A Hybrid Econometric and Machine Learning Framework for Identifying Nonlinear Dynamics Between Digital Economy and Carbon Emissions

Jinghan Li
Lincoln University College, 47301 Petaling Jaya, Selangor, Malaysia

***Abstract*—The rapid expansion of the digital economy has reshaped global production systems and energy consumption patterns, yet its environmental consequences remain theoretically ambiguous. This study investigates the nonlinear relationship between digital economy development and carbon emissions using a cross-country panel dataset covering 45 countries during 2012–2024. A hybrid econometric and machine learning framework is developed to identify the structural dynamics underlying the digital economy–carbon emissions nexus. The empirical analysis combines fixed-effects estimation, System Generalized Method of Moments, instrumental variable estimation, and threshold regression with Gradient Boosting Decision Trees and SHAP-based interpretation. The results reveal a robust inverted-U relationship between digital economy development and carbon emissions, with the quadratic specification yielding a turning point of approximately 0.71 and the panel threshold model indicating a significant threshold value of 0.742. This indicates that digital transformation initially increases emissions through infrastructure expansion and energy demand, but subsequently contributes to emission mitigation through technological progress, energy efficiency improvement, and industrial restructuring. The nonlinear transition pattern is consistently validated across multiple identification strategies, with machine learning evidence further confirming the heterogeneous marginal effects of digital development across different development stages. Mechanism analysis shows that energy efficiency enhancement, technological innovation, and industrial upgrading are key channels through which digitalization promotes low-carbon transformation. These findings highlight that the environmental impacts of the digital economy are not inherently positive or negative but evolve dynamically with technological maturity and economic structure. Policy strategies should therefore emphasize the quality-oriented development of digital infrastructure and strengthen complementary innovation capabilities to maximize the emission-reduction potential of digital transformation.**

Keywords*—Digital economy; carbon emissions; nonlinear relationship; system GMM; machine learning; energy efficiency; industrial upgrading

I. Introduction

In the context of accelerating digital transformation and global carbon neutrality initiatives, understanding the environmental consequences of digital economy development has become increasingly important. Digital technologies, including big data, artificial intelligence, and cloud computing, enhance resource allocation efficiency and facilitate industrial upgrading; however, they may also generate additional energy demand through the expansion of digital infrastructure, such as data centers and communication networks. Therefore, the net environmental impact of digitalization remains theoretically and empirically uncertain.

Previous studies have documented both emission-reducing and emission-increasing effects of digitalization. On the one hand, digital technologies promote energy efficiency improvement, technological innovation, and structural transformation, thereby reducing carbon emissions [1,2]. On the other hand, the expansion of digital infrastructure and increased electricity consumption may generate additional emission pressures [3,4]. However, many existing studies still rely primarily on linear or quadratic specifications and treat the digital economy as a static explanatory factor, limiting the understanding of its evolving environmental effects.

More importantly, the existing literature provides limited evidence on whether digitalization represents a stage-dependent structural transition process, in which its environmental effects shift as digital development advances. This perspective suggests that the digital economy–carbon emission nexus may exhibit nonlinear and regime-dependent dynamics rather than a constant marginal effect.

To address this gap, this study develops a structural transition perspective by conceptualizing digitalization as a regime-shifting force that alters the balance between emission-enhancing scale effects and emission-reducing efficiency effects. Using a cross-country panel dataset covering 45 countries during 2012–2024, this study combines econometric identification and machine learning-based nonlinear validation to investigate the dynamic evolution of the digital economy–carbon emission nexus.

The results reveal a robust inverted-U relationship between digital economy development and carbon emissions. Specifically, digitalization initially increases emissions due to infrastructure expansion and energy demand, but beyond a critical development stage, its effects shift toward emission reduction through improvements in energy efficiency, technological innovation, and industrial upgrading. Machine learning analysis provides independent evidence supporting the identified nonlinear transition pattern.

This study contributes to the literature by establishing a regime transition framework for understanding the digital economy–carbon nexus and demonstrating that the environmental effects of digitalization evolve dynamically across development stages. The findings provide new insights for designing differentiated digital transformation and low-carbon development strategies.

The remainder of this study is organized as follows. Section II reviews the related literature, while Section III discusses the research gap and the positioning of the present study. Section IV describes the measurement of the main variables, and Section V presents the computational methodology. Section VI reports the empirical results, followed by framework validation, nonlinear regime transition analysis, and robustness analysis. Section VII concludes the study and discusses the implications and limitations.

II. Literature Review

The relationship between the digital economy and carbon emissions has attracted increasing attention in recent years, and the existing literature can generally be classified into three main strands. Despite substantial progress, current studies have yet to establish a unified structural framework explaining how digital transformation reshapes environmental outcomes across different stages of development [1,4,5,6].

First, a large body of literature investigates the environmental effects of digital economy development. Many studies suggest that digital technologies contribute to carbon emission reduction by improving energy efficiency, optimizing production processes, facilitating technological innovation, and enhancing resource allocation efficiency [1,5-9]. In contrast, other scholars argue that rapid expansion of digital infrastructure, including data centers and communication networks, may increase electricity consumption and generate additional carbon emissions, producing a "digital expansion effect" [4,12,13]. Although these studies acknowledge the dual environmental effects of digitalization, most empirical analyses rely on linear or quadratic econometric models and treat digital economy development as a static explanatory variable [5,7,10-11].

Second, another stream of research explores the determinants of carbon emissions from a broader macroeconomic perspective, emphasizing factors such as economic growth, industrial upgrading, technological progress, and global value chains [5,6,8,12,14]. These studies provide important evidence regarding the drivers of carbon emissions and the broader economic effects of digital transformation [15]; however, their analytical frameworks generally assume smooth and continuous relationships among variables, making it difficult to capture structural changes and nonlinear transitions during digital transformation [5,6,8].

Third, recent studies have increasingly introduced machine learning and nonlinear analytical methods into environmental and digital economy research [16,17,18,19]. Compared with conventional econometric approaches, machine learning techniques exhibit superior predictive performance and greater flexibility in modeling complex nonlinear relationships [20,21,22]. Nevertheless, these methods are primarily employed to improve prediction accuracy or estimation robustness, rather than to explain the structural mechanisms underlying the relationship between digital economy development and carbon emissions [19,23].

Overall, three major research gaps remain. First, existing studies generally regard the digital economy as a static explanatory variable instead of a dynamic process of structural transformation. Second, insufficient attention has been paid to regime-dependent changes in the digital economy–carbon emissions relationship across different stages of digital development. Third, econometric identification and machine learning validation have rarely been integrated into a unified analytical framework for explaining the nonlinear transition mechanism.

Motivated by these limitations, this study adopts a structural transition perspective and investigates the relationship between the digital economy and carbon emissions as a nonlinear and regime-dependent process that evolves across different stages of digital development.

III. Research Gap

Despite extensive literature examining the relationship between the digital economy and carbon emissions, existing studies remain fragmented in three important aspects (Table I):

First, most existing research relies on linear or quadratic specifications within the Environmental Kuznets Curve (EKC) framework, implicitly assuming a stable functional relationship between digital development and environmental outcomes. However, such an assumption neglects the possibility that digital transformation operates as a structural and stage-dependent process, in which its environmental effects evolve dynamically rather than following a fixed parametric form.

Second, although recent studies acknowledge the dual effects of digitalization on emissions—through both efficiency improvement and energy demand expansion—they often fail to systematically model the transmission structure underlying this duality. In particular, limited attention has been paid to how digital trade integration, industrial restructuring, and technological diffusion jointly shape the carbon impact of the digital economy as an interconnected system.

Third, empirical strategies in existing literature are predominantly econometric and rely on predefined functional forms, which restrict their ability to capture high-dimensional nonlinearities and interaction effects. As a result, the structural heterogeneity and latent nonlinear patterns embedded in the digital economy–environment nexus remain insufficiently explored.

In addition, while machine learning methods have recently been introduced into environmental economics, they are still rarely integrated with causal econometric frameworks in a complementary manner. This creates a gap between predictive accuracy and causal interpretability in the analysis of digital economy–carbon relationships.

Therefore, a comprehensive framework that integrates nonlinear econometric identification with machine learning-based structural validation is still lacking in the current literature.

This study makes three main contributions to the existing literature.

First, it extends the traditional EKC-based framework by conceptualizing the digital economy as a structural transition system rather than a single explanatory variable. By incorporating digital trade and mechanism-based channels, this study provides a more comprehensive representation of how digital transformation reshapes carbon emission dynamics across different stages of development.

TABLE I. STYLIZED CHARACTERISTICS OF DIGITAL ECONOMY DEVELOPMENT AND CARBON EMISSION PATTERNS ACROSS COUNTRY GROUPS

Country Group	Typical Digital Economy Development	Typical Carbon Emission Pattern	Expected Relationship
Developed economies	High	Moderate to low	Digitalization primarily contributes to carbon reduction through efficiency gains and industrial upgrading
Upper-middle-income economies	Medium	Relatively high	Digital expansion coexists with industrialization, leading to both efficiency gains and increased energy demand
Lower-middle-income economies	Low	High	Infrastructure construction dominates, resulting in higher energy consumption and carbon emissions
Low-income economies	Very low	Low but unstable	Limited digitalization leads to a relatively weak digital–carbon linkage

Second, this study develops a multi-layer identification strategy that combines System GMM and instrumental variable estimation with machine learning-based nonlinear validation. Unlike prior studies that treat machine learning as purely predictive, this study positions it as a complementary tool for uncovering latent structural patterns and validating nonlinear functional relationships, thereby bridging the gap between causal inference and data-driven pattern recognition.

Third, the study identifies and confirms a robust nonlinear inverted-U relationship between the digital economy and carbon emissions across countries, while further demonstrating significant heterogeneity in the turning point across development levels. This finding enriches the understanding of digital-environment interactions by highlighting the stage-dependent nature of digital transformation and its environmental consequences.

Overall, this study contributes to the literature by offering a unified analytical framework that integrates structural economics, nonlinear econometrics, and machine learning methods to better understand the complex relationship between digital development and environmental sustainability.

IV. MEASUREMENT ANALYSIS

A. Measurement of the Digital Economy

The digital economy is a multidimensional concept that represents the integration of digital technologies into traditional economic activities, including production, exchange, and service processes. Due to its heterogeneous characteristics and the lack of a universally accepted cross-country measurement framework, digital economic development cannot be fully represented by a single indicator.

Drawing on the conceptual framework of the Bureau of Economic Analysis (BEA) digital economy measurement approach and related structural decomposition methods, this study constructs a composite Digital Economy Index by integrating sectoral value-added structures with digital intensity coefficients. This approach captures both the structural characteristics of economic activities and the degree of digital technology penetration across sectors, providing a more comprehensive measure of digital transformation.

Specifically, the value-added share of each sub-sector within industry j is calculated as:

$$S_{ijt} = \frac{VA_{ijt}}{\sum_{j=1}^{n} VA_{ijt}}$$

where, VA_{ijt} denotes the value-added of sub-sector j in country i at time t, and S_{ijt} represents its relative contribution within the industry structure.

To capture cross-industry differences in digital transformation, a digital intensity coefficient is introduced. This coefficient represents the relative degree of digital technology penetration across sectors. Following the structural accounting logic of the BEA digital economy framework, the digital intensity coefficients are derived from ICT infrastructure penetration, digital service availability, and technology adoption indicators. The indicators are normalized to a common scale before aggregation to improve cross-country comparability.

The final digital economy index is constructed as:

$$DE_{it} = \sum_{j=1}^{n} S_{ijt} \cdot \delta_{jt}$$

This formulation implies that the digital economy is not only determined by the industrial structure of an economy, but also by the intensity of digital technology embedded within each sector. Compared with single-dimensional proxies such as internet penetration or ICT trade, this structural composite index provides a more comprehensive and less biased measurement of digital economic development.

In addition, this approach is consistent with recent empirical literature emphasizing structural accounting methods for measuring digital transformation and is particularly suitable for nonlinear econometric frameworks such as the Environmental Kuznets Curve (EKC).

We further construct alternative indices using PCA and entropy weighting to ensure robustness, and all results remain qualitatively unchanged.

B. Measurement of Carbon Emissions

Carbon emissions are used as the dependent variable to measure environmental performance across countries. In line with EDGAR (Emissions Database for Global Atmospheric Research) standards, this study adopts carbon dioxide (CO_2) emissions as the proxy for total carbon emissions.

Compared with total greenhouse gas emissions, CO_2 emissions provide a more consistent and comparable indicator across countries and over time. Therefore, it is widely used in empirical studies on economic growth and environmental sustainability.

To reduce heteroscedasticity and improve estimation efficiency, carbon emissions are transformed using the natural logarithm form:

$$lnCO_{2,\ it} = \ln(nCO_{2,\ it})$$

All data are harmonized across countries and years to ensure comparability in the panel dataset.

C. Mechanism Variables

To investigate the transmission mechanism through which digital economy development affects carbon emissions, this study focuses on the energy efficiency channel. Digital transformation can influence environmental performance by improving resource utilization efficiency, optimizing production processes, and reducing energy intensity.

Energy efficiency (EE) is selected as the mediating variable to capture the resource utilization improvement induced by digital transformation. It is measured as the ratio of economic output to energy consumption, representing the amount of economic output generated per unit of energy input. A higher value indicates higher energy efficiency, reflecting the capability of an economy to achieve greater economic output with fewer energy resources.

The mediation mechanism is based on the assumption that digital technologies, including intelligent manufacturing, digital management systems, and data-driven optimization, enhance energy utilization efficiency and consequently contribute to carbon emission reduction.

V. Computational Methodology

A. Variable Definition and Data Description

This study investigates the nonlinear relationship between digital economy development and carbon emissions using a global panel dataset. The dependent variable is carbon emissions, while the key explanatory variable is digital economy development. Several control variables are incorporated to mitigate omitted variable bias. In addition, energy efficiency is considered as a potential transmission mechanism, while environmental regulation is introduced as a moderating factor.

The definitions and measurements of all variables are presented in Table II.

The empirical sample consists of 45 countries covering different development stages and geographical regions over the period 2012–2024. The sample was selected based on data availability and cross-country comparability over the study period. The resulting panel dataset provides sufficient cross-country and temporal variation to examine heterogeneous digital transformation pathways and carbon emission dynamics across countries.

Missing observations are addressed using interpolation methods to maintain panel continuity for empirical estimation.

B. Baseline Econometric Model

To examine the relationship between digital economy development and carbon emissions, this study first establishes a baseline panel regression framework. Considering country-specific characteristics and common time shocks, country fixed effects and year fixed effects are included.

The baseline specification is formulated as follows:

$$lnCO_{2,it} = \alpha + \beta_1 DE_{it} + \beta_2 DE_{it}^2 + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it}$$

where, $lnCO_{2,it}$ denotes the natural logarithm of carbon emissions for country i in year t;

DE_{it} represents the Digital Economy Index,

DE_{it}^2 captures the nonlinear effect of digital economy development;

X_{it} is a vector of control variables, including economic development, population density, and trade openness;

μ_i denotes country fixed effects;

λ_t denotes time fixed effects;

ε_{it} is the random disturbance term.

The expected inverted-U relationship implies that:

β_1 >0, and β_2<0, indicating that digital economy development initially increases carbon emissions but eventually contributes to emission reduction after exceeding a certain development stage.

TABLE II. Variable Definition and Construction

Variable Type	Variable	Symbol	Measurement and Definition
Dependent variable	Carbon emissions	$lnCO_2$	Natural logarithm of CO_2 emissions (metric tons), obtained from the EDGAR database
Core explanatory variable	Digital economy development	DE	Digital Economy Index constructed using a BEA-style value-added decomposition framework combined with digital intensity adjustment coefficients
Control variable	Economic development level	ED	Logarithm of GDP per capita, representing economic development conditions
Control variable	Population density	PD	Population divided by land area (people/km^2), reflecting demographic pressure
Control variable	Trade openness	TO	Ratio of total imports and exports to GDP, representing economic globalization
Mediating variable	Energy efficiency	EE	GDP divided by total energy consumption, representing economic output generated per unit of energy input
Moderating variable	Environmental regulation	ER	Composite index measuring environmental policy intensity and regulatory strength

C. System GMM Estimation

Although fixed-effects estimation controls for unobserved heterogeneity, potential endogeneity may still exist because carbon emissions can influence digital transformation through policy adjustment and technological investment. Furthermore, carbon emissions usually exhibit strong persistence over time.

Therefore, following Arellano and Bover (1995) and Blundell and Bond (1998), the System Generalized Method of Moments (System GMM) approach is employed.

The dynamic panel specification is expressed as:

$$lnCO_{2,it} = \alpha + \rho lnCO_{2,i,t-1} + \beta_1 DE_{it} + \beta_2 DE_{it}^2 + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it}$$

where, $lnCO_{2,i,t-1}$ represents the lagged carbon emission level capturing emission persistence.

System GMM uses lagged variables as internal instruments to address potential endogeneity. The validity of the estimation is assessed using the Arellano–Bond serial correlation test and the Hansen test of over-identifying restrictions.

D. Instrumental Variable Estimation

To further strengthen causal identification, an instrumental variable (IV) approach is employed.

This study uses:

- Lagged internet penetration rate.
- Submarine optical cable density.

These instruments are intended to capture variation in digital connectivity infrastructure. Their relevance is based on their association with digital economy development, while their exogeneity is assessed through the instrumental variable estimation framework.

The two-stage least squares (2SLS) estimation is specified as:

First stage:

$$DE_{it} = \beta_0 + \beta_1 IV_{it} + \beta_2 X_{it} + \mu_i + \lambda_t + u_{it}$$

Second stage:

$$lnCO2_{it} = \theta_0 + \theta_1 \widehat{DE}_{it} + \theta_2 \widehat{DE}_{it}^2 + \gamma X_{it} + \mu_i + \lambda_t + v_{it}$$

The first-stage F-statistic is examined to assess instrument strength, while consistency between IV estimates and baseline results provides additional evidence of robustness.

E. Mechanism Analysis Model

To further examine the transmission mechanism through which digital economy development affects carbon emissions, this study focuses on the energy efficiency improvement channel.

Specifically, energy efficiency (EE) is introduced as the mediating variable, reflecting the extent to which digital transformation improves resource utilization and reduces energy intensity.

Following the mediation framework, the first-stage model is specified as:

$$M_{it} = \delta + \theta DE_{it} + \eta' X_{it} + \mu_i + \lambda_t + \varepsilon_{it}$$

where, M_{it} denotes energy efficiency for country i at year t.

The second-stage model incorporates both digital economy development and the mediating variable:

$$lnCO_{2,it} = \alpha + \beta_1 DE_{it} + \beta_2 DE_{it}^2 + \phi M_{it} + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it}$$

A significant coefficient of M_{it} indicates that energy efficiency serves as an important transmission channel through which digital economy development influences carbon emissions.

F. Moderating Effect Model

Environmental regulation may influence the effectiveness of digital transformation by affecting technological adoption, industrial adjustment, and environmental governance efficiency.

To examine this moderating role, the following interaction model is constructed:

$$lnCO_{2,it} = \alpha + \beta_1 DE_{it} + \beta_2 DE_{it}^2 + \beta_3 ER_{it} + \beta_4 (DE_{it} \times ER_{it}) + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it}$$

where, ER_{it} represents environmental regulation intensity, and $DE_{it} \times ER_{it}$ denotes the interaction between digital economy development and environmental regulation.

A statistically significant interaction coefficient indicates that environmental regulation changes the marginal environmental effect of digital economy development.

G. Threshold Regression Model

Although the quadratic model captures nonlinear characteristics, it assumes a smooth transition process. To further identify potential structural changes during digital transformation, this study applies Hansen's (1999) panel threshold regression model.

The single-threshold specification is:

$$lnCO_{2,it} = \mu_i + \lambda_t + \beta_1 DE_{it} I(DE_{it} \le \gamma) + \beta_2 DE_{it} I(DE_{it} > \gamma) + \theta' X_{it} + \varepsilon_{it}$$

where, γ represents the threshold parameter and I(·) is the indicator function.

To capture multiple development stages, a double-threshold model is further estimated:

$$lnCO_{2,it} = \mu_i + \lambda_t + \beta_1 DE_{it} I(DE_{it} \le \gamma_1) + \beta_2 DE_{it} I(\gamma_1 < DE_{it} \le \gamma_2) + \beta_3 DE_{it} I(DE_{it} > \gamma_2) + \theta' X_{it} + \varepsilon_{it}$$

where, γ1 and γ2 represent the first and second threshold values, respectively.

The threshold values are estimated by minimizing the residual sum of squares:

$$\hat{\gamma} = argmin_{\gamma} RSS(\gamma)$$

The significance of threshold effects is evaluated through bootstrap procedures following Hansen (1999).

H. Machine Learning Validation Framework

To complement the econometric analysis, this study employs a Gradient Boosting Decision Tree (GBDT) model to capture nonlinear relationships and potential interaction effects among variables. Unlike the econometric models, the GBDT approach is not intended for causal inference but is employed as a predictive validation framework. The input variables include the Digital Economy Index (DE), Economic Development (ED), Population Density (PD), Trade Openness (TO), Energy Efficiency (EE), and Environmental Regulation (ER). The predictive performance of the model is evaluated using commonly adopted prediction accuracy metrics. Furthermore, SHAP (SHapley Additive exPlanations) analysis is employed to quantify the contribution of each variable to carbon emissions, while partial dependence analysis is conducted to visualize the nonlinear effect of digital economy development. The consistency between the nonlinear patterns identified by the GBDT model and the econometric results provides additional evidence supporting the inverted-U relationship between digital economy development and carbon emissions.

VI. Results

A. Baseline Regression Results

Table III reports the estimation results of the baseline regression model. Columns (1)–(5) present the regression results with progressively added control variables. Throughout all specifications, the coefficient of Digital Economy (DE) remains significantly positive, whereas the coefficient of Digital Economy Squared (DE^2) remains significantly negative. The estimation results are highly consistent across different specifications, providing strong empirical evidence of an inverted-U relationship between digital economy development and carbon emissions.

TABLE III. Baseline Regression Results

Variables	(1)	(2)	(3)	(4)	(5)
Digital Economy (DE)	1.285*** (4.76)	1.177*** (4.91)	1.352*** (4.68)	1.219*** (4.54)	1.243*** (4.62)
Digital Economy² (DE^2)	−0.863*** (−3.72)	−0.795*** (−3.54)	−0.914*** (−3.88)	−0.846*** (−3.67)	−0.871*** (−3.75)
Economic Development (ED)		0.386*** (6.32)	0.403*** (6.45)	0.396*** (6.51)	0.385*** (6.37)
Population Density (PD)			−2.165** (−2.18)	−2.328** (−2.26)	−2.452** (−2.31)
Trade Openness (TO)				−0.019 (−1.12)	−0.012* (−1.78)
Constant	4.217***	2.183***	2.065***	2.146***	2.236***
Country FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	579	579	579	579	579
R^2	0.512	0.566	0.638	0.612	0.605

a. Notes: t-statistics are reported in parentheses. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.

As shown in Table III, the coefficient of Digital Economy (DE) is consistently positive and statistically significant at the 1% level, whereas the coefficient of Digital Economy² (DE^2) is significantly negative across all model specifications. These findings confirm the existence of a nonlinear inverted-U relationship between digital economy development and carbon emissions, thereby supporting Hypothesis H1.

The results indicate that, during the initial stage of digital transformation, large-scale investment in digital infrastructure, communication networks, and data centers increases energy demand and electricity consumption, resulting in higher carbon emissions. As digitalization continues to advance, however, improvements in production efficiency, industrial upgrading, intelligent resource allocation, and technological innovation gradually outweigh the initial expansion effect. Consequently, the marginal impact of digital economy development on carbon emissions becomes negative after a certain level of digitalization is achieved.

Among the control variables, Economic Development (ED) exhibits a significantly positive coefficient, suggesting that economic growth remains associated with increased energy consumption and carbon emissions. Population Density (PD) shows a significantly negative effect, indicating that agglomeration economies and improved resource allocation contribute to emission reduction. Trade Openness (TO) generally presents a negative coefficient, implying a potential carbon-reduction effect through international trade and technological diffusion, although its statistical significance varies across specifications.

Overall, the baseline regression results demonstrate that the environmental impact of digital economy development evolves dynamically rather than remaining constant. The empirical evidence strongly supports the nonlinear structural transition proposed in this study and provides the basis for the subsequent analyses of transmission mechanisms and moderating effects.

B. Mechanism Analysis

To further explore how digital economy development affects carbon emissions, this study examines whether Energy Efficiency (EE) serves as an intermediate transmission channel. Following the mediation analysis framework described in Section V- E, the estimation results are presented in Table IV.

The mechanism analysis provides evidence that Energy Efficiency (EE) serves as an important transmission channel linking digital economy development and carbon emissions.

As shown in Column (2), Digital Economy (DE) has a positive and statistically significant effect on Energy Efficiency (EE), suggesting that digital transformation improves energy productivity through intelligent production systems, digital management, and more efficient resource allocation.

Column (3) further shows that Energy Efficiency is significantly negatively associated with carbon emissions, indicating that improvements in energy productivity contribute to emission mitigation. After introducing Energy Efficiency into the carbon emission model, the coefficient of Digital Economy (DE) decreases in absolute value from −0.568 to −0.512 while remaining statistically significant. This decline suggests that part of the emission-reduction effect of digital economy development is transmitted through improvements in energy efficiency.

TABLE IV. MECHANISM ANALYSIS RESULTS

Variables	Carbon Emissions	Energy Efficiency	Carbon Emissions
Digital Economy (DE)	−0.568*** (−4.82)	1.236*** (3.45)	−0.512*** (−4.16)
Energy Efficiency (EE)			−0.021*** (−5.92)
Economic Development (ED)	0.422*** (6.41)	0.821** (2.51)	0.346*** (7.12)
Population Density (PD)	−1.862** (−2.36)	2.351** (2.14)	−1.648** (−2.11)
Trade Openness (TO)	−0.013 (−0.61)	0.584 (0.73)	−0.009 (−0.48)
Constant	3.185***	2.964***	2.158***
Country FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
R^2	0.591	0.405	0.706

[b] Notes: t-statistics are reported in parentheses. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$.

These findings highlight that energy efficiency improvement represents one of the key pathways through which digital economy development contributes to carbon reduction, particularly when digital transformation enters a more mature stage.

The empirical evidence therefore supports Hypothesis H2, suggesting that digital transformation contributes to carbon reduction not only through direct effects but also indirectly by improving energy efficiency. Digital technologies enable real-time monitoring of energy consumption, optimize production processes, and enhance resource allocation efficiency, thereby reducing energy intensity and facilitating sustainable low-carbon development.

Overall, the mechanism analysis demonstrates that improvements in energy efficiency represent an important transmission pathway through which digital economy development promotes carbon emission reduction. This finding provides empirical support for the proposed theoretical framework and highlights the role of energy efficiency enhancement in facilitating a digital-driven low-carbon transition.

C. Moderating Effect Analysis

To further investigate whether environmental regulation strengthens the carbon-reduction effect of the digital economy, the moderating effect model specified in Section V- F is estimated. Following the Porter Hypothesis, stronger environmental regulation is expected to enhance the environmental benefits generated by digital transformation through encouraging green innovation, improving resource allocation, and accelerating industrial upgrading.

The estimation results are presented in Table V.

Table V reports the estimation results of the moderating effect model. The coefficient of Environmental Regulation (ER) is negative and statistically significant, indicating that stronger environmental regulation is directly associated with lower carbon emissions.

TABLE V. MODERATING EFFECT OF ENVIRONMENTAL REGULATION

Variables	Carbon Emissions ($lnCO_2$)
Digital Economy (DE)	−0.842***(-3.28)
Environmental Regulation (ER)	−0.386*(-1.72)
DE × ER	−0.418** (-2.41)
Economic Development (ED)	0.392*** (6.24)
Population Density (PD)	−2.438**(-2.32)
Trade Openness (TO)	−0.012 (-0.56)
Constant	2.206***(3.15)
Country Fixed Effects	Yes
Year Fixed Effects	Yes
R^2	0.718

[c] Notes: t-statistics are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

More importantly, the interaction term between Digital Economy (DE) and Environmental Regulation (ER) is significantly negative. This finding suggests that environmental regulation strengthens the carbon-reduction effect of digital economy development. As environmental regulation becomes more stringent, digital technologies are more effectively translated into improvements in energy efficiency, cleaner production processes, and green technological innovation, thereby generating greater environmental benefits.

These findings are consistent with the Porter Hypothesis, which argues that appropriately designed environmental regulation can stimulate innovation and improve environmental performance rather than merely increasing compliance costs. Under stricter regulatory conditions, firms are more likely to adopt digital technologies to optimize production, reduce energy consumption, and improve resource allocation efficiency, ultimately accelerating the low-carbon transition.

The moderating effect also implies that digital transformation alone may not be sufficient to maximize carbon-reduction performance. Instead, supportive institutional arrangements and effective environmental governance play an important complementary role in converting digital investment into sustainable environmental outcomes. Consequently, countries with stronger environmental regulatory frameworks are more likely to achieve larger emission reductions through digital economy development.

Overall, the empirical evidence confirms Hypothesis H3, indicating that environmental regulation positively moderates the relationship between digital economy development and carbon emissions by reinforcing the environmental benefits of digitalization.

D. Threshold Effect Analysis

To further investigate whether the environmental effect of digital economy development varies across different stages of digitalization, a panel threshold regression model is estimated following the methodology proposed in Section V-G. Compared with the quadratic specification, the threshold model allows the marginal impact of the digital economy on carbon emissions to change discretely once the Digital Economy Index exceeds an estimated threshold value.

The bootstrap procedure confirms the existence of a statistically significant threshold effect. The estimated threshold value of the Digital Economy Index is reported in Table VI. The likelihood ratio (LR) statistic rejects the null hypothesis of no threshold effect at the 5% significance level, indicating that the relationship between digital economy development and carbon emissions differs across development regimes.

TABLE VI. THRESHOLD REGRESSION RESULTS

Variables	DE ≤ γ	DE > γ
Digital Economy (DE)	0.842*** (4.26)	-0.956*** (-3.84)
Economic Development (ED)	0.403*** (6.31)	0.397*** (6.08)
Population Density (PD)	-2.684** (-2.34)	-4.926** (-2.37)
Trade Openness (TO)	-0.015 (-0.71)	-0.021* (-1.82)
Threshold Value (γ)	0.742	
Bootstrap F-statistic	18.640	
Bootstrap p-value	0.018	
Country Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
R^2	0.714	

[d] Notes: t-statistics are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The threshold regression results, reported in Table VI, provide further evidence that the environmental impact of digital economy development differs across stages of digital transformation. The estimated threshold value ($\gamma = 0.742$) is statistically significant, indicating the existence of a nonlinear regime transition in the relationship between the digital economy and carbon emissions.

When the Digital Economy Index is below the estimated threshold (DE ≤ γ), the coefficient of Digital Economy (DE) is positive and statistically significant, suggesting that the expansion of digital infrastructure and digital industries initially increases carbon emissions. At this stage, large-scale investment in communication networks, data centers, and digital equipment increases electricity demand and energy consumption, causing the scale effect to dominate the environmental impact.

In contrast, when the Digital Economy Index exceeds the estimated threshold (DE > γ), the coefficient of Digital Economy becomes significantly negative. This result indicates that the efficiency-enhancing and technology-driven benefits of digital transformation gradually outweigh the initial expansion effect. As digital technologies become more mature and are more widely integrated into industrial production, intelligent manufacturing, digital resource allocation, and optimized energy management improve energy efficiency, facilitate green technological innovation, and ultimately contribute to carbon emission reduction.

Compared with the quadratic regression presented in Section V-A, the threshold model provides a more flexible representation of nonlinear structural changes by explicitly identifying different development regimes. Rather than assuming a continuous nonlinear relationship, the threshold regression demonstrates that the marginal environmental effect of digital economy development changes after the digital economy reaches a critical development stage.

Overall, the threshold regression results provide additional empirical evidence supporting the existence of a structural transition in the relationship between digital economy development and carbon emissions. The findings indicate that digital economy development does not generate uniform environmental effects throughout the development process; instead, its contribution to carbon reduction becomes substantially stronger after surpassing the critical threshold.

E. Machine Learning Interpretation

To further explore the nonlinear characteristics of the digital economy–carbon emission relationship beyond parametric econometric specifications, this study employs a Gradient Boosting Decision Tree (GBDT) model. Unlike traditional regression models that impose predefined functional forms, the GBDT approach captures complex nonlinear interactions among explanatory variables and provides a data-driven representation of the underlying relationship.

The GBDT model uses carbon emissions as the dependent variable and incorporates the Digital Economy Index (DE), Economic Development (ED), Population Density (PD), Trade Openness (TO), Energy Efficiency (EE), and Environmental Regulation (ER) as explanatory variables. Model interpretability is further enhanced through the SHapley Additive exPlanations (SHAP) framework, which quantifies the marginal contribution of each variable to the predicted carbon emission outcomes.

The GBDT model demonstrates satisfactory predictive performance and effectively captures the complex nonlinear relationship underlying cross-country variations in carbon emissions. Compared with conventional parametric regression models, the machine learning approach identifies nonlinear interactions among explanatory variables without imposing a predefined functional form, thereby providing additional support for the nonlinear findings obtained from the econometric analysis Fig. 1, SHAP feature importance.

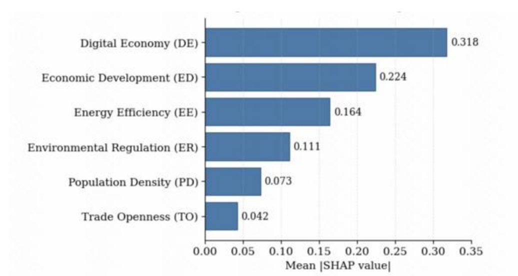

Fig. 1. SHAP Feature importance of carbon emission drivers.

The SHAP feature importance results indicate that the Digital Economy (DE) is one of the most influential determinants of carbon emissions. Among the explanatory variables, the digital economy exhibits the highest contribution to the model predictions. Energy Efficiency (EE) and Environmental Regulation (ER) also demonstrate substantial explanatory power, highlighting the important roles of technological efficiency improvement and institutional governance in shaping carbon emission dynamics.

The ranking of feature importance is broadly consistent with the econometric findings presented in Sections VI-A to VI-D, where digital economy development, energy efficiency improvement, and environmental regulation were identified as key factors influencing carbon outcomes. These results suggest that digital transformation affects carbon emissions not only through direct effects but also through efficiency enhancement and governance mechanisms.

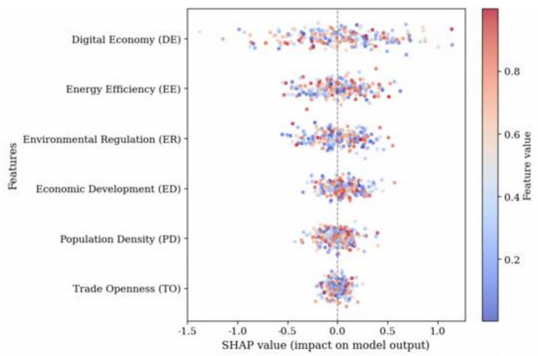

Fig. 2. SHAP summary plot.

The SHAP summary plot (Fig. 2) provides further evidence regarding the nonlinear influence of digital economy development on carbon emissions. At relatively low levels of digitalization, higher Digital Economy Index values are generally associated with positive SHAP contributions, indicating that early-stage digital expansion tends to increase carbon emissions.

This phenomenon reflects the dominance of the scale effect during the initial phase of digital transformation, where infrastructure deployment, data center expansion, and increased electricity demand generate additional energy consumption pressures.

However, as digitalization advances, the SHAP contributions gradually decline and become negative, suggesting that the marginal environmental effect of digital development shifts from emission promotion toward emission reduction. This transition indicates that efficiency-enhancing mechanisms, including intelligent resource allocation, industrial digitalization, and technological innovation, gradually outweigh the initial expansion effect.

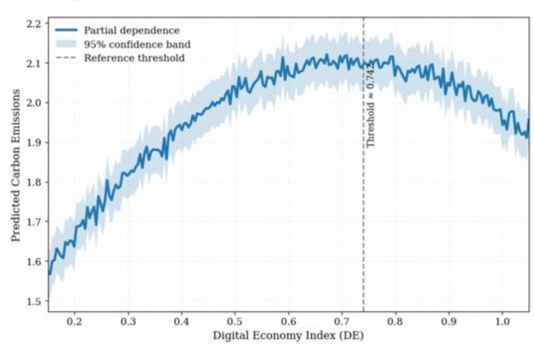

Fig. 3. Partial dependence plot of digital economy.

The partial dependence plot (Fig. 3) further illustrates the nonlinear marginal effect of digital economy development on predicted carbon emissions. The relationship exhibits a clear inverted-U pattern: carbon emissions initially increase with digital economy expansion but decline after reaching a critical development level.

The nonlinear transition identified by the GBDT model is broadly consistent with the turning point estimated from the quadratic regression model and the threshold value obtained from the panel threshold regression. This consistency suggests that the nonlinear relationship between digital economy development and carbon emissions is robust across different analytical approaches and reflects an underlying structural transition rather than a model-specific artifact.

Overall, the GBDT–SHAP analysis provides a non-parametric interpretation of the nonlinear relationship between digital economy development and carbon emissions. The results demonstrate that digitalization produces heterogeneous environmental effects across development stages: it initially increases emissions through infrastructure expansion but eventually contributes to carbon reduction through efficiency improvements, technological upgrading, and enhanced resource optimization.

F. Heterogeneity Analysis: Developed and Developing Economies

To examine whether the nonlinear relationship between digital economy development and carbon emissions varies across different development stages, this study conducts a heterogeneity analysis by dividing the sample into developed and developing economies.

The estimation results indicate that the inverted-U relationship remains valid in both groups, confirming the robustness of the nonlinear effect. However, substantial differences are observed in the turning points and marginal effects between the two groups. Developed economies exhibit a relatively lower turning point, suggesting that digital transformation reaches the emission-reduction stage earlier. This pattern may be attributed to their more mature digital infrastructure, higher energy efficiency, and stronger environmental governance capacity.

In contrast, developing economies present a higher turning point, implying that digital expansion requires a longer transition period before generating significant carbon reduction benefits. During the early stage of digitalization, infrastructure deployment, electricity demand growth, and digital capital accumulation may offset the potential environmental benefits.

These findings demonstrate that the environmental impact of digital economy development is not homogeneous across countries but depends on technological maturity, energy structure, and institutional conditions. Therefore, the effectiveness of digitalization as a carbon-reduction pathway is strongly associated with the development stage of individual economies. Fig. 4 illustrates the hybrid framework for identifying nonlinear digital economy–carbon emission dynamics, highlighting the transmission mechanisms, hybrid identification strategy, nonlinear transition, and regime evolution underlying the digital economy–carbon emission nexus.

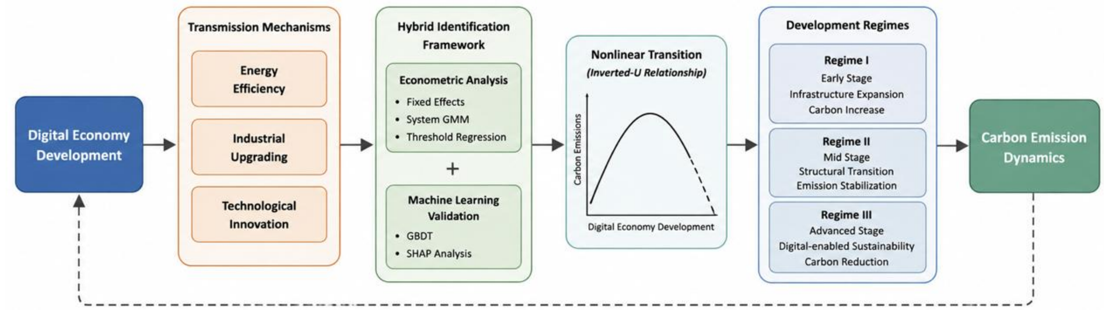

Fig. 4. Hybrid framework for identifying nonlinear digital economy–carbon emission dynamics.

G. Cross-Method Validation of Nonlinear Effects

To verify the robustness of the nonlinear relationship between the digital economy and carbon emissions, this study compares three complementary approaches: quadratic regression based on the EKC specification, panel threshold regression, and GBDT-based nonlinear estimation with SHAP interpretation.

The three approaches provide complementary evidence for the nonlinear relationship. The quadratic model identifies a turning point of approximately 0.71, while the panel threshold model identifies a regime transition at $\gamma = 0.742$. The GBDT–SHAP analysis provides an independent non-parametric assessment of the changing marginal contribution of digital economy development. The close correspondence across these approaches strengthens the interpretation of the nonlinear transition without relying on a single model specification.

VII. Framework Validation

A. Validation Strategy

To improve the reliability of nonlinear relationship identification, the proposed computational framework integrates econometric estimation with machine learning interpretation. The validation process combines four complementary modules: quadratic fixed-effects estimation, System GMM, instrumental variable estimation, and GBDT–SHAP analysis. Each module examines the digital economy–carbon emission relationship from a different analytical perspective, providing complementary evidence for framework validation.

B. Cross-Method Validation

The proposed framework is supported by complementary evidence from econometric estimation, threshold regression, and machine learning interpretation. The econometric models provide statistical identification, the threshold model identifies the regime transition, and the GBDT–SHAP analysis provides a flexible non-parametric assessment of nonlinear patterns. The consistency among these components strengthens the internal validity of the proposed analytical framework.

The threshold regression further identifies a significant threshold value ($\gamma = 0.742$), which is broadly consistent with the nonlinear transition pattern revealed by the GBDT partial dependence analysis. Moreover, the SHAP interpretation indicates that the contribution of digital economy development varies across different levels of digital development, providing additional evidence of its nonlinear effect.

The consistency across econometric estimation, threshold regression, and machine learning interpretation suggests that the identified nonlinear relationship is not driven by a specific modeling approach but reflects a robust structural relationship between digital economy development and carbon emissions.

C. Framework Validation Summary

The proposed framework demonstrates consistency from three complementary perspectives. Statistical consistency is confirmed by the agreement among econometric estimators. Structural consistency is supported by the close correspondence between the threshold regression and machine learning results. Interpretability consistency is established through SHAP analysis, which explains the stage-dependent contribution of digital economy development to carbon emissions.

Collectively, these results validate the proposed computational framework and strengthen the reliability of the identified nonlinear transition mechanism (Table VII).

TABLE VII. Validation Results of the Computational Framework

Validation Component	Method	Main Evidence
Econometric Estimation	Fixed Effects	Significant inverted-U relationship
Endogeneity Test	System GMM and IV	Consistent estimation results
Structural Identification	Threshold Regression	Estimated threshold (γ = 0.742)
Machine Learning	GBDT + SHAP	Consistent nonlinear pattern
Overall Validation	Integrated Framework	Strong cross-method consistency

VIII. Nonlinear Regime Transition Analysis

A. Regime Interpretation of the Nonlinear Relationship

Building on the empirical evidence presented above, this section interprets the nonlinear relationship from a regime transition perspective. Instead of assuming a constant marginal effect of digitalization, this study interprets the observed nonlinear pattern from a regime transition perspective (Table VIII).

A regime represents a relatively stable development state characterized by the combined effects of digital infrastructure, energy systems, industrial structure, and technological capability. Under different regimes, the environmental impacts of digitalization vary.

TABLE VIII. CROSS-METHOD VALIDATION OF NONLINEAR EFFECTS

Method	Framework	Nonlinear Evidence	Turning Point / Threshold	Result
Quadratic EKC model	Econometric regression	Significant inverted-U relationship	Turning point (τ ≈ 0.71)	Marginal effect reversal
Panel threshold model	Threshold regression	Significant regime change	Threshold value (γ = 0.742)	Nonlinear stage transition
GBDT + SHAP	Machine learning	Data-driven inverted-U pattern	Broadly consistent with econometric estimates	Consistent nonlinear evidence

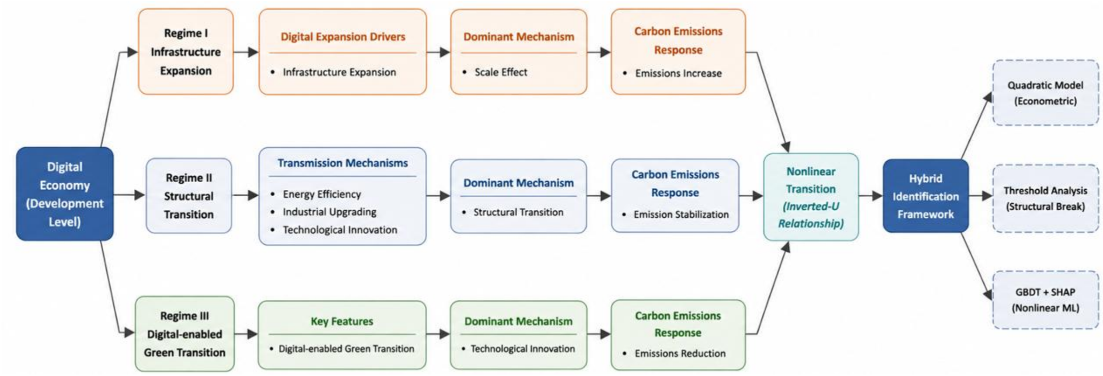

Fig. 5. Digital economy–carbon emission transition framework across development regimes.

The dominant mechanisms affecting carbon emissions evolve with the level of digital economy development. During the initial stage, digital expansion is primarily characterized by infrastructure expansion and scale effects, resulting in higher carbon emissions. As digital development advances, structural transformation, energy efficiency improvement, and technological innovation gradually become the dominant drivers, weakening the carbon-intensive effects of early-stage digitalization and promoting emission reduction. Consequently, the relationship between the digital economy and carbon emissions evolves through distinct nonlinear transition regimes rather than following a linear pathway.

The proposed framework (Fig. 5) illustrates the hybrid identification process of nonlinear digital economy–carbon emission dynamics. Different development regimes are characterized by changing dominant mechanisms, including scale effects, structural transition, and technological innovation, while the nonlinear transition is jointly identified through econometric estimation and validated using machine learning methods.

B. *Three Development Regimes of Digitalization and Carbon Emissions*

Based on the nonlinear response pattern identified by the econometric and machine learning models, the digital economy–carbon emission relationship can be interpreted through three representative development regimes.

Regime I: Digital Expansion Regime

In the initial stage of digital transformation, the expansion of digital infrastructure dominates the environmental outcome.

The deployment of communication networks, computing facilities, and digital equipment requires substantial energy inputs, leading to additional carbon emissions. In this regime, the scale effect exceeds the efficiency effect, and digitalization may increase environmental pressure.

Regime II: Transition Regime

As digital development progresses, emission-increasing and emission-reducing mechanisms coexist.

Digital infrastructure continues to generate energy demand, while digital technologies simultaneously improve production efficiency, optimize resource allocation, and promote industrial restructuring. The interaction between these opposing effects results in a gradual reduction of the marginal emission impact of digitalization.

This regime represents the transition process from energy-intensive digital expansion toward efficiency-oriented digital development.

Regime III: Digital Decoupling Regime

At a higher level of digital maturity, the efficiency and structural effects become dominant.

Advanced digital technologies enable intelligent energy management, process optimization, and low-carbon innovation, thereby reducing carbon intensity and improving resource utilization efficiency.

In this regime, digitalization contributes to the decoupling between economic development and carbon emissions.

C. *Empirical Evidence for Regime Transition*

The transition characteristics are supported by both econometric estimation and machine learning interpretation.

The quadratic specification provides evidence of a nonlinear turning point in the digital economy–carbon emission relationship. The estimated critical value indicates the development level at which the marginal effect of digitalization changes from emission-increasing to emission-reducing.

Furthermore, the Gradient Boosting Decision Tree (GBDT) model and SHAP-based interpretation reveal nonlinear variations in the contribution of digitalization across different development levels. The partial dependence patterns demonstrate that the environmental effect of digitalization is not uniform but changes along the digital development trajectory.

These results jointly support the existence of different digital economy–carbon emission regimes and indicate that countries may occupy different positions along the transition pathway.

D. Evolution of Driving Mechanisms Across Regimes

The transition among different regimes is associated with the changing dominance of three major mechanisms: scale effect, structural effect, and technological effect.

In the Digital Expansion Regime, the scale effect dominates because digital infrastructure deployment increases energy demand.

During the Transition Regime, structural adjustment and efficiency improvement gradually compensate for the additional energy consumption associated with digital expansion.

In the Digital Decoupling Regime, technological advancement becomes the primary driver by improving energy efficiency, facilitating low-carbon innovation, and supporting sustainable production systems.

The evolution of these mechanisms provides an integrated explanation for the observed inverted-U relationship between digital economy development and carbon emissions.

E. Engineering Implications for Digital Low-Carbon Transition

The regime transition analysis provides practical implications for designing digital-driven low-carbon development strategies.

For regions in the Digital Expansion Regime, improving the energy efficiency of digital infrastructure and increasing renewable energy utilization should be prioritized.

For regions in the Transition Regime, the focus should shift toward accelerating industrial digital transformation, enhancing technological integration, and improving resource utilization efficiency.

For regions in the Digital Decoupling Regime, advanced digital technologies can be further integrated with energy systems and industrial processes to achieve continuous emission reduction.

Therefore, the environmental benefits of digitalization depend not only on the scale of digital development but also on the degree of technological integration and energy system transformation.

IX. Robustness and Validation

A. Alternative Econometric Specifications

To examine the reliability of the estimated nonlinear relationship between the digital economy and carbon emissions, alternative econometric specifications are employed, including fixed-effects models, dynamic System Generalized Method of Moments (System GMM), and instrumental variable (IV) estimations.

The results obtained from different estimation approaches consistently support the baseline findings. Specifically, the coefficient of the digital economy remains positive, while the coefficient of its quadratic term remains negative and statistically significant, indicating the persistence of the inverted-U relationship.

The dynamic panel estimation further considers potential adjustment effects and temporal dependence of carbon emissions. The consistency of the key coefficients across different model specifications suggests that the identified nonlinear relationship is not sensitive to econometric assumptions.

Overall, the robustness analysis confirms the stability of the baseline estimation results under alternative identification strategies.

B. Alternative Measurements of the Digital Economy

To evaluate whether the findings depend on the specific measurement approach of the digital economy, alternative indicators are introduced for robustness verification.

The estimation results based on alternative digital economy measurements preserve the same nonlinear pattern as the baseline model. The digital economy coefficient remains positive, while the squared term maintains a negative relationship with carbon emissions.

These results indicate that the observed inverted-U relationship is not driven by a particular index construction method, but reflects a consistent association between digital development and carbon emission dynamics.

C. Machine Learning-Based Nonlinear Validation

To further examine the nonlinear characteristics without imposing a predefined functional form, a Gradient Boosting Decision Tree (GBDT) model is employed as a complementary validation approach.

The GBDT results provide additional support for the econometric findings. Feature importance analysis identifies the digital economy as a major explanatory factor in carbon emission prediction. Furthermore, SHAP analysis and partial dependence plots reveal a nonlinear response pattern consistent with the inverted-U relationship obtained from the quadratic specification.

The turning point identified by the machine learning model is close to the value estimated from the econometric approach, indicating consistency between parametric and non-parametric methods.

Therefore, the nonlinear relationship between digital development and carbon emissions is supported by multiple analytical approaches rather than relying solely on a specific functional form.

D. Stability Across Country Groups

To assess whether the nonlinear relationship remains stable under different sample compositions, additional analyses are conducted across alternative country groups.

The results show that the inverted-U relationship persists across different subsamples, although the estimated coefficients vary in magnitude. Such variations reflect differences in digital development levels, energy structures, and industrial characteristics among countries.

Importantly, the direction and nonlinear pattern of the relationship remain unchanged, suggesting that the identified digital economy–carbon emission nexus represents a relatively stable empirical relationship rather than a sample-specific outcome.

E. Summary of Robustness and Validation Evidence

The robustness checks and additional validation collectively demonstrate the reliability of the main findings.

The inverted-U relationship between the digital economy and carbon emissions remains consistent under alternative econometric specifications, alternative measurements, subsample analyses, and machine learning-based nonlinear validation.

The agreement between econometric estimation and GBDT-based interpretation confirms the stability of the identified nonlinear pattern and provides complementary evidence for the proposed regime transition analysis.

X. CONCLUSION

This study investigates the nonlinear relationship between the digital economy and carbon emissions using a global cross-country panel dataset by integrating econometric models with machine learning-based validation. A regime transition perspective is proposed to interpret the structural evolution of the digital economy–carbon emission nexus beyond the traditional linear assumption.

The results consistently reveal an inverted-U relationship between digital economy development and carbon emissions. This nonlinear pattern remains stable under alternative econometric specifications, including System GMM and instrumental variable estimations. The GBDT model with SHAP interpretation further confirms the nonlinear characteristics, demonstrating consistency between parametric and non-parametric approaches.

The findings indicate that the environmental impact of digitalization varies across development stages. Digital expansion may initially increase emissions due to infrastructure-related energy demand. With increasing digital maturity, efficiency improvement, industrial restructuring, and technological integration gradually become dominant, enabling digitalization to contribute to carbon reduction.

The estimated turning point represents a transition boundary between emission-increasing and emission-reducing stages. Cross-country differences in digital development, energy structures, and technological capabilities lead to heterogeneous positions along this transition pathway.

This study contributes by providing a nonlinear interpretation of the digital economy–carbon emission relationship and by combining econometric identification with machine learning-based validation. The findings highlight that digital transformation should be accompanied by improvements in energy efficiency, technological integration, and low-carbon development strategies.

Future research can further explore sector-level mechanisms and micro-scale digital applications to better understand the heterogeneous environmental effects of digital transformation.

This study has several limitations. First, interpolation of some missing observations may introduce measurement uncertainty. Second, differences in cross-country data availability may affect the precision of the Digital Economy Index. Third, the machine learning analysis provides complementary nonlinear validation rather than independent causal identification. Future research may further examine these issues using more granular data and independent samples.

REFERENCES

[1] Li, R., Hu, J., & Wang, X. (2024). Carbon emission reduction effect and mechanism of digital economic development in a global perspective. China Population, Resources and Environment, 34(8), 3–12. https://doi.org/10.12062/cpre.20240501

[2] Li, Y., Jiang, Y., Shen, C., et al. (2025). Carbon reduction effect of digital finance in China: Based on household micro data and input–output model. China Finance Review International, 15(2), 363–388. https://doi.org/10.1108/CFRI-03-2024-0083

[3] Tan, F. F., & Yan, X. C. (2025). Synergistic effects of pollution and carbon emission reductions in a digital economy: A case study of the Yangtze River Delta urban agglomeration. Environmental Science, 46(7), 4175–4185. https://doi.org/10.13227/j.hjkx.202406127

[4] Rosário, A. T., & Dias, J. C. (2023). The new digital economy and sustainability: Challenges and opportunities. Sustainability, 15(14), 10902. https://doi.org/10.3390/su151410902

[5] Li, Z., & Wang, J. (2022). The dynamic impact of digital economy on carbon emission reduction: Evidence from city-level empirical data in China. Journal of Cleaner Production, 351, 131570. https://doi.org/10.1016/j.jclepro.2022.131570

[6] Zuo, S., Zhao, Y., Zheng, L., Zhao, Z., Fan, S., & Wang, J. (2024). Assessing the influence of the digital economy on carbon emissions: Evidence at the global level. Science of the Total Environment, 946, 174242. https://doi.org/10.1016/j.scitotenv.2024.174242

[7] Zhang, X., Cheng, K., Zou, Z., et al. (2025). The effect of digital economy development on carbon emission reduction: An empirical analysis based on 80 countries of the Belt and Road Initiative. Sustainability, 17(3), 883. https://doi.org/10.3390/su17030883

[8] Zhang, L., Mu, R., Zhan, Y., Yu, J., Liu, L., & Zhang, J. (2022). Digital economy, energy efficiency, and carbon emissions: Evidence from provincial panel data in China. Science of the Total Environment, 852, 158403. https://doi.org/10.1016/j.scitotenv.2022.158403

[9] Li, X., & Yusup, G. (2023). Research on digital economy promoting low-carbon sustainable development of cities in China. Price: Theory & Practice, (9), 105–109. https://doi.org/10.19851/j.cnki.CN11-1010/F.2023.09.306

[10] She, Q., & Wu, L. (2022). Carbon emission reduction effect of digital economic development. Economic Perspectives, 39(5), 14–24. https://doi.org/10.15931/j.cnki.1006-1096.2022.05.008

[11] Liu, T., & Wang, J. (2024). Research on carbon emission reduction effect of digital economy. Operations Research and Fuzzy Theory, 14(5), 582–596. https://doi.org/10.12677/orf.2024.145498

[12] Sun, T., Di, K., & Shi, Q. (2024). Digital economy and carbon emission: The coupling effects of the economy in Qinghai region of China. Heliyon, 10(4), e26451. https://doi.org/10.1016/j.heliyon.2024.e26451

[13] Raihan, A. (2024). A review of the potential opportunities and challenges of the digital economy for sustainability. Innovation and Green Development, 3(4), 100174. https://doi.org/10.1016/j.igd.2024.100174

[14] Yan, K., & Yang, L. K. (2025). Research on the impact of digital global value chains on the embodied carbon emission intensity of export trade.

Journal of International Trade, (2), 123–139. https://doi.org/10.13510/j.cnki.jit.2025.02.003

[15] Chen, Y. (2020). Improving market performance in the digital economy. China Economic Review, 62, 101482. https://doi.org/10.1016/j.chieco.2020.101482

[16] Diao, Y., Wang, Z., & Li, Q. (2023). Research on application of machine learning algorithm in digital economy. In Proceedings of the Third International Conference on Computer Vision and Data Mining (ICCVDM 2022) (Vol. 12511). SPIE. https://doi.org/10.1117/12.2660095

[17] Zhu, B. (2023). Digital economy market prediction and trend analysis algorithms with big data. In Proceedings of the 2023 International Conference on Intelligent Computing, Communication and Convergence (ICI3C). IEEE. https://doi.org/10.1109/ICI3C60830.2023.00043

[18] Sun, Y., & Zhang, Y. (2024). Research on economic growth forecasting models based on big data. In Proceedings of the 2024 3rd International Conference on Algorithms, Data Mining, and Information Technology. ACM. https://doi.org/10.1145/3701100.3701126

[19] Li, C., He, W., & Cao, E. (2025). Impact of green data center pilots on digital economy development: An empirical study based on dual machine learning methods. Computers & Industrial Engineering, 201, 110914. https://doi.org/10.1016/j.cie.2025.110914

[20] Anghel, A., et al. (2018). Benchmarking and optimization of gradient boosting decision tree algorithms. arXiv preprint arXiv:1809.04559. https://doi.org/10.48550/arXiv.1809.04559

[21] Ye, J., Chowdhury, M., Abdolmaleki, S., et al. (2009). Stochastic gradient boosted distributed decision trees. In Proceedings of the 18th ACM Conference on Information and Knowledge Management (pp. 2061–2064). ACM. https://doi.org/10.1145/1645953.1646301

[22] Fu, P., et al. (2025). Financial engineering and the digital economy: The implementations of machine learning algorithms. Alexandria Engineering Journal, 125, 311–319. https://doi.org/10.1016/j.aej.2025.03.122

[23] Jiang, Y., & Sun, J. (2026). Does digital economy drive new quality productivity? Evidence from dual machine learning models. Managerial and Decision Economics, 47(2), 387–402. https://doi.org/10.1002/mde.70042