



## Green supply chain finance credit risk assessment model integrating deep learning - based on the perspective of ESG and multi-source climate data fusion

Yuxin Shen<sup>1,2,\*</sup>

<sup>1</sup> Bank of Beijing Post-Doctoral Research Station, Bank of Beijing, Beijing, 100033, China

<sup>2</sup> School of Finance, Nankai University, Tianjin, 300350, China

**SUMMARY:** *With the increasingly severe global climate change problem, green supply chain finance, as an important financial tool to promote sustainable development, is facing unprecedented challenges. In this context, traditional credit risk assessment methods are unable to effectively address the complex environmental, social, and governance factors in green supply chain finance, as well as the multidimensional impact of climate risk. This study combines financial data, ESG scores, climate risk data, and supply chain network information of enterprises, and innovatively integrates graph neural networks and time series analysis models to construct a new credit risk assessment framework. The research results indicate that the ESG LimateGNN model exhibits superiority in multiple evaluation metrics, particularly in terms of discriminative ability, classification accuracy, and ranking ability, which significantly surpass traditional models and cutting-edge models in the field.*

**KEYWORDS:** *Green Supply Chain Finance, Credit Risk Assessment, Deep Learning, ESG Factors, Climate Risk*

## 1 Introduction

Against the backdrop of the restructuring of the global climate governance system, green finance is undergoing a critical transformation from conceptual advocacy to practical deepening (Bowman & Minas, 2019). As an innovative paradigm for the deep integration of industrial capital and environmental governance, Green Supply Chain Finance (GSCF) constructs a green oriented mechanism for resource allocation through the collaborative innovation of financial instruments and industrial chains (Lai et al., 2022). This mechanism not only reshapes the logic of value creation in the industrial chain, but also drives enterprises to improve resource utilization efficiency through economic incentives, systematically reducing the carbon footprint of the entire chain (Judijanto et al., 2024).

In the supply chain finance ecosystem, the financing difficulties of small and medium-sized enterprises have long constrained the coordinated development of industries (Kinyua et al., 2024). Due to the scarcity of collateral and information asymmetry between banks and enterprises, these companies often face issues such as high financing costs and limited credit limits. Green supply chain finance breaks through the limitations of traditional financial indicators by introducing an environmental, social, and governance (ESG) comprehensive evaluation system, incorporating corporate environmental compliance, social responsibility performance, and governance structure effectiveness into credit evaluation dimensions

\*shenyuxin0625@126.com

<https://doi.org/10.65102/is20261173>

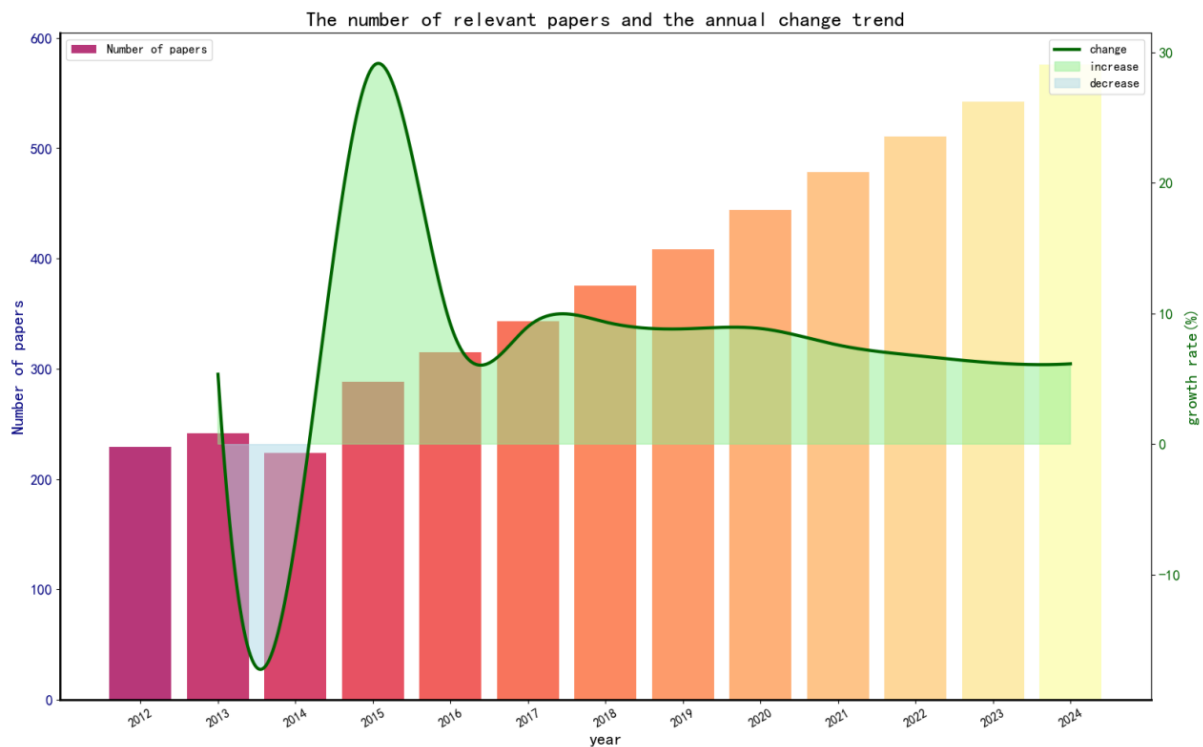
(Agarwal et al., 2024). At the same time, quantitative analysis of corporate climate adaptation capabilities has further expanded the boundaries of credit risk assessment, providing differentiated financing support for upstream and downstream enterprises in the industrial chain, thereby accelerating the green transformation process of the entire supply chain network (Gao et al., 2024).

However, while achieving synergy between environmental and economic benefits, green supply chain finance also faces profound changes in risk structure. Compared to traditional supply chain finance, its risk dimension exhibits significant multidimensional superposition characteristics (Huang et al., 2024). On the basis of traditional financial credit risk, the uncertainty of ESG factors has formed a new type of risk source, covering environmental compliance risk, social reputation risk, and governance structure imbalance risk (Liu & Song, 2025). In addition, the dual nature of climate risks further exacerbates the complexity of risks. Climate transition risks stem from the pressure of asset valuation brought about by policy and regulatory adjustments, technological iterations, and changes in market preferences (Xu et al., 2024); Climate physical risks manifest as direct impacts of extreme weather events on supply chain operations, as well as asset availability declines and logistics network disruptions caused by long-term climate change (Bracco et al., 2025). These risk factors form complex coupling relationships through node associations in the supply chain network, which may trigger a chain transmission of systemic risks (Bertrand et al., 2024).

The traditional credit risk assessment framework has exposed significant theoretical and technical limitations in addressing the complex risks of green supply chain finance (Edunjobi & Odegide, 2024). The credit scoring card model based on financial statement analysis and methods such as logistic regression (Hurlin et al., 2024) have certain applicability in scenarios with sufficient structured data, but they are difficult to meet the risk assessment needs of green finance. This limitation is manifested in four aspects: firstly, insufficient ability to process unstructured data. Traditional models lack effective feature extraction and semantic analysis capabilities in the face of heterogeneous data such as text information, environmental monitoring image data, and supply chain on-site verification records in corporate ESG reports, making it difficult to uncover hidden risk signals (Zhang, 2024). Secondly, the ability to capture dynamic characteristics of climate data is weak. The spatiotemporal evolution characteristics of the climate system require assessment models to have long-term trend prediction and emergency warning functions, but early machine learning models had technical bottlenecks in integrating static financial data and dynamic climate data (Baig et al., 2024). Thirdly, there is a lack of characterization of the transmission mechanism of network risks. Traditional evaluation methods often focus on isolated entities as analysis units, ignoring the multidimensional diffusion paths of risks through transaction relationships, logistics connections, and information transmission in the supply chain network, resulting in underestimation of the risk contagion effect (Khedr, 2024). Fourthly, the modeling of complex risk relationships is ineffective. ESG factors and climate risks exhibit cross modal and multi-scale nonlinear coupling characteristics in supply chain finance scenarios. Traditional linear or shallow models cannot accurately depict the time-varying interaction between risk factors, which can easily lead to risk misjudgment (Billio et al., 2024).

The traditional credit risk assessment paradigm has exposed the dual shortcomings of theoretical explanatory power and technical implementation power in dealing with the complexity of ESG and climate risks. The breakthrough of deep learning technology provides new methodological support for risk assessment innovation. As shown in Figure 1, research papers related to deep learning and multi-source data utilization are emerging like mushrooms after rain. In the field of unstructured data processing, natural language processing technology based on Transformer architecture can achieve semantic and sentiment analysis of ESG reports,

policy texts, and public opinion data, transforming textual information into quantitative risk indicators (Lee et al., 2025); Time series analysis models such as Long Short Term Memory Networks (LSTM) (Zhao et al., 2024) and Time Series Convolutional Networks (TCN) capture the dynamic coupling relationship between climate variables and financial data to construct predictive models for risk evolution (Ghaemi, 2025); The application of Graph Neural Networks (GNNs) enables digital modeling of the topology structure of supply chain networks, enabling visual analysis of risk transmission pathways and identification of key nodes (Wu et al., 2025); The multimodal deep learning framework further integrates structured financial data, unstructured text information, image data, and spatiotemporal climate data, and achieves comprehensive evaluation of complex risk systems through joint representation learning. These technological breakthroughs not only revolutionize the paradigm of risk feature extraction, but also improve the timeliness and accuracy of risk prediction through end-to-end model training.



Data Sources: By utilizing databases such as the Web of Science Core Collection, EBSCOhost, and ProQuest, and using core English search terms like "Green Supply Chain Finance," "ESG Risk Assessment," "Climate-Related Risk," "Deep Learning in Credit Risk," and "Sustainable Finance," it is possible to obtain cutting-edge international research findings. For Chinese research, databases including CNKI, Wanfang Data, and VIP Information are more suitable, with Chinese keywords such as "Lv Se Gong Ying Lian Jin Rong" (Green Supply Chain Finance), "ESG Feng Xian" (ESG Risk), "Qi Hou Feng Xian" (Climate Risk), "Shen Du Xue Xi" (Deep Learning), "Xin Yong Feng Xian Ping Gu" (Credit Risk Assessment), "Shuang Tan Zhan Lue" (Dual Carbon Strategy), and "Jin Rong Feng Xian Guan Li" (Financial Risk Management) for retrieval. Subsequently, the relevant papers should be collected and systematically reviewed.

Figure 1: Statistics of the number of related research papers

In the macro context of accelerating the global "dual carbon" strategy, the risk environment faced by green supply chain finance is undergoing profound changes. Building a credit risk

assessment system based on deep learning and multi-source data fusion is not only an inevitable choice to address ESG and climate risk challenges, but also a key technical support for achieving green allocation of financial resources. This research direction not only helps to improve the risk theoretical framework of green finance, but also provides operational technical paths for financial institutions to optimize risk control strategies, balance sustainable development goals and commercial interests. It has important theoretical value and practical significance for promoting the high-quality development of green supply chain finance.

Based on this, this article mainly studies the credit risk assessment problem in green supply chain finance and proposes an innovative assessment model based on deep learning and multi-source data fusion - ESG ClimateGNN. This model comprehensively enhances the risk prediction capability in green supply chain finance by integrating corporate financial data, ESG scores, climate risk data, and supply chain network information. It is particularly innovative in considering the impact of climate change on corporate credit risk.

## 2 Research design

### (I) Problem Definition and Variables

#### 1. Goal definition

The research focuses on the medium and long-term credit risk assessment of core enterprises in the context of green supply chain finance. The core evaluation objective is to construct a model that predicts the probability of a core enterprise defaulting on its debt during a fixed observation period in the future, based on the fusion of historical financial performance, deeply processed ESG characteristics, quantitative analysis of physical and transitional climate risk exposure, and multidimensional supply chain network information. The "default" events defined here strictly follow financial risk management practices, including but not limited to: substantial delinquency/default of bond principal or interest, significant default of bank loans, or occurrence of debt restructuring events that comply with international/domestic accounting standards (Zhang et al., 2019).

#### 2. Variable system

The data includes enterprise level characteristic data, supply chain relationship data, and multi-source climate data. All data are accurately correlated through a unified social credit code. The specific data and their explanations are shown in Table 1.

*Table 1: Data sources and their explanations*

Data Category	Data Source	Specific Indicators/Features	Description
Enterprise Level	Financial databases like Wind and CSMAR	Quantitative indicators such as debt-to-asset ratio, current ratio, operating profit margin, and revenue growth rate	Reflects short-term solvency, long-term financial health, and operational efficiency of enterprises
	Professional ESG rating agencies (e.g., China Securities, Huazheng, Sustainalytics); unstructured texts such as corporate annual social responsibility reports, environmental penalty announcements, and labor-related news	Comprehensive ESG scores, E/S/G single-dimension scores; key ESG event features extracted from unstructured texts using natural language processing (NLP) technology	Integrates data from professional rating agencies with information from unstructured texts
	Industry characteristics and geographic coordinates of enterprises; historical disaster frequency/intensity data; policy sensitivity scores and industry carbon emission intensity benchmarks	Physical climate risks (comprehensive flood/drought risk indices) and transition climate risks (quantitative scores)	Calculates risk exposure based on industry attributes, geographic coordinates, and climate/policy data
Supply Chain Relationship Data	Listed companies' annual reports, Refinitiv Supply Chain management platform, and enterprise credit databases	Node enterprise associations characterized by directed graph structures; edge attributes include supplier/customer role types, annual transaction amount as a percentage of main business revenue, cooperation duration, and physical distance between both parties	Depicts supply chain associations through graph structures and transaction attribute features
Multi-source Climate Data	Historical observation records of meteorological stations, satellite remote sensing data products, and ERA5-Land high-resolution grid meteorological data	Region/location-specific risk indicators converted via spatial interpolation or grid association	Integrates multi-source meteorological data to generate refined risk indicators
	"Double Carbon" policy regulations issued by governments at all levels, national and industry carbon emission statistical bulletins, carbon trading market quota price fluctuation sequences, and regional green patent databases	Transition risk feature sets reflecting the intensity of low-carbon technology development and regulatory policy impacts	Constructs risk feature systems by integrating policy, market, and technical data

### 3. Label definition

This study will use whether the core enterprises in the supply chain will experience substantial credit default events within a fixed observation period (24 months) in the future as

training and validation labels for the credit risk assessment model. The label variable is a binary discrete variable. If the core enterprise experiences at least one financial instrument default event as defined in International Financial Reporting Standard 9 (IFRS 9) or Chinese Accounting Standard 22- Recognition and Measurement of Financial Instruments within 24 months after the start of the observation window, the label is assigned a value of 1; On the contrary, if the enterprise has not experienced any of the above-mentioned default situations throughout the entire observation period, the label will be assigned a value of 0. The determination of substantive default events is strictly based on public market data and information disclosed by financial regulatory agencies, covering the following three situations: ① The principal or interest payment of public market bonds issued by enterprises is overdue and the overdue period exceeds 90 natural days; ② The loan obtained by the enterprise from commercial banks or other licensed financial institutions has a significant breach of contract terms, and ultimately enters into debt restructuring or non-performing asset disposal procedures; ③ The enterprise enters bankruptcy liquidation or reorganization proceedings under the court's ruling due to the loss of its ability to continue operating or the depletion of cash flow. The data of this tag comes from the national enterprise credit information publicity system, the announcement of default issuers of the Association of Inter bank Market Dealers, the bond default information disclosure platform of Shanghai and Shenzhen Stock Exchanges, and the basic database of the People's Bank of China credit information system, and is cross checked through the Wind financial terminal enterprise default record module.

## (II) Data preprocessing

### 1. Implement differentiated cleaning for heterogeneous data sources

Structured financial data adopts an outlier detection mechanism based on industry quantiles, setting business logic thresholds to directly remove invalid records; Implementing chain equation multivariate imputation with continuous variable missing, constructing a prediction model based on data from same industry enterprises during the same period; If the discrete ESG rating is missing, fill in the median of the industry size grouping. The missing areas of spatial climate grid data are reconstructed using Kriging spatial interpolation algorithm. When there is a lack of supply chain transaction volume, a conservative estimate is made based on the historical cooperation intensity between the buyer and seller and the industry average ratio, and binary identification variables are set to record manual filling behavior.

### 2. Unstructured Data Feature Engineering

After parsing the ESG report text into PDF, the pre trained language model RoBERTa wwm is first used to generate contextualized word vectors and capture phrase level semantic associations; Secondly, a hierarchical attention mechanism is adopted to identify key ESG entities, and the FinBERT financial sentiment analysis model is used to quantify the polarity of text risk simultaneously (Qin, 2023); The final output is a concatenation feature of 768 Uyghur semantic vectors and emotion scores. The core elements of climate policy text were extracted by the BiLSTM CRF entity recognition system and mapped to the industry policy sensitivity matrix to form a structured policy shock indicator (Qiao et al., 2024).

### 3. Quantification of multi-source climate risks

The extraction of physical risk features relies on high-resolution raster data to calculate three core indicators: ① probability of extreme event occurrence; ② Composite Disaster Risk Index; ③ Geospatial Vulnerability Score. The transformation risk signal processing adopts a dynamic quantification framework, and the green technology index is synthesized by the proportion of clean energy patents in the province to which the enterprise belongs and the policy implementation strength (measured by cosine similarity of text vectors), ultimately forming a spatiotemporal aligned transformation pressure feature vector. All climate indicators were standardized by min max and included as model inputs.

(III) Multi source data fusion strategy

As shown in Figure 2, this article systematically integrates enterprise node characteristics, supply chain topology relationships, and climate risk signals through a hierarchical feature fusion architecture.

The first step is node level feature concatenation, which combines the standardized traditional financial indicators of the enterprise ontology, ESG feature vectors including semantic vectors and emotional scores, and climate risk exposure indicators including physical and transformation risk scores along the feature dimensions to generate a joint feature vector  $v_i \in R^d$ , Realize early fusion of static heterogeneous data.

The second step is dynamic integration of neighbor aggregation, relying on the supply chain directed graph structure (adjacency matrix A), and performing neighbor information aggregation through a gated graph convolutional network. As shown in Equation 1:

$$h_i^{(1)} = \text{GCN}(v_i, \sum_{j \in \mathcal{N}(i)} \alpha_{ij} \cdot W \cdot [v_j \parallel e_{ij}]) \tag{1}$$

Among them,  $e_{ij}$  is the edge feature, and  $\alpha_{ij}$  is the attention weight based on the role type (supplier/customer). Output  $h_i$  while encoding the enterprise's own attributes and its supply chain risk propagation path.

The third step is cross modal temporal fusion, which inputs  $h_i$  and climate temporal signals into the dual stream Transformer module. The enterprise flow uses  $h_i$  as the initial hidden state to model the evolution of enterprise characteristics; Climate flow independently encodes climate variables and dynamically captures the driving effects of climate factors on corporate risk through a cross attention mechanism.

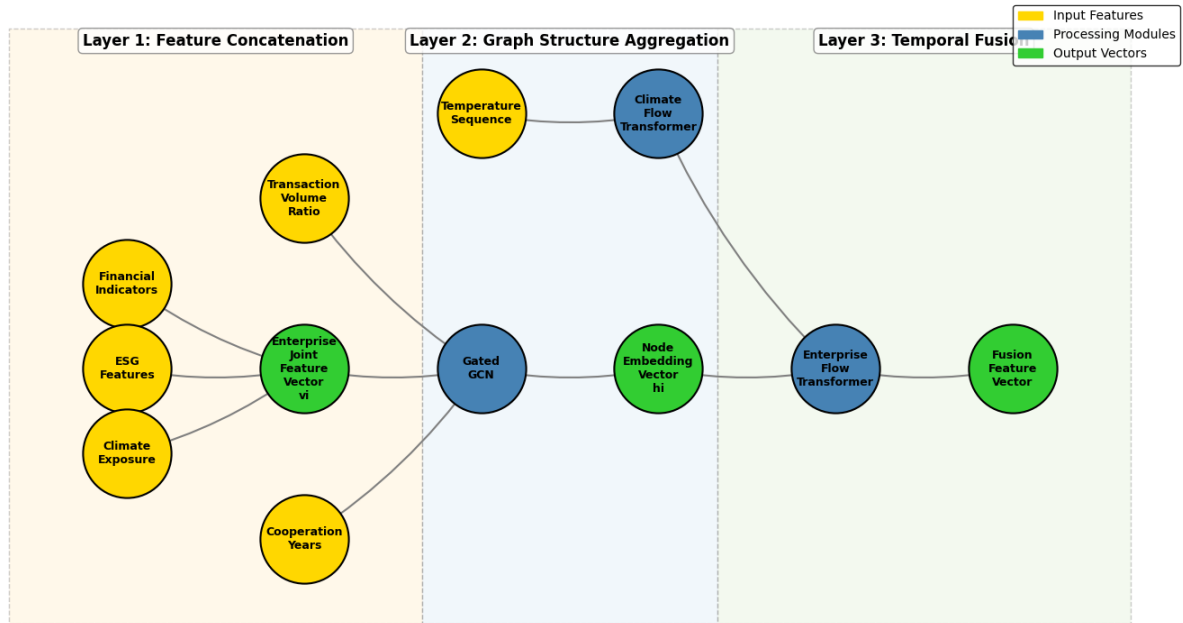


Figure 2: Schematic diagram of data fusion

(IV) Design of Deep Learning Models

This article proposes the ESG ClimateGNN model, which aims to achieve dynamic coupling of "enterprise supply chain climate" risks, and its core operation relies on the orderly promotion of four level modules.

- ① The graph neural network module focuses on modeling the risk transmission path of

supply chain topology. In practical operation, it takes node features  $v_i$  and edge attributes  $e_{ij}$  as input content, undergoes internal operations and processing logic, and finally outputs spatial risk embedding  $h_i^{\text{GNN}}$  to characterize the risk correlation information at the supply chain network structure level. The specific calculation process is as follows:

Firstly, calculate the transaction perception attention coefficient:

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(a^T[W_v v_i \| W_v v_j \| W_e e_{ij}]))}{\sum_{k \in \mathcal{N}(i)} \exp(\cdot)} \quad (2)$$

$W_v \in \mathbb{R}^{128 \times d}$  and  $W_e \in \mathbb{R}^{128 \times 3}$  are learnable weight matrices;  $a \in \mathbb{R}^{384}$  is an attention parameter vector;  $\alpha_{ij}$  quantifies the risk transmission intensity from node  $j$  to  $i$ , modulated by transaction amount and role type.

Next, substitute the result of equation (2) into equation (1) for gated neighbor information aggregation.

② The temporal processing module is designed to capture the time-varying effects of endogenous characteristics and climate shocks within the enterprise. By analyzing and modeling relevant time-series data, this module can output dynamic risk characterization  $h_i^{\text{TS}}$ , accurately capturing risk factors and their impacts that change over time. The specific stages are as follows:

Firstly, the characteristic flow of the enterprise:

Input the enterprise joint feature slice  $V_i \in \mathbb{R}^{12 \times 172}$  with a time window of 12 months, and extract long-term dependencies through bidirectional LSTM layering:

$$\vec{h}_t = \text{LSTM}(v_t, \vec{h}_{t-1}); \quad \overleftarrow{h}_t = \text{LSTM}(v_t, \overleftarrow{h}_{t+1}) \quad (3)$$

At each time step  $t$ , concatenate the forward/reverse hidden states:

$$H_{\text{firm}}^t = [\vec{h}_t \| \overleftarrow{h}_t] \in \mathbb{R}^{256} \quad (4)$$

Output the enterprise dynamic coding tensor  $H_{\text{firm}} \in \mathbb{R}^{12 \times 256}$ , with row vectors encoding the enterprise risk status for the current month and column vectors representing the feature evolution path.

Secondly, climate shock currents:

Input spatially aligned climate grid data  $X_{\text{grid}} \in \mathbb{R}^{12 \times 12}$ .

Introduce industry seasonal modulation layer:

$$\widetilde{x}_k = x_k + \text{Wind} \cdot \text{seas}(k) \quad (5)$$

Among them,  $\text{seas}(k) \in \mathbb{R}^8$  is the monthly sensitivity vector of the industry to which the enterprise belongs, and  $W_{\text{ind}} \in \mathbb{R}^{12 \times 8}$  is the learnable projection matrix. The modulated data is encoded by 2 layers of Transformer:

$$H_{\text{climate}} = \text{Transformer}(\widetilde{X}; \text{head} = 4, d_k = 64) \in \mathbb{R}^{12 \times 256} \quad (6)$$

Establishing a time-varying correlation between enterprise status and climate shocks through attention mechanisms, ultimately generating a fused temporal risk vector:

$$h_i^{\text{TS}} = \text{MaxPool}(A_{\text{cross}} H_{\text{climate}}) \in \mathbb{R}^{256} \quad (7)$$

Among them, the matrix element  $A_{cross}[m,n]$  quantifies the contribution strength of climate events in the  $n$ th month to enterprise risk in the  $m$ th month.

③ The feature fusion module is responsible for cross modal alignment of graph structures and temporal features. It integrates the spatial risk embedding output from the graph neural network module with the dynamic risk representation obtained from the temporal processing module, and generates a joint risk signal  $z_i$  to effectively integrate different modal risk information. The specific stages are as follows:

Gate control vector generation mechanism:

$$g = \sigma W_g \cdot [h_i^{GNN} \parallel h_i^{TS}] + b_g \in [0,1]^{384} \quad (8)$$

Concatenate the 128 - dimensional graph embedding  $h_i^{GNN}$  (representing supply chain topological risks) and the 256 - dimensional temporal embedding  $h_i^{TS}$  (representing climate dynamic shocks) into a 384 - dimensional vector.

Then, calculate the fusion weights using the learnable parameter  $W_g \in \mathbb{R}^{384 \times 384}$  and the bias  $b_g$ . The  $\sigma$  function ensures that the weight  $g$  is strictly restricted to the interval  $[0,1]$ , forming a “risk contribution quantifier”.

Bidirectional feature fusion calculation:

$$z_i = g \odot U_{GNN} h_i^{GNN} + (1 - g) \odot U_{TS} h_i^{TS} \in \mathbb{R}^{384} \quad (9)$$

Ultimately obtaining a joint risk signal  $z_i$ .

④ The predictive output module is responsible for converting the previously generated joint risk signal  $z_i$  into a default probability scalar through a specific mapping relationship  $\hat{y}_i \in [0,1]$ .

(V) Performance indicators

1. Discriminant ability indicator group

The area under the curve (AUC) is used to measure the overall ability of the model to distinguish between default and non default samples. It is calculated based on the ROC curve drawn from the true positive rate and false positive rate, with a value range of 0.5 to 1. In practice, it has practical value when it is higher than 0.8.

The Kolmogorov Smirnov (KS) statistical quantification model quantifies the maximum separation distance of risk scores on the cumulative distribution of positive and negative samples. The calculation method is to take the absolute maximum difference between the cumulative distribution functions of good customer (non default) and bad customer (default) scores. In practice, the threshold should be  $\geq 0.4$ .

F1 Score is the harmonic average of precision (Precision=true positive/predicted positive) and recall (Recall=true positive/actual positive), reflecting the robust performance of the model in class imbalance scenarios. The calculation formula is  $2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$ .

2. Classification accuracy indicator group

Accuracy refers to the proportion of correctly predicted samples in the total sample, calculated by the formula:  $(\text{true positive} + \text{true negative}) / \text{total number of samples}$ . Due to the sparse default sample, other indicators need to be combined when interpreting.

Precision represents the proportion of actual defaults predicted as default samples, reflecting the ability to control the cost of risk misjudgment.

Recall is the proportion correctly identified in the actual default sample, reflecting the completeness of risk coverage.

3. Sorting ability indicator group

The difference in group default rate is the ratio of the actual default rate between the highest

risk group and the lowest risk group after grouping according to the tenth percentile of the model score. It is required to reach a difference of  $\geq 5$  times before it has commercial value.

The Information Value (IV) value calculates the predictive information of default status based on the risk score, evaluates the feature discrimination, and is expressed as  $\sum$  (proportion of good customers - proportion of bad customers)  $\times$  WOE value (evidence weight). When the benchmark value is greater than 0.5, it indicates significant predictive power.

#### 4. Stability index group

The rolling window AUC volatility, which calculates the standard deviation of AUC values for each period during the rolling prediction process, is required to be controlled within  $\leq 0.03$ .

Sample perturbation sensitivity, using Bootstrap resampling 100 times to calculate the coefficient of variation of F1 Score, with a threshold of  $\leq 15\%$ .

This article mainly uses Python 3.9 as the core statistical analysis and modeling tool.

### 3 Empirical Analysis

#### (I) Sample selection

This study focuses on the green supply chain entities among Chinese A-share listed companies in Shanghai and Shenzhen from 2015 to 2023. In terms of industry correlation, based on the "Green Industry Guidance Catalog", six high carbon emission reduction pressure industries including manufacturing, energy, and logistics were selected. To clarify the industry affiliation of enterprises, the mandatory standard is that the proportion of main business income of enterprises is  $\geq 60\%$ . Only enterprises that meet this condition will be included in the sample pool to ensure that the research sample has significant industry characteristics and common emission reduction pressures.

For the determination of green supply chain labeling, enterprises are required to meet at least the following two conditions. Firstly, companies need to release independent ESG reports and disclose carbon emission data within the scope of three carbon emissions. At the same time, verification should be conducted based on the CDP supply chain questionnaire to ensure the reliability of carbon emission data and the standardization of disclosure; Secondly, enterprises need to join national level green supply chain pilot projects and be registered with the Ministry of Industry and Information Technology to prove their official recognition and support in green supply chain construction; Thirdly, the main products of the enterprise need to pass ISO 14040 lifecycle certification, reflecting their green and environmental attributes at the product level. Through these conditions, select enterprises that truly possess green supply chain characteristics.

In terms of data integrity, sample companies are required to disclose complete financial data, supply chain relationship maps, and climate risk response plans for three consecutive years. Complete financial data is used to evaluate the economic condition and risk tolerance of enterprises, supply chain relationship diagrams help build supply chain network structures, and climate risk response plans reflect the enterprise's understanding and management capabilities of climate risks. After strict screening, it finally covered 487 core enterprises and 1932 upstream and downstream enterprises, forming a representative sample set.

In terms of time window design, a dynamic panel dataset is constructed with a semi annual observation unit, aiming to capture the dynamic characteristics of changes in enterprises and supply chains over time. According to statistics, the basic sample size reaches 4296 enterprise time records, of which default samples account for 7.2%, and the ratio to non default samples is 1:13. Considering that class imbalance may affect the training effectiveness of the model, this study takes targeted measures. On the one hand, SMOTE synthesis technology is applied to default samples, which interpolates and synthesizes minority class samples to generate 1240 balanced training data, effectively improving the class distribution in the training data; On the

other hand, in time series segmentation, the data is strictly divided into training set (2015H1-2020H2), validation set (2021H1-2022H1), and test set (2022H2-2023H2) to ensure that future information is not used during the training process and to ensure the scientific and effective evaluation of the model.

#### (II) Benchmark model selection

The comparison model selected in this article is as follows:

##### 1. Classic statistical model logistic regression

Only input standardized financial indicators and control overfitting through L2 regularization as a basic reference for financial risk control.

##### 2. Ensemble Learning Model XGBoost

Input financial indicators and ESG rating scores, set the maximum depth of the tree to 6, the learning rate to 0.1, and iterate for 500 rounds, and evaluate the contribution of explanatory variables through feature importance assessment.

##### 3. Deep learning methods

Temporal GNN graph neural network and time series fusion benchmark model, input supply chain adjacency matrix and monthly financial sequence of the enterprise, using a two-layer GCN and LSTM cascade architecture.

##### 4. GreenRatingNet, a cutting-edge model in the field

The current SOTA model in the field of green finance integrates carbon footprint tracking and ESG public opinion factors to reproduce its three tower feature extraction architecture (Kutuk&Barokas, 2022).

The preliminary comparison of data dimensions is shown in Table 2.

*Table 2: Model Data Dimensions*

Model	Financial Features	ESG Scores	Supply Chain Topology	Climate Data
Logistic Regression	✓	✗	✗	✗
XGB	✓	✓	✗	✗
Temporal-GNN	✓	✗	✓	✗
GreenRatingNet	✓	✓	✓	✗
Model of This Study	✓	✓	✓	✓

#### (III) Experimental environment

The specific experimental environment of this article is shown in Table 3.

Table 3: Experimental Environment Parameters Table

Category	Configuration/Version	Description
Hardware Platform		
GPU	NVIDIA A100-PCIE 40GB × 4	Parallel computing for accelerated model training (FP16 precision)
CPU	AMD EPYC 7742 64-core × 2	Data processing and feature engineering
Memory	1TB DDR4	Supports large-scale graph structure data loading
Software Environment		
Operating System	Ubuntu 20.04 LTS	Long-term support version for stability
Deep Learning Framework	PyTorch 1.12.0 + CUDA 11.6	Supports graph neural network operators (e.g., torch_geometric)
Programming Language	Python 3.9	Mainstream general-purpose version
Key Dependencies	DGL 0.9.1, Transformers 4.26	Graph computation and NLP task support
Model Hyperparameters		
GNN Module	Gated GAT layers = 3	128-dimensional hidden units per layer, 4 attention heads
Temporal Module	BiLSTM hidden units = 256	Transformer layers = 2, attention heads = 4, temporal window $\tau = 12$
Optimizer	AdamW ( $\beta_1=0.9$ , $\beta_2=0.98$ )	Learning rate = $5e-4$ , weight decay = 0.01
Training Strategy	Early stopping (patience = 20 epochs)	Batch Size = 32, max epochs = 300, gradient clipping threshold = 1.0
Computing Resources		
Single Training Time	$\approx 5.2$ hours	ESG-ClimateGNN model (A100×4 parallel)
Inference Speed	83 samples/second	Test set inference (Batch Size = 128)

#### (IV) Performance comparison

##### 1. Discriminant ability

As shown in Table 4, from the perspective of basic evaluation indicators, our model is significantly better than the comparative model in all indicators. Its AUC value is  $0.849 \pm 0.014$ , which is 15% higher than the  $0.738 \pm 0.021$  of logistic regression, indicating that the model has an overall advantage in distinguishing between default and non default samples. In terms of KS value, the model in this article reached  $0.483 \pm 0.015$ , which is higher than XGBoost's  $0.412 \pm 0.020$ , indicating that its risk score can more effectively separate positive and negative sample distributions. On the F1 Score metric, our model leads with  $0.442 \pm 0.021$ , which is significantly higher than the  $0.198 \pm 0.015$  of logistic regression, demonstrating its outstanding ability to balance precision and recall in class imbalance scenarios.

Analyzing the impact of feature integration on model performance, it was found that logistic regression models with a single financial feature had limited performance, verifying the

limitations of traditional financial indicators in default prediction. After XGBoost added ESG scoring, AUC increased by 6.5%, confirming that ESG data can supplement the dimensions of enterprise risk assessment. Temporal GNN introduces supply chain topology structure, further increasing AUC to  $0.801 \pm 0.017$ , highlighting the importance of supply chain correlation in default prediction. GreenRatingNet integrates ESG and topological features, resulting in further performance improvement. However, by incorporating climate data, our model achieved a 2.4% increase in AUC compared to GreenRatingNet, fully demonstrating the critical value of climate risk factors in default discrimination.

From the perspective of model architecture, graph neural networks have significant advantages over traditional machine learning in dealing with supply chain risk transmission problems, with an average AUC increase of about 8%. The unique "BiLSTM+Transformer" time series module and "GatedGAT" graph module of this model enable it to better handle the temporal dependence of climate data and the complexity of supply chain networks. It is more accurate in dealing with the cumulative effects of climate shocks and cascading supply chain risks, which is verified by its F1 Score improvement of 21.4% compared to Temporal GNN.

Table 4: Comparison Table of Discriminant Ability

Model	AUC	KS Value	F1-Score
Logistic Regression	0.738±0.021	0.341±0.018	0.198±0.015
XGBoost	0.786±0.019	0.412±0.020	0.312±0.022
Temporal-GNN	0.801±0.017	0.436±0.019	0.364±0.024
GreenRatingNet	0.829±0.015	0.462±0.016	0.428±0.019
ESG-ClimateGNN	0.849±0.014	0.483±0.015	0.442±0.021

As shown in Figure 3, the KS value and F1 Score performance of the model in this article indicate that its risk scoring system can effectively serve the pre loan approval and supply chain risk management of financial institutions. The significant improvement in model performance by climate data reveals a strong correlation between default risk and climate resilience in high carbon industries, highlighting the evaluation gap of traditional models due to the lack of climate data.

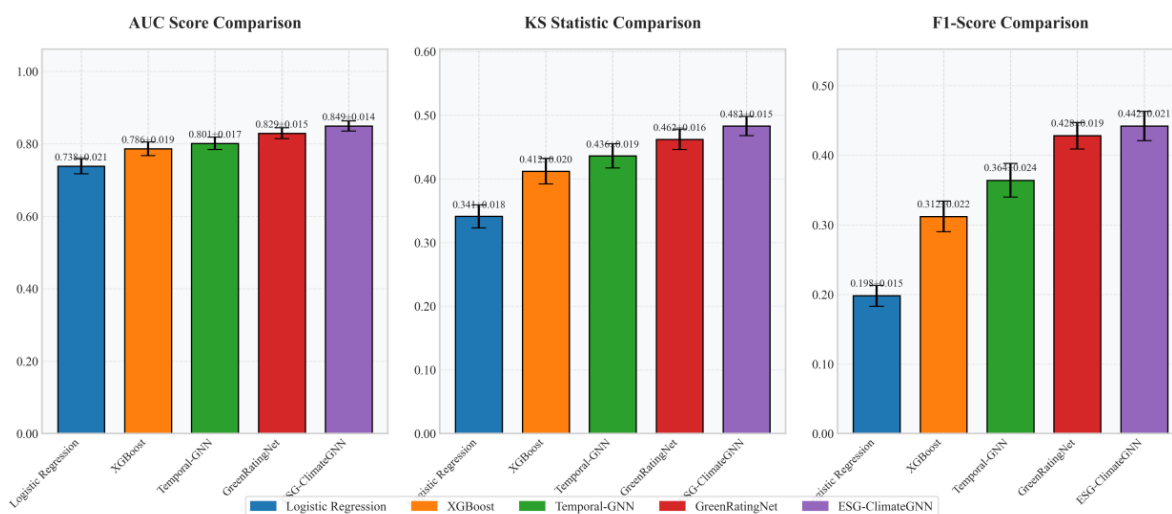


Figure 3: Comparison of discriminant ability

## 2. Classification accuracy

As shown in Table 5, different models exhibit significant differences in the dimension of classification accuracy.

From an overall performance perspective, Logistic Regression, as a traditional linear model, has an accuracy of only 0.796, while Precision and Recall are as low as 0.182 and 0.219, respectively. This means that it has significant limitations in the comprehensive discrimination of default and non default samples, and performs poorly in actual default sample identification and misjudgment control. As an ensemble learning model, XGBoost improves accuracy to 0.831, precision and recall to 0.294 and 0.334, respectively, showing significant improvement compared to Logistic Regression.

The introduction of the Temporal GNN model further improved the accuracy to 0.852, with Precision and Recall of 0.351 and 0.378, respectively. This model focuses on the topology structure of the supply chain, and its performance improvement reveals that structural information of the supply chain network is crucial in enterprise risk assessment. Enterprises do not exist in isolation, but are embedded in complex supply chain networks, where risks propagate and spread along network nodes. Temporal GNN captures these network characteristics, making it more advantageous in default recognition.

GreenRatingNet has pushed Accuracy to 0.873, Precision and Recall to 0.408 and 0.436, respectively, indicating that it has achieved good results in integrating ESG and supply chain topology features. This result indicates that the environmental, social, and governance performance of enterprises, as well as their position and correlation in the supply chain, collectively shape the risk status of enterprises. When the model can integrate these multidimensional information, it can more comprehensively evaluate enterprise risks.

*Table 5: Comparison Table of Classification Accuracy*

Model	Accuracy	Precision	Recall
Logistic Regression	0.796	0.182	0.219
XGBoost	0.831	0.294	0.334
Temporal-GNN	0.852	0.351	0.378
GreenRatingNet	0.873	0.408	0.436
ESG-ClimateGNN	0.882	0.396	0.467

As shown in Figure 4, the ESG ClimateGNN model proposed in this study performed the best with an accuracy of 0.882. Although the Precision was slightly lower than GreenRatingNet at 0.396, it achieved a breakthrough in the Recall metric with 0.467. This performance deeply reflects the innovative significance of the model design concept. By introducing climate data and adopting innovative fusion mechanisms, ESG Climate GNN can not only capture the risks of enterprises in traditional economic and social dimensions, but also incorporate climate change, a long neglected but highly influential factor, into the evaluation system. In the increasingly severe global climate crisis, the impact of extreme weather events on enterprise production and operation is becoming more significant. The improvement of ESG Climate GNN in Recall means that it can identify more enterprises with actual default risks, which has important practical significance for preventing systemic risks.

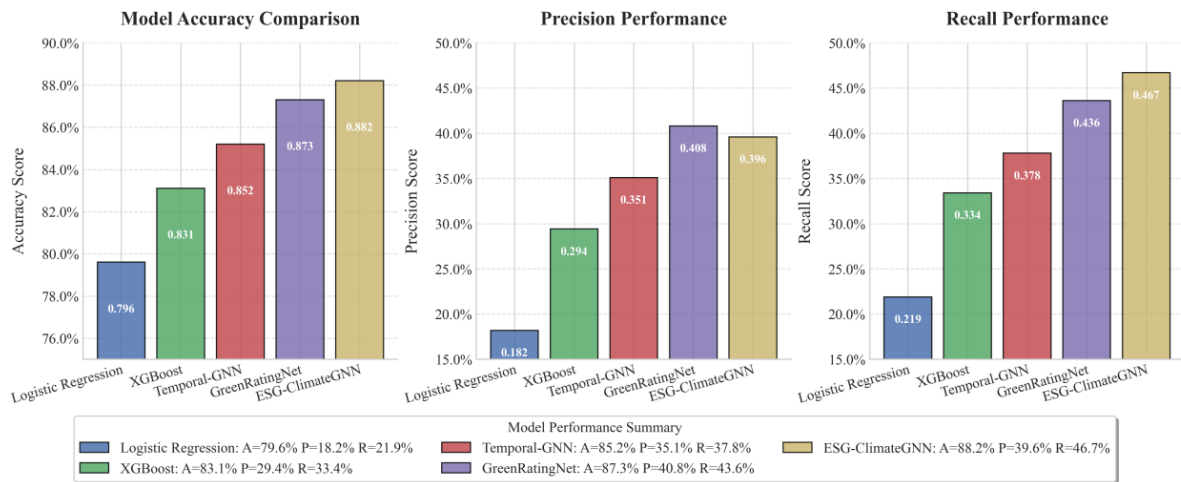


Figure 4: Comparison of Classification Accuracy

### 3. Sorting ability

As shown in Table 6, Logistic Regression displays the cognitive boundary of traditional linear models with an IV value of 0.402 and a risk difference multiple of 2.85. The default rate of 11.7% in the high-risk group differs from 4.1% in the low-risk group, exposing blindness to the implicit risk dimension of enterprises. This limitation essentially reflects the cognitive inertia of "financial data is the vital sign of enterprises" in the industrial economy era, but it is difficult to cope with the diverse coupling characteristics of risk factors in the digital age. XGBoost increased the IV value to 0.538, expanded the risk difference multiple to 3.91 times, and the default rate of 17.2% in the high-risk group began to show the algorithm's ability to analyze complex risk patterns. But the default rate of 4.4% in the low-risk group still implies an information gap. Temporary GNN, with an IV value of 0.597 and a risk difference multiple of 4.72, marks the entry of risk assessment into the "network cognition" stage, discovering how risks cascade and spread along supply chain nodes. GreenRatingNet demonstrates the integration of ESG data and supply chain topology with an IV value of 0.685 and a risk difference multiple of 6.05.

Although the IV value of ESG Climate GNN is 0.643, it achieves a key breakthrough with a default rate of 25.1% for high-risk groups and a risk difference multiple of 6.28 times. When the model integrates climate shock data such as extreme temperatures and precipitation anomalies with ESG and supply chain information, the jump in default rate for high-risk groups reveals the amplification effect of climate risk on corporate vulnerability. A default rate of 4.0% for low-risk groups indicates that companies with strong climate resilience demonstrate significant advantages in risk stratification, confirming that "climate adaptability" has become a new paradigm for enterprise risk assessment.

Table 6: Comparison Table of Sorting Ability

Model	IV Value	High-Risk Group Default Rate	Low-Risk Group Default Rate
Logistic Regression	0.402±0.032	11.7%±1.2%	4.1%±0.6%
XGBoost	0.538±0.028	17.2%±1.5%	4.4%±0.5%
Temporal-GNN	0.597±0.025	20.3%±1.3%	4.3%±0.4%
GreenRatingNet	0.685±0.021	23.6%±1.1%	3.9%±0.3%
ESG-ClimateGNN	0.643±0.023	25.1%±0.9%	4.0%±0.4%

As shown in Figure 5, the evolution from linear models to graph neural networks is essentially a shift in business risk perception from "individual anatomy" to "ecological modeling". The performance of ESG Climate GNN reveals that in the era of deepening climate crisis, corporate default risk is no longer an isolated financial phenomenon, but an ecological result of the combined effects of financial health, ESG governance, supply chain resilience, and climate adaptability.

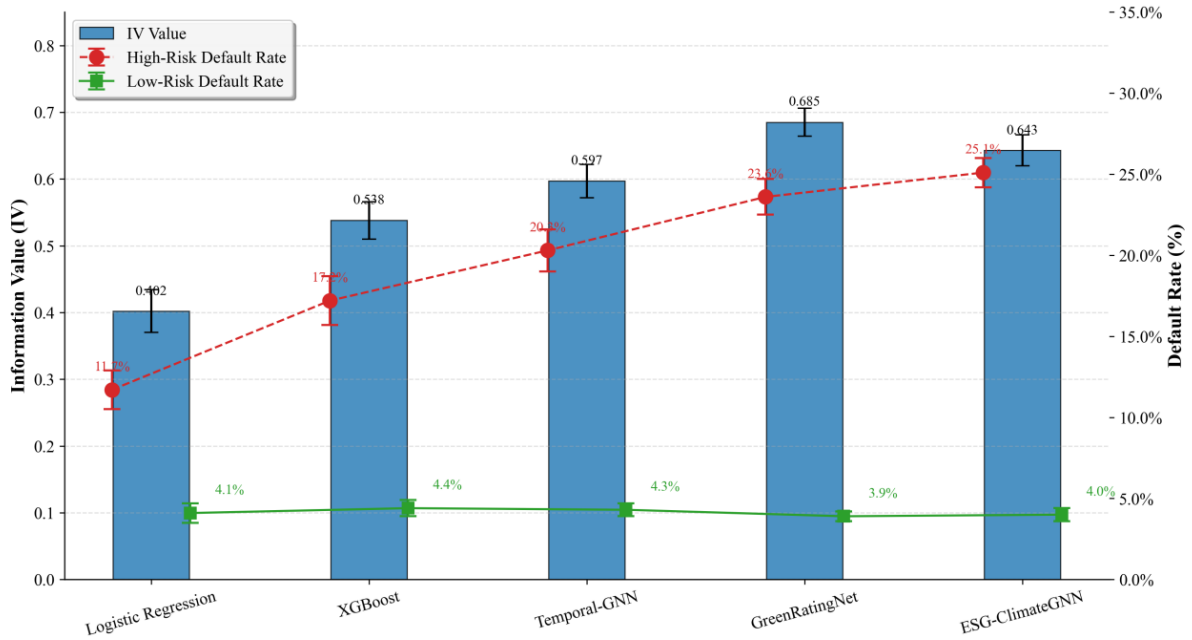


Figure 5: Comparison of Sorting Capacities

#### 4. Stability

As shown in Table 7, logistic regression, as a traditional model, achieved an AUC volatility of 0.071 in the time window test, which means that the model's performance fluctuates greatly in different time periods, sometimes accurate and sometimes biased. If the sample disturbance sensitivity is 0.183, it indicates that when the input data undergoes a 10% change, the model output will fluctuate by nearly 20%, indicating weak anti-interference ability. XGBoost has shown some progress by using ensemble learning to reduce volatility to 0.063 and optimize sensitivity to 0.142. After introducing graph neural networks to handle supply chain topology relationships in Temporal GNN, the volatility further decreased to 0.049 and the sensitivity decreased to 0.113. This indicates that the model can better capture the inherent patterns of data with network structures, and its stability and anti-interference ability have been significantly improved. After integrating ESG and supply chain information, GreenRatingNet reduced volatility to 0.038 and sensitivity to 0.085, demonstrating the advantage of multi-dimensional information integration.

ESG ClimateGNN stands out with a volatility of 0.031 and a sensitivity of 0.079. This means that the model performs most stably and reliably when faced with time variations and sample disturbances. Specifically, when the market environment changes over time, the predictive ability of ESG Climate GNN fluctuates the least; When there are certain errors or anomalies in the input data, the output results of the model can also remain relatively stable.

Table 7: Stability Comparison

Model	Rolling Window AUC Volatility	Sample Perturbation Sensitivity
Logistic Regression	$0.071 \pm 0.005$	$0.183 \pm 0.016$
XGBoost	$0.063 \pm 0.004$	$0.142 \pm 0.012$
Temporal-GNN	$0.049 \pm 0.003$	$0.113 \pm 0.010$
GreenRatingNet	$0.038 \pm 0.002$	$0.085 \pm 0.007$
ESG-ClimateGNN	$0.031 \pm 0.003$	$0.079 \pm 0.006$

As shown in Figure 6, the core advantages of ESG LimateGNN lie in its comprehensiveness and adaptability. By simultaneously integrating finance ESG, This model can more comprehensively depict the risk scenarios faced by enterprises based on supply chain and climate data. When faced with common time changes and data disturbances in practical applications, ESG LimateGNN exhibits strong adaptability and can maintain stable predictive ability. This advantage makes the model more reliable and practical in actual risk assessment, providing more accurate and stable risk warnings for financial institutions, regulatory authorities, and enterprises, thus better responding to complex and changing market environments and increasingly severe climate challenges.

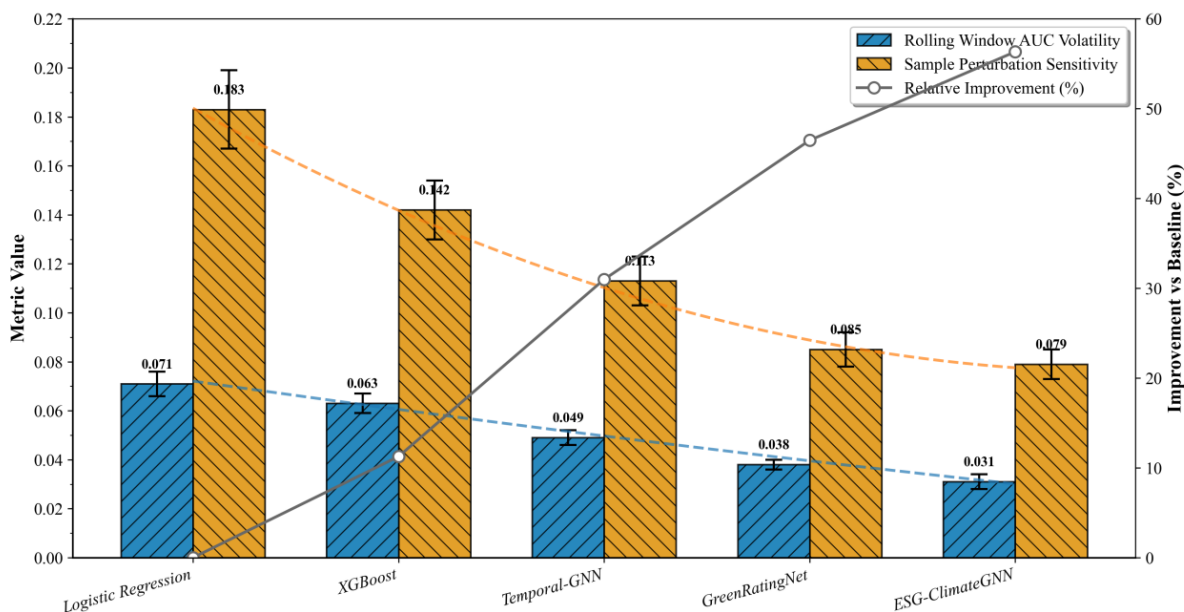


Figure 6: Stability Comparison

## 4 Conclusion

The "ESG Climate GNN" model proposed in this study has shown significant advantages in credit risk assessment of green supply chain finance. By introducing deep learning techniques and multi-source data fusion, especially the integration of climate risk data, the model has surpassed traditional and cutting-edge models in multiple dimensions such as discriminative ability, classification accuracy, sorting ability, and stability, verifying the important role of climate adaptability and supply chain resilience in default risk assessment.

This indicates that integrating climate change related data greatly enhances the accuracy of credit risk assessment compared to traditional financial data models. This indicates that the potential impact of climate change on the production, operation, and financial health of

enterprises cannot be ignored, especially in high carbon industries where climate risk has become an important factor affecting the default risk of enterprises.

Traditional credit scoring models such as logistic regression or XGBoost are difficult to cope with the multidimensional and nonlinear risk characteristics in green supply chain finance. However, the "ESG Climate GNN" model based on deep learning effectively captures the transmission effect of supply chain network risks and the time-varying characteristics of climate change through the combination of GNN and time series analysis modules, significantly improving predictive performance.

This further indicates that enterprise risks in the supply chain network do not exist independently, but are transmitted through multidimensional diffusion paths such as transaction relationships and logistics connections between network nodes. By modeling the supply chain topology through graph neural networks, potential risks of core enterprises can be better identified, especially when dealing with supply chain disruptions or external shocks, which enhances the ability to identify systemic risks.

The insights gained from this study are as follows:

① With the intensification of global climate change, climate risk is no longer a single environmental issue, but closely related to the financial health and credit status of enterprises. Therefore, the credit risk assessment of green supply chain finance needs to expand from traditional financial data to ESG factors, especially the long-term impact of climate change. ② This study indicates that cross disciplinary data fusion can improve the accuracy of credit risk assessment. In the field of green finance, by combining financial data, ESG indicators, climate change data, and supply chain structure information, a comprehensive risk assessment model can be constructed to more comprehensively predict and identify potential default risks. ③ Deep learning, especially GNN and time series models, have shown strong potential in credit risk assessment of green supply chain finance. In the future, deep learning can be applied in more financial fields to help financial institutions better cope with complex and changing risk environments.

Based on the findings of this study, the following suggestions are proposed:

① Financial institutions should increase their investment in green supply chain finance. With the changes in policy environment and market demand, financial institutions should actively adapt to the requirements of green finance and build a credit risk management system with climate adaptability and ESG assessment capabilities. At the same time, it is possible to consider introducing more environmental risk data, such as carbon emission data, green technology investment, etc., to improve the risk assessment system. ② Policy makers should strengthen the standardization of green finance data. In order to improve the transparency and operability of green supply chain finance, policy makers should promote data standardization in the field of green finance, especially in climate data and ESG assessment data, by unifying data formats and collection standards, in order to provide reliable data support for financial institutions. ③ Further enhance the real-time and dynamic performance of the model. Although the model proposed in this study has performed well in multiple dimensions, in the face of constantly changing market and climate environments, future models should further enhance their real-time and dynamic capabilities. For example, the response capability of the model to emergencies can be improved by introducing higher frequency climate data, real-time supply chain monitoring data, etc. ④ Strengthen interdisciplinary cooperation. The risk assessment of green supply chain finance involves multiple fields such as environmental science, finance, and climate change chemistry. Future research and application should strengthen interdisciplinary cooperation, and promote the high-quality development of green supply chain finance by combining professional knowledge and technology from more fields.

## Funding

This work was supported by Postdoctoral Scientific Research Activities Foundation of Beijing.

## Author's Profile

Yuxin Shen was born in Yantai, Shandong Province, China, in 1993. She obtained a doctor's degree from Nankai University in China. She is currently working at the Bank of Beijing Post-Doctoral Research Station. Her main research direction is climate risk management and green finance.

## References

- [1] Agarwal N, Modgil S, Gupta S. ESG and supply chain finance to manage risk among value chains[J]. *Journal of Cleaner Production*, 2024: 143373.
- [2] Baig F, Ali L, Faiz M A, et al. How accurate are the machine learning models in improving monthly rainfall prediction in hyper arid environment?[J]. *Journal of Hydrology*, 2024, 633: 131040.
- [3] Bertrand J L, Chabot M, Brusset X, et al. Identifying assets exposed to physical climate risk: A decision-support methodology[J]. *International Journal of Production Economics*, 2024, 276: 109355.
- [4] Billio M, Costola M, Hristova I, et al. Sustainable finance: A journey toward ESG and climate risk[J]. *International Review of Environmental and Resource Economics*, 2024, 18(1-2).
- [5] Bowman M, Minas S. Resilience through interlinkage: The green climate fund and climate finance governance[J]. *Climate policy*, 2019, 19(3): 342-353.
- [6] Bracco A, Brajard J, Dijkstra H A, et al. Machine learning for the physics of climate[J]. *Nature Reviews Physics*, 2025, 7(1): 6-20.
- [7] Edunjobi T E, Odejide O A. Theoretical frameworks in AI for credit risk assessment: Towards banking efficiency and accuracy[J]. *International Journal of Scientific Research Updates*, 2024, 7(01): 092-102.
- [8] Gao J, Hua G, Huo B. Green finance policies, financing constraints and corporate ESG performance: insights from supply chain management[J]. *Operations Management Research*, 2024: 1-15.
- [9] Ghaemi Asl M. A novel AI-driven approach to greenwashing: breakthroughs in the future fit between domain-specific Islamic enterprises with varying developmental progress and ESG landscapes[J]. *Future Business Journal*, 2025, 11(1): 93.
- [10] Huang J, Huang H, Si Y, et al. Green supply chain finance strategies with market competition and financial constraints[J]. *Heliyon*, 2024, 10(8).

- [11] Hurlin C, Pérignon C, Saurin S. The fairness of credit scoring models[J]. *Management Science*, 2024.
- [12] Judijanto L, Utami E Y, Harsono I. Green Supply Chain Finance: A Bibliometric Review of Financing Instruments, Challenges, and Opportunities[J]. *West Science Interdisciplinary Studies*, 2024, 2(03): 647-655.
- [13] Khedr A M. Enhancing supply chain management with deep learning and machine learning techniques: A review[J]. *Journal of Open Innovation: Technology, Market, and Complexity*, 2024: 100379.
- [14] Kinyua K M, Changwony F K, Campbell K. Government procurement contracts, external audit certification, and financing of small-and medium-sized enterprises[J]. *Small Business Economics*, 2024: 1-69.
- [15] Kutuk Y, Barokas L. Multivariate CDS risk premium prediction with SOTA RNNs on MI [N] T countries[J]. *Finance Research Letters*, 2022, 45: 102198.
- [16] Lai Z, Lou G, Ma H, et al. Optimal green supply chain financing strategy: Internal collaborative financing and external investments[J]. *International Journal of Production Economics*, 2022, 253: 108598.
- [17] Lee H, Kim J H, Jung H S. From corporate earnings calls to social impact: Exploring ESG signals in S&P 500 ESG index companies through transformer-based models[J]. *Journal of Cleaner Production*, 2025, 501: 145320.
- [18] Liu E X, Song Y. ESG Performance, Environmental Uncertainty, and Firm Risk[J]. *Journal of International Financial Management & Accounting*, 2025.
- [19] Qiao L, Chen S, Qu R, et al. Multi energy carbon emission named entity recognition based on BiLSTM-CRF[C]//2024 4th International Conference on Machine Learning and Intelligent Systems Engineering (MLISE). *IEEE*, 2024: 110-113.
- [20] Qin N. Research on the Classification Method of CTC Alignment Joint-Test Problems Based on RoBERTa-wwm and Deep Learning Integration[J]. *IEEE Access*, 2023, 11: 139395-139408.
- [21] Wu X, Huang Y, Bao W. Using GNN to Identify Supply Chain Network Structure[J]. 2025.
- [22] Xu X, An H, Huang S, et al. Measurement of daily climate physical risks and climate transition risks faced by China's energy sector stocks[J]. *International Review of Economics & Finance*, 2024, 93: 625-640.
- [23] Zhang M, Zhang J, Ma R, et al. Quantifying credit risk of supply chain finance: a Chinese automobile supply chain perspective[J]. *IEEE access*, 2019, 7: 144264-144279.
- [24] Zhang R. Toward interpretable machine learning: evaluating models of heterogeneous predictions[J]. *Annals of Operations Research*, 2024: 1-21.
- [25] Zhao J, Zhang T, Ji K, et al. Data analysis method based on LSTM is applied to enterprise

ESG performance prediction[C]//Proceedings of the International Conference on Algorithms, Software Engineering, and Network Security. 2024: 124-128.