Zeus Press
Your Gateway to Scholarly Success

ISSN 3006-3965 (PRINT)
ISSN 3006-3973 (ONLINE)

ESG

Environment, Social and Governance

Environment, Social and Governance

Vol. 3 No. 3 (2026)

Published by

Zeus Press

About the Journal

Environment, Social and Governance (ESG) is a double-blind peer-reviewed, open-access journal dedicated to the exploration and dissemination of research that intersects environmental sustainability, social responsibility, and corporate governance. The journal seeks to foster a deeper understanding of how these critical factors influence and are influenced by corporate behavior, investment strategies, and regulatory frameworks worldwide.

Focus and Scope

ESG publishes high-quality research on a wide array of topics related to environmental, social, and governance issues. Areas of interest include, but are not limited to:

- Environmental impact assessment and sustainability practices
- Corporate social responsibility (CSR) and ethical business practices
- Governance structures and board composition
- Sustainable finance and investment strategies
- Social impact investing and impact measurement
- Environmental regulations and policy analysis
- Stakeholder engagement and reporting practices
- Risk management and corporate transparency
- Climate change mitigation and adaptation strategies
- Human rights, labor standards, and diversity in the workplace
- ESG metrics, frameworks, and reporting standards
- The role of technology in advancing ESG goals

Types of Contributions

The journal publishes the following types of contributions: (1) full-length articles, (2) review articles, (3) book reviews, (4) discussion notes and (5) articles in translation.

More Information

Email: service@zeus-press.org
Website: www.zeuspress.org
Address: FLAT/RM 91_3 9/F CHINACHEM GOLDEN PLAZA NO.77 MODY ROAD TSIM SHA TSUI KL, HONG KONG

Contents

Environment, Social and Governance; Vol.3, No.3; 2026
E-ISSN: 3006-3973 P-ISSN：3006-3965
DOI: https://doi.org/10.70267/ esg2026.v3n3.0112
Published by: Zeus Press

Mid- to Long-Term Runoff Forecasting in the Huaihe River Basin Using a Spatiotemporal Graph Neural Network Coupled with LSTM

Yanzhu Xu* **and Lijiaxing Cheng**

College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China

**Corresponding author: Yanzhu Xu.*

Abstract

To address the difficulty of simultaneously characterizing runoff temporal variation and spatial connectivity among stations under the complex river network conditions of the Huaihe River Basin, a spatiotemporal graph neural network model integrating temporal feature extraction and graph-based spatial modeling was developed for mid- to long-term runoff forecasting. Based on multi-station daily precipitation, air temperature, and runoff data from 2011 to 2020 in the Huaihe River Basin, data preprocessing was conducted, including missing-value imputation, outlier detection, standardization, and basin graph construction, and an adjacency matrix was established according to the upstream–downstream hydrological connectivity among stations. The model employed LSTM to extract dynamic features from historical sequences and graph convolution to capture spatial dependencies within the basin. Compared with LSTM, K-nearest neighbor, support vector regression, decision tree, and AdaBoost models, the proposed model achieved better overall predictive performance on the basin-wide test set, with an average R^2 of 0.915, average MAE of 30.019, and average RMSE of 50.243, demonstrating strong runoff process fitting ability and good error control at the basin scale. The results indicate that incorporating basin spatial connectivity can effectively improve the accuracy of mid- to long-term runoff forecasting and provide methodological support for water resources regulation and flood-control decision-making in complex river basins.

Keywords

spatiotemporal graph neural network, Huaihe River Basin, mid- to long-term runoff forecasting, spatial connectivity, complex river-network basin

1. Introduction

Runoff forecasting serves as a critical foundation for water resources allocation, flood control scheduling, and integrated basin management. Among these, mid- to long-term runoff forecasting holds significant importance for basin operation decision-making, water supply security, and the formulation of risk response plans. Influenced by multiple factors such as meteorological drivers, underlying surface conditions, and human activities, runoff processes typically exhibit pronounced nonlinearity, non-stationarity, and time-varying characteristics, which increase the complexity of predictive modeling.

Therefore, improving the accuracy of mid- to long-term runoff forecasting under complex conditions has become an important research issue in the field of hydrology and water resources [1-2].

Regarding runoff forecasting, domestic and international research has followed a developmental path from traditional statistical models and machine learning models to deep learning models [2-3]. Traditional statistical models and shallow machine learning models have certain application foundations in runoff forecasting, but their ability to characterize complex nonlinear hydrological processes is limited. Although traditional machine learning methods such as Random Forest (RF) and Support Vector Regression (SVR) can improve prediction accuracy to some extent, they usually rely on manual feature engineering and struggle to adequately capture the long-term temporal dependencies and dynamic cumulative effects in runoff evolution. Shallow models such as Back Propagation Neural Networks also tend to be limited by insufficient feature representation capability when addressing mid- to long-term runoff problems [3-6]. In contrast, Long Short-Term Memory (LSTM) networks and related deep learning models demonstrate stronger advantages in time series modeling and have been widely applied in runoff forecasting research [7-9].

Although existing studies have made some progress in improving runoff prediction accuracy, most methods still primarily focus on mining temporal sequence features while paying insufficient attention to spatial dependency relationships in complex river-network basins, such as upstream–downstream transmission, tributary inflows, and multi-station coordinated responses [1-2]. For basin systems like the Huaihe River Basin, which feature dense river networks and significant spatial connectivity, modeling based solely on single stations or purely temporal dimensions makes it difficult to fully reflect the spatiotemporal interaction characteristics during runoff evolution. In recent years, Graph Neural Networks (GNN) have provided new approaches for modeling the spatial topological relationships of basins. Combining them with Long Short-Term Memory networks is expected to more comprehensively characterize both the temporal dependency features and spatial propagation features of runoff changes in complex river-network basins, thereby enhancing mid- to long-term runoff forecasting capability [10-13].

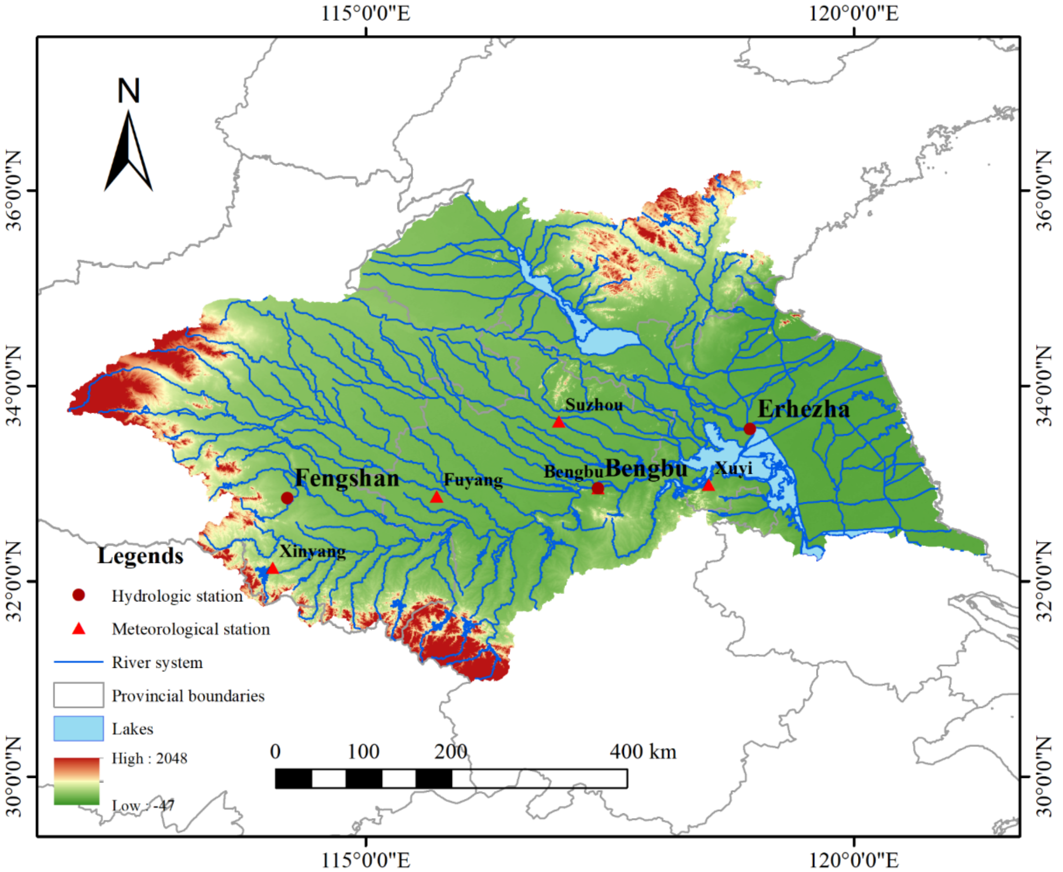
Based on the above considerations, this paper takes the Huaihe River Basin as the study area and constructs a spatiotemporal joint runoff forecasting model that integrates LSTM and GNN. By incorporating basin spatial connectivity information, the model jointly extracts temporal dependency features and spatial propagation features in the runoff evolution process. The proposed model is compared with LSTM, K-nearest neighbor, support vector regression, decision tree, and AdaBoost models [14-17] to evaluate its applicability and accuracy advantages in mid- to long-term runoff forecasting, with the aim of providing new modeling ideas for runoff prediction in complex river-network basins.

2. Study Area and Data Preprocessing

2.1 Overview of the Study Area

The Huaihe River Basin is located in the north-south transitional zone of eastern China and serves as a key region for water resources allocation and flood control scheduling (Figure 1). Affected by the monsoon climate, precipitation within the basin is unevenly distributed throughout the year, with concentrated inflow during the flood season and pronounced fluctuations in the runoff process. According to multi-station daily-scale data from 2011 to 2020, the multi-year average temperature in the study area ranges from 15.6 to 15.8 °C, and the multi-year average annual precipitation ranges from 919 to 981 mm, of which precipitation from June to September accounts for 58%–65% of the annual total. The maximum daily precipitation can reach 92.4–112.03 mm, and the maximum daily discharge at representative hydrological stations can reach 7340.0 m^3/s, indicating significant runoff fluctuations under extreme weather conditions. Meanwhile, the Huaihe River Basin has numerous tributaries and a dense river network, with close hydrological connections between upstream and downstream areas. Runoff changes are simultaneously influenced by local meteorological conditions, upstream inflow transmission, and tributary confluence processes, exhibiting typical spatiotemporal evolution characteristics of a complex river-network basin. Therefore, selecting this basin for mid- to long-term runoff forecasting research is highly representative and practically necessary.

Figure 1: Distribution map of meteorological and hydrological stations in the Huaihe River Basin



2.2 Data Sources and Dataset Construction

2.2.1 Data Sources

The raw data used in this study include meteorological and hydrological data, covering the period from 2011 to 2020 at a daily scale, for a total of 3653 days. Meteorological data were obtained from observations at multiple meteorological stations within the basin, with main variables including sunshine duration, average temperature, daily maximum temperature, daily minimum temperature, relative humidity, average wind speed, and precipitation. Hydrological data were derived from observations at major hydrological stations in the basin, with the indicator being daily runoff discharge.

To facilitate regional-scale analysis, the study area was divided into three regions—upstream, midstream, and downstream—based on the spatial distribution of stations and basin sectional characteristics, and corresponding datasets were constructed. The specific grouping is as follows: For the upstream region, meteorological data from Xinyang and Fuyang stations were selected and matched with runoff data from Fengshan station. For the midstream region, meteorological data from Bengbu and Fuyang stations were selected and matched with runoff data from Bengbu station. For the downstream region, meteorological data from Bengbu, Suzhou, and Xuyi stations were selected and matched with runoff data from Erhezha station. This grouping method takes into account both basin sectional division and data availability, and can better characterize the meteorological–runoff response relationships in different river reaches.

2.2.2 Data Cleaning and Preprocessing

To ensure the integrity and reliability of the modeling data, the raw meteorological and hydrological data were cleaned and preprocessed [1, 14]. For meteorological data, the time format was first unified and arranged in ascending chronological order. Outliers were then handled using a quantile-based method: values below the 1% quantile and above the 99% quantile were regarded as outliers and corrected using linear interpolation to reduce noise interference while preserving the characteristics of extreme hydrometeorological events as much as possible. For hydrological data, missing values in the runoff discharge series were primarily addressed.

After unifying the time format and sorting, linear interpolation was applied to fill missing discharge values, forming a cleaned hydrological station dataset. This approach maintains temporal continuity while mitigating the impact of missing data on prediction results.

2.2.3 Construction of Regional Datasets

After completing data cleaning for meteorological and hydrological stations, an overall dataset covering the upstream, midstream, and downstream areas of the Huaihe River Basin was constructed. The dataset contains 9 variables and 3653 daily-scale samples, with complete temporal coverage and a unified structure, making it suitable for subsequent statistical analysis, correlation analysis, and prediction model construction.

As shown in Table 1, the main meteorological variables in the Huaihe River Basin are generally stable, with a mean temperature of 15.7 °C and a mean precipitation of 2.6 mm. In contrast, the mean runoff discharge is 379.8 m³/s, with a standard deviation of 716.0 m³/s, indicating more significant fluctuations. This suggests that the runoff process has greater complexity and uncertainty. These findings demonstrate objective differences in inflow conditions and hydrological response characteristics across different river reaches, further validating the necessity of partitioned modeling and regional comparative analysis.

Table 1: Statistical table of meteorological and hydrological characteristics in the Huaihe River Basin

Feature	Unit	Mean	Standard Deviation	Minimum	Median	Maximum
Sunshine duration	h	5.0	3.8	0.0	5.3	12.6
Temperature	°C	15.7	9.5	-8.1	17.0	33.7
Humidity	%	71.2	14.2	18.0	72.5	99.5
Wind speed	m/s	2.2	0.9	0.4	2.1	7.2
Precipitation	mm	2.6	8.2	0.0	0.0	112.0
Evaporation	mm	2.7	1.5	0.4	2.5	7.2
Runoff discharge	m³/s	379.8	716.0	-65.6	190.0	7340.0

2.2.4 Dataset Division and Standardization

Given the strong temporal correlation of runoff time series, samples were not randomly divided to ensure the prediction process aligns with real-world application scenarios. Instead, they were divided chronologically, with the first 70% of samples used as the training set and the remaining 30% as the test set [1, 14]. For models sensitive to data scale, the Z-score standardization method was applied to both input features and the target variable, converting the data to a standardized form with a mean of 0 and a standard deviation of 1 using the sample mean and standard deviation. For models insensitive to feature magnitude, the original data were used for modeling and comparison [14].

2.3 Model Evaluation Metrics

To quantitatively evaluate the fitting performance and prediction accuracy of different models on the runoff process, this study selected Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2) as model evaluation indicators [1, 16-17]. These indicators have been widely used in runoff forecasting and hydrological time series simulation research. Among them, RMSE and MAE reflect the magnitude of deviation between predicted and observed values; smaller values indicate better model prediction performance. R^2 is used to measure the model's ability to explain the variation characteristics of the observed series; values closer to 1 indicate better model fitting [1, 16-17].

Let the observed runoff value be y_i, the predicted runoff value be $\hat{y}_i$, the number of samples be n, and the mean of the observed values be $\bar{y}$. The calculation formulas for each evaluation indicator are as follows:

$$R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{1}$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \tag{2}$$

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}\left|y_i-\hat{y}_i\right| \tag{3}$$

By integrating the above indicators, the predictive performance of different models can be compared from the perspectives of error level and fitting degree.

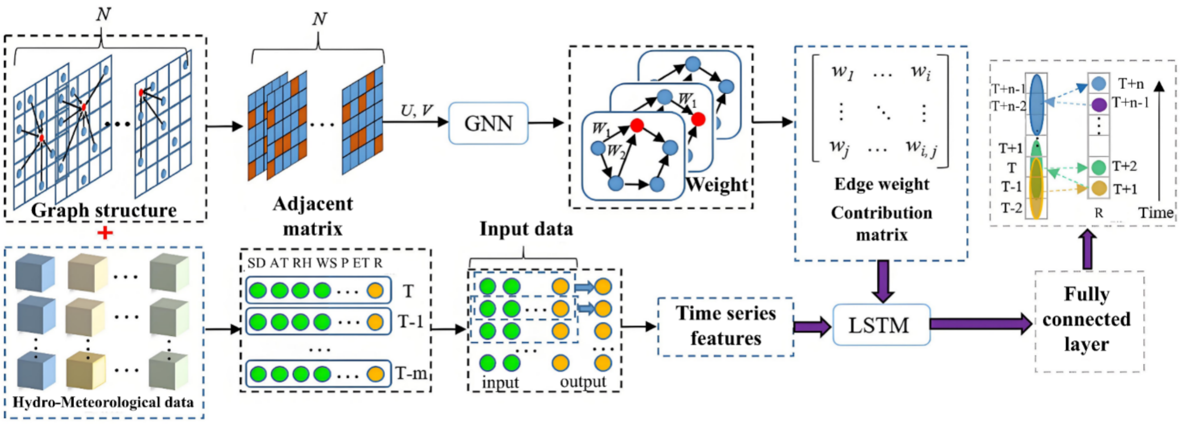
3. Construction of the Spatiotemporal Graph Neural Network Runoff Forecasting Model

To fully characterize the temporal dependency features and spatial connectivity features in the runoff evolution process of the Huaihe River Basin, this study constructs a spatiotemporal joint forecasting model that integrates Long Short-Term Memory (LSTM) and Graph Neural Network (GNN) [9-13]. The model is based on a multi-station meteorological–runoff joint dataset and the basin graph structure. It extracts dynamic features from historical sequences in the temporal dimension and learns propagation relationships between upstream and downstream nodes in the spatial dimension, thereby achieving joint modeling of mid- to long-term runoff processes in complex river-network basins (Figure 2).

3.1 Construction of the Basin Graph Structure

To represent the spatial connectivity relationships among different sections of the Huaihe River Basin, a basin graph structure was constructed based on the regional datasets [10-13]. In this graph structure, river sections and their corresponding control stations are treated as nodes, and the upstream–downstream relationships of rivers are abstracted as edges to describe the spatial transmission paths of the runoff process [10-13]. In accordance with the data organization method used in this study, the upstream, midstream, and downstream regions are regarded as three primary nodes, corresponding respectively to the control areas of Fengshan station, Bengbu station, and Erhezha station. Node features are composed of multidimensional meteorological elements and runoff information from the corresponding region at previous time steps. Based on the natural flow direction in the river network, directed connections are established from upstream to midstream and from midstream to downstream. Self-loops are introduced to retain each node's own information, thereby forming the basin adjacency matrix.

Figure 2: Architecture diagram of the GNN-LSTM model



3.2 Temporal Feature Extraction Module

The runoff process is jointly influenced by multiple factors such as antecedent precipitation, temperature, evaporation, and historical discharge, exhibiting significant temporal correlation and hysteresis. Although traditional regression models or shallow machine learning methods can characterize some nonlinear relationships, they are insufficient in capturing cumulative effects and long-term dependency information in long time series [7-9]. To address this, the Long Short-Term Memory (LSTM) network is introduced in the temporal dimension to extract features from the input sequences [7-9].

As shown in Figure 3, LSTM realizes selective retention and updating of historical information through the forget gate, input gate, and output gate. It can retain key temporal information while suppressing irrelevant disturbances, thereby effectively overcoming the gradient vanishing problem that commonly occurs in ordinary recurrent neural networks during long-sequence training [8-9]. Let x_t be the input vector at time t, and h_{t-1} and C_{t-1} be the hidden state and cell state of the previous time step, respectively. The computation process of LSTM is as follows:

$$f_t=\sigma(W_f[h_{t-1},x_t]+b_f) \tag{4}$$

$$i_t=\sigma(W_i[h_{t-1},x_t]+b_i) \tag{5}$$

$$\widetilde{C}_t=\tanh(W_c[h_{t-1},x_t]+b_c) \tag{6}$$

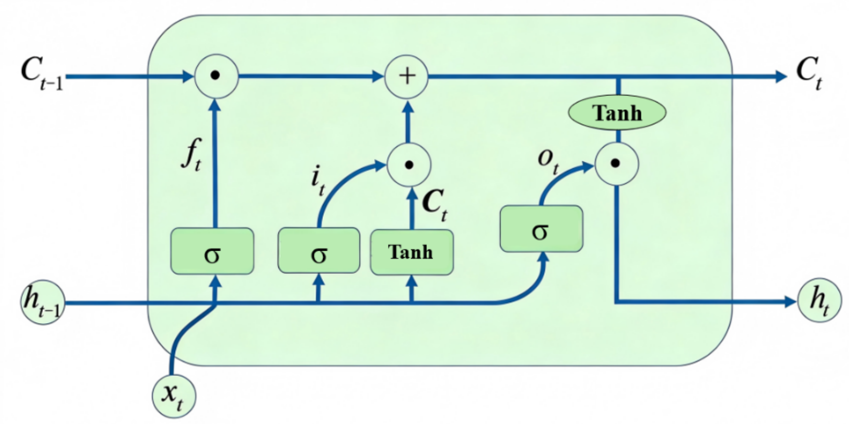
$$C_t=f_t\odot C_{t-1}+i_t\odot\widetilde{C}_t \tag{7}$$

$$o_t=\sigma(W_o[h_{t-1},x_t]+b_o) \tag{8}$$

$$h_t=o_t\odot \tanh(C_t) \tag{9}$$

where f_t, i_t, and o_t represent the forget gate, input gate, and output gate, respectively; σ denotes the Sigmoid activation function; and ⊙ represents the Hadamard product. Through its gating mechanism, LSTM retains historical information useful for current prediction while suppressing irrelevant disturbances, enabling it to effectively capture the long-term temporal dependencies among precipitation, temperature, evaporation, and runoff [8-9].

Figure 3: Structure diagram of the Long Short-Term Memory network model



3.3 Spatial Feature Extraction Module

After constructing the basin graph structure, Graph Neural Network (GNN) is introduced in the spatial dimension for feature extraction [10-13]. Let the basin graph structure be G=(V,E,A), where V is the set of nodes, E is the set of edges, and A is the adjacency matrix. To retain node self-information, self-loops are added to the adjacency matrix:

$$\widetilde{A}=A+I \tag{10}$$

The corresponding degree matrix is:

$$\widetilde{D}_{ii}=\sum_j \widetilde{A}_{ij} \tag{11}$$

The graph convolution layer can be expressed as:

$$H^{(l+1)}=\sigma\left(\widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \tag{12}$$

where $H^{(l)}$ and $H^{(l+1)}$ are the node representations at the 1-th and (l+1)-th layers, respectively, and $W^{(l)}$ is the trainable weight matrix. This process propagates neighborhood information across the graph structure, allowing the model to capture not only the historical variation patterns at individual stations but also the spatial influences brought by upstream inflow, tributary confluence, and coordinated responses from adjacent regions.

3.4 Model Hyperparameter Settings

To ensure model stability and predictive performance, hyperparameters of the constructed spatiotemporal model were set and optimized. The main hyperparameters include the number of graph convolution layers, hidden layer dimension, time step length, learning rate, and number of training epochs. In the spatial modeling component, the number of graph convolutional network layers is set to 2 to balance feature representation capability and overfitting risk; the hidden layer dimension is set to 64 to extract basin spatial correlation features. In the temporal modeling component, the number of LSTM layers is set to 1–2, with 64 hidden units, to capture the temporal dependencies in the runoff sequence. The time window length is set to 7 days to characterize short-term hydrological response processes. During model training, the Adam optimizer is used for parameter updating, with an initial learning rate of 0.001 that is dynamically adjusted according to validation set performance. The batch size is set to 32, the number of training epochs is set to 100, and an early stopping strategy is introduced to prevent overfitting.

3.5 Settings of Comparison Models

To verify the effectiveness of the proposed GNN-LSTM model, LSTM, K-nearest neighbor, support vector regression, decision tree, and AdaBoost models are selected as baseline models for comparative analysis [3-6, 15-17]. Among them, LSTM is primarily used to characterize time series modeling capability, while K-nearest neighbor, support vector regression, decision tree, and AdaBoost represent different types of machine learning methods [3-6, 15-17]. By comparing with the above models, the applicability and accuracy advantages of GNN-LSTM in mid- to long-term runoff forecasting for complex river-network basins can be further evaluated.

In the comparative experiments, all models use the same input variables, identical training set and test set division methods, and a unified evaluation indicator system to ensure fairness and consistency in result comparison [1, 16-17]. The predictive performance of different models is analyzed by comparing their results on RMSE, MAE, and R^2, thereby assessing the accuracy advantages and applicability in mid- to long-term runoff forecasting for complex river-network basins.

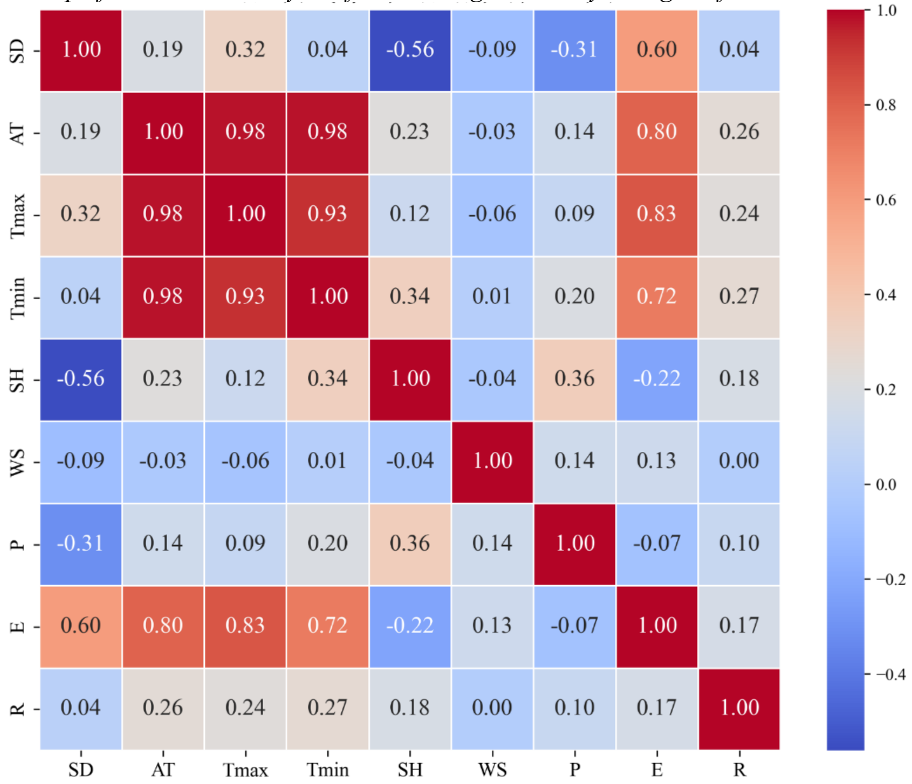
4. Results and Discussion

4.1 Correlation Analysis Between Meteorological Variables and Runoff

To identify the statistical relationships between meteorological elements and runoff in the Huaihe River Basin, a correlation analysis was conducted based on the overall basin data, with results shown in Figure 4. Overall, temperature-related variables exhibited strong correlations with each other; air temperature showed significant positive correlations with both maximum and minimum temperatures. Evaporation also had a high correlation with temperature-related variables ($r=0.72$–0.83). Sunshine duration and humidity displayed a clear negative correlation ($r=-0.56$). In comparison, wind speed showed generally weaker correlations with other variables.

Regarding the relationship between runoff and various meteorological factors, the linear correlations between each variable and runoff were generally weak, showing only a certain degree of weak or low correlation. This indicates that runoff variation is not directly controlled by a single meteorological factor but is jointly influenced by multiple factors and possesses a certain degree of complexity. Considering the high correlation between evaporation and temperature-related variables, as well as its relatively weak correlation with runoff, evaporation was excluded from the final input variables. Ultimately, variables such as temperature, precipitation, lagged precipitation terms, and lagged runoff terms were selected for model construction.

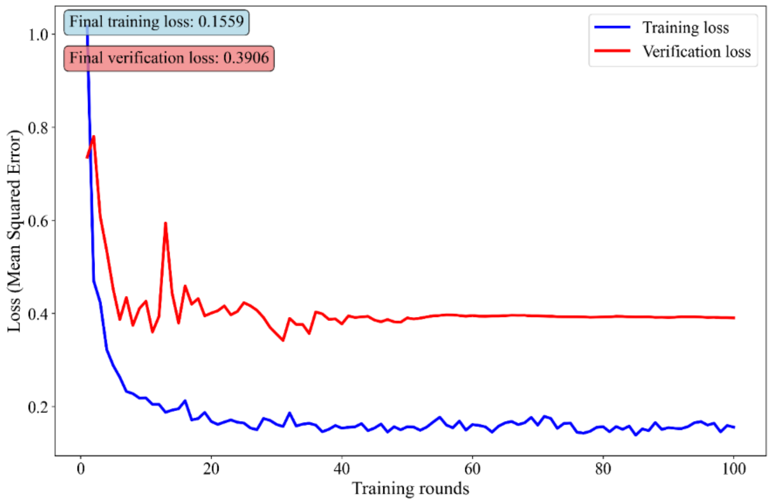
Figure 4: Heatmap of correlation analysis of meteorological and hydrological factors in the Huaihe River Basin



4.2 Comparison of Prediction Accuracy Among Different Models

To compare the performance differences of various models in runoff forecasting for the Huaihe River Basin, this study adopted a unified data division method and evaluation indicator system to analyze the GNN-LSTM, LSTM, K-nearest neighbor, support vector regression, decision tree, and AdaBoost models. The results are presented in Tables 2–4. Overall, the GNN-LSTM model achieved the highest coefficient of determination (R^2) on the test sets across all three regions, reaching 0.915, 0.952, and 0.878 for the upstream, midstream, and downstream regions, respectively. This demonstrates that the model has a strong advantage in fitting the overall variation trend of the runoff process. In addition, Figure 5 shows that both the training loss and validation loss of the GNN-LSTM model decreased steadily and gradually stabilized during training, indicating good convergence and training stability.

Figure 5: Variation curves of the training loss function for the GNN-LSTM model



From the comprehensive results on the overall basin test set, the GNN-LSTM model achieved average R^2, MAE, and RMSE values of 0.915, 30.019, and 50.243, respectively. Compared with the other models, its average R^2 improved by 3.62%–24.49%, the average MAE decreased by 3.62%–42.74%, and the average RMSE was also superior to most comparison models. These results indicate that the model exhibits good comprehensive performance in overall trend fitting and error control. In summary, the LSTM and K-nearest neighbor models produced relatively stable prediction results, though their overall accuracy remained slightly lower than that of GNN-LSTM. Support vector regression and AdaBoost showed relatively weaker performance on the test sets across regions. Although the decision tree model fitted the training set well, its test set metrics declined noticeably, reflecting a certain tendency toward overfitting. Overall, the GNN-LSTM model demonstrates good applicability in mid- to long-term runoff forecasting for the Huaihe River Basin. This suggests that combining temporal feature extraction with spatial topology modeling helps improve the accuracy of runoff prediction in complex river-network basins.

Table 2: Performance comparison of models in the upstream region during training and testing periods

Model	Training Set			Test Set		
	R^2	MAE	RMSE	R^2	MAE	RMSE
GNN-LSTM	0.985	15.321	28.654	0.915	13.842	22.104
LSTM	0.978	8.545	11.923	0.911	15.205	22.621
K-nearest neighbor	0.983	13.336	25.931	0.913	14.152	22.297
Support vector regression	0.972	21.536	51.499	0.78	24.031	35.459
Decision tree	0.989	7.436	10.827	0.866	16.662	27.661
AdaBoost	0.982	28.017	40.918	0.764	28.517	36.795

Table 3: Performance comparison of models in the midstream region during training and testing periods

Model	Training Set			Test Set		
	R^2	MAE	RMSE	R^2	MAE	RMSE
GNN-LSTM	0.962	85.214	70.338	0.952	38.563	65.742
LSTM	0.96	24.609	35.525	0.871	39.181	60.778
K-nearest neighbor	0.963	37.148	67.299	0.853	42.585	64.89
Support vector regression	0.923	42.132	97.307	0.835	44.991	68.731
Decision tree	0.988	25.287	39.019	0.805	48.217	74.802
AdaBoost	0.928	71.662	94.09	0.74	61.802	86.356

Table 4: Performance comparison of models in the downstream region during training and testing periods

Model	Training Set			Test Set		
	R^2	MAE	RMSE	R^2	MAE	RMSE
GNN-LSTM	0.935	48.215	82.441	0.878	37.652	62.883
LSTM	0.939	24.979	36.743	0.868	39.057	61.55
K-nearest neighbor	0.964	36.659	67.057	0.854	41.461	64.619
Support vector regression	0.923	43.422	97.579	0.853	39.729	64.992
Decision tree	0.986	27.275	41.821	0.812	46.609	73.329
AdaBoost	0.915	79.501	102.455	0.701	66.962	92.576

4.3 Mid- to Long-Term Runoff Forecasting Based on the GNN-LSTM Model

To further verify the applicability and accuracy advantages of the GNN-LSTM model in mid- to long-term runoff forecasting for the Huaihe River Basin, Fengshan station, Bengbu station, and Erhezha station were selected as representative stations, and prediction analysis was performed on the runoff process during the test period. The results show that the predicted values and observed values maintained good consistency in overall variation trend and could accurately reflect the phased fluctuation characteristics of the runoff process in different sections. This indicates that the constructed model has good applicability to the runoff evolution process in complex river-network basins. The relevant results are shown in Figures 6–8.

From the prediction results at different stations, the predicted curves for Fengshan station, Bengbu station, and Erhezha station were generally consistent with the observed curves and effectively reflected the phased fluctuation characteristics of the runoff process. The model demonstrated good responsiveness to major flood peak processes, accurately capturing the timing and variation trends of peak values. During the dry season, the predicted values also maintained good consistency with the observed values, indicating strong simulation capability for low-flow processes as well. Overall, the GNN-LSTM model can effectively extract temporal dependency features from runoff sequences and, by incorporating basin spatial topological relationships, enhance its ability to represent the runoff variation process in complex river-network systems.

Figure 6: Rainfall-runoff prediction results at Fengshan station during the test period based on GNN-LSTM

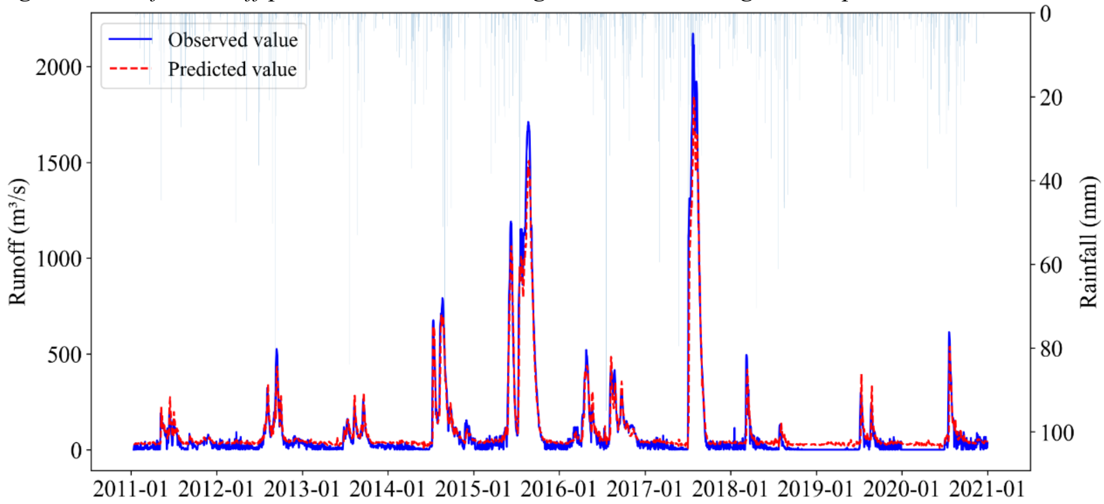


Figure 7: Rainfall-runoff prediction results at Bengbu station based on GNN-LSTM

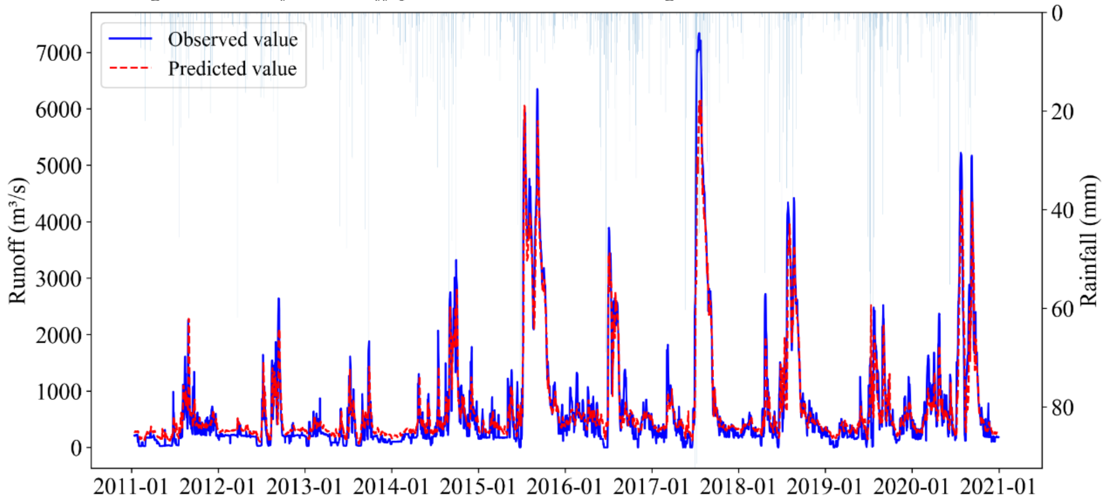
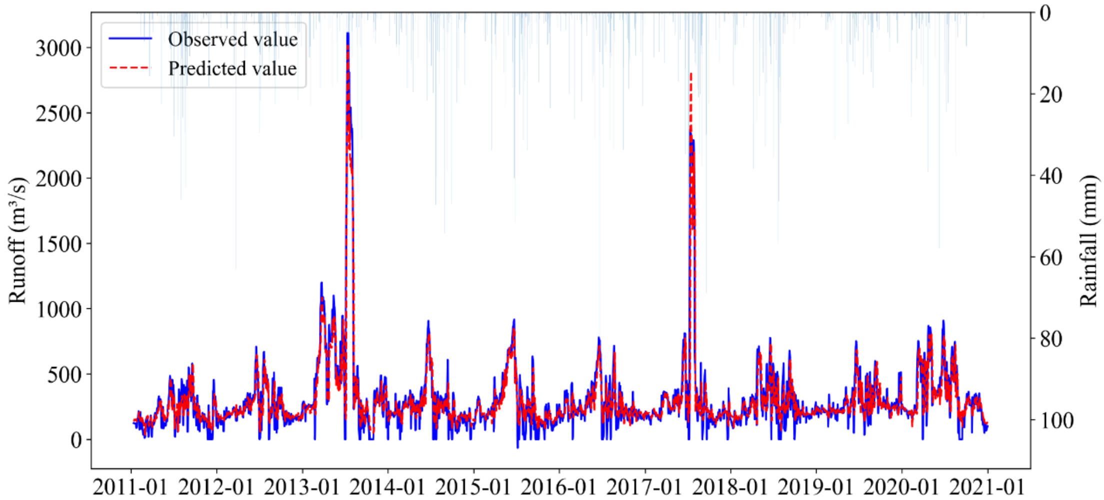


Figure 8: Rainfall-runoff prediction results at Erhezha station based on GNN-LSTM



5. Conclusion

Aiming to address the challenge that temporal dependency features and spatial connectivity relationships in the runoff evolution process under complex river-network conditions in the Huaihe River Basin are difficult to characterize simultaneously, this study constructed a GNN-LSTM mid- to long-term runoff forecasting model integrating LSTM and Graph Neural Network, based on multi-station daily meteorological and runoff data from 2011 to 2020. Through data preprocessing, basin graph structure construction, and multi-model comparative analysis, the applicability and accuracy of the model for runoff forecasting in complex river-network basins were examined. The main conclusions are as follows:

(1) Runoff variation in the Huaihe River Basin is jointly influenced by multiple meteorological factors and underlying surface conditions, exhibiting pronounced nonlinearity and complexity. Correlation analysis shows that the linear correlation between various meteorological elements and runoff is generally weak. This indicates that relying solely on a single factor or purely time-series methods is insufficient to fully characterize the runoff evolution patterns in complex river-network basins.

(2) Compared with LSTM, K-nearest neighbor, support vector regression, decision tree, and AdaBoost models, the GNN-LSTM model demonstrates superior comprehensive performance in mid- to long-term runoff forecasting for the Huaihe River Basin. Test results show that the coefficient of determination on the three regional test sets reached 0.915, 0.952, and 0.878, respectively. The model exhibits stronger fitting capability for the overall runoff variation trend, indicating that the introduction of temporal feature extraction and spatial topology modeling helps improve the accuracy of runoff prediction in complex river-network basins.

(3) From the prediction results at typical stations, the predicted values of the GNN-LSTM model are generally consistent with the observed values. The model not only reflects the phased fluctuation characteristics of the runoff process well but also demonstrates good simulation performance for the timing of flood peaks, trends in peak value changes, and low-flow processes during the dry season. This indicates that the model can effectively capture the spatiotemporal variation characteristics of runoff processes in complex river-network basins and possesses certain engineering application potential.

(4) The model constructed in this study provides a new modeling approach for mid- to long-term runoff forecasting in complex river-network basins, but there remains room for further optimization. Future research can incorporate more hydrometeorological influencing factors, more refined river network topological structures, and longer time-scale data to further improve the model's ability to characterize extreme hydrological events and regional differences in runoff response, thereby enhancing the model's generalization capability and practical application value.

References

[1] Chen, J., Xu, Q., & Cao, D. X. et al. (2024). A medium- to long-term runoff forecasting model based on multi-factor and multi-model ensemble. *Advances in Water Science*, 35(3), 408-419.

[2] Sun, Z. L., Liu, Y. L., & Zhang, J. Y. et al. (2023). Research progress and prospects of medium- to long-term runoff forecasting. *Water Resources Protection*, 39(2), 1-11.

[3] Su, H. D., Jia, Y. W., & Ni, G. H. et al. (2018). Application of machine learning in runoff forecasting. *China Rural Water and Hydropower*, (6), 40-43.

[4] Hu, L. Y., Fu, X. L., & Jiang, X. L. et al. (2024). Runoff forecasting based on three machine learning methods: LSTM, RF, and SVR. *Hydrology*, 44(5), 17-24.

[5] Xie, S., Huang, Y. F., & Li, T. J. et al. (2018). Medium- to long-term runoff forecasting using coupled LASSO regression and support vector regression. *Journal of Basic Science and Engineering*, 26(4), 709-722.

[6] Li, L. J., Wang, Y. T., & Hu, Q. F. et al. (2020). Long-term reservoir runoff forecasting based on random forest and support vector machine. *Journal of Water Resources and Water Engineering*, (4), 33-40.

[7] Zou, H. M., & Zhu, C. T. (2024). Comparative study on medium- to long-term reservoir inflow runoff prediction based on LSTM and BP neural network. *Hydrology*, 44(4), 27-31+37.

[8] Zheng, K. F., Ma, X. H., & Cui, G. T. et al. (2026). Runoff forecasting in the Beiliu River based on SAO-optimized LSTM model. *Water Resources and Hydropower Engineering* (Chinese and English), 57(2), 83-94.

[9] Kratzert, F., Klotz, D., Brenner, C., et al. (2018). Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*, 22, 6005-6022.

[10] Wang, M. Y., Wang, E. Z., & Luo, H. Q. et al. (2025). Graph neural network model for runoff forecasting in small and medium-sized basins: A case study of the Shaxi River Basin in Fujian. *Journal of Hydroelectric Engineering*, 44(6), 50-61.

[11] Yang, S., Zhang, Y., & Zhang, Z. (2023). Runoff prediction based on dynamic spatiotemporal graph neural network. *Water*, 15(13), Article 2463.

[12] Liu, Y., Hou, G., Huang, F., et al. (2022). Directed graph deep neural network for multi-step daily streamflow forecasting. *Journal of Hydrology*, 607, Article 127515.

[13] Xue, L., & Zhu, Y. (2025). Dynamic hydrological flow prediction with self-iterative spatiotemporal graph neural network: Modeling long- and short-period topological dynamics. *Journal of Hydrology*, 663, Article 134122.

[14] Yuan, R., Cai, S., Liao, W., et al. (2021). Daily runoff forecasting using ensemble empirical mode decomposition and long short-term memory. *Frontiers in Earth Science*, 9, Article 621780.

[15] Xu, B., Yang, F. G., & Li, Y. J. (2020). Application of two ensemble learning algorithms in medium- to long-term runoff forecasting. *Water Power*, 46(4), 21-24+34.

[16] He, F., Zhang, H., Wan, Q., et al. (2023). Medium term streamflow prediction based on Bayesian model averaging using multiple machine learning models. *Water*, 15(8), Article 1548.

[17] Kaur, S., & Chavan, S. R. (2025). Comparative analysis of deep learning and machine learning models for one-day-ahead streamflow forecasting in the Krishna River basin. *Journal of Hydrology: Regional Studies*, 60, Article 102549.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Environment, Social and Governance; Vol.3, No.3; 2026
E-ISSN: 3006-3973 P-ISSN：3006-3965
DOI: https://doi.org/10.70267/ esg2026.v3n3.1321
Published by: Zeus Press

Research on the Coupled and Coordinated Development of Green Finance and Low-Carbon Economy: Evidence from Central and Eastern Provinces in China

Yike Zhang*

Research Institute of Economics and Management (RIEM), Southwestern University of Finance and Economics, Chengdu 611130, China

**Corresponding author: Yike Zhang.*

Abstract

Against the backdrop of intensifying global climate change and the continued advancement of China's *Carbon Peaking and Carbon Neutrality Goals*, green finance and the low-carbon economy have become two core systems driving green economic transformation and high-quality development. Taking 16 provinces in eastern and central China as the research objects, this paper focuses on the interactive relationship between green finance and the low-carbon economy, constructs a comprehensive evaluation framework, and employs the entropy method and the coupling coordination degree model to analyze the development levels and synergistic status of the two systems in the sample regions from 2010 to 2022. The results show that: (1) the coupling coordination degree between green finance and the low-carbon economy generally exhibits an evolutionary pattern characterized by continuous improvement with slight local fluctuations, and the overall level in 2022 is significantly higher than that in 2010; (2) from a temporal perspective, the coupled and coordinated development of green finance and the low-carbon economy has roughly undergone three stages, namely the initial cultivation stage, the steady growth stage, and the rapid advancement stage, among which the improvement after 2020 is the most significant; (3) from a spatial perspective, Zhejiang, Tianjin, and Beijing are at relatively high levels, whereas Henan exhibits characteristics of relative stagnation in the low-carbon economy, rapid growth in green finance, but a mismatch between the two systems. Based on the findings, this paper proposes promoting green finance innovation by category, optimizing the layout of low-carbon industries, and strengthening regional policy coordination, so as to provide support for achieving the *Dual Carbon Goals*.

Keywords

green finance, low-carbon economy, coupling coordination degree, regional comparison, dual carbon goals

1. Introduction

Global warming is profoundly reshaping the modes of production and daily life in human society. The resulting extreme weather events, such as floods, droughts, heatwaves, and wildfires, have not only caused substantial economic losses but also posed long-term threats to ecological security and human well-being [1]. As a key factor contributing to global warming, carbon emissions have become a central governance target in the process of green transformation worldwide [2, 3]. As a responsible major country, China officially proposed in 2020 the *Carbon Peaking and Carbon Neutrality Goals* of peaking carbon emissions by 2030 and achieving carbon neutrality by 2060. Under this strategic background, green finance and the low-carbon economy have gradually become important instruments for promoting China's green transformation, high-quality development, and Chinese modernization. From a practical perspective, significant differences exist across regions in China in terms of resource endowments, levels of economic development, industrial structures, infrastructure conditions, and carbon emission characteristics. Benefiting from locational advantages, a strong foundation for industrial upgrading, and technological innovation capacity, the eastern region has long maintained a leading position in the development of green finance and the low-carbon economy. In contrast, the central region, relying on its energy resources and industrial transfer advantages, has accelerated green transformation under the impetus of the *Dual Carbon Goals* and has become an important hub connecting eastern and western China [4]. For this reason, examining the coupled and coordinated development of green finance and the low-carbon economy from a comparative perspective between eastern and central China is not only of strong practical significance but also of important policy value.

Therefore, taking 16 provinces in eastern and central China as the research objects, this paper systematically analyzes the overall trends, stage-specific characteristics, and regional echelon patterns of the coupled and coordinated development of green finance and the low-carbon economy based on a review of existing studies, and puts forward corresponding policy recommendations, with a view to providing references for optimizing green transformation pathways in different regions.

2. Literature Review

2.1 Research on the Measurement of Green Finance

Existing domestic studies generally adopt comprehensive evaluation systems to construct measurement indicators for green finance from dimensions such as scale, structure, and efficiency, with a focus on core financial instruments including green credit and green bonds, while gradually introducing a spatial analysis perspective to reveal regional differences [6, 7]. Existing studies show that the development level of green finance in China has generally exhibited an upward trend. The eastern region has improved rapidly, while the central region has also achieved steady development in the context of industrial transformation. The central and western regions as a whole demonstrate fluctuating growth, with regional disparities gradually narrowing. Foreign studies, by contrast, tend to place greater emphasis on cross-country comparative analyses under international standard frameworks, using ESG ratings and the issuance scale of green financial instruments as key measures [8, 9], which provides a reference for the continuous optimization of China's green finance measurement system.

2.2 Research on the Measurement of Low-Carbon Economy

Regarding the measurement of the low-carbon economy, existing studies have developed an analytical approach combining "single indicators and composite indices" [10]. Single-indicator studies usually focus on carbon emission intensity and carbon productivity, whereas comprehensive evaluation studies integrate multidimensional factors such as economic growth, industrial structure, technological innovation, and the ecological environment, and widely employ methods such as the entropy method and TOPSIS for estimation. From a practical perspective, the eastern region has achieved relatively outstanding progress in low-carbon economic development. The central region, relying on its industrial foundation and resource conditions, has accelerated low-carbon development under the guidance of the *Dual Carbon Goals*. However, due to the incomplete adjustment of industrial structure, a certain gap remains between its low-carbon economic efficiency and that of the eastern region.

2.3 Research on the Relationship Between Green Finance and Low-Carbon Economy

Existing studies generally suggest [11, 12] that green finance has a significant promoting effect on the low-carbon economy and can facilitate emission reduction and efficiency improvement through mechanisms such as capital allocation, incentives for technological innovation, and industrial structure optimization. The eastern region has accumulated substantial experience in cross-border coordination in green finance, targeted allocation of green credit, and financing support for low-carbon projects. Although the green finance system in the central region has been gradually improving, its driving effect has not yet been fully released due to constraints such as the concentration of financial resources, policy coordination, and industrial transfer capacity. Some studies also point out that in regions with underdeveloped financial systems and greater pressure for industrial transformation, the positive effect of green finance on the low-carbon economy may exhibit a certain time lag.

2.4 Literature Review and Commentary

Overall, existing studies have gradually expanded from the single measurement of green finance or the low-carbon economy to analyses of their relationship and interaction effects. Research methods have become increasingly mature, providing a solid foundation for this study. Meanwhile, several shortcomings remain in the existing literature. First, the measurement of green finance does not comprehensively cover instruments such as green insurance and green funds. Second, detailed studies on differences between eastern and central China remain relatively weak, particularly lacking comparative analyses that integrate resource endowments, industrial structure, and financial development capacity. Based on these gaps, this paper attempts to further investigate the coupled and coordinated development of green finance and the low-carbon economy from the comparative perspective of 16 provinces in eastern and central China.

3. Research Design

3.1 Research Objects and Sample Scope

This study selects 16 provinces, namely Zhejiang, Tianjin, Beijing, Jiangsu, Shandong, Hubei, Hunan, Hainan, Guangdong, Shaanxi, Fujian, Anhui, Shanghai, Hebei, Jiangxi, and Henan, as research samples, with the study period spanning from 2010 to 2022.

3.2 Construction of the Indicator System

This study defines green finance and the low-carbon economy as two relatively independent yet interrelated subsystems. A comprehensive indicator system for the green finance subsystem can be constructed from dimensions such as financial scale, financial structure, and financial efficiency. The low-carbon economy subsystem can be measured from dimensions including the quality of economic development, industrial structure optimization, technological innovation capability, energy utilization efficiency, and ecological-environmental constraints.

3.3 Entropy Weight Method

The entropy weight method is an objective weighting approach that assigns weights to indicators according to the magnitude of their information entropy. Based on the concept of information entropy, it measures the amount of information contained in each indicator by calculating its entropy value and then derives the corresponding indicator weights. This study adopts the entropy weight method, an objective weighting method, to determine indicator weights and further measure the levels of green finance and the low-carbon economy.

To eliminate the influence of different indicator dimensions, the range normalization method is used to process the raw data:

$$X'_{ij} = \frac{X_{ij} - \min(X_j)}{\max(X_j) - \min(X_j)} \tag{1}$$

where X'_{ij} denotes the standardized value of indicator j for province i, X_{ij} denotes the original value, and $\max(X_j)$and $\min(X_j)$ represent the maximum and minimum values of indicator j, respectively.

First, the proportion of province i under indicator j is calculated as:

$$P_{ij} = \frac{X'_{ij}}{\sum_{i=1}^{n} X'_{ij}} \tag{2}$$

Then, the information entropy of indicator j is calculated as:

$$e_j = \frac{1}{\ln n} \sum_{i=0}^{n} P_{ij} \ln P_{ij} \quad (3)$$

Further, the weight of indicator j is calculated as:

$$w_j = \frac{1-e_j}{\sum_{j=1}^{m} (1-e_j)} \quad (4)$$

where n is the total number of provinces, m is the total number of indicators, and wj denotes the entropy weight of indicator j, satisfying $\sum_{j=1}^{m} w_j = 1$.

Finally, the composite index is calculated as:

$$Y = \sum_{i=1}^{n} w_j \times X'_{ij} \quad (5)$$

where Y is the composite index, wj is the entropy weight of indicator j, and n is the total number of indicators. Its value ranges from [0,1], with a larger value indicating a higher development level.

3.4 Coupling Coordination Degree Model

The measurement of coupling coordination degree includes three levels: coupling degree, coordination degree, and coupling coordination degree. The coupling degree reflects the interaction intensity between the two systems:

$$C = 2 \times \sqrt{\frac{DE \times GLC}{(DE+GLC)^2}} \quad (6)$$

where C denotes the coupling degree, ranging from [0,1]. A larger value of C indicates a stronger interaction between the two systems.

The coordination degree reflects the synchronicity of development between the two systems:

$$T = \alpha \times DE + \beta \times GLC \quad (7)$$

where T is the coordinated development index, and α and β are weights to be determined. Considering the equal importance of the two systems, this study sets α=β=0.5.

The coupling coordination degree comprehensively reflects the synergistic development level of the two systems:

$$D = \sqrt{C \times T} \quad (8)$$

where D denotes the coupling coordination degree, ranging from [0,1]. A larger value of D indicates a higher level of coupled and coordinated development between the two systems.

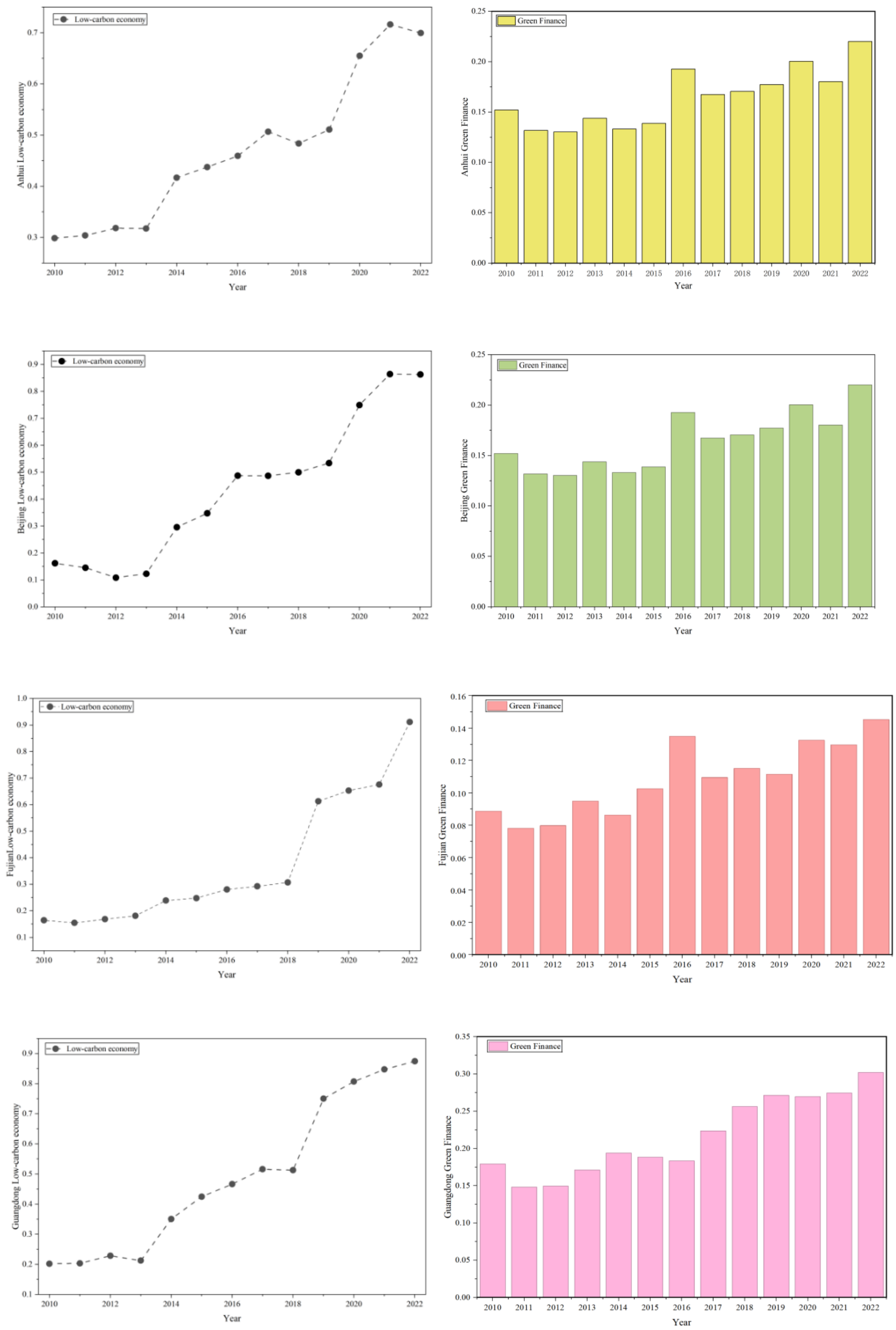
4. Analysis of Empirical Results

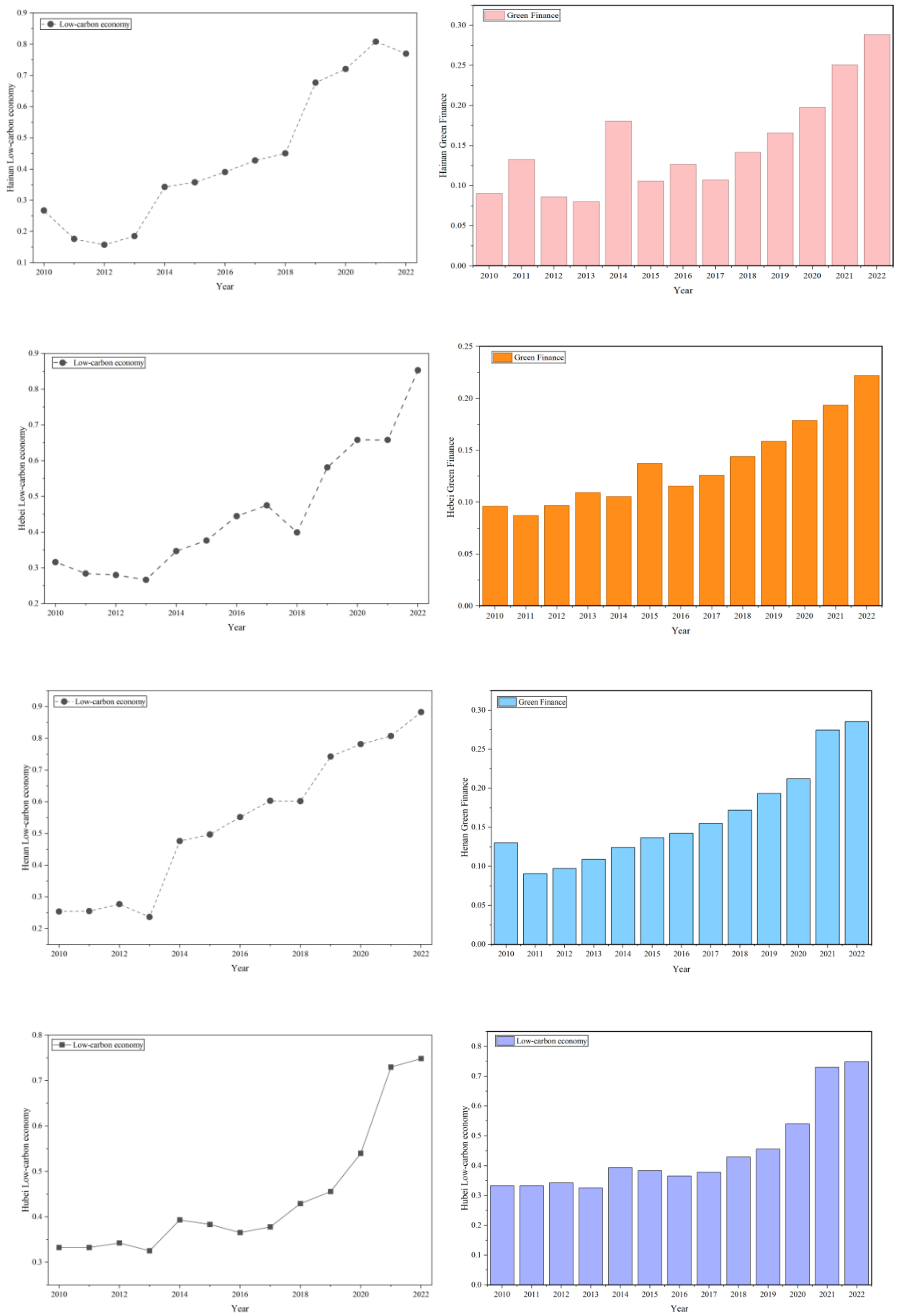
4.1 Evolutionary Characteristics of the Green Finance and Low-Carbon Economy Systems

From the overall trend, during the study period, the coupling coordination degree between green finance and the low-carbon economy in all provinces exhibited the core characteristic of “continuous and steady improvement with slight local fluctuations,” and the overall level in 2022 was generally significantly higher than that in 2010. This indicates that with the continuous deepening of the green development concept, the connection between the two systems of green finance and the low-carbon economy has become increasingly close, gradually moving from a relatively disconnected initial state toward a development trajectory characterized by mutual support and coordinated advancement. Due to space limitations, the main text presents the system evolution trends for eight provinces.

The development of the low-carbon economy and green finance across provinces exhibits differentiated characteristics while sharing overall commonalities. In most provinces, including Anhui, Beijing, Fujian, Guangdong, Hainan, Hebei, Hubei, Jiangsu, Jiangxi, Shandong, Shanghai, Shaanxi, Zhejiang, and Tianjin, the low-carbon economy generally shows a sustained upward trend, with only slight declines in a few individual years. Green finance also demonstrates overall steady growth with minor local fluctuations. Both systems have risen from relatively low levels at the beginning of the study period to higher levels at the end, with a significant improvement in coordination, generally reflecting a development pattern of “overall improvement plus local fluctuations.” Only Hunan and Tianjin exhibit the characteristic of “overall improvement plus differentiated local fluctuations.” Specifically, green finance in Hunan was slightly lower at the end of the period than at the beginning, while green finance in Tianjin experienced slight growth followed by stabilization. Henan is the only province showing a distinct pattern. Its low-carbon economy remained basically stable with minor fluctuations during 2010–2022, without obvious increases or declines, whereas green finance continued to rise significantly. The development pace of the two systems was clearly differentiated, and the overall coordination level exhibited a characteristic of “fluctuating divergence,” failing to achieve synchronized improvement.

Figure 1: Development Indices of the Low-Carbon Economy and Green Finance Across Provinces





4.2 Analysis of the Stage-Specific Characteristics of Coupling Coordination Degree

From a temporal perspective, the evolution of the coupling coordination degree can be broadly divided into three stages. The main text presents the changing trends in coupling coordination degree for nine provinces.

The first stage is the initial cultivation stage (2010–2014). During this period, both green finance and the low-carbon economy were in their nascent stage of development. The policy system had not yet been fully established, and market participants had limited awareness. The coupling coordination degree in most provinces remained within the low-level coordination range, indicating a weak development foundation for the two systems, limited inter-system linkage, and the insufficient manifestation of synergistic effects.

The second stage is the steady growth stage (2015–2019). With the gradual implementation of green finance pilot policies and the increasing diffusion of low-carbon transition concepts, provinces began to simultaneously promote industrial low-carbon transformation and the development of green finance systems. The overall coupling coordination degree rose to the moderate coordination range. The financial support role of green finance for the low-carbon economy gradually strengthened, and the synergy between the two systems improved significantly.

The third stage is the rapid advancement stage (2020–2022). Following the formal proposal of the *Dual Carbon Goals*, governments at all levels intensively introduced policies related to green development and low-carbon transition, while resources such as capital, technology, and human talent accelerated their concentration in green and low-carbon sectors. In most provinces, the coupling coordination degree exceeded 0.8, entering a stage of relatively high coordination or even high coordination. Zhejiang and Tianjin, in particular, approached the level of high-quality coordination, reflecting the significant outcomes generated by the combined efforts of policy support and market forces.

Figure 2: Coupling Coordination Degree Across Provinces

5. Conclusions and Policy Recommendations

5.1 Research Conclusions

Based on the estimation results for 16 provinces in eastern and central China from 2010 to 2022, this paper systematically analyzes the coupled and coordinated development of green finance and the low-carbon economy. The results show that the coupling coordination degree between green finance and the low-carbon economy in the sample regions has continuously improved overall. Although local fluctuations exist, the long-term upward trend is clear, and the two systems are gradually moving from relative disconnection toward coordinated advancement.

From a temporal perspective, the evolution of coupling coordination in the sample regions has undergone three stages: the initial cultivation stage, the steady growth stage, and the rapid advancement stage. Among these, the improvement after 2020 is the most significant, indicating that the *Dual Carbon Goals* and the improvement of the green development policy system have played a decisive role in promoting synergy between the two systems.

From a spatial perspective, there is a clear echelon differentiation among the sample provinces. Zhejiang and Tianjin are in leading positions, followed closely by Beijing and Jiangsu. Shandong, Hubei, and Guangdong fall within the medium-to-high coordination range, whereas Henan remains at a relatively low level due to industrial structure inertia and system mismatch. This suggests that the coupling coordination between green finance and the low-carbon economy is not achieved automatically, but is jointly influenced by differences in regional policies, industrial structures, energy endowments, and financial innovation capabilities.

5.2 Policy Recommendations

First, for regions with high and relatively high levels of coordination, efforts should continue to consolidate first-mover advantages by further deepening breakthroughs in green financial product innovation, carbon finance market development, cross-regional capital allocation, and research and development of core low-carbon technologies, thereby continuously enhancing high-level synergistic capacity and playing a national demonstration and leading role.

Second, for regions with medium-to-high levels of coordination, efforts should accelerate the low-carbon and clean transformation of traditional high-energy-consuming industries, strictly control the expansion of high-energy-consumption and high-emission projects, and vigorously foster green and low-carbon industries such as new energy, new materials, and high-end manufacturing. In the financial sector, innovative products such as transition finance and carbon account loans should be enriched to promote the deeper integration of green finance into the industrial transformation process.

Third, for regions with medium levels of coordination, development should be based on local resource endowments and industrial foundations, with accelerated industrial structure optimization, strengthened cultivation of low-carbon industries and project reserves, and active efforts to obtain national policy support for green development. Meanwhile, foundational systems such as green credit and green bonds should be improved to guide more financial resources toward low-carbon sectors and promote progress toward a higher stage of coordination.

Fourth, for regions with low levels of coordination, priority should be given to overcoming industrial structure inertia and insufficient project absorption capacity. Through formulating specialized low-carbon transformation plans, strengthening investment promotion, fostering low-carbon industrial clusters, and improving green finance incentive mechanisms, the chain of "financial support–industrial implementation–economic transformation" should be opened up, thereby gradually enhancing the systemic synergy between green finance and the low-carbon economy.

References

[1] Gan, Z. K., Yao, J. Y., & Chen, A. G. (2025). Research on atmospheric pollution hazards and environmental governance strategies under the background of global warming. *Heilongjiang Environmental Bulletin, 38*(6), 74–76.

[2] Werning, M., Hooke, D., Krey, V., Riahi, K., van Ruijven, B., & Byers, E. A. (2024). Global warming level indicators of climate change and hotspots of exposure. *Environmental Research: Climate, 3*(4), 045015. https://doi.org/10.1088/2752-5295/ad8300

[3] Li, X., Zhang, L. L., & Chen, H. B. (2025). Exploring pathways for achieving China's carbon peaking and carbon neutrality goals: Based on system dynamics simulation analysis. *Price: Theory & Practice*, (10), 177–183. https://doi.org/10.19851/j.cnki.CN11-1010/F.2025.10.347

[4] Zhao, Y., & Jia, L. P. (2026). Research on the impact of the synergy between fintech and green finance on economic resilience. *Modern Finance and Economics (Journal of Tianjin University of Finance and Economics)*, (4), 111–130. https://doi.org/10.19559/j.cnki.12-1387.2026.04.007

[5] Li, J., & Meng, Q. (2026). Financing green innovation: When digitalization amplifies green finance. *Finance Research Letters, 98*, 109837. https://doi.org/10.1016/j.frl.2026.109837

[6] Wang, J., Ye, B., He, Z., Pu, H., Su, B., & Lu, Y. (2026). Does green finance ensure energy security while achieving low-carbon transformation of listed electricity firms? Evidence from China. *Energy Economics, 153*, 109092. https://doi.org/10.1016/j.eneco.2025.109092

[7] Zou, X., Liu, S., & Yang, L. (2025). Green finance-driven and low-carbon energy transition: A tripartite game-theoretic and spatial econometric analysis based on evidence from 30 Chinese provinces. *Sustainability, 17*(21), 9474. https://doi.org/10.3390/su17219474

[8] Habib, Y., Abd Rahman, N. R., Hashmi, S. H., & Ali, M. (2025). Green finance and environmental decentralization drive OECD low-carbon transitions. *Scientific Reports, 15*(1), 28140. https://doi.org/10.1038/S41598-025-11967-Y

[9] Ru, J., Gan, L., & Yusufu, G. (2025). Ripple effect or spatial interaction? A spatial analysis of green finance and carbon emissions in the Yellow River Basin. *Sustainability, 17*(10), 4713. https://doi.org/10.3390/su17104713

[10] Monasterolo, I., Mandel, A., Battiston, S., Mazzocchetti, A., Oppermann, K., Coony, J., ... Dunz, N. (2024). The role of green financial sector initiatives in the low-carbon transition: A theory of change. *Global Environmental Change, 89*, 102915. https://doi.org/10.1016/j.gloenvcha.2024.102915

[11] Freire, V. D., Vieira, R. G., & Dhimish, M. (2026). Green electrification of agroindustry: Technical, economic and environmental performance of solar photovoltaic integration in Brazilian rural production. *Energy for Sustainable Development, 93*, 101992. https://doi.org/10.1016/j.esd.2026.101992

[12] Nnabuife, S. G., Hamzat, A. K., Darko, C. K., Udemu, C., Quainoo, K. A., Yang, J., & Kuang, B. (2026). Low-carbon energy storage via liquid organic hydrogen carriers: From new developments to global initiatives. *Next Energy, 11*, 100533. https://doi.org/10.1016/j.nxener.2026.100533

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Environment, Social and Governance; Vol.3, No.3; 2026
E-ISSN: 3006-3973 P-ISSN：3006-3965
DOI: https://doi.org/10.70267/ esg2026.v3n3.2234
Published by: Zeus Press

Digital Carbon Management and ESG Greenwashing Risk: A Case Study of BYD's Sustainability Disclosure

Rongrong Zhang*

School of Professional Education and Executive Development, The Hong Kong Polytechnic University, Hong Kong, China

**Corresponding author: Rongrong Zhang.*

Abstract

This paper explores the effect of digital carbon management on ESG greenwashing in corporate sustainability reporting. The paper uses a qualitative single case research design and preliminary textual analysis based on the text of BYD's 2025 Sustainability Report. The results indicate that digital carbon management can lessen the risk of greenwashing through increase in data traceability, disclosure specificity, carbon accounting capability, embedding of governance, and readiness for verification. BYD's iDi Carbon Chain platform is an example of a pathway to transparency, achieved by organizational carbon management, product carbon footprint accounting and compliance reporting. However, the analysis also shows that symbolic sustainability narratives remain central in ESG communication. The paper thus seeks to show that the anti-greenwashing effect of digitalisation is not self-evident. It is dependent on the embedding of digital tools in external standards and board-level governance and verification mechanisms.

Keywords

ESG disclosure, greenwashing risk, digital carbon management, sustainability reporting, new energy vehicles

1. Introduction

1.1 Research Background

Environmental, social and governance (ESG) disclosure has emerged as a critical means for companies to convey their sustainability plans, environmental impact and governance obligations to investors, regulators, consumers and other stakeholders. Christensen, Hail and Leuz [1] state that sustainability reporting has the potential to mitigate information asymmetries between a company and outsiders. However, the usefulness of information reported in sustainability reports depends on the credibility, comparability and decision usefulness of the information disclosed. The global survey carried out by KPMG [2] also indicates a shift from voluntary sustainability reporting to a normal part of corporate life. This implies that ESG disclosure isn't just supposed to be a communication tool, it should be a way for companies to show accountability.

However, the rapid growth of ESG reporting has also intensified concerns about greenwashing. Delmas and Burbano [3] describe greenwashing as a situation in which firms present a positive environmental image while their actual environmental performance may not fully support such claims. Lyon and Montgomery [4] further conceptualize greenwash as a broad category of potentially misleading environmental communication, including selective disclosure, vague claims and symbolic language. In this regard, the disclosure of ESG can have an informational role as well as a role to legitimize the strategy. In Suchman's legitimacy theory [5], the author also argues that companies can obtain social legitimacy by turning the information they communicate to the public into socially acceptable information.

This is especially topical when considering the new energy vehicle industry. Many new energy vehicle companies are seen as inherently linked to environmental transition as their products advance the transition towards electrification of transport and carbon reduction. But this “green” brand also can lead to increased stakeholder expectations. For these companies, ESG credibility is not only about the ESG impact of their products and services, but about how transparent their operations and supply chains are, as well as their carbon accounting systems and sustainability governance.

1.2 Case Selection and Research Question

BYD provides an interesting use case to explore this relationship. BYD is a leading Chinese new energy vehicle company with business segments of automobiles, electronics, new energy and rail transit and is headquartered in Shenzhen. In 2025, more than 4.6 million new energy vehicles were sold, and the company's 2025 Sustainability Report shows that it has reduced the carbon footprint by an estimated 46.6 million tonnes in comparison with the carbon footprint of traditional fuel vehicles within the scope of the Sustainability Report; it also indicates that the company has consumed 7.29 billion kWh of clean electricity and invested RMB 2.04 billion in environmental protection in 2025 [6]. These disclosures form a bedrock to discuss the connection between substantive ESG evidence and symbolic sustainability communication.

Digital carbon management could be one solution to the greenwashing risk. There are several options available to companies to enhance the traceability, specificity and verifiability of ESG information through digital platforms, automated data collection and carbon accounting tools and product carbon footprint systems. However, going digital doesn't make greenwashing risk go away. Digital tools can likewise contribute to the development of a more compelling sustainability story from companies, a more complex visual reporting and more appealing ESG communication. Therefore, digital carbon management may have a dual effect. It can reduce greenwashing risk by strengthening evidence-based disclosure, but it may also coexist with symbolic ESG storytelling.

This paper asks how does digital carbon management influence ESG greenwashing risk in corporate sustainability disclosure? The study does not claim that BYD engages in greenwashing. Instead, it uses BYD as an exploratory case to understand how digital carbon management may reshape the relationship between ESG disclosure credibility and greenwashing risk. The central argument is that digital carbon management reduces greenwashing risk when it is supported by data traceability, external standards, governance structures and verification mechanisms, while symbolic ESG narratives remain important in corporate sustainability communication.

2. Literature Review

This section reviews three streams of research that are relevant to the study: ESG disclosure and greenwashing, digital transformation and ESG disclosure, and reporting standards, governance and verification. These streams provide the theoretical foundation for the dual-path framework developed in the next section.

2.1 ESG Disclosure and Greenwashing

ESG disclosure is widely expected to improve transparency between corporations and stakeholders. In principle, sustainability reports help stakeholders assess firms' environmental and social impacts, governance systems and long-term risk exposure. However, prior research warns that sustainability reporting may not always produce substantive accountability. Cho, Laine, Roberts and Rodrigue [7] argue that sustainability reports can become organizational facades when firms face contradictory institutional expectations and attempt to manage public impressions. Boiral [8] similarly suggests that some highly rated sustainability reports may reproduce idealized representations rather than fully reflect organizational realities .

The distinction between substantive disclosure and symbolic disclosure is central to greenwashing research. Substantive disclosure provides specific, measurable and verifiable information. Symbolic disclosure relies more heavily on broad values, narratives and future-oriented commitments. Michelon, Pilonato and Ricceri [9] argue that CSR reporting practices, including stand-alone reports, assurance and reporting guidance, do not automatically improve disclosure quality; their effect depends on how firms use them . Hummel and Schlick [10] further show that the relationship between sustainability performance and sustainability disclosure is theoretically ambiguous because voluntary disclosure theory and legitimacy theory imply different reporting incentives .

Greenwashing risk can therefore be considered a disclosure credibility problem. ESG disclosure does not in itself prove the authenticity of ESG. It's not about whether disclosure is general or whether it's detailed, it's about whether disclosure is specific, balanced, comparable and verifiable, and whether there are real governance and performance systems that relate to the disclosure. This paper is an extension of this perspective and explores the potential of digital carbon management as a means to enhance the credibility of disclosures.

2.2 Digital Transformation and ESG Disclosure

Digital transformation has the potential to transform the way ESG data is produced and governed. Digital platforms facilitate the businesses to gather data on the environment from factories, supply chains, product life cycle processes and energy systems. Burritt and Schaltegger [11] stress that sustainability information is more significant if linked to tangible information and management decisions. Based on this logic, digital carbon management can create more value for the management usefulness of ESG disclosure by linking the environmental claims with measurable operational data.

Digital technologies are also facilitating transparency of supply chains. Bai and Sarkis [12] claim that technology like blockchain can enhance the transparency and sustainability assessment of the supply chain, making it more traceable and boosting the confidence of the stakeholders. This paper does not specifically address blockchain, but the principle is the same: digital systems are better suited to make data more traceable, auditable and comparable in order to significantly improve the transparency of the situation. Digital carbon management could thus be a disclosure infrastructure of interest for companies with extensive manufacturing processes and global value chains.

But digitalization can also lead to a more complex communication situation. The ESG reports are becoming more and more visual, structured and narrative. The claims of sustainable products can be more convincing with digital communication, but not necessarily more verifiable. However, the connection between digital transformation and ESG greenwashing risk should not be taken for granted. The digitalization can have a positive effect on data quality but can also enhance the symbolic communication. This tension is the impetus behind the framework of dual-paths in this paper.

2.3 Reporting Standards, Governance and Verification

Reporting standards and governance systems contribute to the credibility of the ESG disclosure. According to Ioannou and Serafeim [13] mandatory sustainability reporting can have an impact on the disclosure practices and orientation to social responsibility of a firm. Global Reporting Initiatives (GRI) Standards are based on the principles of accuracy, balance, clarity, comparability and verifiability [14]. The orientation of sustainability disclosures and climate disclosures towards investors is echoed in the ISSB's IFRS S1 and IFRS S2 [15]. These frameworks facilitate the consistency of ESG reporting in terms of language and structure.

In the case of carbon related disclosure, verification is especially critical. ISO 14064-1 specifies the requirements and guidelines for quantifying and reporting GHGs for an organization and ISO 14067 specifies the requirements and guidelines for quantifying the GHGs of a product [16, 17]. Digital carbon management systems aligned with such standards can enhance the credibility of carbon data by providing more structure and verifiability. But it's not just standards alone. These should be part of internal governance systems, external verification processes and meaningful scrutiny by stakeholders.

3. Conceptual Framework

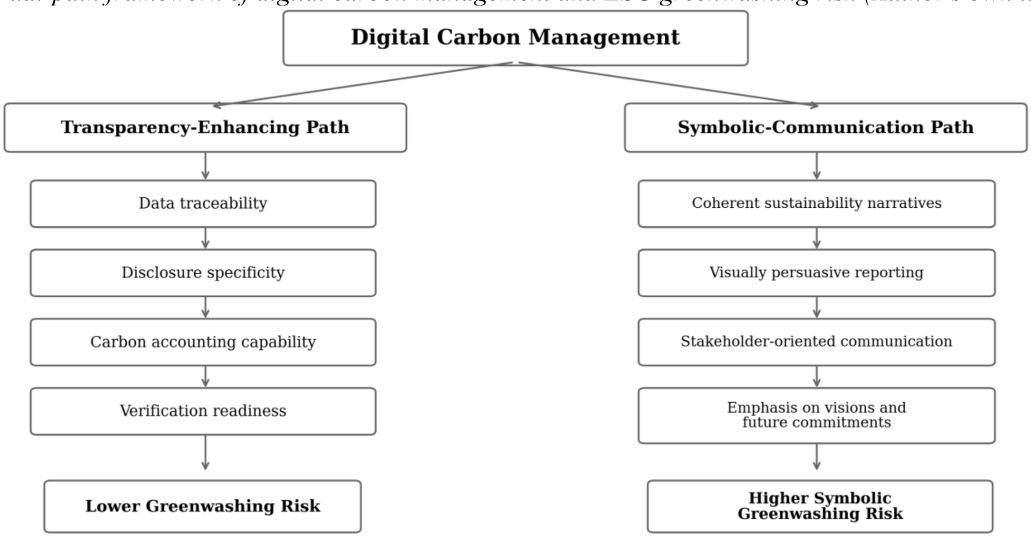
This section develops a dual-path framework that links digital carbon management to ESG greenwashing risk. The framework explains why digitalization may both strengthen disclosure accountability and support symbolic communication.

3.1 A dual-path Perspective

The first one is the transparency-enhancing path. Digital carbon management can mitigate the risk of greenwashing by enhancing data traceability, specificity of disclosures, carbon accounting capacity and preparedness for verification. This route is based on using digital tools to assist with accountability as they enable measurable evidence of environmental claims. The second one is the symbolic-communication path. Digital ESG communication can contribute to bolstering corporate legitimacy as it renders ESG communication more coherent, persuasive and visually appealing for stakeholders. As part of this route, digital instruments can also be used alongside symbolic disclosure if the reporting focuses on visions, values and commitments rather than on measurable performance.

The question isn't really whether symbolic narratives exist or not, but rather whether there is any evidence that balances the symbolic narratives. Corporate reporting is mostly symbolic storytelling, as the companies have to share values, missions and long-term strategies. When symbolic communication is however not correlated with measurable performance, it can be an additional risk for greenwashing. Thus, the trustworthiness of digital ESG disclosure hinges on the balance between the intensity of the narrative and the specificity of the evidence.

Figure 1: Dual-path framework of digital carbon management and ESG greenwashing risk (Author's own illustration)



3.2 Propositions

Based on the dual-path framework, this study develops three propositions that guide the textual analysis of BYD's sustainability disclosure.

Proposition 1: Digital carbon management reduces ESG greenwashing risk by improving the traceability, specificity and verifiability of environmental disclosure.

Proposition 2: Digitalized ESG communication may increase symbolic greenwashing risk when narrative intensity exceeds evidence-based disclosure.

Proposition 3: The greenwashing-reducing effect of digital carbon management is stronger when digital disclosure is supported by external standards, board-level governance and third-party verification.

These propositions are not tested through causal econometric methods in this study. Instead, they provide a conceptual framework for interpreting preliminary textual evidence from BYD's sustainability disclosure.

4. Methodology

This section explains the research design, data source, coding scheme and scoring method used to analyze the case. The study is designed as an exploratory qualitative case study with preliminary textual evidence.

4.1 Research Design and Data Source

The research design of this study is a qualitative single-case research, using preliminary textual analysis. A single-case design is suitable as the research purpose is exploratory, aimed at gaining an understanding of the representation of digital carbon management in sustainability disclosure and how it could be associated with ESG greenwashing risk. BYD is chosen as an information-rich case, as it is a front runner in the new energy vehicle sector, has high sustainability visibility and provides a specific digital carbon management platform.

The main source of data is BYD's Sustainability Report 2025. The report covers the time frame from 1st January 2025 to 31st December 2025. The report is acceptable for textual analysis as it has substantive as well as symbolic aspects. Substantive elements involve carbon emission information, utilization of clean electricity, environmental investments, zero carbon industrial parks, green factories, governance structures and references to ISO related carbon accounting. There are symbolic elements like the visions of 'cooling the Earth by 1 degree Celsius', three green dreams and the DREAMS sustainability concept. The study is not intended to provide verification of BYD's environmental performance. Rather, it considers how disclosure attributes can enhance or diminish greenwashing risk.

4.2 Coding Scheme and Scoring Method

There are five categories for coding the texts. The categories aim to differentiate between disclosure that is backed by evidence and symbolic sustainability communication. A summary of the coding scheme is provided in Table 1. The keyword hits listed in Table 3 were created via a text extraction and keyword matching procedure using Python, to make the keyword-assisted stage replicable. PyMuPDF (fitz) was used to convert the PDF report to machine-readable text, and category-specific keyword lists were applied to the text. Since the source report is mainly a Chinese language document, both the English and the Chinese equivalent of the key words were included. The counts were not used to provide independent statistical evidence, but to inform the identification of the textual presence of appropriate themes and to inform qualitative coding.

Table 1: Coding scheme for preliminary textual analysis

Code	Category	Definition	Examples in the BYD case
A	Digital carbon management	Disclosure related to digital carbon platforms, automated accounting, carbon data monitoring and product carbon footprint systems.	iDi Carbon Chain; automated carbon data calculation; product carbon footprint accounting.
B	Substantive ESG disclosure	Quantified or evidence-based ESG information, including emissions, energy use, investment, targets and performance indicators.	Clean electricity use; environmental investment; Scope 1 and Scope 2 emissions; carbon intensity.
C	Standards and verification	References to reporting standards, ISO certification, third-party verification or external compliance mechanisms.	GRI; ISSB; ESRS; ISO 14064; ISO 14067; assurance mechanisms.
D	Governance embedding	Texts showing ESG governance structures, board oversight, committees, internal control or accountability mechanisms.	Board responsibility; Strategic and Sustainability Committee; ESG Management Committee.
E	Symbolic ESG narrative	Broad sustainability language, vision statements, legitimacy claims and corporate identity narratives.	Three green dreams; cooling the Earth by 1 degree Celsius; DREAMS concept.

In addition, the study applies an interpretive disclosure specificity score. The purpose of this score is not to produce a statistically representative measure, but to provide a transparent way to compare different types of disclosure. The scoring system is shown in Table 2.

Table 2: Disclosure specificity score

Score	Meaning	Greenwashing-risk implication
0	Pure symbolic narrative with no concrete action or data.	Higher symbolic ambiguity.
1	Action description exists, but without quantitative evidence.	Some substantive content, but limited verifiability.

Score	Meaning	Greenwashing-risk implication
2	Quantitative evidence is provided, but external verification or governance linkage is limited.	Lower ambiguity, but credibility still depends on scope and method.
3	Quantitative evidence is supported by standards, certification, verification or governance mechanisms.	Stronger evidence base and lower greenwashing risk.

5. Case Background

This section presents the case background of BYD and explains why its sustainability disclosure provides a useful setting for analyzing digital carbon management and greenwashing risk.

5.1 BYD and the New Energy Vehicle Industry

BYD describes itself as a technology-driven company committed to green development and innovation. The company was established in 1994 and now the company is based in Shenzhen. It operates in the automotive industry, electronics, new energy and rail transportation and is listed both in Hong Kong and in Shenzhen. The report describes a company which combines electric mobility, renewable energies, storage and technological solutions within a larger sustainability plan: BYD.

The new energy vehicle industry is closely related to carbon reduction and transportation electrification, which are key aspects of ESG research. But, when this link occurs it can be dangerous because it gives the appearance of legitimacy. Automatic labelling of new energy vehicle companies as green, could also make stakeholders less aware of operational emissions and supply chain risks, battery recycling, resource usage and manufacturing impacts. Thus, disclosures on sustainability at this sector should not only concern the environmental value of the products, but also evidence on organizational carbon management and governance.

5.2 BYD Sustainability Disclosure

BYD's 2025 Sustainability Report emphasizes formal standardization. It states that it has been prepared based on the Hong Kong Stock Exchange ESG Reporting Code and Shenzhen Stock Exchange sustainability reporting guidance, and is consistent with ESRS, GRI Standards, ISSB IFRS S1 and IFRS S2, SDGs and ISO IWA 48:2024. This means that it sits in a number of domestic and international disclosure frameworks.

The report also highlights the reporting principles such as materiality, quantification, balance and consistency. Furthermore, it also mentions GRI principles including accuracy, balance, clarity, comparability, completeness and sustainability context, as well as timeliness and verifiability. The principles are also pertinent to greenwashing risk, as they directly tackle recognised weaknesses of the symbolic disclosure, such as low verifiability, selective reporting and vagueness.

5.3 iDi Carbon Chain Platform

The iDi Carbon Chain digital carbon management platform is a key component of BYD's sustainability reporting. The report says that the platform was created to meet the needs of the digital and intelligent transformation of corporate carbon management. It emphasizes carbon modules of an organization and of products and facilitates monitoring operational carbon emissions, accounting for product life cycle carbon footprints, automated calculation of vehicle carbon data, compliance reporting on an international scale, etc. [6].

The report also claims that BYD adopted the iDi Carbon Chain platform to do intelligent data statistics and calculation on greenhouse gases, and that the platform passed the ISO14064 certification. It also says that the platform was ISO 14067 certified with regards to accounting the product carbon footprint. These assertions are important as they take ESG disclosure beyond a generalised sustainability story into a more data-centric approach to carbon management.

6. Textual Analysis and Preliminary Evidence

This section strengthens the empirical component of the paper by moving beyond a small set of representative examples. The analysis treats BYD's 2025 Sustainability Report as a structured disclosure text and examines it through systematic qualitative coding, keyword-assisted evidence mapping and interpretive disclosure specificity scoring. The purpose is not to prove whether BYD engages in greenwashing, but to evaluate how different types of disclosure may reduce or leave residual ESG greenwashing risk. In line with the conceptual framework, the analysis distinguishes transparency-enhancing evidence from symbolic sustainability communication.

6.1 Coding Procedure and Keyword-assisted Screening

The textual analysis is based on a three-step qualitative procedure that is assisted by Python. First, the PDF version of BYD's 2025 Sustainability Report was converted into machine-readable text by using PyMuPDF (fitz). Second, Python keyword matching was used with five category-specific keyword lists to the extracted text. The categories have been listed according to the coding scheme: Digital carbon management; substantive environmental metrics; standards and verification; governance embedding; and symbolic sustainability narratives. Third, the textual segments located were qualitatively interpreted and scored using the disclosure specificity scale (0 – 3). The higher the score, the more evidence, stronger connections to standards or verification, and stronger governance support are found in the disclosure.

The number of keywords is employed as a screening tool and not as a statistical test. The list of keywords was tightened as this would otherwise overcount very broad terms like data, platform or sustainability, as general reference. Some terms are common in more than one section and translation equivalents may be found in a variety of forms; the count should be considered as suggestive of the visibility of a theme, rather than as a precise measure of the quality of disclosure. The use of Python in screening, however, enhances the transparency of the text analysis process as the evidence selection procedure becomes more explicit and replicable.

Table 3: Keyword-assisted textual screening of BYD's 2025 Sustainability Report

Analytical category	Main keywords used for screening	Keyword hits	Interpretation
Digital carbon management	iDi Carbon Chain; digital carbon management; intelligent carbon management; automated carbon calculation; product carbon footprint; organizational carbon management	123	The stricter Python screening shows that digital carbon management is a visible but focused theme, supporting the analysis of carbon data infrastructure rather than broad digital rhetoric.
Substantive environmental metrics	carbon emission; carbon reduction; clean electricity; green electricity; environmental investment; zero-carbon park; green factory; Scope 1; Scope 2; carbon neutrality	275	Environmental metrics are strongly represented, suggesting that the report contains measurable disclosure rather than only general sustainability claims.
Standards and verification	ISO 14064; ISO 14067; GRI; ISSB; ESRS; TCFD; third party; assurance; certification; verification	416	Verification- and standards-related language is highly visible, supporting the argument that formal standards are central to disclosure credibility.
Governance embedding	board; Strategic and Sustainability Committee; ESG Management Committee; risk management; internal control; performance assessment; remuneration; training	390	Governance-related language is prominent, indicating that ESG disclosure is repeatedly connected with oversight, accountability and incentive mechanisms.
Symbolic sustainability narratives	green dreams; cooling the Earth by 1 degree Celsius; DREAMS; vision; mission; responsibility; shared value; green innovation	414	Symbolic and legitimacy-oriented narratives remain highly prominent, supporting the symbolic-communication path in the conceptual framework.

Note. Keyword hits were generated through a Python-based keyword-assisted textual screening procedure using the extracted text of BYD's Chinese-language 2025 Sustainability Report. The keyword list included both English terms and Chinese equivalents, but the table reports English labels only. The counts are used only to guide qualitative coding and should not be interpreted as independent statistical evidence.

6.2 Evidence of Transparency-enhancing Disclosure

BYD's iDi Carbon Chain platform is the best proof of the transparency-enhancing route. The report describes the platform as a digital carbon management infrastructure and not a communications platform. It's said to enable organizational carbon management and product carbon modules, operational carbon emission monitoring, product life cycle carbon footprint accounting, automated vehicle carbon data calculation and international compliance reporting.

There are two reasons why this evidence is significant. First, it takes digitalization and makes it carbon accounting capable. The platform is not only used to present ESG information, but also to collect, calculate and manage carbon-related data. Secondly the platform is connected to the external standards. The report mentions a platform that is certified to ISO 14064 and ISO 14067 for better readiness in verification. Thus, iDi Carbon Chain platform evidence is that digital carbon management can help mitigate the risk of greenwashing through enhanced traceability and verifiability.

But this conclusion should be noted with caution. The fact that there is a digital platform does not mean that all environmental claims are fully verified. Instead, it indicates that BYD is building a disclosure infrastructure that is capable of more evidence-based ESG disclosure.

Detailed textual evidence and disclosure specificity scores are provided in Appendix A to support the interpretation of the transparency-enhancing and symbolic-communication paths.

6.3 Quantified Environmental Disclosure and Specificity

BYD's report contains multiple quantified environmental indicators. These include clean electricity use, environmental protection investment, carbon emissions, carbon intensity, green factories, zero-carbon parks and carbon reduction estimates [6]. These quantified indicators improve disclosure specificity because they allow stakeholders to evaluate ESG claims through measurable information. Compared with broad statements such as a general commitment to green development, quantified indicators reduce ambiguity and make the disclosure more comparable over time.

Nevertheless, quantified disclosure does not automatically eliminate greenwashing risk. Numbers can create a perception of objectivity, but their credibility still depends on reporting boundaries, calculation assumptions, data sources and verification mechanisms. For example, carbon reduction estimates depend on comparison baselines and usage assumptions. Therefore, the textual evidence suggests a reduction in symbolic ambiguity, but not the complete removal of greenwashing risk.

6.4 Governance Embedding and Accountability Mechanisms

The report provides evidence that carbon management is embedded in governance structures. BYD describes a climate governance architecture involving the board of directors, the Strategic and Sustainability Committee, the ESG Management Committee, the ESG Sustainability Department and ESG personnel within business groups and divisions. This structure indicates that sustainability disclosure is connected to internal governance rather than remaining only a public communication exercise.

The report also mentions that climate-related information is distributed to the board members, that board members and management receive training on climate change and that sustainability performance indicators are tied to the remuneration of the executive and senior managers. This is because, incentive alignment increases the relationship between disclosure and the accountability of the managers. Governance embedding has a moderating effect from a greenwashing-risk perspective. Carbon data connected to board oversight, risk management, internal reporting and performance incentives will be more likely to mitigate greenwashing risk in digital carbon management.

6.5 Symbolic Sustainability Narratives and Legitimacy Building

The report also has a powerful symbolic sustainability story. There is also recurring language used to mark identity building which include concepts like the three green dreams, DREAMS, cooling the Earth by 1 degree Celsius, green innovation, responsibility and shared value. These can be expressions in a positive way. Symbols narratives are frequently used as a method of communicating values, mission and long-term strategic orientation in corporate sustainability reports.

But when symbolizations are not complemented by quantifiable proof, they do have relevance in the context of the risk of greenwashing. The BYD case demonstrates that there can be both symbolic and substantive disclosure. On the other hand, the report assures evidence-based disclosure via carbon data, ISO-related standards, product carbon footprint accounting and governance structures. It also, however, applies potent legitimacy-building narratives in building a green corporate identity. This is a finding that corroborates the symbolic-communication path as digital ESG communication can lead to increased corporate legitimacy due to the more coherent and convincing sustainability narratives.

6.6 Disclosure Boundary and Residual Greenwashing Risk

There are also disclosure limitations that remain, with a more critical reading of the report. First, there is not a complete coverage of the environmental KPI boundary for all overseas bases and sales stores. This boundary disclosure is transparent, but also indicates that this data on the environment cannot be interpreted as a full global operational footprint. Second, other information related to climate change transactions, for example costs and revenues from carbon trading, is not reported because of business sensitivity. Third, a number of carbon-related disclosures are tied to standards and verification, but not all environmental disclosures are created equal and not all are equally assured by a third party.

These restrictions are not to be interpreted as greenwashing by BYD. Instead, they illustrate why the paper should refer to greenwashing risk, rather than greenwashing behavior. The textual evidence indicates that digital carbon management can lead to a reduction of the risk of greenwashing, but that risks persist in cases of incomplete disclosure boundaries, assumptions or verification coverage.

Table 5: Pathway-level synthesis of the textual analysis

Pathway	Supporting evidence	Risk-reduction mechanism	Residual concern
Transparency-enhancing path	iDi Carbon Chain; ISO 14064/14067; Scope 1 and Scope 2 emissions; clean electricity use; climate governance; ESG-linked remuneration	Improves traceability, specificity, carbon accounting capability, governance embedding and verification readiness.	Disclosure still depends on boundary, calculation method and assurance coverage.
Symbolic-communication path	Three green dreams; DREAMS; cooling the Earth by 1 degree Celsius; responsibility and shared value language	Strengthens corporate legitimacy and communicates sustainability identity.	Symbolic narratives can increase ambiguity when not sufficiently anchored in measurable evidence.
Overall implication	The report contains both substantive evidence and symbolic narratives.	Digital carbon management can reduce greenwashing risk when linked to standards, governance and verification.	Digitalization should not be treated as automatically anti-greenwashing.

7. Discussion

This section discusses the theoretical and practical implications of the findings. It explains how the BYD case contributes to the literature and what it suggests for companies, regulators and investors.

7.1 Digitalization is Not Automatically Anti-greenwashing

The BYD case suggests that digitalization should not be treated as inherently anti-greenwashing. Digital carbon management can minimise the risk of greenwashing if it enhances data traceability, carbon accounting and readiness to analyse and verify carbon impact of products. But digitalized ESG communication can also help to make sustainability narratives more convincing. The difference lies in the fact that digital tools are being used mainly for data governance or impression management.

This distinction enables to reconcile two strands in the literature. On the other hand, research in the field of sustainability accounting underscores the need to have measurable environmental information to use in decision making [11]. However, greenwashing and legitimacy research make it known that sustainability reports can also serve to manage stakeholder perception. Both mechanisms can occur at the same time as illustrated in the BYD case. The symbolic sustainability narrative remains influential in terms of legitimacy, and digital carbon management can enhance accountability.

7.2 From ESG Storytelling to ESG Accountability

A key theoretical implication of this study is the shift from ESG storytelling to ESG accountability. ESG storytelling refers to the use of narratives, visions and values to communicate sustainability identity. ESG accountability refers to the provision of specific, measurable, comparable and verifiable evidence. Effective sustainability disclosure needs both, but the balance matters. Storytelling without accountability creates greenwashing risk. Accountability without narrative may be technically accurate but less understandable to stakeholders.

Digital carbon management can help bridge this gap. It allows companies to support sustainability narratives with carbon data, product footprint results, automated calculations, governance processes and external standards. In BYD's case, the iDi Carbon Chain platform provides a mechanism through which carbon claims can be connected to operational data and compliance reporting [6]. The platform therefore represents more than a technological tool. It is a potential accountability infrastructure.

7.3 Practical Implications

The findings have practical implications for firms in the new energy vehicle industry. First, companies need to not just try to build their brand on their product level green identity. It is important to remember that although EVs have the potential of helping to decarbonize transportation, stakeholders' demands for information regarding manufacturing emissions, supply chain effects, resource utilization and recycling are also growing. Second, companies need to invest in digital carbon management systems that are connected to recognized standards and internal governance. Third, there need to be clear and open techniques, data limits, and verification procedures, to accompany sustainability stories.

The results indicate that regulators and investors should consider not only the length and quality of the narrative of an ESG report, but also the specificity of the evidence, in determining the extent of ESG disclosure. Some questions to consider: Are the boundaries of reporting clear? Is there any correlation between carbon data and methodologies? Do the products have their carbon footprints calculated? Is there external evidence? Are the board members on board? Are there any management incentives for targets? These questions can be used to identify substantive disclosure from symbolic communication.

8. Conclusion

This paper focused on the impact of digital carbon management on ESG greenwashing risks concerning corporate sustainability disclosure. The study used BYD's Sustainability Report 2025 as an exploratory case and came up with a dual-path framework, adopting a preliminary text analysis approach. When digital carbon management is able to increase the traceability of data, disclosure specificity, the ability to account for carbon, readiness to embed into governance and readiness for verification, the findings indicate that there may be a reduced risk of greenwashing. The transparency-enhancing pathway is evident in BYD's iDi Carbon Chain platform, which is linked to organizational carbon management, product carbon footprint accounting, automated carbon data calculation, international compliance reporting and ISO-related standards.

The study makes three contributions to works. Secondly, it adds to greenwashing research the perspective on the production and verification of ESG information, not just the mere disclosure of sustainability information by firms. Second, it contributes to the research on digital transformation by demonstrating that digitalization can have a dual effect, on the one hand making it possible to control and report on decision-making processes, and on the other hand symbolically communicating decisions. Third, it introduces a pragmatic analytical framework to evaluate the credibility of ESG disclosures using coding categories and specificity scores.

There are also some limitations to the study. Firstly, it is a single case study on BYD, thus the findings are not generalizable. Second, analysis is primarily based on corporate sustainability disclosure, which does not necessarily represent the real ESG performance. Thirdly, the textual analysis is preliminary and not based on large scale computational text analysis. Future studies can be expanded by comparing several new energy vehicle companies, using third-party ESG ratings and applying quantitative textual analysis methods to systematically explore the connection between digital carbon management and greenwashing risk.

The limitations aside, the research offers some practical guidelines for corporate ESG governance. It recommends that companies transition from ESG storytelling to ESG accountability, by connecting sustainability storytelling with data systems, governance, external standards and verification. This is particularly relevant for new energy vehicle companies as a green product identity is not enough to prove ESG credibility. Digital carbon management can help diminish the risk of 'greenwashing' if it is not just a communication tool, but a real governance and accounting tool.

References

[1] Christensen HB, Hail L, Leuz C. Mandatory CSR and sustainability reporting: economic analysis and literature review. Review of Accounting Studies. 2021; 26(3):1176-1248.

[2] KPMG International. Big shifts, small steps: Survey of sustainability reporting 2022. KPMG International; 2022.

[3] Delmas MA, Burbano VC. The drivers of greenwashing. California Management Review. 2011; 54(1):64-87.

[4] Lyon TP, Montgomery AW. The means and end of greenwash. Organization & Environment. 2015; 28(2):223-249.

[5] Suchman MC. Managing legitimacy: strategic and institutional approaches. Academy of Management Review. 1995; 20(3):571-610.

[6] BYD Company Limited. 2025 Sustainability Report. BYD Company Limited; 2026.

[7] Cho CH, Laine M, Roberts RW, Rodrigue M. Organized hypocrisy, organizational facades, and sustainability reporting. Accounting, Organizations and Society. 2015; 40:78-94.

[8] Boiral O. Sustainability reports as simulacra? A counter-account of A and A+ GRI reports. Accounting, Auditing & Accountability Journal. 2013; 26(7):1036-1071.

[9] Michelon G, Pilonato S, Ricceri F. CSR reporting practices and the quality of disclosure: an empirical analysis. Critical Perspectives on Accounting. 2015; 33:59-78.

[10] Hummel K, Schlick C. The relationship between sustainability performance and sustainability disclosure: reconciling voluntary disclosure theory and legitimacy theory. Journal of Accounting and Public Policy. 2016; 35(5):455-476.

[11] Burritt RL, Schaltegger S. Sustainability accounting and reporting: fad or trend? Accounting, Auditing & Accountability Journal. 2010; 23(7):829-846.

[12] Bai C, Sarkis J. A supply chain transparency and sustainability technology appraisal model for blockchain technology. International Journal of Production Research. 2020; 58(7):2142-2162.

[13] Ioannou I, Serafeim G. The consequences of mandatory corporate sustainability reporting. Management Science. 2017; 63(5):1306-1329.

[14] Global Reporting Initiative. GRI 1: Foundation 2021. Global Reporting Initiative; 2021.

[15] IFRS Foundation. IFRS S1 General Requirements for Disclosure of Sustainability-related Financial Information and IFRS S2 Climate-related Disclosures. International Sustainability Standards Board; 2023.

[16] International Organization for Standardization. ISO 14064-1:2018 Greenhouse gases - Part 1: Specification with guidance at the organization level for quantification and reporting of greenhouse gas emissions and removals. ISO; 2018.

[17] International Organization for Standardization. ISO 14067:2018 Greenhouse gases - Carbon footprint of products - Requirements and guidelines for quantification. ISO; 2018.

Appendix. Detailed textual evidence and disclosure specificity scores

Table A1: Detailed textual evidence and disclosure specificity scores

Evidence item	Code / score	Evidence summarized from the report	Analytical interpretation
Reporting boundary	B/C; 2	The report discloses that economic and social KPIs cover the group, while environmental KPIs cover operating sites under BYD's operational control and exclude overseas bases and sales stores.	Boundary disclosure improves transparency, but the exclusion of overseas bases and sales stores should be treated as a limitation for global environmental assessment.
Reporting standards	C; 3	The report follows HKEX and SZSE requirements and refers to ESRS, GRI, ISSB IFRS S1/S2, SDGs and ISO IWA 48.	Multi-framework alignment supports formal comparability and reduces purely symbolic ambiguity.
Reporting principles	B/C; 3	The report states principles including materiality, quantification, balance, consistency, accuracy, comparability, completeness and verifiability.	These principles directly address common sources of greenwashing risk such as vagueness and selective disclosure.
Board approval	D/C; 3	The board supervises the report and states responsibility for the authenticity, accuracy and completeness of the content.	Board-level approval strengthens accountability for sustainability disclosure.
iDi Carbon Chain	A/C; 3	The platform supports organizational carbon management, product carbon footprint accounting, automated carbon data calculation and compliance reporting.	This is the strongest evidence of the transparency-enhancing path.
ISO 14064 and ISO 14067	C; 3	The report states that the iDi Carbon Chain platform passed ISO 14064 and ISO 14067 certification.	External standards improve verification readiness and carbon accounting credibility.
Clean electricity use	B; 2	The report discloses 7.29 billion kWh of clean electricity use.	Quantified evidence improves specificity, but its interpretation still depends on procurement boundary and accounting method.
Environmental investment	B; 2	The report discloses RMB 2.04 billion in environmental protection investment.	Quantified financial input supports substantive disclosure, but input data does not directly prove environmental outcome.
Scope 1 and Scope 2 emissions	B/C; 3	The report discloses Scope 1 and Scope 2 emissions for the automobile sector and related carbon intensity indicators.	Scope-based disclosure increases measurability and comparability.
Carbon intensity target	B/D; 3	The report states a target to reduce operational carbon emission intensity by 50% by 2030 from a 2023 baseline.	A time-bound quantitative target strengthens accountability.
Climate governance	D; 3	The report describes a climate governance structure involving the board, Strategic and Sustainability Committee, ESG Management Committee and ESG personnel in business units.	Governance embedding links digital carbon management to organizational accountability.
ESG-linked remuneration	D; 3	The report states that sustainability indicators are linked to 10% of executive and senior management remuneration, with climate indicators accounting for more than 40% of ESG indicators for ESG-related executives.	Incentive alignment strengthens the credibility of ESG governance.
Green vision narratives	E; 0-1	The report uses concepts such as the three green dreams, DREAMS and cooling the Earth by 1 degree Celsius.	Symbolic narrative supports legitimacy but requires evidence-based anchoring.
Withheld climate transaction data	B/C; 1-2	Some carbon trading-related cost and income information is not disclosed due to business sensitivity.	This shows residual information asymmetry and should be discussed as a disclosure limitation.

Note. The specificity score is interpretive: 0 = symbolic narrative only; 1 = action description without quantitative evidence; 2 = quantitative evidence with limited verification or boundary detail; 3 = quantitative evidence supported by standards, certification, governance or verification mechanisms.

Funding
This research received no external funding.

Conflicts of Interest
The authors declare no conflict of interest.

Acknowledgment
This paper is an output of the science project.

Environment, Social and Governance; Vol.3, No.3; 2026
E-ISSN: 3006-3973 P-ISSN：3006-3965
DOI: https://doi.org/10.70267/ esg2026.v3n3.3548
Published by: Zeus Press

The Reverse Pressure Effect of ESG Rating Divergence on High-Quality Corporate Development

Xin Li* and Wei Jiang

Business School, Guangdong Ocean University, Yangjiang, 529500, China

**Corresponding author: Xin Li.*

Abstract

As ESG concepts become increasingly integrated into China's capital market, the divergence in ESG ratings across different rating agencies has become an increasingly prominent phenomenon. This paper examines the Chinese A-share listed companies from 2012 to 2024 empirically to investigate the influence of the differences in ESG ratings on the excellent development of enterprises (assessed by total factor productivity). The research results show that the differences in ESG ratings can remarkably improve the total factor productivity of the enterprise, exerting a "stimulating effect" instead of a suppressing effect. This conclusion is also confirmed by several robustness tests, such as variable replacement, modification of model specification and instrumental variables method. The heterogeneity analysis indicates that the "stimulating effect" is more obvious in the companies with high analyst coverage, low managerial ownership and a well-structured organization. This study implies that the disagreement in ESG ratings may motivate companies to improve the quality of their development. It gives a new insight into the economic impacts of the differences in ESG ratings and proposes some feasible recommendations for regulators and enterprises.

Keywords

ESG rating divergence, high-quality corporate development, total factor productivity, reverse pressure effect

1. Introduction

During the “14th Five-Year Plan” period, China’s economy is in a stage of high-quality development, in which improving total factor productivity (TFP) becomes an important driving force for sustainable corporate growth and competitiveness improvement [1]. At the same time, ESG (Environmental, Social, and Governance) principles, serving as a benchmark for evaluating the sustainability capability of companies, receive more and more attention from regulators, investors and the general public. Under the influence of policies such as the "dual carbon" target and the Guidelines for Sustainable Development Reports of Listed Companies, the ESG evaluation results are gradually incorporated into investment decisions, financing assessments and risk management procedures. However, it is a neglected fact that different rating agencies usually give significantly different assessments of the ESG performance of the same enterprise. This leads to the problem of "ESG rating divergence". Will this divergence ultimately affect the performance of enterprises or will it instead urge them to improve their total factor productivity? This issue is closely related to the effectiveness of the ESG system and the way towards high-quality corporate development, and therefore urgently needs empirical confirmation.

In fact, the differences in ESG ratings are not only technical problems. According to the China ESG Development Report (2023), the average correlation coefficient among the ratings given by major domestic ESG rating agencies to the same listed company is less than 0.6, indicating significant divergence. From the perspective of information economics, the divergence increases the uncertainty of external stakeholders' perception of a company's true ESG level, which may result in negative effects such as investor confusion and higher financing costs [2-4]. It also causes stock prices to be synchronized and decreases the pricing efficiency of the financial market [5, 6]. When enterprises encounter conflicting external evaluations, they may have the motivation to carry out "legitimacy repair" to rebuild market trust by improving governance, enhancing efficiency and strengthening information disclosure. In fact, previous research shows that the differences in ESG ratings can compel companies to increase their green innovation levels [7-9]. It also prompts them to improve the quality of their ESG disclosures and attract analyst attention, thereby reducing their debt financing costs [3]. The differences in ESG ratings may not be just "noise", but rather an external pressure that forces companies to make self-improvements. However, existing research has not sufficiently explored this positive effect; especially, there is still a lack of empirical verification at the level of the core production efficiency of the enterprises.

A review of the existing literature reveals that scholarly research on divergences in ESG ratings primarily focuses on two directions: First, the causes of divergence—such as differences among rating agencies regarding the scope of evaluation, measurement methodologies, and the assignment of weights. [10, 11] Insufficient Standardization and Transparency in Corporate ESG Information Disclosure[2, 12]. The negative economic consequences of divergence—such as increased stock price synchronicity and reduced capital market information efficiency. [5, 6] Raises debt financing costs [2, 4] and triggers a reduction in holdings by institutional investors [13], It weakens an enterprise's bargaining power within its supply chain and the stability of its customer base [14, 15], and may even stifle its continuous innovation and green collaborative innovation [16, 17]. The aforementioned studies generally view disagreement as an "external burden." Scant literature explores whether divergence might compel enterprises to take substantive action to enhance the quality of their development. Although some studies have identified the positive incentive effect of ESG rating divergence on green innovation [7-9], no research has directly examined the causal relationship between ESG rating divergence and the total factor productivity of listed companies, which is widely regarded as the core indicator of high-quality corporate development. This theoretical gap provides important research space for the present study.

To solve the above problems, this paper takes Chinese A-share listed companies from 2012 to 2024 as the research objects to empirically investigate the impact of the difference in ESG ratings on the total factor productivity of the firms. It also examines the heterogeneity of this effect under different external supervisory environments, internal governance structures and the corporate life-cycle stages. The results indicate that: (1) The difference in ESG ratings can improve the total factor productivity of the firms. The coefficient of ESG_rank in the basic regression is 0.076 (with a significant level of 1%). This means that an increase of one standard deviation in the difference will lead to an approximately 0.014 standard deviation increase in the TFP of the firm, showing a "forcing effect". (2) The conclusions remain valid after the robustness tests, such as replacing the dependent variable with the TFP calculated by the OP method and replacing the independent variable with the rank difference. (3) The analysis of heterogeneity shows that this "forcing effect" is more obvious in the firms with high analyst coverage, low managerial ownership and mature corporate stage. The results obtained through instrumental variable method have the same direction and support the robustness of the basic conclusions.

The main contributions of this paper are mainly reflected in three aspects. From a theoretical perspective, this study expands the previous understanding that 'the disagreement in ESG ratings is completely negative'. By considering its economic effects from the positive side of 'encouraging enterprises to achieve high-quality development', this research enriches the literature about the informational value of ESG ratings. In terms of mechanism identification, heterogeneity was examined from three aspects: analyst attention, management shareholding ratio and company age. It reveals the conditions under which the differences in ratings exert their effects on external supervision, internal governance and organizational abilities. It provides empirical evidence to understand the context dependence of the reverse-forcing mechanism. From a practical standpoint, this research offers strategic advice to enterprises on how to handle the differences in ESG ratings, especially suggesting that they should actively transform external pressure into improvement of internal efficiency. At the same time, it provides some policy recommendations for regulatory authorities to regulate the ESG rating market and prevent 'bad ratings pushing out good ones'.

2. Theoretical Mechanisms and Research Hypotheses

2.1 ESG Rating Divergence and High-Quality Corporate Development

Several rating agencies evaluate the environmental, social and governance performance of the same company differently. From the viewpoint of information economics, these differences cause more uncertainty for the external stakeholders about the actual ESG level of the company, which may result in negative effects like investor confusion and higher financial costs. Based on the concept of legitimacy, if a company considers that some important external resources providers (such as investors, creditors and regulators) doubt its legitimacy, it will take 'legitimacy repair' measures to maintain or restore the social legitimacy of the organization [18]. The inconsistency in ESG ratings acts as an 'unfavorable signal', thus increasing the legitimacy pressure faced by the companies, leading them to adopt effective improvement strategies.

Specifically, the difference in ESG ratings can improve the corporate total factor productivity through the following three ways:

Firstly, the impact of strengthened external supervision. The differences in ratings will attract more attention from analysts, the media and regulators. Studies show that the differences in ESG ratings can increase the attention of analysts and the media, forcing enterprises to participate in green innovative activities [7, 8]. In order to eliminate the differences and resolve market doubts, companies have the motivation to provide more detailed information about their ESG practices and internal management procedures. This 'forced disclosure' is usually accompanied by the standardization of the process and the improvement of resource utilization efficiency, contributing to the growth of TFP.

Secondly, the effect of incentives and constraints on managers. When the management's ownership is low and the alignment with shareholders' interests is insufficient, the rating difference can serve as a 'warning signal' prompting board intervention or shareholder activism, thereby correcting the inefficient decisions and guiding the allocation of resources to profitable projects. Previous research has found that the management's shareholding has a negative moderating effect on the relationship between the ESG rating difference and the company's behavior. A lower management ownership strengthens the external governance role of the rating difference [19]. This mechanism is especially important in companies encountering serious management conflicts of interest.

Lastly, the effect of resource redistribution. The rating difference makes it harder or more expensive for companies to obtain funds in the capital market, thus worsening the financing constraints of the company [2, 20, 21]. This compels the companies to reconsider their internal resource allocation and reduce unnecessary and inefficient expenditures. Redirecting investments to fields which can significantly enhance productivity, such as green technological innovation [7-9]. This 'effort to improve efficiency under pressure' directly results in the enhancement of TFP.

H1: ESG rating divergence exerts a reverse pressure effect on total factor productivity.

2.2 The Joint Moderating Effects of Analyst Coverage, Managerial Ownership, and Firm Age

The degree to which the above-mentioned reverse coercive effect depends on the external information environment, internal governance structure and its own organizational capacity of the enterprise's operation. This article comprehensively examines its boundary conditions from three dimensions.

The analyst coverage reflects the intensity of the external review faced by the company. High analyst coverage amplifies the market visibility of ESG rating differences. Accelerated the internal transmission of negative signals within the company. The forced pressure has increased. Some studies believe that analyst coverage is the key communication channel for the positive impact of ESG rating differences [7]. On the contrary, in the case of low attention. Differences may be overwhelmed by information noise, making it difficult to generate effective incentives.

The shareholding ratio of the management measures the degree of consistency between the interests of shareholders and the interests of the management in the internal governance mechanism. When the shareholding ratio is low, the management lacks the motivation to take the initiative to improve productivity. So external rating differences can be used as an alternative governance mechanism. Let the management take action by drawing the attention of the board of directors and investors. Empirical studies shows that management shareholding negatively adjust the relationship between ESG rating differences and high-quality audit demand [19]. The lower the shareholding ratio, the stronger the external governance effect of the difference. When the shareholding ratio is high, the management already has strong incentives. So the marginal effect of divergence is weakened.

Enterprise age is a characteristic of organizational experience, resource accumulation and process maturity. Mature companies have richer management experience, more stable cash flow and more sound internal control system. They are more capable of converting external pressure into improved system efficiency. Studies have found that the negative impact of ESG rating differences on enterprise innovation is more obvious in enterprises in the growth and maturity stage, and not in enterprises in the recession stage[16]. This shows that mature enterprises are more sensitive to the pressure of differentiation. Younger enterprises are constrained by limited resources and lack of management experience, and it is difficult to effectively take advantage of the pressure brought by differences.

Although these three operate through different channels. External supervision, internal governance and organizational ability. But they all have the same basic logic. Under the condition of stronger external supervision, more obvious internal institutional problems, and more mature organizational capacity. The mandatory impact of ESG rating differences is more significant. So this paper incorporates the three into a comprehensive hypothetical framework:

H2: The positive effect of ESG rating divergence on total factor productivity is stronger among firms with higher analyst coverage, lower managerial ownership, and greater firm age.

3. Research Design

3.1 Sample Selection and Data Sources

This paper selects Chinese A-share listed companies from 2012 to 2024 as the research subjects, and, drawing upon existing studies, applies the following screening criteria to the sample: (1) ST and *ST firms are excluded; (2) firms in the financial industry are excluded; (3) firms that have been delisted or suspended from listing are excluded; (4) observations with abnormal or missing values for relevant variables are excluded. Ultimately, a total of 47,883 valid observations were obtained. The ESG rating data in this article comes from the scores published by six rating agencies (China Securities, Wind, SynTao Green Finance, Bloomberg, FTSE Russell, and CNRS). The remaining raw data is sourced from the CSMAR database. In addition, to reduce the potential impact of extreme values on the research results, this paper applies Winsorization to all continuous variables at the 1% and 99% quantiles.

3.2 Variable Definitions

3.2.1 Dependent Variable: Total Factor Productivity (TFP_LP)

Following prior studies [22, 23], this paper adopts firm-level total factor productivity estimated using the LP method as a quantitative indicator of high-quality corporate development. Total factor productivity reflects firms' resource allocation efficiency and technological progress beyond factor inputs and is widely recognized as a key indicator for measuring the level of high-quality corporate development.

3.2.2 Independent Variable: ESG Rating Divergence (ESG_rank)

Following existing literature [24], ESG rating divergence is measured by directly calculating the standard deviation of the rankings assigned by all rating agencies to the same firm in a given year. A larger value indicates greater inconsistency among rating agencies in evaluating the firm's ESG performance, implying a higher degree of ESG rating divergence.

3.2.3 Control Variables

Drawing on prior studies [24, 25], this paper includes the following control variables: (1) fixed asset ratio (Fixed), used to control for the impact of asset structure on production efficiency; (2) ownership concentration (First), reflecting the governance effect of major shareholders; (3) firm size (Size), controlling for economies of scale; (4) market value (TobinQ), measuring firms' growth opportunities and market valuation; (5) return on assets (Roa), controlling for differences in profitability; (6) leverage ratio (Lev), reflecting firms' financial leverage and risk level; and (7) operating revenue growth rate (Growth), controlling for the influence of firm growth on total factor productivity. Detailed definitions of all variables are presented in Table 1.

Table 1: Definitions and Descriptions of Variables

Variable Type	Variable Name	Variable Definition
Dependent Variable	High-Quality Corporate Development (TFP_LP)	Firm-level total factor productivity estimated using the LP method
Independent Variable	ESG Rating Divergence (ESG_rank)	Standard deviation of the rankings assigned by all rating agencies to the firm in a given year
Control Variables	Fixed Asset Ratio (Fixed)	Net fixed assets / total assets
	Ownership Concentration (First)	Shareholding ratio of the largest shareholder
	Firm Size (Size)	Natural logarithm of total assets
	Market Value (TobinQ)	Market value / total assets
	Return on Assets (Roa)	Net profit / total assets
	Leverage Ratio (Lev)	Total liabilities / total assets
	Revenue Growth Rate (Growth)	(Current quarter operating revenue − previous quarter operating revenue) / previous quarter operating revenue

3.3 Model Specification

Based on the above discussion and hypotheses regarding ESG rating divergence and high-quality corporate development, this paper constructs the following model to examine their relationship:

$$TFP\_LP_{i,t} = \alpha + \beta_1 ESG\_rank_{i,t} + \gamma Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \quad (1)$$

where $TFP_LP_{i,t}$ denotes the total factor productivity of firm *i* in year *t*; $ESG\_rank_{i,t}$ represents the degree of ESG rating divergence for firm *i* in year *t*; $Controls_{i,t}$ denotes the set of control variables; μ_i represents firm fixed effects controlling for time-invariant firm characteristics; λ_t denotes year fixed effects controlling for macroeconomic shocks; and $\varepsilon_{i,t}$ is the random error term. Standard errors are clustered at the industry level.

4. Empirical Results

4.1 Descriptive Statistics

As shown in Table 2, the sample consists of 47,883 observations. The mean value of firms' total factor productivity (TFP_LP) is 8.967, with a standard deviation of 1.092, indicating substantial variation in development quality among sample firms. The mean value of ESG rating divergence (ESG_rank) is 0.438, with a standard deviation of 0.193, ranging from a minimum of 0.07 to a maximum of 0.88, suggesting significant heterogeneity in the degree of ESG rating divergence faced by different firms during the sample period. The distributions of the remaining control variables are all within a reasonable range.

Table 2: Descriptive Statistics

	Observations	Mean	Standard Deviation	Minimum	Median	Maximum
TFP_LP	47883	8.967	1.092	6.747	8.857	11.877
ESG_rank	47883	0.438	0.193	0.070	0.425	0.880

Fixed	47883	0.201	0.154	0.002	0.169	0.673
First	47883	3.409	0.478	2.111	3.439	4.313
Size	47883	22.144	1.351	19.274	21.963	26.273
TobinQ	47883	2.003	1.283	0.830	1.588	8.469
Roa	47883	0.034	0.068	-0.275	0.036	0.204
Lev	47883	0.412	0.207	0.053	0.400	0.911
Growth	47883	0.337	0.893	-0.751	0.118	6.098

4.2 Baseline Regression Results

To examine the impact of ESG rating divergence on high-quality corporate development, this paper constructs the following two-way fixed effects model:

$$TFP_LP_{i,t}=\alpha+\beta_1ESG_rank_{i,t}+\gamma Controls_{i,t}+\mu_i+\lambda_t+\varepsilon_{i,t} \tag{2}$$

where μ_i denotes firm fixed effects and λ_t denotes year fixed effects. The regression results are reported in Table 3.

Table 3: Baseline Regression Results

	(1)	(2)
	TFP_LP	TFP_LP
ESG_rank	0.355***	0.076***
	(0.045)	(0.025)
Fixed		-0.799***
		(0.097)
First		-0.033
		(0.024)
Size		0.573***
		(0.012)
TobinQ		0.024***
		(0.003)
Roa		1.900***
		(0.129)
Lev		0.479***
		(0.066)
Growth		-0.009
		(0.010)
Constant	8.822***	-3.866***
	(0.020)	(0.264)
Company	Yes	Yes
Year	Yes	Yes
N	40,099	39,561
R^2	0.878	0.935

Notes: The standard error of industry-level gathering is reported in parentheses. *, * and * represent statistical significance at the levels of 10%, 5% and 1% respectively. The same symbol is also applicable below.*

Column (1) of Table 3 includes explanatory variables and company/year fixed effects. The coefficient for ESG_rank is 0.355, which is significant at the 1% level. In Column (2), after including all control variables, the coefficient for ESG_rank is 0.076, remaining positive at the 1% level. This result indicates that for every one-standard-deviation increase (0.193) in ESG rating divergence, the firm's total factor productivity increases by approximately 0.0147 (0.076 × 0.193) on average. Relative to the TFP_LP mean of 8.967, the economic significance is substantial. The benchmark regression results validate the core hypothesis of this paper: ESG rating divergence exerts a significant positive forcing effect on the high-quality development of enterprises.

4.3 Robustness Tests

4.3.1 Replacing the Dependent Variable

To avoid the influence of different TFP estimation methods on the results, this paper replaces TFP_LP estimated using the LP method in the baseline regression with total factor productivity estimated using the OP method (TFP_OP). As shown in Column (1) of Table 4, the coefficient of ESG rating divergence (ESG_rank) is 0.056 and positive at the 5% level, which is consistent with the baseline findings.

4.3.2 Replacing the Independent Variable

Following the measurement approach of prior literature [26], this paper uses the ESG ratings issued by four rating agencies—Huazheng ESG Rating, Wind, SynTao Green Finance, and Bloomberg. Each institution was assigned a rating, and then the standard deviation of the four institutions' ratings was calculated. Construct the ESG rating divergence indicator (ESGdif4) to replace the original explanatory variable in the regression analysis. As shown in Column (2) of Table 4, the coefficient for ESGdif4 is 0.005, which is positive at the 5% level; the conclusion is robust.

4.3.3 Replacing the Independent Variable

In addition to the four existing rating agencies, FTSE Russell's ESG rating has been added. Calculate the standard deviation of the scores from five institutions, construct the ESG rating divergence indicator (ESGdif5), and re-run the regression. As shown in Column (3) of Table 4, the coefficient of ESGdif5 is 0.006, which is positive at the 5% level, further validating the robustness of the benchmark conclusion.

Table 4: Robustness Tests

	(1)	(2)	(3)
	TFP_OP	TFP_LP	TFP_LP
ESG_rank	0.056**		
	(0.027)		
Fixed	-0.818***	-0.800***	-0.800***
	(0.092)	(0.097)	(0.097)
First	-0.036	-0.030	-0.030
	(0.024)	(0.024)	(0.024)
Size	0.366***	0.576***	0.576***
	(0.011)	(0.013)	(0.012)
TobinQ	0.019***	0.024***	0.024***
	(0.003)	(0.003)	(0.003)
Roa	1.885***	1.904***	1.904***
	(0.131)	(0.129)	(0.129)
Lev	0.393***	0.472***	0.472***
	(0.064)	(0.067)	(0.067)
Growth	-0.003	-0.009	-0.009
	(0.010)	(0.010)	(0.010)
ESGdif4		0.005**	
		(0.003)	
ESGdif5			0.006**
			(0.003)
Constant	-1.533***	-3.916***	-3.910***
	(0.227)	(0.267)	(0.266)
Company	Yes	Yes	Yes
Year	Yes	Yes	Yes
N	39,561	39,561	39,561
R^2	0.897	0.934	0.934

4.4 Endogeneity Treatment: Instrumental Variable Approach

Considering that the disagreements in ESG ratings may have an inverse causal relationship with high-quality corporate development or be due to endogeneity issues caused by omitted variables. Drawing upon the approach adopted in reference [27], this paper employs ESG fund holdings data as an instrumental variable (IV) for corporate ESG performance. The IV is the percentage of shares held by ESG funds in a company in that year. The correlation lies in the fact that ESG funds exhibit a preference for firms with consistent ESG performance; the exogeneity stems from the fact that fund holdings primarily exert influence by affecting the divergence in ESG ratings, rather than by directly determining a firm's Total Factor Productivity (TFP). The regression results are presented in Table 5.

Table 5: Instrumental Variable Results

	(1) ESG_rank	(2) TFP_LP
ESG_rank		0.812
		(0.430)
IV	-0.001***	
	(0.000)	
Fixed	-0.027	-0.780***
	(0.017)	(0.094)
First	0.031***	-0.056*
	(0.007)	(0.028)
Size	0.048***	0.539***
	(0.006)	(0.025)
TobinQ	-0.004**	0.028***
	(0.001)	(0.004)
Profitabil~y	0.044	1.871***
	(0.026)	(0.126)
Lev	-0.079***	0.539***
	(0.014)	(0.080)
Growth	-0.003**	-0.006
	(0.001)	(0.010)
Company	Yes	Yes
Year	Yes	Yes
N	39561	39561
F	30.698	417.495

In the first-stage regression, the coefficient of IV was -0.001, which was significant at the 1% level, and the F-statistic was 30.698, which was greater than 10. Rule out the problem of weak instrumental variables. In the second-stage regression, the coefficient of ESG_rank was 0.812, which was significant at the 10% level (p=0.059). The direction and magnitude of the coefficients continue to support the benchmark conclusions. ESG rating divergence has a positive impact on TFP_LP. The instrumental variable results further validate the causal robustness of the "forcing effect."

5. Heterogeneity Analysis

5.1 Grouping Based on Analyst Coverage

Research indicates that analysts, as important information intermediaries in capital markets, can mitigate information asymmetry between companies and external investors through their tracking and coverage activities. Exert external supervisory pressure on corporate business practices[28, 29]. When companies face high analyst scrutiny, the negative signals generated by disagreements in ESG ratings are more easily captured and interpreted by market participants. Impose stronger reputation constraints and performance pressure on the management. Management is more motivated to deal with external problems by improving productivity and optimizing resource allocation. Expand the mandatory effect of ESG rating differences. But in companies with less attention from analysts, the information environment is relatively closed. The difference in ESG ratings is difficult to effectively translate into internal decision-making and is unlikely to be convincing. This paper follows the method of existing literature [30], and divides the sample into high-attention group and low-attention group according to the medin of analyst attention. The results of the grouped regression are presented in Table 6.

Table 6: Heterogeneity Test: Analyst Coverage

	(1)	(2)
	Low Analyst Coverage	High Analyst Coverage
	TFP_LP	TFP_LP
ESG_rank	0.029	0.074***
	(0.033)	(0.025)
Fixed	-0.920***	-0.753***
	(0.124)	(0.109)
First	-0.037	-0.014
	(0.041)	(0.025)
Size	0.528***	0.585***
	(0.022)	(0.013)
TobinQ	0.019**	0.024***
	(0.008)	(0.003)
Roa	1.820***	1.821***
	(0.202)	(0.128)
Lev	0.562***	0.404***
	(0.088)	(0.073)
Growth	-0.004	-0.009
	(0.015)	(0.009)
Constant	-2.842***	-4.156***
	(0.485)	(0.301)
Company	Yes	Yes
Year	Yes	Yes
N	11,543	26,989
R^2	0.943	0.942

The results indicate that, in the high-attention group, the coefficient for ESG_rank is 0.074, which is positive at the 1% level. In the low-attention group, the coefficient is 0.029 and not significant. This result validates the aforementioned theoretical expectation: the external monitoring function of analysts serves as a crucial transmission channel through which ESG rating divergence exerts its coercive influence. When analyst coverage is high, the governance-related information embedded in rating divergence is fully disseminated, thereby prompting firms to improve their performance. But when external informational pressure is insufficient, rating divergence struggles to translate into effective governance constraints. Discrepancies in ESG ratings do not directly impact firms; rather, they amplify the effects of external pressure through information intermediaries.

5.2 Grouping Based on Managerial Ownership

Some people think that management shareholding is an important internal governance mechanism. When management holds a considerable proportion of shares, their personal interests are consistent with those of shareholders, which reduces agency costs[31, 32]. The management itself already has a strong motivation to create value. The marginal utility of the additional pressure exerted by external ESG rating differences is low, and the resulting coercive effect may not be significant. In companies with a low shareholding ratio of management, the conflict of interest between management and shareholders is more obvious. The management lacks sufficient intrinsic motivation to improve the long-term value of the enterprise. Differences in ESG ratings can make up for the lack of internal incentives and trigger a defensive response from management. By enhancing production efficiency, the uncertainty associated with external evaluations is reduced, thereby generating a more pronounced "reverse-forcing" effect.

This paper divides the sample into two groups of low-management shareholding and high-management shareholding according to the medin of the management shareholding ratio. The regression results are presented in Table 7.

Table 7: Heterogeneity Test: Managerial Ownership

	(1)	(2)
	Low Managerial Ownership	High Managerial Ownership
	TFP_LP	TFP_LP

ESG_rank	0.100***	0.038
	(0.035)	(0.027)
Fixed	-0.678***	-1.062***
	(0.086)	(0.134)
First	-0.000	-0.084***
	(0.032)	(0.031)
Size	0.552***	0.567***
	(0.019)	(0.016)
TobinQ	0.019***	0.031***
	(0.005)	(0.003)
Roa	2.032***	1.789***
	(0.132)	(0.215)
Lev	0.434***	0.538***
	(0.080)	(0.076)
Growth	-0.001	-0.024**
	(0.011)	(0.009)
Constant	-3.479***	-3.554***
	(0.412)	(0.343)
Company	Yes	Yes
Year	Yes	Yes
N	20,089	19,170
R^2	0.936	0.940

In the low-shareholding group, the coefficient for ESG_rank is 0.100, significant at the 1% level. In the high-holding group, the coefficient is 0.038 and not significant. There exists a substitutive relationship between ESG rating divergence and internal governance structures. When the internal incentive mechanism is weak, the external rating difference can play a "compensation" role, so that the management can improve the company's performance. When the management has gained sufficient incentives through stock holdings, the marginal impact of external differences is limited. This provides new empirical evidence for understanding the differentiated effect of ESG rating differences in different governance environments.

5.3 Grouping Based on Firm Age

The age of the organization reflects the differences in resource accumulation, management experience and adaptability of the organization [33]. Mature enterprises usually have a perfect governance structure and standardized operation process; They possess an exceptional ability to absorb and process information, can turn external pressure into improved internal efficiency [34]. Partners are more likely to trust established companies. Research suggests that when established companies supply used equipment to younger companies, it leads to regional economic prosperity [35]. Domestic scholars have made similar findings. For example, there is an inverted U-shaped relationship between firm age and government subsidies [36], and national innovation systems tend to favor established firms [37]. Mature companies are subject to stricter market oversight and reputational constraints, and are more concerned about discrepancies in external evaluations. Young companies have limited resources. Their organization is undergoing rapid change, and management is placing greater emphasis on market expansion and survival. The response to relatively “indirect” external signals (such as differences in ESG ratings) is weaker; so the “crowding-out effect” is difficult to materialize.

This paper divides firms into young firm and mature firm groups according to the annual industry median of firm age. The regression results are reported in Table 8.

Table 8: Heterogeneity Test: Firm Age

	(1)	(2)
	Young Firm	Mature Firm
	TFP_LP	TFP_LP
ESG_rank	0.026	0.102***
	(0.029)	(0.034)
Fixed	-1.090***	-0.669***
	(0.149)	(0.082)
First	-0.064**	-0.014

	(0.032)	(0.024)
Size	0.559***	0.576***
	(0.021)	(0.014)
TobinQ	0.025***	0.022***
	(0.004)	(0.006)
Roa	1.944***	1.861***
	(0.201)	(0.128)
Lev	0.538***	0.436***
	(0.076)	(0.072)
Growth	-0.021**	-0.005
	(0.010)	(0.011)
Constant	-3.449***	-3.981***
	(0.450)	(0.311)
Company	Yes	Yes
Year	Yes	Yes
N	17,152	22,409
R^2	0.938	0.926

The results shows that in mature companies, the ESG_rank coefficient is 0.102 and is statistically significant at the 1% level. For young enterprises, the coefficient is 0.026, which is not statistically significant. This suggests that the positive spillover effects of ESG rating differences do not apply to all companies, it depends on their "absorption capacity." Mature companies rely on robust governance structures and substantial resources. More effectively transform external pressures into a driving force for high-quality development. Due to their limited capabilities, young companies are unable to derive similar incentives from rating divergences. If a company relies solely on differences in ESG ratings to drive high-quality corporate development. That may need to be supplemented by capacity-building efforts within the company. Otherwise, the expected "compulsory effect" may be absent among young companies that are most in need of improvement.

6. Conclusions and Policy Implications

This article is based on data from Chinese A-share listed companies from 2012 to 2024. This study systematically examined the impact of ESG rating divergence on the high-quality development of enterprises; the main findings are as follows: (1) Discrepancies in ESG ratings are not merely "informational noise," rather, it serves as a key catalyst for driving improvements in total factor productivity. Even after addressing endogeneity, using proxy variables, and modifying the model specifications. The "compulsion effect" is also strong. And this confirms the core hypothesis of the "compulsion mechanism." (2) From the perspective of the external oversight environment, the visibility of rating discrepancies is greater among companies that are closely monitored by analysts. Greater external oversight pressure has amplified the enforcement effect. High analyst coverage can amplify the positive incentive effect of divergent views. (3) From the perspective of internal governance structure, agency problems are more pronounced when management holds a low percentage of shares. Different external ratings can serve as a substitute for internal governance mechanisms and create more positive incentives. On the contrary, a high degree of managerial ownership mitigates the marginal effects of disagreements. (4) In terms of the company's maturity, mature companies have better management expertise, cash flow, and internal control systems. They can better channel external pressure into improvements in system efficiency. Young companies are constrained by limited resources and a lack of management experience. Struggling to effectively harness the pressure created by disagreements. This reveals the fundamental logic behind how differences in ESG ratings drive the high-quality development of enterprises through three distinct channels. "strengthened external oversight—management incentives and constraints—resource reallocation."

So this paper proposes the following policy recommendations: (1) Regulatory authorities should objectively recognize the "double-edged sword" effect of differences in ESG ratings; While making ratings and disclosures more standardized, they do not need to go to great lengths to eliminate all these differences. Appropriate "constructive disagreement" can create a mechanism for external pressure. Encourage companies to actively improve their governance and efficiency. A mechanism for monitoring and issuing early warnings regarding credit rating discrepancies could be established, with preferential policies offered to companies that show significant but sustained improvement in their credit ratings. (2) Investors should pay attention to the various signals conveyed by ESG ratings and use them as supplementary indicators for assessing corporate governance risks and future sustainability potential. Actively engage with management through channels such as shareholder meetings and investor engagement platforms to drive improvements in the company's ESG performance and operational efficiency. For companies that are controversial but fundamentally sound, they may be viewed as potential "value pockets." (3) Companies should recognize the value of the external pressure inherent in rating divergences. Use rating feedback to identify gaps in environmental, social, and governance (ESG) performance. Transform external pressures into an endogenous driver for improving total factor productivity. Established companies can leverage their organizational capabilities to be the first to improve efficiency. Young companies should establish an ESG management system in phases to lay the groundwork for better development.

This study also has some limitations and suggests directions for future research. (1) Differences in ESG ratings are primarily measured using the standard deviation of the rankings assigned by different rating agencies. It doesn't specify the exact source of the differences. For example, whether the divergence stems primarily from the environmental, social, or governance dimension; The mechanisms by which differences across various dimensions affect total factor productivity may vary. And the researchers can explore this further later. (2) Due to limitations in data availability, the sampling period did not cover a longer time window. It is challenging to characterize the long-term dynamic effects of ESG rating divergence on high-quality corporate development, as well as its potential nonlinear features (such as U-shaped or inverted U-shaped relationships). In the future, the researchers can combine data from longer time periods for tracking and analysis. (3) This article primarily focuses on three heterogeneity dimensions: analyst attention, management shareholding, and company age. Other moderating factors that may affect the boundary of the forcing effect, such as industry attributes (heavy pollution vs. non-heavy pollution), regional institutional environment, and the level of enterprise digital transformation, have not yet been explored. In the future, the heterogeneity analysis framework can be expanded to better understand the mechanisms underlying the differential impacts of ESG rating variations.

References

[1] Xi Jinping. (2022). *Hold high the great banner of socialism with Chinese characteristics and strive in unity to build a modern socialist country in all respects: Report delivered at the 20th National Congress of the Communist Party of China*. Beijing, China: People's Publishing House.

[2] Zhang, Y. Q., Yang, H. Y., & Zhang, X. Y. (2023). ESG rating divergence and the cost of debt capital. *Financial Theory*, 15(4), 22–43+124.

[3] Chen, P. C., Li, Z., & He, Q. L. (2024). Does ESG rating divergence necessarily generate negative effects? Evidence from the perspective of debt financing costs. *Technology Economics*, 43(5), 117–136.

[4] Huang, X. W., An, X. X., & Lv, W. M. (2025). Does ESG rating divergence affect corporate credit ratings? *Journal of Financial Economics Research*, 40(3), 124–139.

[5] Liu, X. Q., Yang, Q. Q., & Hu, J. (2023). ESG rating divergence and stock price synchronicity. *China Soft Science*, (8), 108–120.

[6] Chen, H. T., Yin, H. F., Zhang, T. S., & Huang, J. (2025). Does ESG rating divergence affect capital market pricing efficiency? Evidence from stock price synchronicity of listed companies. *Journal of Finance and Economics*, 51(5), 65–80. https://doi.org/10.16538/j.cnki.jfe.20240116.101

[7] Chen, S. F., Guo, W. L., & Li, H. X. (2025). The impact of ESG rating divergence on corporate green technological innovation. *Journal of Xi'an University of Finance and Economics*, 38(6), 51–65. https://doi.org/10.19331/j.cnki.jxufe.20250520.002

[8] Dong, C., Dong, X. C., Jiang, Q. Z., & Tian, J. Y. (2024). The impact and mechanism of ESG rating divergence on green innovation of listed companies. *China Population, Resources and Environment*, 34(8), 103–113.

[9] Guo, J. Z., Xie, Y., & Liu, X. L. (2026). Research on the impact of ESG rating divergence on green innovation in manufacturing enterprises. *Statistics and Management*, 41(2), 133–147. https://doi.org/10.16722/j.issn.1674-537x.2026.02.013

[10] Berg, F., Koelbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, 26(6), 1315–1344. https://doi.org/10.1093/rof/rfac033

[11] Christensen, D. M., Serafeim, G., & Sikochi, A. (2022). Why is corporate virtue in the eye of the beholder? The case of ESG ratings. *The Accounting Review*, 97(1), 147–175. https://doi.org/10.2308/TAR-2019-0506

[12] Kimbrough, M. D., McVay, S. E., & Wang, L. (2024). Does voluntary ESG reporting resolve disagreement among ESG rating agencies? *European Accounting Review*, 33(1), 15–47. https://doi.org/10.1080/09638180.2022.2088588

[13] Wang, L., Li, K., & Ding, L. L. (2025). ESG rating agency divergence and fund investment strategy choice. *Journal of Shanghai University of Finance and Economics*, 27(5), 93–107. https://doi.org/10.16538/j.cnki.jsufe.2025.05.007

[14] Jiang, B., & Ma, C. (2025). ESG rating divergence and supply chain bargaining power. *Journal of the University of Jinan (Social Science Edition)*, 35(3), 137–148. https://doi.org/10.20004/j.cnki.ujn.2025.03.012

[15] Huo, R. (2024). *Does corporate ESG rating divergence affect customer stability?* (Master's thesis, Dongbei University of Finance and Economics). https://doi.org/10.27006/d.cnki.gdbcu.2024.001647

[16] Lu, M. F., Yang, Z., & Chen, J. ESG rating divergence and corporate persistent innovation. *Studies in Science of Science*, 1–15. https://doi.org/10.16192/j.cnki.1003-2053.20260130.002

[17] Yang, C. (2025). Does ESG rating divergence affect green collaborative innovation? The moderating effect of stakeholder relationships. *Journal of Financial Development Research*, (7), 15–25. https://doi.org/10.19647/j.cnki.37-1462/f.2025.07.002

[18] Mark C. Suchman. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, 20(3), 571–610. https://doi.org/10.5465/amr.1995.9508080331

[19] Dong, Y. Q. (2025). *Research on the impact of ESG rating divergence on corporate demand for high-quality auditing* (Master's thesis, Yangtze University). https://doi.org/10.26981/d.cnki.gjhsc.2025.000464

[20] Wang, R., Wang, W. H., & Song, S. S. (2025). ESG rating divergence and corporate financing constraints: Impact effects, transmission channels, and heterogeneity characteristics. *Collected Essays on Finance and Economics*, (5), 55–66. https://doi.org/10.13762/j.cnki.cjlc.20250303.001

[21] Zhao, Y. H., Sun, Y., Feng, T. W., & Liu, Y. (2024). How does supplier ESG rating divergence affect corporate operational resilience? *China Industrial Economics*, (11), 174–192. https://doi.org/10.19581/j.cnki.ciejournal.2024.11.010

[22] Huang, B., Li, H. T., Liu, J. Q., & Lei, J. H. (2023). Digital technological innovation and high-quality development of Chinese enterprises: Evidence from corporate digital patents. *Economic Research Journal*, 58(3), 97–115.

[23] Shen, K. R., Qiao, G., & Lin, J. W. (2024). Intelligent manufacturing policies and high-quality development of Chinese enterprises. *The Journal of Quantitative & Technical Economics*, 41(2), 5–25. https://doi.org/10.13653/j.cnki.jqte.20231214.008

[24] Zhou, Z. J., Ding, X. J., & San, Z. Y. (2023). ESG rating divergence and audit risk premium. *Auditing Research*, (6), 72–83.

[25] Fang, X. M., & Hu, D. (2023). Corporate ESG performance and innovation: Evidence from A-share listed companies. *Economic Research Journal*, 58(2), 91–106.

[26] He, T. M., Li, Y. P., Wang, Z., & Tan, Z. D. (2023). Does ESG rating divergence improve voluntary information disclosure of listed companies? *Accounting and Economics Research*, 37(3), 54–70. https://doi.org/10.16314/j.cnki.31-2074/f.2023.03.006

[27] Xie, H. J., & Lv, X. (2022). Responsible international investment: ESG and China's outward foreign direct investment. *Economic Research Journal*, 57(3), 83–99.

[28] Zhu, H. J., He, X. J., & Tao, L. (2007). Can securities analysts improve the efficiency of China's capital market? Empirical evidence based on stock price synchronicity and stock price information content. *Journal of Financial Research*, (2), 110–121.

[29] Pan, Y., Dai, Y. Y., & Lin, C. Q. (2011). Information opacity, analyst coverage, and stock price crash risk. *Journal of Financial Research*, (9), 138–151.

[30] Xi, L. S., & Wang, Y. (2022). Corporate ESG information disclosure and stock price crash risk. *On Economic Problems*, (8), 57–64. https://doi.org/10.16011/j.cnki.jjwt.2022.08.003

[31] Morck, R., Wolfenzon, D., & Yeung, B. (2005). Corporate governance, economic entrenchment, and growth. *Journal of Economic Literature*, 43(3), 655–720. https://doi.org/10.1257/002205105774431252

[32] Zhu, D. S. (2010). Controlling shareholders, equity balance, and corporate dividend policy choice. *Journal of Shandong University (Philosophy and Social Sciences Edition)*, (3), 80–87.

[33] Desai, P. S., Kalra, A., & Murthi, B. P. S. (2008). When old is gold: The role of business longevity in risky situations. *Journal of Marketing*, 72(1), 95–107. https://doi.org/10.1509/jmkg.72.1.95

[34] Jiang, C. X., Chua, R. Y. J., Kotabe, M., & Murray, J. Y. (2011). Effects of cultural ethnicity, firm size, and firm age on senior executives' trust in their overseas business partners: Evidence from China. *Journal of International Business Studies*, 42(9), 1150–1173. https://doi.org/10.1057/jibs.2011.35

[35] Ma, S., Murfin, J., & Pratt, R. (2022). Young firms, old capital. *Journal of Financial Economics*, 146(1), 331–356. https://doi.org/10.1016/j.jfineco.2021.09.008

[36] Yuan, D. Y., Li, D. G., & Song, X. N. (2018). Industrial agglomeration, firm age, and government subsidies. *Finance & Trade Economics*, 39(9), 39–56. https://doi.org/10.19795/j.cnki.cn11-1166/f.2018.09.003

[37] Yang, L. Q., Liang, Q. Z., & Kang, H. (2016). National innovation systems and corporate R&D investment based on the moderating effect of firm characteristics. *Chinese Journal of Management*, 13(5), 707–714.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Environment, Social and Governance; Vol.3, No.3; 2026
E-ISSN: 3006-3973 P-ISSN：3006-3965
DOI: https://doi.org/10.70267/ esg2026.v3n3.4957
Published by: Zeus Press

The Impact of Corporate ESG Performance on Corporate Tax Avoidance

Zixi Xu*

School of Finance and Taxation, Capital University of Economics and Business, Beijing 100070, China

**Corresponding author: Zixi Xu.*

Abstract

ESG performance has increasingly become an important means for firms to reconcile short-term financial objectives with long-term sustainable development. Nevertheless, the informational role and economic consequences of ESG performance have not yet been sufficiently examined in existing theoretical research. This study finds that firms with better ESG performance are associated with significantly lower levels of tax avoidance. Further mechanism analysis shows that strong ESG performance helps ease firms' financing constraints, thereby weakening the incentive to rely on tax avoidance for capital accumulation. Additional heterogeneity analyses indicate that this inhibitory effect is more evident in non-state-owned enterprises, firms with higher-quality internal controls, and firms receiving greater analyst attention. The findings enrich the literature on the economic consequences of ESG performance and provide empirical support for promoting corporate ESG governance as a means of restraining aggressive tax avoidance behavior.

Keywords

ESG performance, corporate tax avoidance, financing constraints, mechanism analysis

1. Introduction

The 15th Five-Year Plan identifies carbon peaking and carbon neutrality as core guiding objectives, aiming to coordinate efforts in carbon reduction, pollution control, ecological expansion, and economic growth, while strengthening ecological security and enhancing the internal drivers of green development. In April 2026, the General Office of the CPC Central Committee and the General Office of the State Council issued the Opinions on Advancing Energy Conservation and Carbon Reduction at a Higher Level and with Higher Quality, further underscoring the Party and government's strong commitment to energy conservation, carbon peaking and neutrality, and the green transformation of development patterns. In this policy context, ESG performance has become increasingly important in corporate strategic planning and investment decision-making, and is now widely regarded as an important driver of green transformation and high-quality economic development.

Taxation constitutes the primary source of fiscal revenue, underpins government operations, and functions as an important tool for macroeconomic regulation and income redistribution [1]. Corporate tax avoidance has long been a focal issue for various stakeholders. However, few studies have examined tax avoidance from an ESG perspective. China's current tax audits primarily focus on compliance and the authenticity of financial reporting, with limited attention paid to firms' performance in environmental, social, and governance dimensions. Incorporating ESG factors into the assessment of corporate tax behavior may help improve existing frameworks for evaluating tax compliance. Against this background, this study examines whether corporate ESG performance affects firms' tax avoidance behavior and further explores the mechanisms through which such effects may arise.

The marginal contributions of this paper are threefold. First, this paper extends the literature on the determinants of corporate tax avoidance by introducing ESG performance into the analytical framework. By showing that firms with stronger ESG performance tend to engage in less tax avoidance, this study broadens existing research on the drivers of tax behavior and provides a useful reference for tax authorities seeking to incorporate ESG information into tax risk identification. Second, it complements the literature on the economic consequences of ESG performance. Prior studies have largely focused on the effects of ESG on firm value, stock returns, investment efficiency, financing costs, and green development [2–6]. In contrast to existing studies, this paper investigates the economic implications of ESG performance from the perspective of corporate tax behavior, thereby extending the scope of ESG-related research. Third, in the context of promoting green, low-carbon, and sustainable development, ESG has become a vital metric for assessing corporate value and social responsibility. While enhancing ESG performance entails governance costs, this study provides policy insights—viewed through the lens of tax avoidance—on how to better integrate ESG into corporate governance. The findings may also provide support for better coordination between tax governance and sustainable development policies.

2. Literature Review and Research Hypotheses

2.1 Literature Review

2.1.1 Determinants of Corporate Tax Avoidance

Corporate tax avoidance refers to the reduction of tax liabilities to the minimum level permitted by law. Existing studies have examined the determinants of corporate tax avoidance from both internal corporate factors and external institutional environments. Internal determinants include corporate governance mechanisms [7], managerial characteristics [8], and financing constraints [9]. External studies have primarily explored inter-jurisdictional tax competition among local governments [10] and the efficiency of tax enforcement. Notably, few studies have investigated the impact of ESG governance on corporate tax avoidance.

2.1.2 Economic Consequences of Corporate ESG Performance

ESG is a sustainability evaluation framework proposed by the United Nations Principles for Responsible Investment, encompassing environmental, social, and governance dimensions. It serves as a key indicator of non-financial performance. The literature has systematically documented various economic outcomes associated with ESG performance, such as reduced financing costs [3], enhanced firm value [2], facilitation of green transformation [4], and improved liquidity through ESG disclosure [11]. Although the literature on the economic consequences of ESG performance has expanded rapidly in recent years, relatively little attention has been paid to its relationship with corporate tax avoidance. This study therefore investigates the relationship between ESG performance and corporate tax avoidance, offering a meaningful extension to the literature and exploring the underlying mechanisms.

2.2 Theoretical Analysis and Research Hypotheses

2.2.1 Corporate ESG Performance and Tax Avoidance

Corporate ESG performance is closely linked not only to sustainable development but also to tax avoidance behavior. First, firms that actively embrace ESG principles maintain higher standards in environmental management, social responsibility, and internal governance. Firms with strong ESG performance are generally more likely to obtain green credit, attract ESG-oriented investment, and gain support from long-term institutional investors, which helps reduce both debt and equity financing costs. In addition, as regulators and stock exchanges place greater emphasis on ESG disclosure in financing reviews, firms with better ESG performance may find it easier to access external financing at lower cost, thereby weakening incentives to rely on tax avoidance. Second, strong ESG performance signals operational stability, which mitigates information asymmetry between the firm and capital providers. As a result, firms are more capable of meeting financing needs through conventional channels instead of adopting aggressive tax avoidance strategies that may damage corporate legitimacy and reputation. Third, superior ESG performance builds positive market reputation and moral capital. Once tax avoidance behavior is exposed, inconsistencies between a firm's ESG commitments and its actual tax practices may lead to substantial reputational damage. A firm that actively promotes environmental and social responsibility while simultaneously engaging in aggressive tax practices may face accusations of inconsistency or hypocrisy, thereby undermining its reputational capital. To maintain reputational credibility and preserve stakeholder trust, firms with strong ESG reputations are more likely to restrain tax avoidance behavior and keep their tax practices consistent with their public ESG commitments. Based on the above analysis, this paper proposes the following hypothesis:

H1: Ceteris paribus, firms with better ESG performance engage in lower levels of tax avoidance.

3. Research Design

3.1 Sample Selection and Data Sources

This study takes A-share listed companies on the Shanghai and Shenzhen Stock Exchanges from 2009 to 2024 as the research sample to examine the relationship between corporate ESG performance and tax avoidance behavior. To improve the reliability of the empirical results, the sample is screened using the following criteria: (1) exclusion of firms with ST or *ST designations; (2) exclusion of financial firms; and (3) exclusion of observations with missing data or extreme values for key variables. After applying these filters, the final sample consists of 36,586 firm-year observations. To alleviate the influence of extreme observations, all continuous variables are winsorized at the 1% and 99% levels. The data are sourced from the CSMAR and Wind databases.

3.2 Variable Definitions

3.2.1 Dependent Variable

The dependent variable is the degree of corporate tax avoidance (BTD). Existing studies generally measure corporate tax avoidance using either book-tax differences (BTD) or effective tax rates (ETR). Following Li Qingyuan and Chen Shilai [12], this study measures tax avoidance using the book-tax difference. Specifically, BTD is calculated as (pre-tax accounting profit − taxable income) / beginning-of-period total assets. A larger BTD value indicates a higher level of corporate tax avoidance.

3.2.2 Independent Variable

Corporate ESG performance is measured using the ESG rating data developed by Huazheng Index Information Service Co., Ltd. (Huazheng). This indicator serves as the core explanatory variable (ESG). Huazheng's ESG rating system draws on international ESG evaluation standards while taking into account the characteristics of China's capital market, and provides comprehensive assessments of firms' environmental, social, and governance performance.

3.2.3 Control Variables

Following Wang Yao [13], Li Shigang [14], Wang Linlin [2], and others, this study controls for the following variables: firm size (Size), leverage (LEV), profitability (ROA), fixed asset intensity (Fixed), inventory intensity (INV), cash flow (CFO), growth (GRO), ownership concentration (FIR), and ownership nature (SOE). Detailed variable definitions are presented in Table 1.

3.3 Model Specification

To test the effect of ESG performance on corporate tax avoidance, this study constructs the following baseline regression model, drawing on Liu Xing and Ye Kangtao [15]:

$$BTD_{i,t} = \beta_0 + \beta_1 ESG_{i,t} + \beta_2 Controls_{i,t} + \varepsilon_{i,t} \quad (1)$$

In Model (1), ESG represents the firm's ESG score and BTD denotes the degree of tax avoidance. If the coefficient β_1 is significantly negative, it suggests that stronger ESG performance is associated with lower levels of tax avoidance, thereby supporting Hypothesis H1.

Table 1: Variable Definitions

Variable Name	Symbol	Description
Corporate Tax Avoidance	BTD	(Pre-tax accounting profit − Taxable income) / Total assets
ESG Performance	ESG	Huazheng comprehensive ESG score
ESG Rating	ESGpj	Huazheng ESG rating
Firm Size	Size	Natural logarithm of total assets
Leverage	LEV	Total liabilities / Total assets
Profitability	ROA	Net profit / Total assets
Fixed Asset Intensity	Fixed	Fixed assets / Total assets
Inventory Intensity	INV	Net inventory / Total assets
Cash Flow	CFO	Net cash flow from operating activities / Total assets
Growth	GRO	Operating revenue growth rate
Ownership Concentration	FOR	Shareholding percentage of the largest shareholder
Ownership Nature	SOE	1 for state-owned enterprises, 0 otherwise

4. Empirical Results and Analysis

4.1 Descriptive Statistics

Table 2 reports the descriptive statistics of the main variables. For the tax avoidance measure (BTD), the mean is 0.003, the standard deviation is 0.028, and the median is 0.001, suggesting that the overall level of tax avoidance among sample firms is relatively low with moderate variation and a roughly symmetric distribution. The mean ESG score is 74.09, with a standard deviation of 5.093 and a median of 74.10, indicating that ESG performance is generally moderate and concentrated in the 75–80 range. ESG scores range from a minimum of 41.19 to a maximum of 100, reflecting substantial variation across firms. The distributions of the control variables are generally reasonable and consistent with those reported in previous studies.

Table 2: Descriptive Statistics

Variable	N	Mean	Std. Dev.	Min	Median	Max
BTD	25608	0.003	0.028	-0.083	0.001	0.171
ESG	25608	74.09	5.093	41.19	74.10	100
SOE	25608	0.337	0.473	0	0	1
CFO	25608	0.060	0.072	-0.742	0.058	0.850
Fixed	25608	0.212	0.158	0	0.180	0.954
LEV	25608	0.412	0.196	0.008	0.408	1.255
ROA	25608	0.059	0.048	-0.015	0.049	2.637
FIR	25608	34.89	15.12	2.197	32.88	89.99
GRO	25608	3.010	371.5	-2.897	0.117	59412
INV	25608	0.139	0.126	0	0.110	0.943
Size	25608	22.50	1.368	15.71	22.30	28.79

4.2 Baseline Regression Analysis

Table 3 presents the regression results examining the effect of ESG performance on corporate tax avoidance. The coefficient on the key explanatory variable ESG is significantly negative at the 1% level ($t = -2.65$). The results suggest that firms with better ESG performance tend to engage in less tax avoidance, possibly because stronger ESG governance improves information transparency, enhances external supervision, and alleviates agency problems. These findings provide empirical support for Hypothesis H1. The coefficients of the control variables are generally in line with expectations and consistent with prior research.

4.3 Endogeneity and Robustness Tests

4.3.1 Instrumental Variable Approach

To alleviate potential endogeneity concerns, this study adopts an instrumental variable approach. Management shareholding ratio (MSL) is chosen as the instrumental variable because it may affect ESG performance through managerial incentive mechanisms, while having no direct theoretical relationship with corporate tax avoidance. Table 4 reports the two-stage least squares (2SLS) results. In the first stage, MSL is significantly positively related to ESG (coefficient = 0.0235, $p < 0.01$), indicating that the relevance requirement for the instrumental variable is satisfied. In the second stage, the coefficient on ESG remains significantly negative at the 1% level (-0.0024), suggesting that the baseline conclusions remain robust after controlling for potential endogeneity.

Table 3: Baseline Regression Results

Variable	BTD
ESG	-0.000***
	(-2.65)
SOE	0.001
	(0.74)
CFO	-0.010
	(-1.59)
Fixed	0.006***
	(2.77)
LEV	-0.014***
	(-5.42)
ROA	0.185***
	(6.89)
FIR	-0.000***
	(-6.02)
GRO	-0.000***
	(-21.78)
INV	-0.008***
	(-3.12)
Size	0.004***
	(11.96)
Constant	-0.070***
	(-9.96)
N	25,608
R^2	0.1105

*Notes: t-statistics in parentheses; ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*

Table 4: Instrumental Variable Regression

Variable	ESG	BTD
MSL	0.0235***	
	(13.04)	
ESG		-0.0024***
		(-5.33)
Control variables	Yes	Yes
N	25,608	25,608
R^2	0.1045	0.02807

4.3.2 Fixed Effects Models

To further control for unobservable firm-level factors, this study additionally employs firm fixed effects and two-way fixed effects models. As shown in Table 5, the coefficient on ESG remains significantly negative at the 1% level in both specifications, consistent with the baseline results and supporting the robustness of our findings.

Table 5: Fixed Effects Models

Variable	(1) ESG	(2) ESG

ESG	-0.000**	-0.000**
	(-2.33)	(-2.17)
Control variables	Yes	Yes
N	25,608	25,608
R^2	0.113	0.137

4.3.3 Alternative Measure of the Independent Variable

As an additional robustness check, the continuous ESG score is replaced with the categorical ESG rating (ESGpj) issued by Huazheng. The rating consists of nine grades: AAA (score ≥ 95), AA (90–94), A (85–89), BBB (80–84), BB (75–79), B (70–74), CCC (65–69), CC (60–64), and C (score < 60). As reported in Table 6, the coefficient on ESGpj is significantly negative at the 1% level, further supporting the robustness of the empirical findings.

Table 6: Robustness Test: Alternative ESG Measure

Variable	BTD
ESGpj	-0.000**
	(-2.25)
Control variables	Yes
N	25,608
R^2	0.1103

4.3.4 Alternative Sample Period

To eliminate possible interference from the COVID-19 pandemic, the sample period is further restricted to the years before 2020. The results in Table 7 show that the coefficient on ESG remains significantly negative at the 1% level, indicating that our findings are robust.

Table 7: Robustness Test: Pre-Pandemic Sample

Variable	BTD
ESG	-0.000***
	(-3.05)
Control variables	Yes
N	16,208
R^2	0.0895

4.4 Mechanism Test

According to the theoretical analysis, ESG performance may reduce corporate tax avoidance by easing firms' financing constraints (SA). Firms with stronger ESG performance are generally able to obtain external financing at lower cost, which in turn reduces the incentive to rely on tax avoidance to ease financing pressure. Table 8 presents the mechanism test results following the "ESG performance–financing constraints–tax avoidance" pathway. Column (1) replicates the baseline result. Column (2) shows that ESG is significantly negatively associated with financing constraints (coefficient = -0.002, $p < 0.01$). In Column (3), when both ESG and SA are included, the coefficient on ESG remains significantly negative, while SA is significantly positive. The above results indicate the existence of a mediating effect: firms with better ESG performance tend to face fewer financing constraints, thereby weakening their incentives to obtain additional internal funds through tax avoidance activities.

Table 8: Mechanism Test

Variable	(1) BTD	(2) SA	(3) BTD
ESG	-0.000***	-0.002***	-0.000**
	(-2.65)	(-6.16)	(-2.11)
SA			0.013***
			(12.35)
Control variables	Yes	Yes	Yes
N	25,608	25,608	25,608
R^2	0.1105	0.0014	0.1185

4.5 Heterogeneity Analysis

4.5.1 Ownership Type

The effect of ESG performance on tax avoidance may vary across ownership types. State-owned enterprises face both economic and political costs of tax avoidance [16], whereas non-state-owned enterprises encounter greater resource constraints and competitive pressure [17]. Accordingly, the restraining effect of ESG performance on tax avoidance is expected to be more pronounced among non-SOEs. As shown in Table 9, in non-SOEs, the ESG coefficient is significantly negative at the 1% level, consistent with the baseline regression results. However, in SOEs, the ESG coefficient is insignificant. These results suggest that ownership structure moderates the relationship between ESG performance and corporate tax avoidance.

Table 9: Heterogeneity Analysis: Ownership Type

Variable	BTD	
	SOEs	Non-SOEs
ESG	-0.000	-0.000**
	(-0.85)	(-2.30)
Control variables	Yes	Yes
N	8,627	16,981
R^2	0.1128	0.1134

4.5.2 Internal Control Quality

Corporate tax avoidance involves substantial tax costs and risks, which require robust internal control systems for effective risk management. Existing studies show that internal control quality exerts a constraining effect on the extent of tax avoidance [18]. Accordingly, the constraining effect of ESG performance on tax avoidance is expected to be stronger among firms with effective internal controls. As reported in Table 10, the coefficient on ESG is significantly negative at the 1% level for firms with effective internal controls, consistent with the baseline regression results. However, the coefficient becomes insignificant for firms with internal control weaknesses. These findings suggest that the effectiveness of ESG performance in curbing tax avoidance depends, to some extent, on the quality of firms' internal control systems.

Table 10: Heterogeneity Analysis: Internal Control Quality

Variable	BTD	
	No IC Weakness	IC Weakness
ESG	-0.000***	0.000
	(-3.30)	(0.49)
Control variables	Yes	Yes
N	18,789	6,819
R^2	0.1089	0.1218

4.5.3 Analyst Coverage

Analyst coverage enhances corporate information transparency [19] and increases both the difficulty and risk of tax avoidance. Following the approach of Wang Yao et al. [20], this study measures analyst coverage using the number of analyst teams following the firm during the sample period. A dummy variable is constructed that equals 1 if the number of analysts exceeds the sample median, and 0 otherwise. Table 11 reports the subsample regression results. In the high analyst coverage group, the coefficient on ESG is significantly negative at the 1% level, consistent with the baseline findings. In contrast, the coefficient is insignificant in the low analyst coverage group. These findings imply that analyst coverage strengthens the restraining effect of ESG performance on corporate tax avoidance.

Table 11: Heterogeneity Analysis: Analyst Coverage

Variable	BTD	
	High Coverage	Low Coverage
ESG	-0.000**	-0.000
	(-2.33)	(-0.91)
Control variables	Yes	Yes
N	13,132	12,476
R^2	0.1170	0.1110

5. Conclusions and Policy Implications

Using a sample of A-share listed companies on the Shanghai and Shenzhen Stock Exchanges from 2009 to 2024, this study examines the relationship between corporate ESG performance and tax avoidance, as well as the mechanisms underlying this relationship. The empirical results indicate that: first, firms with superior ESG performance exhibit significantly lower levels of tax avoidance; second, the mechanism test shows that firms with better ESG performance face lower financing constraints, which in turn reduces their incentive to raise funds through tax avoidance; and third, the heterogeneity analysis reveals that the negative impact of ESG performance on tax avoidance is more pronounced among non-state-owned enterprises, firms with higher internal control quality, and firms with greater analyst coverage. A series of robustness checks, including alternative measures of the key explanatory variable, further support the reliability of the main findings.

Based on the above conclusions, this study offers the following implications:

(1) For listed companies: Strong ESG performance can weaken the incentive to retain internal cash flows through tax avoidance by alleviating financing constraints, while also improving internal control quality and attracting greater analyst scrutiny. Firms should therefore incorporate ESG practices into their long-term strategic planning. Specifically, companies should increase investments in environmental protection, social responsibility, and governance practices, strengthen internal control systems, and enhance information transparency to reduce information asymmetry. Greater scrutiny from analysts and institutional investors may also help discourage overly aggressive tax avoidance practices. Continuous improvement in ESG performance can help firms reduce financing costs, improve access to external capital, and enhance their reputation for tax compliance, thereby supporting more sustainable long-term development.

(2) For investors: ESG performance can serve as a valuable indicator when assessing potential investees' tax behavior and compliance risks. Investors should prioritize firms with strong ESG ratings, reasonable and stable effective tax rates, and sound governance. By contrast, firms with consistently poor ESG performance may face higher risks associated with aggressive tax practices and should therefore be evaluated with caution.

(3) For government authorities: The finding that ESG performance is associated with lower tax avoidance suggests that regulators may strengthen tax governance indirectly through improved ESG-related incentive and supervisory mechanisms. One possible approach is to incorporate ESG indicators into tax administration practices by granting firms with consistently strong ESG performance certain compliance conveniences, while increasing scrutiny of firms with weak ESG records and abnormal tax behavior. In addition, regulators should continue improving ESG disclosure standards and strengthen information sharing among tax authorities, environmental regulators, and third-party rating agencies in order to enhance the effectiveness of tax supervision and policy implementation.

References

[1] State Taxation Administration. (2013, April 18). The functions and roles of taxation. China Economic Net.

[2] Wang, L. L., Lian, Y. H., & Dong, J. (2022). Research on the influence mechanism of ESG performance on corporate value. *Securities Market Herald*, (05), 23–34.

[3] Qiu, M. Y., & Yin, H. (2019). Corporate ESG performance and financing cost under the background of ecological civilization construction. *Journal of Quantitative & Technological Economics*, 36(03), 108–123.

[4] Hu, J., Yu, X. R., & Han, Y. M. (2023). Can ESG rating promote corporate green transformation? Verification based on the multi-phase difference-in-differences method. *Journal of Quantitative & Technological Economics*, 40(07), 90–111.

[5] Li, Z. F. (2023). Corporate ESG performance and the use of short-term debt for long-term investment. *Journal of Quantitative & Technological Economics*, 40(12), 152–171.

[6] Anselmi, G., & Petrella, G. (2024). ESG ratings: Disagreement across providers and effects on stock returns. *Journal of International Financial Markets, Institutions and Money*, 102133.

[7] Bassyouny, H., & Aboud, A. (2024). Tax avoidance and narrative disclosure tone: An ethical dimension of corporate governance. *European Journal of Finance*, 31(14), 1713–1738.

[8] Li, J. Y., Deng, Y. W., & Zhang, M. (2020). Local CEOs and corporate tax avoidance: Hometown identity or rent-seeking? *Accounting Research*, (07), 119–130.

[9] Chen, Z. H., & Fang, H. X. (2018). Financial constraints, internal control and corporate tax avoidance. *Journal of Management Sciences in China*, 31(03), 125–139.

[10] Fan, Z. Y., & Tian, B. B. (2013). Tax competition, tax enforcement and corporate tax avoidance. *Economic Research Journal*, 48(09), 99–111.

[11] Krueger, P., Sautner, Z., & Tang, D. Y. (2024). The effects of mandatory ESG disclosure around the world. *Journal of Accounting Research*, 62(5), 1795–1847.

[12] Li, Q. Y., & Chen, S. L. (2021). Corporate listing and tax avoidance. *Journal of World Economy*, 44(11), 169–193.

[13] Wang, Y., Zhang, Y. M., & Hou, D. S. (2022). Does corporate ESG performance affect audit opinions? *Audit and Economic Research*, 37(05), 54–64.

[14] Li, S. G., Zeng, C., Huang, Y. S., et al. (2024). Common institutional shareholders and corporate tax avoidance. *Journal of Management Sciences in China*, 37(04), 126–146.

[15] Liu, X., & Ye, K. T. (2014). Financial development, property rights and corporate tax burden. *Management World*, (03), 41–52.

[16] Xie, J., Tang, G. P., & Xiang, Y. R. (2016). Management ability, property right nature and corporate tax avoidance. *Journal of Jiangxi University of Finance and Economics*, (02), 43–59.

[17] Wu, F., Hu, H. Z., Lin, H. Y., et al. (2021). Corporate digital transformation and capital market performance: Empirical evidence from stock liquidity. *Management World*, 37(07), 130–144+10.

[18] Bauer, A. M. (2016). Tax avoidance and the implications of weak internal controls. *Contemporary Accounting Research*, 33(2), 449–486.

[19] Zhang, Z. X., & Zhou, J. J. (2019). Can analyst coverage improve the information transparency of listed companies? Based on the perspective of earnings management. *Research on Financial and Economic Issues*, (12), 49–57.

[20] Wang, Y., Zhang, Y. M., & Hou, D. S. (2022). Does corporate ESG performance affect audit opinions? *Audit and Economic Research*, 37(05), 54–64.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Environment, Social and Governance; Vol.3, No.3; 2026
E-ISSN: 3006-3973 P-ISSN：3006-3965
DOI: https://doi.org/10.70267/ esg2026.v3n3.5863
Published by: Zeus Press

A Study on the Application of Google Earth Engine for Precipitation Monitoring

Weicheng Lin*

School of Atmospheric Sciences, Nanjing University of Information Science and Technology, Nanjing, China

**Corresponding author: Weicheng Lin.*

Abstract

Monitoring precipitation is an important basis for climate change studies and for the prevention and control of hydrological disasters. There are currently skewed distributions of ground-based meteorological observation stations, and satellite-based precipitation products still lack sufficient spatial resolution and systematic accuracy, directly influencing the accuracy of runoff models and early flood and drought warning systems. In this paper, the advancement in precipitation monitoring on the Google Earth Engine (GEE) platform was reviewed. It methodically evaluates the synthesis and use of mainstream precipitation products, including IMERG, CHIRPS, and ERA5-Land, as well as complementary information on topography and vegetation, and ground-truth validation data on the GEE cloud platform. The article focuses on the essence of technical workflow procedures for the precise spatial resolution of precipitation in complex terrain and urban environments by combining spatial-decay techniques with machine-learning algorithms, including Random Forest and XGBoost, on the platform. It also examines existing issues, such as inefficiencies in the parameter optimisation of prebuilt algorithms at GEE, the lack of deep learning functionality, and the inability to validate multi-source precipitation products under extreme precipitation conditions. Moreover, the paper outlines future development directions, including deep integration with external AI frameworks and the implementation of higher-resolution spatiotemporal observational data.

Keywords

satellite remote sensing, precipitation monitoring, Google Earth Engine

1. Introduction

Precipitation plays an essential role in the global water cycle, and its spatiotemporal variations significantly affect the distribution of water resources in the region, ecological processes, and the survival of natural disasters, including floods and droughts. In the context of climate change, the spatiotemporal distribution of precipitation is characterised by high uncertainty and regional variability, and the proportion of extreme precipitation events increases in certain regions. Thus, it is particularly important to conduct long-term, continuous, and high-precision precipitation monitoring studies to understand climate change processes better and enhance disaster prevention and mitigation [1].

In the modern world, the main sources of precipitation monitoring are ground-based weather station measurements and satellite-based remote sensing data. This is because ground-based meteorological stations are useful for providing credible precipitation data, but have very limited spatial coverage. The use of satellite-based precipitation products has steadily become a valuable source of data with the evolution of remote sensing technology; such missions include TRMM (Tropical Rainfall Measuring Mission) and GPM (Global Precipitation Measurement mission), which have provided continuous global precipitation measurements [2]. Moreover, there are multi-source fusion products, including the CHIRPS precipitation dataset and the ERA5 reanalysis dataset, which are actually used in climate change and hydrological studies [3]. Nevertheless, such precipitation products have limited spatiotemporal resolution and systematic biases (e.g., high-intensity rainfall is overstated and solid-state and trace precipitation are underestimated). This can lead to an inversion in the spatial distribution of precipitation within a watershed, thereby undermining the reliability of runoff simulation, early flood and drought warnings, and water resource distribution. GEE, with its built-in big-data and cloud-computing solutions, can be used to merge, evaluate the quality of, and analyse multi-source precipitation data, which could help overcome barriers such as high data acquisition costs, data-heavy preprocessing, and limited computing power and throughput [4]. Thus, it is of great importance to synthesise the research achievements in the domain of GEE for precipitation monitoring. The current paper reviews recent precipitation-monitoring research conducted on the GEE platform and serves as a reference for related work.

It is against this development backdrop that this paper will take a logical structure of Platform and Data Sources, Key Methods, Applications and Outlook. First, it presents an overview of the background and methodological value of the GEE platform, second, it presents the precipitation products and validation data systems that are accessible using GEE, third, it summaries typical workflows that include remote sensing inversion, spatial downscaling, and machine learning, and lastly, it explains current challenges and the future research directions to become an academic source of high resolution precipitation monitoring studies using GEE.

2. Methodological Value of Gee in Precipitation Monitoring Research

2.1 Platform Background and Technical Framework

GEE is a distributed geospatial big-data analytics platform, a cloud-based system created and managed by Google to support research on environmental change at a global scale. It is a platform that integrates multi-source remote sensing, reanalysis, and meteorological observation data and provides online computing services based on a distributed parallel computing framework.

In hydrological studies, the usefulness of GEE goes beyond the ease of data access; it is characterised by providing a common computational platform for examining and processing precipitation data across long time series at a variety of spatial scales. This homogeneity is crucial for watershed-scale hydrological modelling and climate change response studies.

2.2 Precipitation Monitoring Data Sources of Gee

The essence of precipitation monitoring lies in selecting and processing data sources. Since GEE is a cloud-based centre intended to enable global-scale geospatial analysis, its key benefit is the combination of multi-source Earth observation data and a range of environmental information spanning more than 40 years, which offers a one-stop database for the study of precipitation values on a large-scale, long-term basis. The platform not only efficiently combines long-term image series from mainstream imagers such as Landsat, MODIS, and Sentinel, but also methodically presents multidimensional data on topography, meteorology, land cover, and socioeconomic variables. This allows scientists to go beyond the constraints of individual data sources and easily retrieve and jointly process multiscale, multiplex data on precipitation processes through an integrated cloud platform [5]. According to this, the precipitation monitoring data ecosystem created by this paper via GEE comprises three main tiers: core satellite precipitation products, key environmental covariates, and ground-based validation data. The remainder of this paper provides a systematic overview and explanation of the data sources for each tier.

The use of satellite-based precipitation products as the central data source in precipitation monitoring research also requires that the most frequently used products have their own focus and unique research interests. With the GPM IMERG series, a third-generation advancement of the TRMM mission, the latest state-of-the-art multi-satellite combined precipitation retrieval exists. It combines passive microwave, infrared remote sensing, and ground-based rain-gauge data through calibration, fusion, and interpolation algorithms to generate a world-gridded precipitation dataset [6]. In general, studies on the GEE platform apply the Final Run version (V06 or V07) calibrated by GPCC, which has high data reliability, the CHIRPS dataset is specifically designed to support drought monitoring and long-term climate change analysis and is based on high-resolution 0.05 o infrared data which combines observations made by world weather stations to deliver continuous global products of precipitation since 1981 [7]. Its longer length of the time series, and lower spatial resolution than IMERG, has led to its acceptance as a fundamental contributor to the long-term series of precipitation trends and interdecadal drought research: the land component of the ERA5 product, ERA5-Land, is based on superior data assimilation systems and numerical models, in achieving physically coherent inversion of land-surface variables like precipitation. It can exhibit spatiotemporal continuity and physical consistency due to its 0.1° spatial resolution and 1-hour temporal resolution, which is why it is widely used as a reference dataset to assess the plausibility of satellite precipitation products and to compare multi-source data.

Important environmental covariates can provide crucial assistance in downscaling precipitation patterns or even aid in the mechanistic analysis of the phenomenon. However, topography is the most important covariate when defining the spatial variation in precipitation conditions. To properly derive topographic parameters, i.e., elevation, slope, and aspect, the 30-meter-resolution SRTM Digital Elevation Model (DEM) data made available by the GEE platform can be used to correct systematic biases in satellite-based precipitation products, such as IMERG, in complex terrain, such as mountainous regions [8]. The vegetation and surface thermal state data (MODIS NDVI/EVI and LST) are also considered crucial: the vegetation indices are relevant indicators of surface moisture conditions and the potential for evapotranspiration, whereas the surface temperature is directly proportional to the severity of local convective activity. Both are lagging or instantaneous measures of precipitation and are fundamental input characteristics for high-resolution precipitation forecasting. In addition, ancillary information, e.g., land cover types and proximity to water bodies, also contributes to a more in-depth characterisation of the complex effects of surface characteristics on the spatial distribution of precipitation. A combination of all these covariates allows the coarse-resolution precipitation fields to be downsampled to the kilometre or even hundred-meter scale, thereby realising greater utility for fine-tuning hydrological management and disaster anticipation and control.

Ground-based validation data provide direct support for the analysis and calibration of the accuracy of satellite precipitation products. GEE has an embedded Global Summary of the Day (GSOD) database that includes basic meteorological parameters, such as daily precipitation. Direct code can be written to retrieve station data for specific locations and centuries to compute accuracy metrics, including correlation coefficients, root-mean-square error, and bias, and to contrast them with satellite precipitation products. This is one of the most widely used data sources for confirming precipitation products. In areas where ground observation networks are well established (as is the case in China), on top of the above global data, there is frequently a combination of high-density station data issued by organizations like the National Meteorological Information Centre (it can be accessed by uploading user-ownable content in GEE or connecting to other APIs) to more stringent and granular local accuracy measurements [9].

3. Monitoring Precipitation Using Gee

3.1 Remote Sensing Inversion Method

One of the most popular approaches to the monitoring of precipitation is remote sensing inversion. In the quest to improve inversion accuracy, research usually integrates downscaling approaches with machine learning algorithms. Satellites such as TRMM and GPM provide precipitation products characterised by temporal continuity and extensive spatial coverage, thereby supporting global and regional precipitation estimates and playing important roles in monitoring both short-duration heavy precipitation and extreme weather [2, 10]. To enhance the reliability of local precipitation forecasts, scientists typically combine satellite data and ground-based rain gauges. They construct precipitation prediction models using machine learning or statistical methods, incorporating environmental factors such as elevation, vegetation indices, and land cover to improve inversion performance in complex terrain and urban areas [11, 12]. With the GEE platform, researchers will have direct access to various satellite-based precipitation products and the ability to fuse them with other auxiliary data, including MODIS vegetation indices and SRTM elevation data, for spatial and temporal matching and preprocessing. This facilitates spatial downscaling and high-resolution retrieval, as well as the joint operation of various methods.

3.2 Spatial Downscaling Method

There are also significant benefits to the spatial downscaling of precipitation products provided by GEE. Using the example of TRMM 3B43 V7, the instrument's 0.25° spatial resolution is not very favourable for imaging the spatial heterogeneity of precipitation in small-watershed or urban-scale studies due to topographic and land-cover differences. TRMM precipitation data, vegetation indices (NDVI, EVI, and LAI) from MODIS, SRTM elevation, and latitude/longitude can be accessed online using the GEE platform. Multi-source data are unified via spatiotemporal registration in the cloud, followed by the construction of predictor datasets at two resolutions: 25 km and 1 km. On this basis, nonlinear models of precipitation influenced by geographic environmental factors are developed, with coarse-resolution precipitation data downsampled to 1 km. The residual correction and scale decomposition methods are then used to generate finer-resolution spatiotemporal precipitation products, and ground station information is utilised to assess accuracy and conduct spatial pattern analysis. The complete workflow leverages GEE's strength in processing massive volumes of remote sensing data, cloud computing, and multi-scale raster operations to bring TRMM precipitation down to 1 km resolution [13].

3.3 Machine Learning and Artificial Intelligence Methods

Machine learning and artificial intelligence technologies are now available through the GEE platform, offering new directions in precipitation monitoring, with considerable benefits for complex terrain, multi-source data fusion, and high-resolution inversion. Using combinations of satellite precipitation products, ground-based observation data, and auxiliary variables (elevation, vegetation indices, and land cover), regression and classification models can be developed in the cloud to provide nonlinear estimates of precipitation. Random Forest, XGBoost, and CNN models are also efficient at capturing spatial features of precipitation, and the LSTM model is better placed to capture temporal changes [11, 12]. These are trainable, valid and predictable with an in-built EE in the GEE. Python or Classifier API, and take advantage of the benefits of cloud computing to process very large, all-around data in time-series [14]. It has been established that precipitation estimation is more accurate through machine learning than conventional models or single-satellite products. It is very effective in augmenting the detail and reliability of spatial distribution in mountainous locations, small watersheds and urban regions. Moreover, in conjunction with downscaling and residual-correction techniques, it may produce high-resolution, temporally continuous precious information that can reliably support disaster prevention and mitigation, water resource management, and climate-change studies [11, 12, 15].

4. Conclusion

This paper presents a systematic review of advances in precipitation monitoring using the Google Earth Engine cloud platform. The review states that by leveraging its strong cloud-based, distributed computing capabilities and its ability to integrate large volumes of multi-source data, GEE has radically changed the conventional research paradigm in precipitation monitoring. At the data-source level, GEE combines long-term precipitation forecasts, including GPM IMERG, CHIRPS, and ERA5-Land. With these, coupled with environmental covariates such as SRTM topography, MODIS vegetation indices, and GSOD ground-truth validation data, it has defined a multidimensional data ecosystem that has been proven using satellite, Earth-surface, and model data, known as satellite-ground-model validation. This is quite efficient in overcoming the weaknesses of very high data acquisition barriers and tedious preprocessing experienced in conventional studies.

Methodologically, research studies are currently moving away from single-technique remote sensing inversion and towards multi-technique fusion. More specifically, the joint use of spatial downscaling techniques and machine learning procedures is becoming more mature, thanks to the GEE platform. Studies have shown that, with appropriate algorithms such as Random Forest and XGBoost, coarse-resolution satellite precipitation products can be downsampled to 100-meter or kilometre resolution, and the spatial extent of precipitation characterisation in compressed terrain and urban regions can be improved by more than an order of magnitude.

Even though the GEE platform is very convenient at monitoring precipitation, the existing research still presents some challenges: first of all, the machine learning algorithms implemented into the GEE platform are still limited about hyperparameter optimization and called to support deep learning structures, second, the dissimilarities between the applicability of the multi-source precipitation products in the context of the various climatic regions and the severe precipitation events themselves need to be cross-validated more critically. Eventually, as GEE is further integrated with more dynamic overseas AI systems (such as TensorFlow) and local satellite data with denser spatiotemporal coverage are added, cloud-based precipitation monitoring studies will provide greater value for finer hydrological predictions and real-time emergency response to disasters.

References

[1] K.E. Trenberth, Changes in precipitation with climate change. Climate Res. 47, 123-138 (2011)

[2] G.J. Huffman, D.T. Bolvin, D. Braithwaite, et al., Integrated multi-satellite retrievals for the global precipitation measurement (GPM) mission (IMERG). In: V. Levizzani, C. Kidd, D.B. Kirschbaum, et al. (Eds.), Satellite Precipitation Measurement (Springer, Cham, 2020), pp. 343-353.

[3] H.E. Beck, E.F. Wood, M. Pan, et al., MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. Hydrol. Earth Syst. Sci. 21, 589-615 (2017)

[4] X. Wang, J. Tian, X. Li, et al., The transformation of remote sensing development by Google Earth Engine cloud platform. Natl. Remote Sens. Bull. 26, 299-309 (2022) Mutanga, L. Kumar, Google Earth Engine applications. Remote Sens. 11, 591 (2019)

[5] C. Wang, Y. Hong, A review of inversion, verification and application of satellite remote sensing precipitation. Water Resour. Hydropower Eng. 49, 1-9 (2018)

[6] C.C. Funk, P.J. Peterson, M.F. Landsfeld, et al., The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. Sci. Data 2, 150066 (2015)

[7] T.G. Farr, P.A. Rosen, E. Caro, et al., The Shuttle Radar Topography Mission. Rev. Geophys. 45, RG2004 (2007) Smith, N. Lott, R. Vose, The integrated surface database: recent developments and partnerships. Bull. Am. Meteorol. Soc. 92, 704-708 (2011)

[8] A.Y. Hou, G. Skofronick-Jackson, R.K. Kakar, The global precipitation measurement (GPM) mission: overview and early results. Bull. Am. Meteorol. Soc. 95, 701-722 (2014)

[9] Y. Tian, Y. Chen, M. Zhang, et al., Improving satellite precipitation estimates using random forest regression and topographic information. Remote Sens. 11, 1906 (2019)

[10] A.Y. Sun, G. Tang, Downscaling satellite and reanalysis precipitation products using attention-based deep convolutional neural nets. Front. Water 2, 536743 (2020). Elnashar, H. Zeng, B. Wu, et al., Downscaling TRMM monthly precipitation using Google Earth Engine and Google cloud computing. Remote Sens. 12, 3860 (2020)

[11] X. Hu, Y. Zhang, J. Wang, et al., A machine learning approach for high-resolution daily precipitation estimation using satellite and gauge data. Remote Sens. 12, 3200 (2020)

[12] M.A.E. Bhuiyan, E.I. Nikolopoulos, E.N. Anagnostou, Machine learning–based blending of satellite and reanalysis precipitation datasets: A multiregional evaluation of tropical complex terrain. J. Hydrometeorol. 20, 2147-2161 (2019).

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment
This paper is an output of the science project.

Environment, Social and Governance; Vol.3, No.3; 2026
E-ISSN: 3006-3973 P-ISSN：3006-3965
DOI: https://doi.org/10.70267/ esg2026.v3n3.6472
Published by: Zeus Press

Carbon Constraints and Corporate Green Innovation: Evidence from China's National Carbon Market

Xiaoyuan Liu*

School of Economics, Hebei University, Baoding, China

**Corresponding author: Xiaoyuan Liu.*

Abstract

Against the background of China's continuing pursuit of the dual-carbon goals, whether the national carbon emissions trading market can promote corporate green innovation through a market-based mechanism has become an important question for evaluating the effectiveness of environmental regulation. Using A-share listed companies from 2018 to 2023 as the research sample, this study treats the official launch of the national carbon market in 2021 as a quasi-natural experiment. Because the first compliance cycle mainly covered the power generation sector, firms in industry code D44, namely electricity and heat production and supply, are identified as the treatment group. A difference-in-differences model is then constructed to examine the impact of the national carbon market on corporate green innovation. The results show that, after the launch of the national carbon market, the total number of green patent applications by treated firms increased significantly relative to control firms. Mechanism tests indicate that the policy significantly reduced carbon emission intensity among treated firms, and carbon emission intensity is negatively associated with green innovation, suggesting that carbon constraints are an important channel through which the carbon market affects green innovation. Heterogeneity analyses further show that firms with higher pre-policy attention to green transformation respond more strongly to the policy. In this sample, firms with higher financing constraints exhibit stronger policy responses than firms with lower financing constraints. Dynamic DID estimates show no significant pre-policy differential trend between treated and control firms, while the policy effect gradually emerges after 2021. Placebo tests using the pre-policy period, an alternative post-policy window, an alternative dependent variable, and PSM-DID tests generally support the baseline conclusion. The findings suggest that China's national carbon market not only imposes emission-reduction constraints but may also influence corporate green innovation through carbon-cost pressure and expectations of low-carbon transition.

Keywords

national carbon market, green innovation, carbon emission intensity, difference-in-differences, financing constraints

1. Introduction

Climate governance is reshaping the external institutional environment and internal strategic choices of firms under China's carbon-peaking and carbon-neutrality policy framework [1]. Compared with traditional command-and-control environmental regulation, carbon emissions trading schemes transform the scarcity of emission rights into observable operating costs and asset opportunities through emission allowances, carbon price signals, and compliance obligations [2]. Firms that continue to rely on high-carbon production modes face greater compliance pressure. By contrast, firms that reduce carbon emissions per unit of output through energy-saving retrofits, process optimization, and green technology research and development may save allowance costs and even obtain additional carbon-asset returns. In this sense, a carbon market is not only an emission-reduction instrument, but also a market-based policy tool that may stimulate corporate green innovation.

China's national carbon market officially began trading on 16 July 2021. Its first compliance cycle started with the power generation industry [3,4]. Selecting the power sector as the initial coverage industry has clear policy significance. On the one hand, the sector has large-scale carbon emissions and a relatively mature emissions accounting foundation, making it suitable for piloting a unified national carbon pricing mechanism [5]. On the other hand, power enterprises are central actors in energy transition and green technology diffusion, and changes in their innovation behavior may affect the electricity structure, energy efficiency, and low-carbon transition of downstream industries.

From the perspective of corporate decision-making, the effect of a carbon market on green innovation is not merely a question of whether firms reduce emissions. It involves a comprehensive adjustment of cost constraints, technological routes, organizational cognition, and financing conditions. In response to carbon-market regulation, firms may adopt short-term compliance strategies, such as purchasing allowances or reducing output, but they may also adopt long-term transformation strategies, such as equipment upgrading, process optimization, and green R&D. Which strategy dominates depends on the intensity of carbon constraints, the firm's original emission level, managerial attention to green transformation, and financing constraints.

This paper contributes to the literature in three ways. First, it focuses on the official launch of China's national carbon market, rather than regional pilot programs, and therefore helps explain the micro-level policy effects of China's transition from regional carbon pilots to a unified national market. Second, beyond estimating the baseline DID effect, the study examines the mechanism of carbon emission intensity and explores heterogeneity by pre-policy green-transformation attention and financing constraints. Third, the paper conducts dynamic DID analysis, a pre-policy placebo test, and PSM-DID estimation to strengthen the credibility of causal identification.

2. Literature Review and Hypothesis Development

2.1 Environmental Regulation, Carbon Markets, and Green Innovation

The relationship between environmental regulation and technological innovation has long been debated. The Porter hypothesis argues that well-designed environmental regulation may stimulate innovation and offset part of the compliance cost [6]. Empirical studies, however, have found that regulatory effects depend on the type of regulation, policy intensity, and firm characteristics [7]. In the field of energy and environmental economics, induced innovation theory suggests that changes in relative factor prices can redirect technological progress toward technologies that save the now more costly or scarcer factor [8]. Directed technical change models further show that policy incentives can affect the direction of innovation in clean and dirty technologies [9].

Carbon emissions trading is a typical market-based environmental regulatory instrument. By assigning scarcity and transaction value to carbon emission allowances, a carbon market can change firms' expected costs and returns. Evidence from the European Union Emissions Trading System shows that carbon pricing can guide directed technological change and low-carbon patenting [10]. In China, recent studies have examined the effects of carbon emissions trading schemes on green technological innovation, carbon performance, and carbon-market mechanisms [11,12,13,14]. However, compared with regional pilots, the national carbon market has broader institutional significance and requires further firm-level evidence.

2.2 Carbon Constraints and Innovation Incentives

Carbon emissions trading incorporates corporate carbon emissions into a market-based cost-accounting system through allowance constraints, carbon price signals, and compliance mechanisms. For power-sector enterprises, carbon allowances gradually become production factors with opportunity costs. The higher a firm's emission intensity, the greater the compliance pressure and potential carbon costs it faces. After the national carbon market is launched, high-carbon production modes are constrained, and the importance of green technology R&D and low-carbon transformation increases.

From the cost-constraint perspective, the carbon market changes the relative prices of production factors. When carbon emission allowances become scarce and tradable, firms maintaining high-emission production modes need to bear higher compliance costs. In contrast, firms that reduce emissions per unit of output through low-carbon combustion, energy-saving dispatch, pollution co-control, and digital energy-efficiency management can mitigate carbon-cost pressure. From the incentive perspective, carbon markets also increase the expected returns of green innovation. If green innovation lowers carbon emission intensity, firms may reduce allowance purchases and compliance expenditures, while potentially generating carbon-asset gains from saved allowances.

Therefore, compared with purely administrative regulation, emissions trading leaves firms with more autonomy to choose among emission reduction, technology investment, and allowance trading. The carbon market may promote green R&D and patenting through the dual mechanism of cost constraints and expected returns. Accordingly, this paper proposes the following hypothesis:

H1: The launch of the national carbon market significantly promotes green innovation among listed firms in the power sector.

2.3 Mechanism and Heterogeneity

Carbon emission intensity is an important mechanism linking the national carbon market and green innovation. The carbon market first affects firms' emission behavior. To reduce emissions per unit of output and lower compliance pressure, firms may increase investment in energy conservation, equipment upgrading, and process optimization. These low-carbon transformation activities often require technical support and may further appear as increases in green patent applications. Therefore, this study proposes:

H2: The national carbon market promotes green innovation by reducing corporate carbon emission intensity.

Organizational attention theory suggests that firms facing external institutional shocks differ in their response speed and intensity. Firms that already paid more attention to green transformation, environmental responsibility, and low-carbon development before the policy shock are likely to have better information foundations, managerial cognition, and resource preparation. After the launch of the national carbon market, these firms may be better able to recognize strategic opportunities embedded in carbon price signals and transform policy pressure into green R&D action. Thus:

H3: The higher a firm's pre-policy attention to green transformation, the stronger the promoting effect of the national carbon market on green innovation.

Financing constraints may also shape firms' innovation responses to carbon-market regulation. In general, firms with lower financing constraints are better able to support long-term green R&D. However, under a carbon-market regime, highly constrained firms that maintain high-carbon production may face stronger compliance costs and cash-flow pressure. They may therefore have stronger incentives to reduce future carbon costs through technological upgrading and green patent deployment. This leads to the following hypothesis:

H4: The promoting effect of the national carbon market on green innovation is stronger among firms with higher financing constraints than among firms with lower financing constraints.

3. Methodology

3.1 Sample Selection and Data Sources

This study uses A-share listed companies from 2018 to 2023 as the initial sample. The data are obtained from the CSMAR database and text-based indicators related to green transformation disclosed in listed companies' annual reports. Industry classification follows the Guidelines for the Industry Classification of Listed Companies issued by the China Securities Regulatory Commission [15]. After data cleaning, the sample contains 35,628 firm-year observations covering 5,938 listed firms. Among them, 96 firms in industry code D44, namely electricity and heat production and supply, are classified as the treatment group.

Because some financial control variables and carbon emission intensity variables are missing, the baseline regression, which jointly controls for firm fixed effects, year fixed effects, and firm-level controls, uses 20,425 effective observations. The mechanism test for carbon emission intensity uses 14,248 effective observations. To reduce the influence of extreme values, continuous variables are winsorized where appropriate.

3.2 Variable Definitions

Table 1: Variable Definitions

Variable type	Variable name	Definition
Green innovation	ln_green_total	ln(1 + green invention patent applications + green utility-model patent applications)
Treatment group	Treat	Equals 1 for firms in industry code D44, i.e., electricity and heat production and supply enterprises; 0 otherwise
Post-policy period	Post	Equals 1 for 2021 and subsequent years; 0 otherwise
Core DID term	DID	Treat x Post
Carbon emission intensity	carbon_intensity	Corporate carbon emission intensity; used as the mechanism variable
Green attention	attention_base	Equals 1 if the firm's pre-policy green-transformation word frequency is above the median
Low financing constraints	lowfc_base	Equals 1 if the absolute value of the pre-policy SA index is below the median
Control variables	Controls	Firm size, growth, cash-to-assets ratio, ownership concentration, ROA, and Tobin's Q

3.3 Model Specification

The baseline DID model is specified as follows:

$$GreenInnovation_{it} = \alpha + \beta DID_{it} + \gamma Controls_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (1)$$

where GreenInnovation_it denotes corporate green innovation; DID_it is the interaction between the treatment group and the post-policy period; Controls_it represents firm-level control variables; μ_i denotes firm fixed effects; λ_t denotes year fixed effects; and ε_it is the error term. Standard errors are clustered at the firm level to address within-firm serial correlation in DID estimation [16]. A significantly positive β indicates that, after the launch of the national carbon market, green innovation among treated firms increased significantly relative to control firms.

The mechanism test is conducted in two steps. First, carbon emission intensity is used as the dependent variable to test whether the DID term significantly reduces the carbon emission intensity of treated firms:

$$CarbonIntensity_{it} = \alpha + \beta DID_{it} + \gamma Controls_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (2)$$

Second, carbon emission intensity is added to the green innovation model to examine the relationship between carbon intensity and green innovation:

$$GreenInnovation_{it} = \alpha + \beta DID_{it} + \delta CarbonIntensity_{it} + \gamma Controls_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (3)$$

For heterogeneity tests, this study further includes interaction terms between DID and pre-policy green attention, and between DID and the low-financing-constraint group:

$$GreenInnovation_{it} = \alpha + \beta DID_{it} + \theta(DID_{it} \times Moderator_i) + \gamma Controls_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (4)$$

4. Empirical Results and Analysis

4.1 Descriptive Statistics

Table 2 presents the descriptive statistics. The mean value of ln(1 + total green patents) is 0.683, and the raw count of total green patents varies substantially across firms, indicating obvious heterogeneity in green innovation. The mean of the DID variable is relatively small because the treated firms account for a limited share of all A-share listed companies. The carbon emission intensity variable also shows marked dispersion, which provides empirical variation for the mechanism test.

Table 2: Descriptive Statistics

Variable	N	Mean	SD	Min.	Max.
ln(1 + total green patents)	35,628	0.683	1.079	0.000	4.466
Total green patents	35,628	5.785	36.666	0.000	1,445.000
DID	35,628	0.007	0.084	0.000	1.000
Carbon emission intensity	18,337	0.462	0.587	0.038	2.267
Firm size	21,963	22.425	1.406	19.616	26.684
Growth	26,130	0.118	0.381	-0.678	2.189
Cash-to-assets ratio	33,088	0.166	0.132	0.006	0.640
Ownership concentration	26,975	0.325	0.147	0.081	0.733
ROA	33,093	0.039	0.092	-0.411	0.265
Tobin's Q	26,607	1.904	1.240	0.813	8.310

4.2 Baseline Results, Mechanism Tests, and Heterogeneity

Table 3: Baseline Regression, Mechanism Tests, and Heterogeneity Analyses

Variable	Baseline	Mechanism 1	Mechanism 2	Attention	Financing constraints
Dependent variable	Green innovation	Carbon emission intensity	Green innovation	Green innovation	Green innovation
DID	0.321*** (0.077)	-0.131*** (0.013)	0.326*** (0.078)	-0.129 (0.134)	0.411*** (0.099)
Carbon emission intensity			-0.082* (0.044)		
DID x High attention				0.470*** (0.155)	
DID x Low financing constraints					-0.313** (0.135)
Observations	20,425	14,248	14,248	20,425	20,425
Firm clusters	4,639	3,378	3,378	4,639	4,639
Firm/year fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Within R^2	0.023	0.038	0.026	0.023	0.023

*Note: Standard errors clustered at the firm level are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*

Column (1) of Table 3 reports the baseline DID result. The coefficient of DID is 0.321 and is significant at the 1% level, indicating that, after the launch of the national carbon market, the total number of green patent applications by D44 treated firms increased significantly relative to firms in other industries. Because the dependent variable is ln(1 + x), the result has not only statistical significance but also clear economic meaning: the carbon market strengthens firms' incentives to conduct green technology R&D and patent deployment through allowance constraints and carbon price signals. H1 is therefore supported.

Columns (2) and (3) report the mechanism test based on carbon emission intensity. In the first step, the DID coefficient for carbon emission intensity is significantly negative, with a coefficient of -0.131, indicating that the national carbon market reduces the carbon emission intensity of treated firms. In the second step, the coefficient of carbon emission intensity is -0.082 and significant at the 10% level, suggesting that lower carbon emission intensity is associated with higher green innovation. This result indicates that the carbon market does not simply increase firms' compliance burden; rather, it promotes green innovation by lowering emission intensity and forcing technological upgrading and low-carbon R&D. H2 is broadly supported.

Column (4) examines the moderating role of pre-policy attention to green transformation. The coefficient of DID x High attention is 0.470 and significant at the 1% level, indicating that firms with greater pre-policy attention to green transformation show stronger green innovation responses after the launch of the national carbon market. H3 is supported. Column (5) reports the interaction result for financing constraints. The coefficient of DID x Low financing constraints is -0.313 and significant at the 5% level. Because the reference group consists of firms with higher financing constraints, this finding means that the policy effect is weaker for firms with lower financing constraints, or conversely, that firms with higher financing constraints experience a stronger forced-innovation effect. H4 is supported.

5. Robustness Checks

5.1 PSM-DID Matching Balance Test

Table 4: Balance Test after PSM-DID Matching

Variable	Treatment mean	Control mean	SMD before matching	SMD after matching
Green innovation	1.569	1.568	0.514	0.001
Firm size	23.581	23.509	0.875	0.050
Growth	0.081	0.088	-0.022	-0.036
Cash-to-assets ratio	0.083	0.087	-0.693	-0.075
Ownership concentration	0.387	0.394	0.395	-0.042
ROA	0.024	0.021	-0.006	0.081
Tobin's Q	1.156	1.168	-0.784	-0.040
SA financing constraints	3.937	3.920	0.149	0.052
Green attention	0.971	0.971	0.587	0.000

Table 4 reports the balance test after PSM-DID matching. The matching variables include pre-policy green innovation, firm size, growth, cash-to-assets ratio, ownership concentration, ROA, Tobin's Q, the SA financing-constraint index, and green attention. Nearest-neighbor matching is conducted within the common support region following the propensity-score matching literature [17,18]. After matching, the absolute values of standardized mean differences for all covariates are below 0.1, indicating that the treated and control firms are relatively comparable in observable characteristics and that the matched sample provides a reasonable basis for subsequent PSM-DID estimation.

5.2 Dynamic DID and Additional Robustness Tests

Table 5: Dynamic DID and Robustness Tests

Variable	Dynamic DID	After 2022	Invention patents	Placebo	PSM-DID
Treat x 2018	-0.116 (0.121)				
Treat x 2019	-0.096 (0.100)				
Treat x 2021	0.176* (0.106)				
Treat x 2022	0.325*** (0.102)				
Treat x 2023	0.277** (0.117)				
DID (2022 and after)		0.301*** (0.083)			
DID (green invention patents)			0.075 (0.062)		
DID (2019 placebo; pre-policy sample)				0.050 (0.083)	
DID (PSM matched sample)					0.245** (0.107)
Observations	20,425	20,425	20,425	8,894	795
Firm clusters	4,639	4,639	4,639	3,365	136
Firm/year fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

*Note: Standard errors clustered at the firm level are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*

Table 5 first reports the dynamic DID results. Taking 2020 as the base year, the coefficients of Treat x 2018 and Treat x 2019 are not significant, suggesting that the treatment and control groups did not exhibit significant differential trends before the policy. The coefficient of Treat x 2021 is positive and significant at the 10% level, while the coefficients of Treat x 2022 and Treat x 2023 are significantly positive. These findings support the parallel-trends assumption and indicate that the green innovation response emerged gradually after the national carbon market began operating.

Several robustness checks are then conducted. First, when the post-policy period is redefined as 2022 and after, the DID coefficient remains significantly positive. This indicates that the baseline conclusion does not depend entirely on whether the policy took effect in 2021, the initial year of trading. Second, when the dependent variable is replaced with green invention patents, the coefficient remains positive but is not statistically significant. This finding does not reject the baseline result; instead, it suggests structural differences in the policy effect. After the launch of the national carbon market, firms may first respond through the quantity of green innovation, while green invention patents, which represent higher-quality innovation, require longer R&D cycles and greater technical difficulty and may not become statistically significant in the short run.

Third, a placebo test is performed using only the pre-policy sample from 2018 to 2020 and setting 2019 as a pseudo-policy year. The placebo DID coefficient is insignificant, indicating that the baseline result is not mechanically driven by pre-existing trends. Fourth, after PSM matching, the matched sample contains 68 treated firms and 68 control firms, and the DID coefficient remains significantly positive. Therefore, after accounting for possible sample-selection bias, the core conclusion remains robust.

6. Conclusion and Policy Implications

Using panel data on A-share listed companies from 2018 to 2023, this paper treats the 2021 launch of China's national carbon market as a quasi-natural experiment and applies a difference-in-differences model to examine its effect on corporate green innovation. The results show that the national carbon market significantly increases the total number of green patent applications by D44 treated firms. The reduction of carbon emission intensity is an important mechanism. Firms with higher pre-policy attention to green transformation respond more strongly, and firms with higher financing constraints show stronger policy effects than those with lower financing constraints. Dynamic DID estimates, a pre-policy placebo test, and PSM-DID analysis further support the robustness of the findings.

The findings have several policy implications. First, China should continue improving the price-discovery and compliance mechanisms of the national carbon market, thereby strengthening the stability of carbon price signals and helping firms form long-term expectations for green R&D. Second, regulators should enhance the quality management of corporate carbon emissions data and assessment of carbon emission intensity, guiding firms to shift from passive compliance to active technological upgrading. Third, firms should be encouraged to systematically disclose green transformation targets in annual reports, social responsibility reports, and strategic plans, so as to strengthen organizational attention and implementation capacity for low-carbon transition. Fourth, the green innovation needs of firms facing financing constraints should receive more policy support. Green credit, transition finance, and carbon-asset pledge financing may help reduce the financing cost of technological transformation [19,20].

This study also has limitations. First, the sample period covers only the first three years after the national carbon market was launched, and high-quality innovations such as green invention patents may have longer time lags. Second, the treatment group is identified mainly by industry code, and the analysis does not further distinguish the actual compliance status of listed companies and their subsidiaries. Future studies may use firm-level allowance, compliance, and carbon-trading data to identify more precisely the causal mechanisms linking carbon prices, allowance constraints, and green innovation.

References

[1] State Council of the People's Republic of China. (2021, October 27). Action plan for carbon dioxide peaking before 2030.

https://english.www.gov.cn/policies/latestreleases/202110/27/content_WS6178a47ec6d0df57f98e3dfb.html

[2] Ministry of Ecology and Environment. (2020, December 31). Measures for the administration of carbon emissions trading (Trial). [In Chinese]. https://www.mee.gov.cn/xxgk2018/xxgk/xxgk02/202101/t20210105_816131.html

[3] Ministry of Ecology and Environment. (2021, December 31). The first compliance cycle of the national carbon market was successfully completed. [In Chinese]. https://www.mee.gov.cn/ywgz/ydqhbh/wsqtkz/202112/t20211231_965906.shtml

[4] Ministry of Ecology and Environment. (2022). Report on the first compliance cycle of the national carbon emissions trading market. [In Chinese]. https://www.mee.gov.cn/ywgz/ydqhbh/wsqtkz/202212/P020221230799532329594.pdf

[5] Ministry of Ecology and Environment. (2022, December 19). Guidelines for accounting and reporting greenhouse gas emissions from power generation facilities. [In Chinese]. https://www.mee.gov.cn/xxgk2018/xxgk/xxgk06/202212/t20221221_1008430.html

[6] Porter, M. E., & van der Linde, C. (1995). Toward a new conception of the environment-competitiveness relationship. *Journal of Economic Perspectives*, *9*(4), 97–118. https://doi.org/10.1257/jep.9.4.97

[7] Jaffe, A. B., & Palmer, K. (1997). Environmental regulation and innovation: A panel data study. *Review of Economics and Statistics*, *79*(4), 610–619. https://doi.org/10.1162/003465397557196

[8] Popp, D. (2002). Induced innovation and energy prices. *American Economic Review*, *92*(1), 160–180. https://doi.org/10.1257/000282802760015658

[9] Acemoglu, D., Aghion, P., Bursztyn, L., & Hémous, D. (2012). The environment and directed technical change. *American Economic Review*, *102*(1), 131–166. https://doi.org/10.1257/aer.102.1.131

[10] Calel, R., & Dechezleprêtre, A. (2016). Environmental policy and directed technological change: Evidence from the European carbon market. *Review of Economics and Statistics*, *98*(1), 173–191. https://doi.org/10.1162/REST_a_00470

[11] Liu, J. Y., & Liu, X. (2023). Effects of carbon emission trading schemes on green technological innovation by industrial enterprises: Evidence from a quasi-natural experiment in China. *Journal of Innovation & Knowledge*, *8*(3), Article 100410. https://doi.org/10.1016/j.jik.2023.100410

[12] Wang, P., Huang, S., Wang, Y., & Li, W. (2023). Research on the impact of carbon emission rights trading on carbon performance of enterprises [In Chinese]. *Science Research Management, 44*(12), 158–169. https://doi.org/10.19571/j.cnki.1000-2995.2023.12.016

[13] Jia, L. J., Zhang, X., Wang, X. N., Chen, X. L., Xu, X. F., & Song, M. L. (2024). Impact of carbon emission trading system on green technology innovation of energy enterprises in China. *Journal of Environmental Management*, *360*, Article 121229. https://doi.org/10.1016/j.jenvman.2024.121229

[14] Cao, H., Yi, Y., & Jiang, J. (2024). The driving mechanism of green innovation in carbon markets: Property-right definition or market trading? [In Chinese]. *Journal of Finance and Economics, 50*(7), 80–94. https://doi.org/10.16538/j.cnki.jfe.20240511.301

[15] China Securities Regulatory Commission. (2012). Guidelines for the industry classification of listed companies. [In Chinese]. https://www.csrc.gov.cn/csrc_en/c102030/c1370756/content.shtml

[16] Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, *119*(1), 249–275. https://doi.org/10.1162/003355304772839588

[17] Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

[18] Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies*, *64*(4), 605–654. https://doi.org/10.2307/2971733

[19] People's Bank of China et al. (2016, August 31). Guidelines for establishing the green financial system. https://www.pbc.gov.cn/english/130721/3133045/index.html

[20] Wang, X., & Wang, Y. (2021). Research on green credit policy to enhance green innovation [In Chinese]. *Management World, 37*(6), 173–188, 11. https://doi.org/10.19744/j.cnki.11-1235/f.2021.0085

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Data Availability

The data used in this study are available from the CSMAR database and publicly disclosed annual reports of listed companies, subject to the database provider's terms of use.

Acknowledgment

This paper is an independent research output by the author.

ISSN 3006-3965