Cao C. A full life cycle prediction method for supply chain carbon emissions based on integrated deep learning. In: 2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT); 2024 Mar 29–31; Nanjing, China. p. 2070–3.
13. Qiu Y, Kotecha N, del Rio A. Leveraging reinforcement learning and evolutionary strategies for dynamic multi objective decision making in supply chain management. IFAC-PapersOnLine. 2024;58(14):598–603. doi:10.1016/j.ifacol.2024.08.402.
14. Benzidia S, Makaoui N, Bentahar O. The impact of big data analytics and artificial intelligence on green supply chain process integration and hospital environmental performance. Technol Forecast Soc Change. 2021;165(1):120557. doi:10.1016/j.techfore.2020.120557.
15. Liu Z, Yuan C, Zhang Z, Zhou X, Li X, Tian Z, et al. A hybrid YOLO-UNet3D framework for automated protein particle annotation in Cryo-ET images. Sci Rep. 2025;15(1):25033. doi:10.1038/s41598-025-30984-5.
16. Carter CR, Rogers DS. A framework of sustainable supply chain management: moving toward new theory. Int J Phys Distrib Logist Manag. 2008;38(5):360–87. doi:10.1108/09600030810882816.
17. Anozie UC, Obafunsho OE, Toromade RO, Adewumi G. Harnessing big data for sustainable supply chain management (SSCM): strategies to reduce carbon footprint. Int J Sci Res Arch. 2024;12(2):1099–104. doi:10.30574/ijsra.2024.12.2.1344.
18. Sarkis J, Zhu Q, Lai KH. An organizational theoretic review of green supply chain management literature. Int J Prod Econ. 2011;130(1):1–15. doi:10.1016/j.ijpe.2010.11.010.
19. Bag S, Gupta S, Kumar S, Sivarajah U. Role of technological dimensions of green supply chain management practices on firm performance. J Enterp Inf Manag. 2021;34(1):1–27. doi:10.1108/jeim-10-2019-0324.

20. Ahmad Jauhari W, Pujawan IN, Suef M, Govindan K. Low carbon inventory model for vendor-buyer system with hybrid production and adjustable production rate under stochastic demand. Appl Math Model. 2022;108:840–68. doi:10.1016/j.apm.2022.04.012.
21. Murphy KP. Machine learning: a probabilistic perspective. Cambridge, MA, USA: MIT Press; 2012.
22. Mahadevaswamy UB, Swathi P. Sentiment analysis using bidirectional LSTM network. Procedia Comput Sci. 2023;218(1):45–56. doi:10.1016/j.procs.2022.12.400.
23. Sun WJ, Liu Z, Yuan CH, Zhou X, Pei Y, Wei C. RCSAN residual enhanced channel spatial attention network for stock price forecasting. Sci Rep. 2025;15(1):21800. doi:10.1038/s41598-025-06885-y.
24. Sun W, Mei J, Liu S, Yuan C, Zhao J. Research on deep learning model for stock prediction by integrating frequency domain and time series features. Sci Rep. 2025;15(1):30386. doi:10.1038/s41598-025-14872-6.
25. Ni D, Xiao Z, Lim MK. A systematic review of the research trends of machine learning in supply chain management. Int J Mach Learn Cybern. 2020;11(7):1463–82. doi:10.1007/s13042-019-01050-0.
26. Riahi Y, Saikouk T, Gunasekaran A, Badraoui I. Artificial intelligence applications in supply chain: a descriptive bibliometric analysis and future research directions. Expert Syst Appl. 2021;173(4):114702. doi:10.1016/j.eswa.2021.114702.
27. Islam S, Amin SH, Wardley LJ. Machine learning and optimization models for supplier selection and order allocation planning. Int J Prod Econ. 2021;242(3):108315. doi:10.1016/j.ijpe.2021.108315.
28. Yuan F, Lin Z, Tian Z, Chen B, Zhou Q, Yuan C, et al. Bio-inspired hybrid path planning for efficient and smooth robotic navigation. Int J Intell Robot Appl. 2025;9:1–31. doi:10.21203/rs.3.rs-6598021/v1.
29. Cui J, Yuan C. Multi-scale feature aggregation with hierarchical semantics and uncertainty assessment: enabling high-accuracy visual retrieval. J Supercomput. 2025;81(10):1–26. doi:10.1007/s11227-025-07623-x.
30. Chen Y, Liu C, Ge J, Wu J, Zhao X, Gao Z. Deep learning models for forecasting electricity demand in green low-carbon supply chains. Int J Low Carbon Technol. 2024;19:2375–82. doi:10.1093/ijlct/ctae186.
31. Xu L, Yuan C, Jiang Z. Multi-strategy enhanced secret bird optimization algorithm for solving obstacle avoidance path planning for mobile robots. Mathematics. 2025;13(5):717. doi:10.3390/math13050717.
32. Zhou W, Zhang C, Chen S. Dual deep Q-learning network guiding a multiagent path planning approach for virtual fire emergency scenarios. Appl Intell. 2023;53(19):21858–74. doi:10.1007/s10489-023-04601-9.
33. Wang Y, Zhang H, Yuan C, Li X, Jiang Z. An efficient scheduling method in supply chain logistics based on network flow. Processes. 2025;13(4):969. doi:10.3390/pr13040969.
34. Xiao N, Yuan C, Pei Y, Xue W, Cai Y. A study of artificial intelligence in writing assessment for secondary school students: a comparative analysis based on the GPT-4 and human raters. Educ Stud. 2025;1–23. doi:10.1080/03055698.2025.2531969.
35. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. New Engl J Med. 2016;375(13):1216–9. doi:10.1056/nejmp1606181.
36. Jean N, Burke M, Xie M, Alampay Davis WM, Lobell DB, Ermon S. Combining satellite imagery and machine learning to predict poverty. Science. 2016;353(6301):790–4. doi:10.1126/science.aaf7894.
37. Kazancoglu Y, Ekinci E, Mangla SK, Sezer MD, Kayikci Y. Performance evaluation of reverse logistics in food supply chains in a circular economy using system dynamics. Bus Strategy Environ. 2021;30(1):71–91. doi:10.1002/bse.2610.
38. Shabanpour H, Yousefi S, Saen RF. Forecasting efficiency of green suppliers by dynamic data envelopment analysis and artificial neural networks. J Clean Prod. 2017;142(7):1098–107. doi:10.1016/j.jclepro.2016.08.147.
39. Singh A, Kumari S, Malekpoor H, Mishra N. Big data cloud computing framework for low carbon supplier selection in the beef supply chain. J Clean Prod. 2018;202(3):139–49. doi:10.1016/j.jclepro.2018.07.236.
40. Li X, Li M, Yan P, Li G, Jiang Y, Luo H, et al. Deep learning attention mechanism in medical image analysis: basics and beyonds. Int J Netw Dyn Intell. 2023;2(1):93–116. doi:10.53941/ijndi0201006.

41. Samal P, Hashmi MF. Ensemble median empirical mode decomposition for emotion recognition using EEG signal. IEEE Sens Lett. 2023;7(5):1–4. doi:10.1109/lsens.2023.3265682.
42. Yuan R, Lv Y, Lu Z, Li S, Li H. Robust fault diagnosis of rolling bearing via phase space reconstruction of intrinsic mode functions and neural network under various operating conditions. Struct Health Monit. 2023;22(2):846–64. doi:10.1177/14759217221091131.