DOI: 10.2478/eces-2026-0020 ECOL CHEM ENG S. 2026;33(3):375-398

Zirui GUO[1*] and Hanyu LU[2]

ARTIFICIAL INTELLIGENCE POLICY AND URBAN CARBON EMISSION PERFORMANCE: EVIDENCE FROM THE YANGTZE RIVER DELTA, CHINA

Abstract: Whether artificial intelligence (AI) policy translates into measurable urban decarbonisation remains uncertain: most evidence is firm- or robot-level, and transparent city-level policy evidence is scarce. This study treats the State Council's "New Generation Artificial Intelligence Development Plan" and subsequent local responses as a quasi-natural experiment, using panel data on 26 Yangtze River Delta cities over 2010-2023. Four core cities (Shanghai, Nanjing, Hangzhou, Suzhou) are defined as the AI-policy treatment group, and a two-way fixed-effects difference-in-differences (DID) model estimates the effect of AI policy exposure on urban low-carbon performance, measured as a standardised inverse carbon-intensity indicator. The baseline DID coefficient is 0.493 and significant at the 5 % level, and the positive sign is stable when the onset year is set to 2018, the treatment group is expanded, the window is shortened, and Shanghai is excluded. The estimate is, however, sensitive to city-specific linear trends and to over-control specifications, so its magnitude should be read cautiously; event-study results are consistent with parallel pre-trends, with effects emerging gradually after implementation. Stepwise mechanism tests give indicative - not formal - evidence that industrial structure upgrading and green technological innovation act as transmission channels. Heterogeneity across cities is suggestive but not statistically confirmed, and cross-city spillovers are not robust across spatial weight matrices. AI policy thus appears associated with better urban carbon performance mainly through local channels, and implementation should prioritise practical low-carbon applications, green-innovation commercialisation, and locally differentiated strategies.

Keywords: artificial intelligence policy, carbon emission performance, difference-in-differences, quasi-natural experiment, industrial structure upgrading

Introduction

Since reform and opening-up, China's rapid economic growth has intensified the tension between economic development and environmental sustainability, especially with respect to carbon emissions. The Yangtze River Delta (YRD) region is one of China's most economically concentrated, industrially complete and technologically dynamic urban agglomerations. At the same time, CEADs city-level carbon-emission inventories [1], complemented by recent city-level carbon-emission evidence for 2023 [2], and statistical yearbook data [3] indicate that the region faces substantial energy-consumption and

[1] School of Economics and Management, North University of China, No. 1 Xueyuan Road, Jiancaoping District, Taiyuan 030000, China, phone +86 17582808688, email: 2026942647@qq.com, ORCID: 0009-0000-1361-873X
[2] School of Mechanical and Energy Engineering, Beijing University of Technology, No. 100 Pingleyuan, Nanmofang Subdistrict, Chaoyang District, Beijing 100124, China, phone +86 15611882312, email: keanu_0507@qq.com, ORCID: 0009-0000-4582-6541
[*] Corresponding author: 2026942647@qq.com

carbon-emission pressures. Therefore, emission reduction in the YRD is not merely a matter of scaling back production, but of improving carbon performance per unit of economic activity through technological progress, industrial upgrading and enhanced governance capacity.

As a key driver of the new wave of technological revolution and industrial transformation, artificial intelligence is being widely applied in industrial energy conservation, intelligent scheduling, equipment maintenance, energy management and environmental monitoring. According to the International Federation of Robotics (IFR), approximately 542,000 industrial robots were newly installed worldwide in 2024, while the global operational stock reached approximately 4.66 million units. China remained the world's largest industrial-robot market, with approximately 295,000 new installations in 2024, accounting for about 54 % of global deployments; its operational stock exceeded 2 million units [4]. These updated figures indicate that the diffusion of intelligent manufacturing and automation technologies has continued to deepen in recent years. They are used here only as background evidence of the broader development of industrial intelligence rather than as direct city-level evidence for the YRD. Existing studies have shown that intelligent energy management systems can reduce energy overconsumption and related carbon emissions by enabling real-time energy monitoring, data-driven control, predictive optimisation and demand-side management. Such systems provide technical support for AI-driven low-carbon transformation in manufacturing-intensive regions such as the Yangtze River Delta [5].

Existing studies link digital technologies to carbon emissions through production efficiency, factor reallocation and structural change, but relatively few focus on AI itself at the regional level. Recent city-level evidence further shows that digital-economy development can promote coordinated pollution and carbon reduction, with potential spatial spillover effects across Chinese cities [6]. The YRD provides an appropriate empirical setting because it combines a strong industrial base, considerable low-carbon transition pressure and relatively advanced AI and digital-innovation foundations [3]. This study uses panel data from 26 YRD cities over 2010-2023 to examine the effect of AI policy exposure on urban low-carbon performance, proxied by a standardised inverse carbon-intensity indicator, and further investigates potential transmission mechanisms, contextual heterogeneity and spatial extension effects.

Literature review

Global climate change has become a major challenge to the sustainable development of human society, and China's "dual carbon" goals have charted a course for the green and low-carbon transformation of the economy and society. As the core driving force behind the new round of technological revolution and industrial transformation, artificial intelligence (AI), with its powerful data processing, intelligent decision-making and autonomous learning capabilities, is profoundly altering the operational models of key carbon-emitting sectors such as energy production, industrial manufacturing and transport, and has become a key technological means for improving carbon emission performance. In the environmental-efficiency literature, carbon emission performance is generally conceptualised as the ability of an economic unit to generate desirable economic output while constraining undesirable carbon emissions. Zhou et al. [7] incorporated undesirable outputs into environmental DEA technologies to evaluate environmental and

carbon-emission performance under a production framework. Zhou and Ang [8] further developed a production-theoretical approach in which changes in aggregate CO_2 emissions are related to production technology and efficiency components. Building on this literature, Zhou et al. [9] proposed a Malmquist CO_2 emission performance index to evaluate changes in total-factor carbon emission performance over time. These production-frontier approaches differ from simple carbon-intensity indicators because they jointly account for multiple production inputs, desirable outputs and undesirable carbon emissions. Accordingly, carbon intensity and total-factor carbon emission performance should not be treated as conceptually identical measures. In the present study, the term "carbon emission performance" is used in a narrower, intensity-based sense to describe relative carbon pressure per unit of economic activity, rather than as a formally estimated total-factor production-frontier efficiency score.

Existing research generally confirms that AI has a significant overall emission-reduction effect. However, the environmental implications of AI are not uniformly positive and should be understood as context-dependent. Vinuesa et al. [10] showed that AI can support sustainable development by improving resource efficiency, information processing and technological innovation, while also generating adverse effects through additional energy demand, unequal access and insufficient regulatory oversight. From a climate-mitigation perspective, Rolnick et al. [11] identified substantial opportunities for machine learning in electricity systems, transport, buildings, industry and emissions monitoring. Kaack et al. [12] further argued that the climate value of AI depends on whether its applications are aligned with high-impact mitigation priorities and whether indirect effects, including rebound effects and infrastructure-related emissions, are appropriately managed. These studies suggest that AI should not be treated as an inherently low-carbon technology; its net environmental effect depends on the application scenario, energy structure and institutional context. At the urban level, Zhang et al. [13] treated the National New Generation Artificial Intelligence Innovation and Development Pilot Zones as a quasi-natural experiment and, using a difference-in-differences model, reported an average 3.4 % reduction in urban carbon emissions. The estimate remained stable across a series of robustness checks. Yu et al. [14] focusing on the application of industrial robots, further demonstrated that industrial robots significantly reduce urban carbon emissions through improvements in energy efficiency and green technology efficiency. At the firm level, Feng et al. [15] examine China's artificial-intelligence pilot-zone policy as a quasi-natural experiment and find that AI policy exposure is associated with significant reductions in firms' carbon emissions. Their analysis further reveals substantial heterogeneity across firm characteristics, indicating that the environmental effects of AI policy depend on firms' internal capabilities and operating conditions. Related evidence from energy-intensive manufacturing also supports the environmental benefits of digital transformation. He et al. [16] find that digital transformation significantly affects carbon emissions in Chinese steel enterprises, while green technological innovation further shapes this relationship. Their findings indicate that the environmental consequences of intelligent and digital technologies depend not only on technology adoption itself but also on firms' innovation capabilities and stages of digital transformation. From a supply-chain perspective, Ren et al. [17] show that AI adoption by focal firms can generate decarbonisation effects among associated upstream and downstream firms. Their results further suggest that green technological innovation constitutes an important channel through which such supply-chain decarbonisation effects are transmitted. Evidence from

the international literature also supports this conclusion. Wang et al. [18] in a study based on China's carbon total factor productivity perspective, found that industrial digitalisation has significant spatial spillover effects on the improvement of carbon emission performance. This total-factor productivity approach represents a broader production-efficiency concept than the carbon-intensity-based indicator adopted in the present study, because it incorporates multiple production inputs and productivity dynamics rather than relying solely on the ratio of CO_2 emissions to economic output.

As environmental governance enters a new phase of coordinated pollution and carbon reduction efforts, scholars have begun to focus on the synergistic governance effects of artificial intelligence. Chu et al. [19] examined the impact of the digital economy on pollution and carbon reduction in the Yangtze River Delta. They found that digital-economy development has nonlinear direct and spatial effects on pollution and carbon emissions, with green technological progress becoming an important emission-reduction channel at later stages of digital development. They found that both indicators increased between 2011 and 2022 and that greater digital-intelligent integration was associated with stronger coordinated governance performance.

Technological innovation represents an important potential pathway through which AI may improve carbon-emission performance. Haefner et al. [20] showed that AI can reshape innovation management by strengthening information search, knowledge recombination, forecasting and decision support, thereby reducing uncertainty and accelerating innovation processes. These capabilities are particularly relevant to green innovation, where firms and cities must identify energy-saving technologies, evaluate complex environmental information and coordinate knowledge across multiple actors. Recent evidence provides stronger support for the link between AI and green technological innovation. Chen et al. [21] show that AI significantly promotes firms' disruptive green innovation by strengthening internal innovation capabilities and facilitating access to external collaborative resources. Their findings also indicate that the effect varies across environmental-regulation conditions, ownership types and pollution categories, suggesting that the innovation-enhancing role of AI is context-dependent. In the international literature, Lee et al. [22] showed that the integration of artificial intelligence and the digital economy can significantly drive the energy transition and accelerate the research, development and application of renewable energy technologies.

Improving energy efficiency represents an important pathway through which AI may affect environmental outcomes. Zhou et al. [23] show that the adoption of AI-related industrial intelligence can reduce regional pollutant-emission intensity by improving energy efficiency and advancing pollution-control technologies. The study further identified energy efficiency as a potential transmission channel linking AI development to carbon-emission performance. Related firm-level evidence also suggests that digital empowerment can mitigate carbon-related risks in manufacturing enterprises, although the effect may vary across firms and over time [24]. Wang et al. [25] used data from 278 Chinese cities to further validate the spatial spillover effects of improved energy efficiency, demonstrating that AI development in one city can drive improvements in energy efficiency in neighbouring cities. Industrial structure upgrading provides another potential transmission channel. Xia et al. [26] show that industrial intelligence can influence both the optimisation and upgrading of industrial structure by reshaping production organisation, technological adoption and sectoral transformation. Such structural change may contribute to lower

carbon intensity when economic activity shifts towards more technology-intensive and higher-value-added production.

Furthermore, the emission-reduction effects of AI exhibit significant heterogeneity. Regional heterogeneity has also been documented in recent research. Zhou et al. [23] find that the emission-reduction effects associated with industrial intelligence vary across Chinese regions, with more pronounced effects in the eastern and western regions. Focusing specifically on the Yangtze River Delta, Chu et al. [19] show that the effects of digital-economy development on pollution and carbon reduction differ between cities with high and low levels of digital-economy development. At the firm level, Feng et al. [15] document heterogeneous carbon-reduction effects of AI pilot-zone policies across firms with different internal capabilities and operating conditions. Chen et al. [21] further show that the effect of AI on green innovation varies across ownership types, pollution categories and environmental-regulation conditions.

Existing research on the impact of artificial intelligence on carbon emission performance has established a relatively comprehensive analytical framework. It has yielded substantial results in identifying impact effects, analysing mechanisms of action and exploring heterogeneity, providing important theoretical support for understanding the relationship between artificial intelligence and green development. However, there remain several shortcomings that require further refinement. Firstly, the research perspective is relatively narrow; most literature focuses on carbon emissions at the industrial production and urban levels, with insufficient attention paid to other high-carbon sectors such as agriculture, the service industry and transport, as well as carbon emissions at the household consumption level. There is a lack of systematic analysis covering the entire industrial chain and the full life cycle. Secondly, the analysis of mechanisms is not sufficiently in-depth. Existing studies predominantly examine the role of individual mechanisms, such as technological innovation or energy efficiency, in isolation, with limited research into the interactive effects and transmission pathways between different mechanisms. Furthermore, although growing attention has been paid to the environmental costs generated by AI itself, these costs remain insufficiently integrated into empirical assessments of the net carbon effects of AI, particularly at the urban level. The expansion of AI relies on data centres, computing hardware and communication infrastructure, which may increase electricity demand, embodied emissions and electronic waste. Strubell et al. [27] demonstrated that the training and optimisation of large AI models can involve substantial energy consumption and carbon emissions, while Schwartz et al. [28] argued that computational efficiency and environmental cost should be considered alongside predictive accuracy in the development of “Green AI”. Lannelongue et al. [29] further showed that the carbon footprint of computation depends on hardware type, processing time, data-centre efficiency, geographic location and the carbon intensity of electricity supply. From a broader governance perspective, van Wynsberghe [30] distinguished between the use of AI for sustainability and the sustainability of AI itself, emphasising that both dimensions should be assessed simultaneously. However, these foundational studies largely predate the rapid diffusion of generative AI and the recent expansion of large-scale AI models and associated computing infrastructure. The emergence of generative AI has therefore renewed concerns about whether improvements in computational efficiency are sufficient to offset the rapidly growing scale of AI deployment and use. In the context of the recent large-model era, De Vries [31] highlighted the growing energy footprint of AI and argued that the widespread adoption of generative AI could substantially increase demand for specialised

computing resources and electricity. Moving beyond the energy cost of model training, Luccioni et al. [32] further showed that AI deployment and repeated inference can generate substantial energy consumption, with general-purpose generative models tending to require considerably more computational resources than task-specific models. This evidence suggests that the environmental cost of contemporary AI should not be evaluated solely on the basis of one-off training expenditure, but should also account for the cumulative energy demand arising from large-scale and repeated inference. At the infrastructure level, the International Energy Agency [33] projected that global data-centre electricity consumption could more than double to approximately 945 TWh by 2030, with AI identified as a major driver of this growth. This projection indicates that the environmental implications of AI are increasingly extending beyond individual models to the broader electricity demand associated with data-centre and computing-infrastructure expansion. More recent IEA evidence further reveals an important tension between improvements in unit-level efficiency and growth in aggregate energy demand. Although the energy required for individual AI tasks is declining rapidly, global data-centre electricity demand continued to increase substantially in 2025, reflecting the simultaneous expansion of AI adoption, inference workloads and increasingly computationally intensive applications [34]. This pattern can be characterised as an efficiency-scale tension: efficiency improvements reduce the energy required per unit of AI service, whereas the expansion of model scale, usage frequency and supporting infrastructure may increase total electricity demand. Taken together, these studies suggest that the environmental effect of AI should be understood as the net outcome of competing efficiency and scale effects. On the one hand, AI may reduce carbon intensity by improving resource allocation, production efficiency and technological innovation; on the other hand, the rapid expansion of model deployment, inference activity and supporting digital infrastructure may increase aggregate electricity and material demand. Therefore, whether AI ultimately improves carbon-emission performance depends on whether application-side efficiency and structural gains outweigh the additional energy demand associated with scale expansion and digital infrastructure. Thirdly, heterogeneity analysis is not sufficiently detailed; existing studies mostly categorise data at macro levels such as regions or corporate ownership, with insufficient research into the differences in AI's emission-reduction effects across different technology types, development stages, and application scenarios, making it difficult to support targeted policy measures. Finally, causal identification methods require improvement; some studies suffer from endogeneity issues, such as reverse causality and omitted variable bias, making it difficult to accurately reveal the causal relationship between AI and carbon emission performance.

Compared with existing studies, this paper makes three main contributions. (1) It focuses on the Yangtze River Delta and uses panel data from 26 cities during 2010-2023 to examine the effect of AI policy exposure on urban low-carbon performance through a two-way fixed-effects difference-in-differences (DID) model. This provides regional policy evidence beyond existing firm-level and industrial robot-based studies. (2) It examines industrial structure upgrading and green technological innovation as two transmission mechanisms, and further analyses differences in policy effects across cities with different levels of government intervention, digital infrastructure, financial development and manufacturing foundations. (3) It investigates spatial extension effects using multiple spatial weight matrices and shows that, although the local policy effect is generally positive, the spatial spillover effect is not robust. This provides a more cautious understanding of AI-driven regional carbon reduction.

Theoretical analysis and research hypotheses

Direct mechanisms of AI on carbon emission performance

The impact of artificial intelligence (AI) on carbon emission performance (CEP) should be understood as a net effect resulting from both emission-reducing and emission-increasing mechanisms. On the one hand, AI can improve low-carbon performance through intelligent sensing, data analysis, predictive maintenance and automated control. These applications help identify inefficiencies in energy use, reduce equipment idling and unplanned downtime, optimise production scheduling and improve the energy input-output efficiency of industrial processes. When low-carbon performance is measured inversely by carbon intensity, these efficiency improvements are expected to be positively associated with CEP.

On the other hand, AI may also generate negative environmental effects. The deployment of AI systems requires computing infrastructure, data centres, sensors and communication networks, all of which may increase electricity demand. This energy-demand channel has become increasingly important in the large-model and generative-AI era. Recent evidence indicates that the environmental cost of AI arises not only from model training but also from repeated inference and the expansion of specialised computing and data-centre infrastructure [31-33]. At the same time, improvements in hardware, software and model efficiency may reduce the energy required for individual AI tasks without necessarily reducing aggregate energy consumption. As AI becomes more widely deployed, increases in model scale, inference frequency and application intensity may offset part of these unit-level efficiency gains, creating an efficiency-scale tension [34]. In addition, efficiency improvements may induce a rebound effect: lower unit energy costs and higher productivity may encourage firms to expand production scale, thereby partly offsetting the emission-reduction gains brought by technological progress. This mechanism is consistent with the rebound-effect literature. Sorrell [35] shows that improvements in energy efficiency can reduce the effective cost of energy services and thereby induce behavioural, production and economy-wide responses that offset part of the expected energy savings. Gillingham et al. [36] further emphasise that rebound effects may arise through both direct increases in energy-service demand and broader indirect or economy-wide adjustments following improvements in energy efficiency. AI-enabled automation may also accelerate production and logistics activities, leading to higher aggregate energy consumption if energy structures remain carbon-intensive. This efficiency-scale tension within the broader dual-effect framework is consistent with the distinction between AI for sustainability and the sustainability of AI itself [30], as well as with evidence that the climate contribution of AI depends on application choices, electricity sources and the management of indirect effects [12, 29].

In the context of the Yangtze River Delta, the region has relatively strong industrial foundations, digital infrastructure and environmental governance capacity. AI technologies are therefore more likely to be embedded in industrial energy conservation, intelligent scheduling, carbon monitoring and green innovation scenarios. Accordingly, this paper expects the positive efficiency and structural effects of AI policy exposure to dominate, but treats this as an empirical proposition to be tested rather than as an established fact.

Hypothesis H1: Under the condition that AI-induced efficiency improvements, industrial upgrading and green innovation outweigh the additional energy-demand, scale-expansion and rebound effects associated with AI deployment, exposure to AI policies in

core cities is expected to be positively associated with improvements in urban low-carbon performance.

Mediating mechanism: industrial structure upgrading

Technological progress can reshape industrial structure by altering relative production costs, factor allocation and the technological requirements of different sectors. As new technologies diffuse, resources may gradually shift from relatively labour- and energy-intensive activities towards more capital-, knowledge- and technology-intensive industries, thereby promoting structural upgrading. Recent theoretical work further shows that technology choice and factor-endowment changes can drive industries towards more advanced production technologies and facilitate industrial upgrading [37]. Artificial intelligence may reinforce this structural transformation through several channels. First, AI-enabled automation can substitute for routine and low-productivity production tasks, encouraging firms and industries to adopt more technology-intensive production modes. Second, AI can improve information processing, resource allocation and production coordination, thereby lowering the relative cost of knowledge-intensive and high-value-added activities. Third, the diffusion of intelligent technologies may stimulate the expansion of digital and producer-service sectors while accelerating the transformation of traditional manufacturing. Empirical evidence from China indicates that AI development can significantly promote industrial structure upgrading [38], while more recent evidence shows that industrial intelligence affects both the quantity and quality of industrial-structure transformation [26]. Complementary evidence from the environmental-management literature also highlights the joint role of technological investment and institutional conditions in structural transformation. Chen et al. [39] show that R&D investment and environmental regulation can significantly shape industrial transformation and upgrading, with the effect varying across regulatory intensity. This evidence further supports the view that technology-driven structural upgrading is conditional on both innovation inputs and the surrounding institutional environment. From an environmental perspective, such structural upgrading can improve urban low-carbon performance when economic activity shifts away from highly energy-intensive sectors towards activities with higher value added and lower carbon intensity. This provides the theoretical basis for treating industrial structure upgrading as a potential transmission channel between AI policy exposure and carbon-emission performance.

Hypothesis H2: AI improves CEP by promoting industrial structure upgrading.

Mediating mechanism: green technological innovation

AI can improve R&D search efficiency, shorten trial-and-error cycles, and promote the diffusion of green technologies. The broader eco-innovation literature provides additional support for the environmental relevance of this mechanism. Dabuo et al. [40] find that eco-innovation significantly improves green development efficiency in China by reducing resource and energy consumption and improving resource utilisation, while environmental regulation further conditions this relationship. However, there may be a time lag in the conversion of innovation inputs into energy-saving and emission-reduction performance. Therefore, this paper treats green technological innovation as a mediating variable in the mechanism testing, rather than controlling for it prematurely in the baseline model. Hypothesis H3: AI improves CEP by promoting green technological innovation.

Spatial extension and contextual heterogeneity

The dense economic linkages within the YRD may create potential channels for cross-city diffusion of technology and environmental governance practices. Recent evidence also documents substantial spatial heterogeneity in the coordinated green-development performance of YRD cities [41]. Recent evidence from Chinese urban agglomerations further suggests that green innovation may generate cross-city spillovers through knowledge diffusion and collaborative networks. Zhao et al. [42] using spatial econometric models for 272 Chinese cities, show that environmental regulation can reshape green-innovation spillover patterns across urban agglomerations. However, whether such spatial effects are empirically robust remains an open question to be tested in the spatial extension analysis. In addition, the low-carbon effects of AI policies may vary across different urban contexts. Government intervention and environmental regulation may also shape the effectiveness of low-carbon policies. Zhao [43] shows that heterogeneous environmental-regulation instruments generate different carbon-reduction effects, partly through green innovation and productivity improvements; digital infrastructure determines data collection and intelligent scheduling capabilities; financial development influences financing for green innovation; and the manufacturing base determines the application scenarios for AI-driven energy-saving retrofits. Hypothesis H4: The effects of AI policies exhibit heterogeneity across different urban contexts.

Research design

Variable measurement

Carbon emission performance has been measured in different ways in the existing literature. Production-theoretical approaches typically incorporate capital, labour, energy, desirable output and undesirable CO_2 emissions into DEA or related frontier models to construct environmental-efficiency or total-factor carbon-performance indicators [7]. Zhou and Ang [8] further link carbon-emission changes to production technology and efficiency components within a production-theoretical framework, while Zhou et al. [9] develop a Malmquist CO_2 emission performance index for measuring dynamic total-factor carbon performance. Such measures provide a comprehensive representation of production-side carbon efficiency but require consistent multi-input and multi-output data. The present study adopts a more parsimonious intensity-based measure because consistent city-level data on capital stock, labour, energy inputs and other variables required for production-frontier estimation are not uniformly available across all 26 YRD cities over 2010-2023. Specifically, carbon intensity is first calculated as CO_2 emissions per unit of GDP, and the resulting series is standardised and multiplied by −1, yielding $CEP_CI = -Z(CO_2/GDP)$. A higher value therefore indicates lower relative carbon-emission pressure per unit of economic output. This indicator should be interpreted as a standardised inverse carbon-intensity proxy for urban low-carbon performance, rather than as a total-factor carbon-efficiency or DEA-based carbon-performance score.

DID identification strategy

Difference-in-Differences (DID) provides a widely used quasi-experimental framework for identifying policy effects by comparing changes in outcomes between treated and untreated units before and after policy implementation. In panel settings, the

two-way fixed-effects (TWFE) specification can absorb time-invariant unit characteristics and common time shocks. Recent methodological studies, however, have shown that conventional TWFE estimators and event-study specifications require particular caution when treatment timing varies across units or treatment effects are heterogeneous. Sun and Abraham [44] show that conventional event-study coefficients may be contaminated when treatment effects vary across cohorts and treatment timing is staggered.

Table 1

Variable definitions, measurement methods and data sources

Variable	Definition	Measurement methods	Data sources
CEP_CI	Inverse carbon intensity performance indicator	CEP_CI = –Z(CO_2/GDP), where a higher value indicates lower relative carbon-emission pressure	Calculated from GDP and CO_2 in the original panel data
DID	AI Policy Exposure	Treat4 × Post-2017; Treat4 = Shanghai, Nanjing, Hangzhou, Suzhou	State Council, New Generation Artificial Intelligence Development Plan (State Council Document No. 35 [2017]) [45], and relevant local policy documents
GDP	Level of economic development	Per capita regional GDP, deflated to 2010 prices	China Urban Statistical Yearbook
Urb	Level of urbanisation	Proportion of permanent urban population	China Urban Statistical Yearbook
Open	Level of openness	Actual utilised foreign investment	China Urban Statistical Yearbook, China Business Yearbook
Hc	Human capital	Number of students enrolled in regular higher education institutions	China Urban Statistical Yearbook, China Education Statistical Yearbook
IS	Industrial structure upgrading	Value added in the tertiary sector / Value added in the secondary sector	China Urban Statistical Yearbook
GI	Green technological innovation	Number of green patent applications / Permanent resident population	CNRDS Green Patent Research Database, China National Intellectual Property Administration [45, 46]
ES	Energy mix	Share of coal or fossil fuel consumption	China Energy Statistical Yearbook, CEADs (Carbon Emission Accounts and Datasets)
Gov	Intensity of government intervention	Local government general public budget expenditure / GDP	China Urban Statistical Yearbook, China Fiscal Yearbook
Fin	Level of financial development	Balance of deposits and loans in financial institutions / GDP	China Urban Statistical Yearbook, China Financial Yearbook
Digi	Digital infrastructure	Number of broadband internet subscribers / Permanent resident population	China Urban Statistical Yearbook, China Information Yearbook
Robot	Industrial robot density	Number of industrial robots / Number of people employed in manufacturing	IFR World Robotics Database
CO_2	Total carbon emissions	CO_2 emissions at city level	CEADs China Urban Carbon Emissions Inventory, Ecological and Environmental Bulletin

Callaway and Sant'Anna [47] develop DID identification and estimation methods for settings with multiple treatment periods and heterogeneous treatment effects. Borusyak et al. [48] further propose robust and efficient event-study estimators for staggered-adoption designs. In the present study, the four treated cities are assigned

a common policy onset rather than adopting treatment at different times. Therefore, the staggered-adoption concerns emphasised in these studies are less directly applicable to the baseline design, although the parallel-trends assumption and the reliability of statistical inference remain central to identification. This paper designates Shanghai, Nanjing, Hangzhou and Suzhou as the treatment group cities, with the remaining 22 cities forming the control group. The selection of these four core cities is based on their central role in the Yangtze River Delta's AI industry cluster, policy pilot schemes and technology diffusion, enabling them to more accurately reflect the intensity of AI policy impacts. Given that the State Council issued the New Generation Artificial Intelligence Development Plan (State Council Document No. 35 in July 2017 [45], local governments and enterprises may have formed policy expectations and initiated early preparations during that year. Accordingly, this paper defines the period from 2017 onwards as the post-policy period and uses 2018 as an alternative policy onset for robustness testing. Because the baseline treatment definition assigns the same policy onset to all four treated cities, the present design does not involve multiple treatment cohorts with staggered adoption. Accordingly, the conventional TWFE-DID specification is retained as the baseline estimator. This modelling choice avoids the principal staggered-treatment concern discussed by Sun and Abraham [44] differs from the multiple-treatment-period setting considered by Callaway and Sant'Anna [47] and does not require the staggered-adoption correction proposed by Borusyak et al. [48]. The event-study specification is therefore used to examine dynamic coefficients around a common treatment date rather than to aggregate effects across multiple treatment cohorts.

$$\mathrm{CEP_CI}_{it} = \alpha + \beta \mathrm{DID}_{it} + \gamma X_{it} + \lambda_i + \delta_t + \varepsilon_{it}$$

Here, $\mathrm{CEP_CI}_{it}$ denotes the low-carbon performance of city i in year t, measured as the standardised inverse carbon-emission-intensity indicator, CEP_CI = –Z(CO2/GDP), where a higher value indicates lower relative carbon-emission pressure. As clarified above, CEP_CI is an intensity-based proxy for relative urban low-carbon performance and should not be interpreted as a production-frontier or total-factor carbon-efficiency measure. DID_{it} denotes AI policy exposure and equals the interaction between the treatment-group indicator and the post-policy-period indicator. X_{it} is a vector of control variables, including per capita GDP, urbanisation, openness, human capital, energy structure and government intervention. λ_i and δ_t denote city fixed effects and year fixed effects, and ε_{it} is the error term. Industrial structure upgrading (IS) and green technological innovation (GI) are excluded from the baseline total-effect specification and are examined separately in the mechanism analysis to avoid over-control bias. This treatment of the mechanism variables is motivated by the potential post-treatment bias arising from conditioning on variables that may themselves be affected by policy exposure. Acharya et al. [49] show that controlling for variables located on the causal pathway between treatment and outcome can distort the interpretation of treatment effects and may generate biased estimates of direct effects. Accordingly, IS and GI are not included as baseline controls because they are theoretically positioned as potential post-treatment transmission channels rather than predetermined confounders.

Mechanisms, heterogeneity and spatial extension

Mechanism tests employ a panel fixed-effects stepwise regression framework. First, the total effect of DID on $\mathrm{CEP_CI}_{it}$ is estimated. Second, the effects of DID on industrial

structure upgrading (IS) and green technological innovation (GI) are examined separately. Third, IS and GI are introduced into the $\mathrm{CEP_CI}_{it}$ equation to assess whether these variables are associated with improvements in low-carbon performance. It should be noted that changes in the DID coefficient after additional variables are introduced do not, by themselves, constitute a formal decomposition of the treatment effect. Gelbach [50] shows that sequentially adding correlated covariates can produce order-dependent coefficient changes, and proposes a conditional decomposition framework to attribute changes in a focal coefficient to specific covariates. In the present study, the stepwise regressions are therefore used only to examine whether the proposed channel variables display associations consistent with the hypothesised transmission mechanisms, rather than to interpret changes in the DID coefficient as an exact decomposition of the total policy effect. This procedure is used only to provide indicative evidence on possible transmission channels. Because the analysis is based on stepwise fixed-effects regressions rather than SEM, 3SLS, Sobel tests or bootstrap-based mediation analysis, this paper does not report precise mediation shares or treat the products of coefficients as formal causal decompositions. This distinction is also consistent with the causal mediation literature. Imai et al. [51] emphasise that causal mediation effects require explicit identification assumptions, including assumptions governing the treatment-mediator and mediator-outcome relationships, rather than being identified automatically from a sequence of regression coefficients. Because these identifying assumptions cannot be fully verified in the present city-level observational setting, the mechanism analysis is not interpreted as estimating formally identified natural direct or indirect effects.

Heterogeneity analysis is conducted using both subgroup regressions and formal coefficient-difference tests. Specifically, cities are divided into high- and low-level groups according to government intervention, financial development, digital infrastructure and manufacturing foundation. In addition to estimating subgroup DID coefficients, this paper further introduces interaction terms between DID and the corresponding group dummy variables in the full-sample model. The coefficient of the interaction term is used to test whether the policy effect differs significantly across groups. The empirical specification is as follows:

$$\mathrm{CEP_CI}_{it} = \alpha + \beta \mathrm{DID}_{it} + \theta(\mathrm{DID}_{it} \times \mathrm{Group}_i) + \gamma X_{it} + \lambda_i + \delta_t + \varepsilon_{it}$$

where $Group_i$ denotes the high-level group dummy for a given urban characteristic. The coefficient θ captures the conditional incremental difference in the DID effect associated with the high-level group within the pooled TWFE specification. It should not be interpreted as the arithmetic difference between DID coefficients obtained from separately estimated subgroup regressions. If θ is not statistically significant, the subgroup estimates are interpreted only as descriptive heterogeneity rather than evidence of a statistically significant difference between groups. For groups with insufficient DID variation, the coefficient-difference test is not reported, and no strong heterogeneity conclusion is drawn.

Spatial extension is conducted as a supplementary analysis using alternative spatial weight matrices to construct spatial lag terms. All spatial weight matrices are row-standardised prior to estimation so that the non-zero weights in each row sum to one. Given that spatial spillovers may affect the SUTVA assumption of traditional DID, the spatial results are treated solely as supplementary evidence and do not serve as the core basis for benchmark causal identification.

Scope of auxiliary analyses and diagnostic reporting

To improve the transparency of the empirical design, this paper clarifies the scope of several auxiliary analyses. All variables used in the regressions are defined in Table 1, including their measurement methods and data sources. The mechanism analysis is conducted using panel fixed-effects stepwise regressions rather than SEM, 3SLS, Sobel tests or bootstrap-based mediation analysis, and the mechanism results are interpreted only as indicative evidence. The baseline DID and robustness regressions report standard errors clustered at the city level to account for within-city dependence over time. However, treatment variation in the baseline design is concentrated in only four treated cities, which raises an additional finite-sample inference concern. Conley and Taber [52] show that conventional asymptotic inference may be unreliable when policy variation is generated by only a small number of treated groups. Ferman and Pinto [53] further demonstrate that heteroskedasticity can distort inference in DID designs with few treated groups. MacKinnon and Webb [54] show that conventional cluster-robust and bootstrap-based inference may also perform poorly when the number of treated clusters is very small. Accordingly, the city-clustered standard errors reported in this study are treated as conventional reference estimates, and the associated significance levels are interpreted cautiously rather than as eliminating the few-treated-group inference problem. The PSM-DID test is used only as a supplementary robustness check because of the small number of treated cities and limited common support.

This study does not report an instrumental-variable specification for AI application intensity, because available candidate instruments are difficult to justify in terms of the exclusion restriction. In particular, lagged industrial robot density may affect carbon emission performance through production-process adjustment, energy-use patterns and automation-related changes, rather than only through AI application intensity. Therefore, it is not used as a formal instrument in this study.

Empirical analysis

The empirical results are presented in the following order: ‘Sample description - Benchmark identification - Dynamic testing - Robustness and supplementary checks - Mechanisms and heterogeneity - Spatial extension’.

Table 2 shows that, after standardisation, the mean of CEP_CI is close to 0 and the standard deviation is approximately 1, making the indicator suitable for comparing relative carbon-intensity-based low-carbon performance across cities within the sample. The mean of DID is fully matched with the four treatment cities and the 2017 policy time point. Significant differences in robot density, green technological innovation and digital infrastructure indicate a marked divergence among Yangtze River Delta cities in terms of their AI application foundations and innovation capabilities.

Table 3 reports the correlation matrix of key variables. The correlation matrix shows that DID is positively correlated with GDP, IS, GI and digital infrastructure, indicating that core AI cities tend to possess stronger economic foundations and conditions for innovation.

Table 4 reports the baseline regression, robustness checks and sensitivity diagnostics. In the baseline model, the DID coefficient is 0.493 and is significant at the 5 % level, suggesting a positive association between AI policy exposure and low-carbon performance. The coefficient remains positive and statistically significant when the policy year is

changed to 2018, when the treatment group is expanded, when the sample period is shortened and when Shanghai is excluded.

Table 2

Complete descriptive statistics

Variables	*N*	Mean	Standard deviation	Minimum	P25	Median	P75	Maximum
CEP_CI	364	–0.000	1.000	–5.719	–0.301	0.215	0.601	1.426
DID	364	0.077	0.267	0.000	0.000	0.000	0.000	1.000
GDP	364	9.001	3.861	2.000	6.020	8.595	11.580	21.680
Urb	364	66.716	11.552	37.200	58.800	67.250	74.600	89.600
Open	364	28.817	37.082	2.800	10.475	17.950	32.550	255.000
Hc	364	15.533	18.087	1.850	5.157	8.425	13.610	76.150
IS	364	1.088	0.269	0.400	0.900	1.100	1.240	2.050
GI	364	3.388	2.525	0.320	1.530	2.675	4.400	12.660
ES	364	61.696	8.783	38.500	55.600	61.600	68.150	83.500
Gov	364	19.123	4.692	10.800	15.600	18.150	22.125	33.500
Fin	364	2.550	0.499	1.760	2.190	2.450	2.772	4.260
Digi	364	44.482	18.537	11.200	29.025	43.450	58.350	94.700
Robot	364	28.238	22.352	2.800	12.150	21.700	37.725	132.700
CO_2	364	6,357.918	4885.693	1890.000	3579.000	4,766.500	7,480.500	25,089.000

CEP_CI is the standardised inverse carbon emission intensity indicator; the DID mean of 0.077 corresponds to the policy exposure observations for the four treatment cities between 2017 and 2023

Table 3

Correlation matrix of key variables

Variable	CEP_CI	DID	GDP	Urb	Open	IS	GI	ES	Gov	Digi
CEP_CI	1.000	–0.230	0.102	–0.153	–0.700	–0.117	–0.050	–0.109	0.330	0.248
DID	–0.230	1.000	0.466	0.423	0.578	0.488	0.639	–0.384	–0.068	0.489
GDP	0.102	0.466	1.000	0.836	0.468	0.370	0.921	–0.673	–0.255	0.889
Urb	–0.153	0.423	0.836	1.000	0.589	0.364	0.747	–0.581	–0.284	0.697
Open	–0.700	0.578	0.468	0.589	1.000	0.492	0.605	–0.407	–0.326	0.368
IS	–0.117	0.488	0.370	0.364	0.492	1.000	0.530	–0.797	0.327	0.610
GI	–0.050	0.639	0.921	0.747	0.605	0.530	1.000	–0.694	–0.205	0.878
ES	–0.109	–0.384	–0.673	–0.581	–0.407	–0.797	–0.694	1.000	–0.083	–0.826
Gov	0.330	–0.068	–0.255	–0.284	–0.326	0.327	–0.205	–0.083	1.000	0.070
Digi	0.248	0.489	0.889	0.697	0.368	0.610	0.878	–0.826	0.070	1.000

Correlation coefficients are used solely to describe preliminary relationships between variables; formal conclusions are subject to the results of the fixed-effects DID analysis

However, the results are not uniformly robust across all specifications. After the potential mechanism variables are included directly, the DID coefficient declines to 0.199 and becomes statistically insignificant. However, this specification should not be interpreted as a preferred estimate of the total policy effect, because IS and GI may themselves be affected by AI policy exposure. Conditioning on such post-treatment variables can generate over-control or post-treatment bias and may alter the interpretation of the treatment coefficient [48]. The specification is therefore reported only as a sensitivity diagnostic rather than as evidence that the baseline policy effect disappears after "controlling for" the mechanisms. After adding city-specific linear trends, the coefficient further declines to 0.048, while the goodness-of-fit becomes extremely high (R^2 = 0.9996), indicating that city-level differential trends may absorb much of the policy-related

variation. Therefore, the baseline result supports a positive relationship between AI policy exposure and low-carbon performance, but the effect size and robustness should be interpreted cautiously.

Table 4

Results of baseline regression, robustness checks and sensitivity diagnostics (CEP_CI scope)

Model specification	**DID variable**	**Coefficient**	**Standard error**	***t*-value**	***p*-value**	R^2	***N***
Base model: 4 cities × post-2017	did4_2017	0.493	0.210	2.343	0.019	0.985	364
Policy effective date changed to 2018	did4_2018	0.440	0.185	2.372	0.018	0.984	364
Expanded to an 8-city treatment group	did8_2018	0.439	0.105	4.172	3.00e-05	0.986	364
Inclusion of mediating variables (over-control test)	did4_2017	0.199	0.192	1.038	0.299	0.988	364
Shortened sample period: 2015-2023	did4_2017	0.332	0.141	2.361	0.018	0.993	234
Excluding Shanghai	did4_2017	0.484	0.134	3.611	3.00e-04	0.976	350
Add city linear trend	did4_2017	0.048	0.018	2.645	0.008	0.9996	364

The dependent variable is CEP_CI = $-Z(CO_2/GDP)$. Standard errors are cluster-robust at the city level. The specifications including mediating variables and city-specific linear trends are interpreted as sensitivity diagnostics rather than purely confirmatory robustness tests

Recent methodological research has highlighted potential interpretational problems in conventional TWFE event-study regressions when treatment timing differs across units and treatment effects are heterogeneous. Sun and Abraham [46] show that lead and lag coefficients may be contaminated by treatment effects from other relative periods under staggered treatment timing. Callaway and Sant'Anna [47] provide an alternative framework for identifying group-time average treatment effects when units receive treatment at different times. Borusyak et al. [48] further develop robust and efficient event-study estimators for such staggered-adoption settings. In the present study, however, the four treated cities share the same baseline policy onset, while the remaining cities remain untreated under the baseline definition. Therefore, the event-study specification does not compare early-treated with later-treated cohorts. The event-study results are nevertheless used primarily to examine pre-policy dynamics and the temporal evolution of the estimated policy association rather than as independent proof of causal identification.

Table 5 and Figure 1 show that the reported pre-policy coefficients do not exhibit systematic statistically significant deviations from zero, and the joint pre-policy test yields a *p*-value of 0.438. Thus, the null hypothesis that the reported pre-treatment coefficients are jointly equal to zero is not rejected. However, this non-rejection should not be interpreted as proof that the parallel-trends assumption holds. Roth [55] emphasises that conventional pre-trend tests may have limited power and that non-rejection of pre-treatment differences

should therefore be interpreted cautiously. Accordingly, the event-study evidence in this study is interpreted as being consistent with the absence of detectable pre-trends rather than as establishing the parallel-trends assumption. Rambachan and Roth [56] further show that DID conclusions can be evaluated under bounded deviations from exact parallel trends rather than relying exclusively on a binary pre-trend test. In light of this perspective, and given the limited number of treated cities in the present study, the event-study evidence is interpreted cautiously and jointly with the broader robustness and sensitivity analyses.

Table 5

Event study and parallel trends tests

Relative period	Coefficient	Standard error	*t*-value	*p*-value
Pre ≤ –5	–0.165	0.250	–0.660	0.509
Pre-4	–0.213	0.163	–1.306	0.192
Pre-3	–0.182	0.106	–1.717	0.086
Pre-2	–0.099	0.055	–1.793	0.073
2016 (Baseline)	-	-	-	-
Post0	0.249	0.111	2.246	0.025
Post1	0.377	0.141	2.669	0.008
Post2	0.407	0.121	3.369	0.001
Post3	0.410	0.117	3.519	< 0.001
Post4	0.409	0.127	3.218	0.001
Post5	0.388	0.149	2.604	0.009
Post6	0.389	0.194	2.002	0.045
Pre-policy joint inspection	-	-	-	0.438

2016 is taken as the base period. The joint test of the reported pre-policy coefficients yields a *p*-value of 0.438, so the null hypothesis that these coefficients are jointly equal to zero is not rejected. However, this non-rejection should not be interpreted as verification of the parallel-trends assumption. Roth [55] shows that conventional pre-trend tests may have limited statistical power, so a failure to detect significant pre-trends does not by itself establish parallel trends

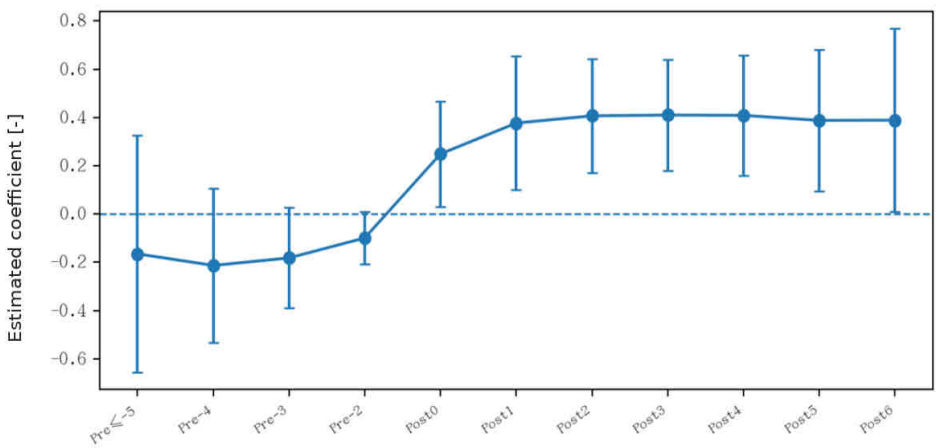


Fig. 1. Coefficients estimated using the event study method (base period: 2016)

Table 6 presents the PSM-DID result as a supplementary check on observable sample comparability. The estimated coefficient remains positive after matching. However, the

matched sample contains only 70 observations, and nearest-neighbour matching with replacement may reuse some control observations. The resulting standard error is unusually small relative to the sample size, producing a large t-statistic. Therefore, the apparent precision of this estimate should not be overinterpreted. The PSM-DID result is used only to examine whether the direction of the baseline association remains positive after matching and is not treated as independent or decisive causal evidence. The main conclusions continue to rely on the baseline DID model, the event-study analysis and the broader sensitivity checks.

Table 6

Supplementary PSM-DID check

Test	Variable	Coefficient	Standard error	*t*-value	*p*-value	*N*
PSM-DID (nearest neighbour matching, with replacement)	DID	0.307	0.018	16.740	< 0.001	70

The PSM-DID analysis is reported only as a supplementary sample-comparability check. Matching was conducted using nearest-neighbour matching with replacement, resulting in 70 matched observations. Given the small matched sample, the reuse of control observations under nearest-neighbour matching with replacement, and the limited number of treated cities, the conventional standard error may understate the true uncertainty surrounding the estimate. Conley and Taber [52] caution that standard DID inference may be unreliable when only a small number of groups experience policy changes. Ferman and Pinto [53] further show that heteroskedasticity can exacerbate inference problems in DID designs with few treated groups. MacKinnon and Webb [54] demonstrate that conventional cluster-based inference can also be problematic when the number of treated clusters is very small. Accordingly, the unusually small standard error and large t-statistic in the PSM-DID specification are interpreted cautiously and are not regarded as evidence of unusually high statistical precision

This study does not further report an instrumental-variable specification for AI application intensity. Although lagged industrial robot density is closely related to a city's automation and intelligent manufacturing foundation, it may also directly affect carbon emission performance through production-process adjustment, equipment replacement, energy-use patterns and automation-related efficiency changes. These channels overlap with the mechanisms through which AI and industrial upgrading may influence low-carbon performance. Therefore, using lagged robot density as an instrument would require an exclusion restriction that is difficult to justify within the theoretical framework of this study. For this reason, the empirical identification in this paper relies primarily on the policy-exposure DID framework, the event-study evidence and a series of robustness checks, while the PSM-DID result is retained as a supplementary check on sample comparability.

Formal causal mediation analysis would require stronger identification conditions than those needed for the baseline DID effect. In particular, causal interpretation of an indirect effect requires assumptions concerning confounding of the mediator-outcome relationship that cannot be guaranteed merely by controlling for observed covariates [51]. Therefore, the products of the DID → mediator and mediator → CEP coefficients are reported only as descriptive arithmetic quantities and are not labelled as average causal mediation effects.

Table 7 shows positive point estimates along both proposed transmission channels. AI policy exposure is positively associated with industrial structure upgrading and green technological innovation, while both channel variables are positively associated with

CEP_CI in the corresponding stepwise regressions. After rounding to one significant digit, the arithmetic products are approximately 0.2 for both transmission channels. Before rounding, however, the two products are approximately 0.155 and 0.156, respectively. Their near equality reflects the specific combination of the component coefficients: the industrial-structure pathway combines a larger DID → IS coefficient with a smaller IS → CEP_CI coefficient, whereas the green-innovation pathway combines a smaller DID → GI coefficient with a larger GI → CEP_CI coefficient. This numerical coincidence should not be interpreted as evidence that the two mechanisms contribute equally to the estimated policy effect.

Table 7

Stepwise mechanism associations

Potential transmission channel	Path	Coefficient	Standard error	Approximate arithmetic product
Industrial structure upgrading (IS)	DID→IS	0.044	0.012	-
Industrial structure upgrading (IS)	IS→CEP_CI	3.534	1.635	≈0.2
Green innovation (GI)	DID→GI	1.114	0.340	-
Green innovation (GI)	GI→CEP_CI	0.140	0.052	≈0.2

The arithmetic products are reported to one significant digit as ≈0.2 for both transmission channels. Before rounding, the products are approximately 0.155 for the industrial-structure channel (3.534 × 0.044) and 0.156 for the green-innovation channel (1.114 × 0.140). Their close numerical values arise from different combinations of component coefficients and should not be interpreted as evidence that the two channels have equal causal indirect effects or equal substantive importance. These values are reported solely as descriptive calculations rather than formally estimated indirect effects

Table 8

Heterogeneity analysis

Grouping dimensions	Grouping	DID coefficient	Standard error	*t*-value	*p*-value	*N*
Government intervention	High group	0.743	0.201	3.700	< 0.001	182
Government intervention	Low group	0.387	0.175	2.220	0.027	182
Financial development	High group	0.238	0.232	1.030	0.303	174
Financial development	Low group	Cannot be reliably estimated	-	-	-	190
Digital infrastructure	High group	0.206	0.122	1.680	0.093	182
Digital infrastructure	Low group	Cannot be reliably estimated	-	-	-	182
Manufacturing base	High group	0.292	0.188	1.550	0.120	178
Manufacturing base	Low group	0.404	0.160	2.520	0.012	186

For some groups with low financial and digital infrastructure, reliable estimates could not be obtained due to insufficient DID variation; this is noted in the table. The high- and low-group DID coefficients are estimated from separate subgroup regressions. Therefore, the arithmetic difference between the two subgroup coefficients is reported only as descriptive evidence and is not interpreted as a formal estimate of the between-group treatment-effect difference. Formal coefficient-difference tests are reported separately in Table 9

Table 9

Coefficient-difference tests for heterogeneity

Grouping dimensions	Interaction term	Coefficient θ	Standard error	*t*-value	*p*-value	Interpretation
Government intervention	DID × HighGov	–0.178	0.084	–2.117	0.034	Significant at 5 %
Financial development	DID × HighFin	Not separately estimable	-	-	-	Insufficient DID variation
Digital infrastructure	DID × HighDigi	Not separately estimable	-	-	-	Insufficient DID variation
Manufacturing base	DID × HighManu	–0.009	0.114	–0.078	0.938	Not significant

The dependent variable is CEP_CI. Each row reports the coefficient of the interaction term between DID and the corresponding high-level group dummy in the full-sample two-way fixed-effects model. Because this interaction coefficient is estimated conditionally within a pooled specification that imposes common coefficients on the remaining covariates, it is not algebraically equivalent to the difference between DID coefficients obtained from two separately estimated subgroup regressions. City fixed effects and year fixed effects are included, and standard errors are clustered at the city level. The high- and low-level groups are defined according to the median-split rule used in Table 8. For financial development and digital infrastructure, the low-level groups contain no treated-post DID observations under the current grouping rule; therefore, the coefficient-difference test cannot be separately estimated for these two dimensions

To avoid inferring heterogeneity solely from differences in subgroup significance levels, this study further conducts coefficient-difference tests using interaction terms in the full sample. Specifically, for each grouping dimension, an interaction term between DID and the high-level group dummy is introduced into the two-way fixed-effects model. The coefficient of DID captures the policy effect in the low-level group, while the coefficient of DID × High Group captures whether the policy effect differs significantly between the high- and low-level groups. If the interaction term is not statistically significant, the subgroup results are interpreted only as descriptive differences rather than as evidence of statistically significant heterogeneity. Because the high- and low-level coefficients reported in Table 8 are obtained from separately estimated subgroup regressions, their arithmetic difference should be interpreted only descriptively and should not be treated as algebraically equivalent to the interaction coefficient estimated from the pooled full-sample model in Table 9.

Table 8 reports the subgroup regression results, while Table 9 further presents the coefficient-difference tests based on DID × High Group interaction terms in the full-sample two-way fixed-effects model. The interaction coefficient θ captures whether the DID effect differs significantly between the high- and low-level groups.

For government intervention, the separate subgroup regressions in Table 8 yield DID coefficients of 0.743 for the high-intervention group and 0.387 for the low-intervention group, implying a descriptive arithmetic difference of +0.356. By contrast, the pooled full-sample specification in Table 9 yields an interaction coefficient of –0.178 for DID × HighGov (p = 0.034). The two quantities, however, are not estimates of exactly the same parameter. The subgroup coefficients are obtained from two independently estimated regressions, which allow the relationships between the control variables, fixed effects and the dependent variable to be re-estimated separately within each subsample. In contrast, the interaction coefficient is estimated within a pooled TWFE specification, in which the remaining covariate coefficients are constrained to be common across the high- and

low-intervention groups while the DID × HighGov term captures the conditional incremental difference associated with the high-intervention group. Consequently, the arithmetic difference between the two independently estimated subgroup coefficients need not coincide with the pooled interaction coefficient and may even differ in sign when the two specifications impose different parameter restrictions. Nevertheless, the sign reversal is substantively informative because it shows that the apparent government-intervention heterogeneity is sensitive to model specification. We therefore do not interpret the larger subgroup coefficient in the high-intervention sample as robust evidence that stronger government intervention amplifies the carbon-reduction effect of AI policy. Instead, the government-intervention result is treated as specification-sensitive and suggestive, rather than as evidence of a stable or monotonic moderating effect.

Therefore, the heterogeneity results should be interpreted cautiously. The subgroup regressions are used mainly to identify possible contextual conditions for AI policy implementation, rather than to claim statistically confirmed differences across all city groups.

Table 10

Spatial extension test

Spatial weight matrix	DID coefficients	DID *p*-value	W×DID coefficient	W×DID *p*-value	ρ coefficient	ρ *p*-value
Inverse square of geographical distance matrix	0.468	< 0.001	0.009	0.926	0.011	0.921
Geographical distance inverse matrix	0.461	< 0.001	0.196	0.405	–0.445	0.066
K-nearest neighbours matrix (k = 4)	0.488	< 0.001	0.047	0.496	0.086	0.303

All spatial weight matrices are row-standardised prior to estimation. The spatial extension is used only as a supplementary analysis. The local DID coefficient remains positive across alternative spatial weight matrices, whereas the spatially lagged policy term (W × DID) is statistically insignificant in all three specifications. The estimated spatial autoregressive parameters also remain within the admissible range. Accordingly, the results do not provide robust evidence of a stable positive cross-city spillover effect

After row-standardising all spatial weight matrices and re-estimating the spatial specifications, Table 10 shows that the local AI-policy coefficient remains positive across all three spatial-connectivity definitions. In contrast, although the estimated W × DID coefficients remain positive across the three specifications, none is statistically distinguishable from zero at conventional significance levels. The estimated spatial autoregressive parameters are 0.011, -0.445 and 0.086, respectively, all of which fall within the admissible range under the row-standardised specifications. These findings therefore do not provide robust evidence of a stable positive cross-city spillover effect. The spatial analysis is retained only as a supplementary extension and does not constitute the core basis for causal identification.

Research conclusions and implications

Main conclusions

Artificial intelligence (AI) policy exposure is associated with improved carbon-emission performance, mainly through local policy effects. The stepwise regressions provide indicative evidence that industrial structure upgrading and green technological innovation may serve as potential transmission channels. These findings

support an integrated local emission-reduction framework of "technology empowerment - structural transformation - green innovation". Importantly, the spatial extension tests do not provide consistent evidence of stable positive cross-city spillover effects across alternative spatial weight matrices. Therefore, this study treats the spatial results only as a supplementary diagnostic extension and does not regard cross-city spillovers as a confirmed pathway through which AI improves carbon emission performance.

This study employs panel data from 26 prefecture-level cities in the Yangtze River Delta region (2010-2023) to empirically examine the impact of AI development on CEP. The main findings are as follows:

First, based on a modified low-carbon performance indicator and a difference-in-differences (DID) specification, this study provides empirical evidence that policy exposure in core AI cities promotes improvements in low-carbon performance. The baseline DID coefficient is 0.493 and is significant at the 5 % level; several alternative specifications support a positive association, but the estimates are sensitive to city-specific linear trends and over-control specifications. Therefore, the magnitude and robustness of the estimated effect should be interpreted cautiously. This result supports a positive association between AI policy exposure and urban low-carbon performance.

Second, the results suggest that the low-carbon effects associated with AI policy exposure may be transmitted partly through industrial structure upgrading and green technological innovation, rather than being interpreted solely as a direct policy effect. The stepwise mechanism tests provide indicative evidence that both IS and GI are relevant transmission channels. Therefore, policy design should focus more on the practical application of AI in energy-intensive sectors, the optimisation of production processes, and the industrialisation of green technologies.

Third, the low-carbon effects of AI are conditional, but the heterogeneity evidence should be interpreted cautiously. Subgroup regressions suggest that the estimated policy effects vary across cities with different levels of government intervention, digital infrastructure and manufacturing foundations. However, formal coefficient-difference tests do not provide consistent evidence that all subgroup coefficients are statistically different from each other. Therefore, this paper does not claim that AI policy effects are definitively stronger in a particular type of city. Instead, the results indicate that local implementation capacity, digital infrastructure and industrial application scenarios may shape the realisation of AI-enabled low-carbon effects and deserve further investigation.

Policy recommendations

Based on the above conclusions, this paper offers the following policy implications. First, a performance evaluation mechanism for low-carbon AI applications should be established by incorporating energy efficiency improvements, reductions in carbon intensity and intelligent energy-saving scenarios into AI policy assessments. Second, key AI cities such as Shanghai, Nanjing, Hangzhou and Suzhou should be encouraged to develop replicable application scenarios in industrial energy conservation, carbon monitoring and intelligent dispatching. Third, mechanisms for the transformation of green technological innovation should be improved, with attention paid not only to the number of green patents but also to patent commercialisation, actual energy savings and emission-reduction outcomes. Fourth, cities with weak digital infrastructure should address shortcomings in industrial internet, energy data collection and intelligent management platforms, so as to enhance the basic conditions for AI-enabled low-carbon transformation.

Fifth, given the heterogeneity of policy effects across different urban contexts, differentiated local implementation strategies should be adopted. Cities with stronger government intervention capacity should strengthen departmental coordination, project selection and performance assessment within their own jurisdictions, while cities with weaker industrial or digital foundations should prioritise practical AI application scenarios, energy-data collection capacity, intelligent energy-management platforms and the commercialisation of green technologies. These recommendations are based on the heterogeneity results and local implementation conditions, rather than on the spatial extension results.

Research limitations

This study is subject to several limitations. First, due to the availability and consistency of city-level data, the empirical analysis is based on a selected sample of 26 cities in the Yangtze River Delta during 2010-2023, and the findings should therefore not be automatically generalised to all cities in the broader regional integration area. Second, the DID treatment variable captures exposure to AI-related policy in selected core cities and may not fully reflect differences in the actual intensity or timing of AI adoption across cities. Third, although the model incorporates several city-level controls and conducts multiple robustness tests, unobserved factors, such as differences in local environmental governance capacity, industrial energy-management practices or other concurrent policies, may still affect low-carbon performance. In addition, this study does not use an instrumental-variable strategy to identify the independent causal effect of AI application intensity. Potential instruments such as lagged industrial robot density may be correlated with AI-related automation, but they may also affect carbon emission performance directly through production-process adjustment, energy-use patterns and equipment-structure changes. Therefore, future research should construct more refined measures of AI application intensity and seek more exogenous instruments or alternative identification strategies to further evaluate the causal relationship between AI adoption and carbon emission performance.

References

[1] Carbon Emission Accounts and Datasets (CEADs). China City-Level CO_2 Emission Inventory (1997-2019) [DB/OL]. Published 2022-10-11. Available from: https://www.ceads.net/data/carbon-inventory/city-level/.

[2] Meng F, Hu H, Sun Y, Zhang L, Hou J, Zhang Z, et al. Full-scope carbon dioxide emission dataset for Chinese cities in 2023. Scientific Data. 2025;12:1672. DOI: 10.1038/s41597-025-05949-y.

[3] National Bureau of Statistics of China and relevant ministerial statistical yearbooks. China Statistical Yearbook, China City Statistical Yearbook, China Energy Statistical Yearbook, China Fiscal Yearbook, China Education Statistical Yearbook, China Financial Yearbook and related statistical yearbooks (various years, 2011-2024) [DB/OL]. Beijing: China Statistics Press. Available from: https://www.stats.gov.cn/.

[4] International Federation of Robotics (IFR). World Robotics 2025: Industrial Robots [R/OL]. Frankfurt: International Federation of Robotics, 2025. Available from: https://ifr.org/worldrobotics/report-2025.

[5] Mischos S, Dalagdi E, Vrakas D. Intelligent energy management systems: a review. Artificial Intelligence Rev. 2023;56(10):11635-74. DOI: 10.1007/s10462-023-10441-3.

[6] Ma N, Deng Y. Synergistic governance of digital economy development and carbon reduction: Evidence from Chinese cities. Ecol Chem Eng S. 2024;31(2):203-13. DOI: 10.2478/eces-2024-0014.

[7] Zhou P, Ang BW, Poh KL. Measuring environmental performance under different environmental DEA technologies. Energy Economics. 2008;30(1):1-14. DOI: 10.1016/j.eneco.2006.05.001.

[8] Zhou P, Ang BW. Decomposition of aggregate CO_2 emissions: A production-theoretical approach. Energy Economics. 2008;30(3):1054-67. DOI: 10.1016/j.eneco.2007.10.005.

[9] Zhou P, Ang BW, Han JY. Total factor carbon emission performance: A Malmquist index analysis. Energy Economics. 2010;32(1):194-201. DOI: 10.1016/j.eneco.2009.10.003.
[10] Vinuesa R, Azizpour H, Leite I, Balaam M, Dignum V, Domisch S, et al. The role of artificial intelligence in achieving the Sustainable Development Goals. Nature Commun. 2020;11:233. DOI: 10.1038/s41467-019-14108-y.
[11] Rolnick D, Donti PL, Kaack LH, Kochanski K, Lacoste A, Sankaran K, et al. Tackling climate change with machine learning. ACM Computing Surveys. 2022;55(2):Article 42. DOI: 10.1145/3485128.
[12] Kaack LH, Donti PL, Strubell E, Kamiya G, Creutzig F, Rolnick D. Aligning artificial intelligence with climate change mitigation. Nature Climate Change. 2022;12:518-27. DOI: 10.1038/s41558-022-01377-7.
[13] Zhang K, Kou ZX, Zhu PH, Qian XY, Yang YZ. How does AI affect urban carbon emissions? Quasi-experimental evidence from China's AI innovation and development pilot zones. Economic Analysis Policy. 2025;85:426-47. DOI: 10.1016/j.eap.2024.12.013.
[14] Yu L, Wang Y, Wei X, Zeng C. Towards low-carbon development: The role of industrial robots in decarbonization in Chinese cities. J Environ Manage. 2023;330:117216. DOI: 10.1016/j.jenvman.2023.117216.
[15] Feng L, Qi J, Zheng Y. How can AI reduce carbon emissions? Insights from a quasi-natural experiment using generalized random forest. Energy Economics. 2025;141:108040. DOI: 10.1016/j.eneco.2024.108040.
[16] He H, Wu S, Wang X. Carbon reduction effect of digital transformation in Chinese steel enterprises. Environ Eng Manage J. 2025;24(1):105-15. DOI: 10.30638/eemj.2025.009.
[17] Ren B, Qiu Z, Liu B. Supply chain decarbonisation effects of artificial intelligence: Evidence from China. Int Rev Economics Finance. 2025;101:104198. DOI: 10.1016/j.iref.2025.104198.
[18] Wang L, Wang H, Cao Z, He Y, Dong Z, Wang S. Can industrial intellectualisation reduce carbon emissions? - Empirical evidence from the perspective of carbon total factor productivity in China. Technol Forecasting Social Change. 2022;184:121969. DOI: 10.1016/j.techfore.2022.121969.
[19] Chu H, Li D, Qie X, Ning Y. Research on the impact of digital economy on pollution and carbon reduction in the Yangtze River Delta urban agglomeration. Sustainability. 2026;18(8):4090. DOI: 10.3390/su18084090.
[20] Haefner N, Wincent J, Parida V, Gassmann O. Artificial intelligence and innovation management: A review, framework, and research agenda. Technol Forecasting Social Change. 2021;162:120392. DOI: 10.1016/j.techfore.2020.120392.
[21] Chen J, Yang L, Su H. Green transformation in the era of intelligence: How artificial intelligence affects disruptive green innovation in firms. Energy Economics. 2026;153:109093. DOI: 10.1016/j.eneco.2025.109093.
[22] Lee CC, Fang Y, Quan S, Li X. Leveraging the power of artificial intelligence toward the energy transition: The key role of the digital economy. Energy Economics. 2024;135:107654. DOI: 10.1016/j.eneco.2024.107654.
[23] Zhou W, Zhuang Y, Chen Y. How does artificial intelligence affect pollutant emissions by improving energy efficiency and developing green technology. Energy Economics. 2024;131:107355. DOI: 10.1016/j.eneco.2024.107355.
[24] Han Y. Digital empowerment and enterprise carbon risk management: evidence from manufacturing sector. Ecol Chem Eng S. 2024;31(2):193-202. DOI: 10.2478/eces-2024-0013.
[25] Wang Y, Zhao W, Ma X. The spatial spillover impact of artificial intelligence on energy efficiency: Empirical evidence from 278 Chinese cities. Energy. 2024;312:133497. DOI: 10.1016/j.energy.2024.133497.
[26] Xia L, Han Q, Yu S. Industrial intelligence and industrial structure change: Effect and mechanism. Int Rev Economics Finance. 2024;93:1494-506. DOI: 10.1016/j.iref.2024.04.002.
[27] Strubell E, Ganesh A, McCallum A. Energy and policy considerations for deep learning in NLP [C]. Proc 57th Annual Meeting Assoc Computational Linguistics. Florence: Assoc Computational Linguistics. 2019:3645-50. DOI: 10.18653/v1/P19-1355.
[28] Schwartz R, Dodge J, Smith NA, Etzioni O. Green AI. Commun ACM. 2020;63(12):54-63. DOI: 10.1145/3381831.
[29] Lannelongue L, Grealey J, Inouye M. Green algorithms: Quantifying the carbon footprint of computation. Adv Sci. 2021;8(12):2100707. DOI: 10.1002/advs.202100707.
[30] van Wynsberghe A. Sustainable AI: AI for sustainability and the sustainability of AI. AI Ethics. 2021;1:213-8. DOI: 10.1007/s43681-021-00043-6.
[31] De Vries A. The growing energy footprint of artificial intelligence. Joule. 2023;7(10):2191-4. DOI: 10.1016/j.joule.2023.09.004.
[32] Luccioni AS, Jernite Y, Strubell E. Power hungry processing: Watts driving the cost of AI deployment? [C]. Proc 2024 ACM Conf Fairness, Accountability Transparency. New York: Assoc Computing Machinery. 2024:85-99. DOI: 10.1145/3630106.3658542.

[33] International Energy Agency (IEA). Energy and AI [R/OL]. Paris: International Energy Agency, 2025. Available from: https://www.iea.org/reports/energy-and-ai.
[34] International Energy Agency (IEA). Key Questions on Energy and AI [R/OL]. Paris: International Energy Agency, 2026. Available from: https://www.iea.org/reports/key-questions-on-energy-and-ai.
[35] Sorrell S. Jevons' Paradox revisited: The evidence for backfire from improved energy efficiency. Energy Policy. 2009;37(4):1456-69. DOI: 10.1016/j.enpol.2008.12.003.
[36] Gillingham K, Rapson D, Wagner G. The rebound effect and energy efficiency policy. Rev Environ Economics Policy. 2016;10(1):68-88. DOI: 10.1093/reep/rev017.
[37] Lin JY, Wang X. Endowment, technology choice, and industrial upgrading. Structural Change Economic Dynamics. 2023;65:364-81. DOI: 10.1016/j.strueco.2023.03.002.
[38] Zou W, Xiong Y. Does artificial intelligence promote industrial upgrading? Evidence from China. Economic Res - Ekonomska Istraživanja. 2023;36(1):1666-87. DOI: 10.1080/1331677X.2022.2092168.
[39] Chen H, Liang M, Wang J, Shi B. Environmental regulation, R&D investment and industrial transformation and upgrading. Environ Eng Manage J. 2025;24(3):449-60. DOI: 10.30638/eemj.2025.036.
[40] Dabuo FT, Xu L, Madzikanda B, Appiah-Twum F. Analyzing the impact of eco-innovation and environmental regulation on green development efficiency. Environ Eng Manage J. 2024;23(11):2247-66. DOI: 10.30638/eemj.2024.181.
[41] Yang Y. Green development synergies and regional disparities in the Yangtze River Delta urban agglomeration. Ecol Chem Eng S. 2025;32(2):189-212. DOI: 10.2478/eces-2025-0010.
[42] Zhao Y, Qiao Y, Zhang L, Fan Q. Green innovation spillover vs. competition in Chinese urban agglomerations: A comparative study of green total factor productivity before and after the 2015 Environmental Protection Law revision. Environ Eng Manage J. 2024;23(12):2593-605. DOI: 10.30638/eemj.2024.208.
[43] Zhao Y. Carbon emission reduction effects of heterogeneous environmental regulation: evidence from the firm level. Ecol Chem Eng S. 2024;31(2):243-52. DOI: 10.2478/eces-2024-0017.
[44] Sun L, Abraham S. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. J Econometrics. 2021;225(2):175-99. DOI: 10.1016/j.jeconom.2020.09.006.
[45] State Council of the People's Republic of China. New Generation Artificial Intelligence Development Plan [EB/OL]. State Council Document No. 35 [2017], 2017-07-20. Available from: https://app.www.gov.cn/govdata/gov/201707/20/408540/article.html.
[46] China Research Data Services Platform (CNRDS). Green Patent Research Database (GPRD) [DB/OL]. Database accessed in 2026. Available from: https://www.cnrds.com/ (accessed 12 August 2026); China National Intellectual Property Administration (CNIPA). Patent Search and Analysis System [DB/OL]. 2023. Available from: https://pss-system.cponline.cnipa.gov.cn.
[47] Callaway B, Sant'Anna PHC. Difference-in-Differences with multiple time periods. J Econometrics. 2021;225(2):200-30. DOI: 10.1016/j.jeconom.2020.12.001.
[48] Borusyak K, Jaravel X, Spiess J. Revisiting event-study designs: Robust and efficient estimation. Rev Economic Stud. 2024;91(6):3253-85. DOI: 10.1093/restud/rdae007.
[49] Acharya A, Blackwell M, Sen M. Explaining causal findings without bias: Detecting and assessing direct effects. Amer Political Sci Rev. 2016;110(3):512-29. DOI: 10.1017/S0003055416000216.
[50] Gelbach JB. When do covariates matter? And which ones, and how much? J Labor Economics. 2016;34(2):509-43. DOI: 10.1086/683668.
[51] Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. Psychological Methods. 2010;15(4):309-34. DOI: 10.1037/a0020761.
[52] Conley TG, Taber CR. Inference with "Difference in Differences" with a small number of policy changes. Rev Economics Statistics. 2011;93(1):113-25. DOI: 10.1162/REST_a_00049.
[53] Ferman B, Pinto C. Inference in Differences-in-Differences with few treated groups and heteroskedasticity. Rev Economics Statistics. 2019;101(3):452-67. DOI: 10.1162/rest_a_00759.
[54] MacKinnon JG, Webb MD. The wild bootstrap for few (treated) clusters. Econometrics J. 2018;21(2):114-35. DOI: 10.1111/ectj.12107.
[55] Roth J. Pretest with caution: Event-study estimates after testing for parallel trends. Amer Economic Rev: Insights. 2022;4(3):305-22. DOI: 10.1257/aeri.20210236.
[56] Rambachan A, Roth J. A more credible approach to parallel trends. Rev Economic Stud. 2023;90(5):2555-91. DOI: 10.1093/restud/rdad018.