

A multi-agent evolutionary game approach to green collaborative strategies in port consolidation and distribution systems

Qi Xu¹, Caixin Li¹, Hailun Deng¹, Tianwei Ma¹ and Song Wu^{2*}

¹ School of Architecture and Transportation Engineering, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China

² Guangxi Laboratory of Modern Canal, Pinglu Canal Group Co., Ltd, Nanning, Guangxi 530012, China

* Correspondence: npcplem372125562@163.com (Wu S)

Abstract

This paper addresses the structural imbalances that hinder the green transition of port consolidation and distribution systems. To tackle this issue, we construct a tripartite evolutionary game model that is integrated with multi-agent reinforcement learning. By mapping game-theoretic logic onto spatial networks, the model reveals the mechanisms of strategy propagation and emergence patterns. Based on the proposed model, key findings are as follows: (1) The three agent types exhibit distinct policy sensitivities: ports show diminishing marginal effects to regulatory intensity, river-sea operators are cost-sensitive, and road operators prioritize reputation. (2) Carbon taxes force road transport to lead transformation by raising environmental costs, while subsidies hedge initial risks for intermodal operators. (3) Collaborative strategies propagate spatially following 'local nucleation cluster expansion global emergence'. (4) Experience-driven agents play a pioneering role initially, while adaptive agents form experience lock-in through dynamic adjustment, enhancing system robustness. This study reveals the intrinsic mechanisms by which structural transformation drives structural emission reduction, providing scientific support for collaborative governance under the 'dual carbon' goals.

Keywords: Port consolidation and distribution system, Multi-party coordination, Evolutionary game, Numerical simulation, Multi-agent

Citation: Xu Q, Li C, Deng H, Ma T, Wu S. 2026. A multi-agent evolutionary game approach to green collaborative strategies in port consolidation and distribution systems. *Digital Transportation and Safety* 5(3): 305–317 <https://doi.org/10.48130/dts-0026-0024>

Introduction

Global climate change and environmental degradation are becoming increasingly severe. In particular, excessive emissions of greenhouse gases, such as carbon dioxide, have posed grave challenges to the sustainable development of humanity. Following the signing of the Paris Agreement, addressing climate change and controlling greenhouse gas emissions have become central concerns of the international community. As the world's largest developing country, China announced its 'dual carbon' strategic goals in 2020, pledging to peak carbon emissions by 2030 and achieve carbon neutrality by 2060. These goals not only reflect China's commitment to fulfilling its responsibilities as a major nation and advancing the building of a community with a shared future for mankind, but also serve as a crucial driver for deep-seated industrial restructuring and the green transformation of its energy systems. Under strong policy guidance from the central government, the transportation sector, as the lifeblood of the national economy, is witnessing its green transition shift from a singular focus on energy efficiency improvement toward structural optimization, thereby positioning it as a key breakthrough in fulfilling the carbon peaking and neutrality mandates.

In today's complex modern transportation and logistics networks, the port consolidation and distribution system serves as a pivotal hub connecting global supply chains with inland hinterlands. Its operational efficiency not only directly affects port throughput capacity and overall competitiveness, but also determines the energy consumption level of the entire regional logistics chain.

However, under the traditional development model, the port consolidation and distribution system suffers from a severe structural imbalance. Statistics show that road transport handles 62.7% of

container throughput at China's major coastal ports, while railway and waterway account for only 3.0% and 34.3%, respectively. This road-dominated modal mix is significantly higher than the approximately 50% road share observed in European and American ports. This excessive reliance on road transport has led to traffic congestion, increased carbon emissions, and severe environmental pollution, locking the system into a long-standing state of structural imbalance dominated by highways.

Although river-sea intermodal transport exhibits considerable green potential due to its advantages of low energy consumption and high volume capacity, its advantages have not been fully realised in actual operations. Poor coordination among port operators, waterway carriers, and road transport providers, coupled with the absence of effective interest coordination mechanisms, has kept overall system emissions at a high level.

Optimizing the consolidation and distribution system under the 'dual carbon' goals is, in essence, a process of promoting structural emission reduction through multi-party collaborative gaming. This requires port operators, river-sea intermodal operators, and road transport operators to move beyond isolated decision-making aimed at maximizing individual profits under the logic of green transformation. Driven jointly by carbon tax pressure, revenue compensation, and policy incentives, the system should evolve from the traditional structurally imbalanced state toward a green, collaborative structure. In the process of multi-party collaborative development of port consolidation and distribution systems, the main stakeholders involved include ports, river-sea intermodal transport operators, and road transport operators. The relationships among the parties in the tripartite evolutionary game are shown in Fig. 1.

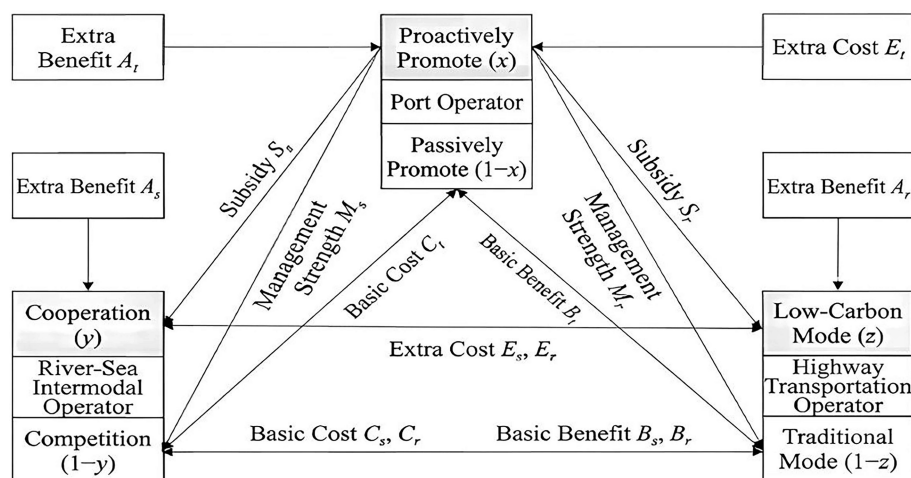


Fig. 1 Logical diagram of the tripartite evolutionary game model.

Literature review

Operational efficiency and network optimization

Early research focused on resource allocation and operational efficiency in port-hinterland networks. Shen & Khoong pioneered system network optimization for empty container repositioning^[1]. Janic developed an integrated cost model comparing transport modes in European freight systems^[2]. Recent studies have refined these approaches: He et al. optimized external truck task scheduling at automated terminals^[3]; Guo et al. advanced carbon emission estimation for hinterland container intermodal networks^[4]; and Xu et al. addressed tractor-trailer scheduling under disrupted events^[5].

Multimodal transport planning under uncertainty

A second stream addresses route planning and network design under uncertainty. Li et al. constructed a multi-objective fuzzy non-linear model for intermodal route selection under time-window constraints^[6]. Xin et al. developed a bi-level model for liner alliance shipping network design with slot allocation and empty container repositioning^[7]. Yan et al. proposed an optimization approach for port integration through capacity adjustments^[8]. Yin et al. examined collaborative effects of rail subsidies and carbon trading^[9]. Hu et al. investigated low-carbon multimodal route optimization from a government-enterprise collaborative perspective^[10]. Wang et al. analyzed port pricing strategies under intra-port coordination and inter-port competition^[11].

Stakeholder interactions and game-theoretic models

Recognizing the limitations above, researchers have turned to game-theoretic approaches. Zhu et al. analyzed shipping company strategies under cap-and-trade mechanisms^[12]. Jiang et al. constructed a marine fuel supply chain model examining green fuel promotion^[13]. Sheng et al. developed a framework incorporating regional port competition dynamics^[14]. Xu et al. employed a Stackelberg game to examine the trade-off between government subsidies and port competition in shore power construction^[15].

Recent tripartite evolutionary game studies are particularly relevant. Xu et al. constructed a three-party model including port operators, river-sea intermodal operators, and road transport operators under the 'dual carbon' goal. Their finding—that intermodal

operators are more cost-sensitive while road operators depend more on policy-derived benefits—aligns closely with our simulation results^[16]. Chen et al. examined tripartite interactions among governments, ports, and shipping companies under carbon trading, demonstrating that subsidies and fines play complementary roles^[17]. Zhang et al. employed tripartite evolutionary game theory with system dynamics to analyze shore power promotion under carbon tax and subsidy policies^[18].

Multi-agent reinforcement learning in transportation

MARL has emerged as a complementary paradigm for modeling adaptive decision-making. Recent applications include: a MARL framework for port recovery investment decisions under extreme weather events demonstrating superior performance over state-dependent evolutionary game benchmarks, an adaptive MARL large model for logistics-energy coordination at container seaports, and a MARL approach for joint block assignment and yard crane redeployment at intermodal terminals. Li et al. systematically examined MARL's evolutionary trajectory across value function decomposition, policy gradient, and actor-critic architectures^[19]. Surveys have explored MARL applications in traffic signal control, UAV navigation, air traffic management, and maritime autonomous navigation.

Research gaps and our contribution

The existing literature exhibits three interconnected gaps: First, centralized optimization models overlook the decentralized, self-interested nature of stakeholder decision-making. Second, game-theoretic models employ static structures that fail to capture dynamic learning and bounded rationality. Third, the spatial dimension of strategy propagation—how collaborative behaviors emerge from local interactions and diffuse across agent networks—remains unaddressed. Fourth, the integration of MARL with evolutionary game theory in the port collection-distribution context is under-explored.

To bridge these gaps, we construct a co-evolutionary model integrating evolutionary game theory with MARL. Unlike existing approaches, our model: (1) captures the bounded rationality of three heterogeneous agent types through dynamic Q-learning with satisfaction-based adjustment; (2) incorporates a spatial grid structure, revealing how local collaboration clusters expand into global patterns; (3) embeds carbon tax and subsidy policies into the reward

function to quantify their interactive effects; and (4) distinguishes between experience-driven and adaptive learning agents to capture real-world heterogeneity in decision-making styles.

Overview of multi-agent systems and reinforcement learning

From a micro perspective, an agent is a computational entity with independent decision-making capabilities. To precisely describe its behavioural logic in complex environments, an agent is usually formally defined as a quadruple:

$$Agent_i = \mathcal{P}_i, \mathcal{S}_i, \mathcal{A}_i, \mathcal{D}_i. \quad (1)$$

The perception function $\mathcal{P}_i$ represents the agent's observational mapping of the external environment state. In games with incomplete information, the agent can often only obtain local observations, i.e., partial observations^[20].

The internal state $\mathcal{S}_i$ refers to the experience, historical payoffs, and current policy inclination stored by the agent, serving as the memory foundation for its 'bounded rationality' learning.

The action space $\mathcal{A}_i$ is the set of all decision alternatives executable by the agent. In the port consolidation and distribution scenario studied in this paper, the action space corresponds to the collaborative strategy choices of each agent.

The decision logic $\mathcal{D}_i$, i.e., the policy function $\pi_i: \mathcal{S}_i \times \mathcal{O} \rightarrow \mathcal{A}_i$, describes how the agent maps its current perception and internal state to the optimal action.

When multiple agents are placed in the same dynamic environment, the core of system evolution shifts to the interactions among agents. The essential characteristic of a multi-agent system lies in the coupling of its joint action space. There are n agents in the system, and their joint action is denoted as $a = (a_1, a_2, \dots, a_n) \in \mathcal{A}$.

Under this framework, the payoff function of any agent is no longer determined solely by its own decision a_i , but is jointly influenced by the joint action of all agents:

$$R_i = f_i(s, a_i, a_{-i}), \quad (2)$$

where, a_{-i} denotes the combination of actions taken by all other agents except agent i . This mutual interdependence constitutes the underlying mathematical logic of multi-player games. Under the constraints of the 'dual carbon' goals, due to the non-alignment of interests among the various agents, the system often faces the challenge of 'non-stationarity', i.e., the strategy optimisation of one agent alters the environmental feedback received by the others^[21].

The rigorous mathematical foundation underlying reinforcement learning is the Markov decision process (MDP), whose core logic is to enable an agent to maximise cumulative rewards through continuous trial and error in an unknown environment. The collaborative process of multiple agents in port consolidation and distribution can be abstracted as a Markov decision process composed of a five-tuple $\mathcal{M} = \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma$:

State space $\mathcal{S}$: encompasses all environmental variables that influence collaborative decision-making, such as the carbon emission strategies around each operator and the current policy parameters.

Action space $\mathcal{A}$: corresponds to the strategy sets of the gaming agents; for example, the policy promotion strategies of port operators, the operational modes, and transport mode choices of operators, etc.

State transition probability $\mathcal{P}$: $\mathcal{P}(s'|s, a) = \mathbb{P}[S_{t+1} = s' | S_t = s, A_t = a]$, denoting the likelihood that the system transitions to the next state after executing a specific action in the current state^[22].

Reward function $\mathcal{R}$: the immediate feedback obtained by the agent from the environment after performing an action. In our model, the reward function integrates dual objectives: economic payoff and payoff satisfaction.

Discount factor $\gamma: \gamma \in [0, 1]$, reflecting the degree of importance the agent attaches to long term collaborative benefits in the future^[23].

The ultimate goal of reinforcement learning is to find the optimal policy $\pi^*(a|s)$, so as to maximise the expected cumulative discounted return G_t from the current time step onward:

$$G_t = \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k} \right] \quad (3)$$

To evaluate the quality of a policy, the state-action value function $Q^\pi(s, a)$ is introduced. According to the principle of dynamic programming, this function can be expanded into a recursive form, namely the Bellman equation:

$$Q^\pi(s, a) = \mathbb{E}_\pi [r_t + \gamma Q^\pi(s_{t+1}, a_{t+1}) | S_t = s, A_t = a] \quad (4)$$

This equation reveals the dynamic relationship between the decision value at the current moment and the expected value of future states, providing a mathematical basis for agents to allocate strategic weights during the gaming process.

The Q-learning algorithm was proposed by Watkins in 1989 and is a value-iteration-based reinforcement learning method. Its main characteristics lie in its two attributes: 'model-free' and 'off-policy'. In simple terms, the agent does not need to know the state transition probabilities of the environment in advance; instead, by continuously interacting with the environment and recording the expected return Q-value for each 'state-action' pair, it can gradually search for the optimal policy. In the context of multi-party gaming in port consolidation and distribution, Q-learning transforms the bounded rational behaviours of participating agents into an incremental update process of Q-values. The Q-value can be understood as the expected long-term discounted return that an agent can obtain by adopting a certain depth of collaboration under a specific policy environment. As the Q-table is continuously updated, the agent can gradually find the optimal behavioural boundary under the constraints of the 'dual carbon' goals.

The mathematical logic of Q-learning is built upon the Bellman equation. At each time step, the agent is in state s_t , executes action a_t , receives an immediate reward r_t from the environment, and transitions to the next state s_{t+1} . The goal of the algorithm is to solve for the optimal action value function $Q^*(s, a)$ such that it satisfies the Bellman optimality equation:

$$Q^*(s, a) = \mathbb{E} [r_t + \gamma \max_{a'} Q^*(s_{t+1}, a') | S_t = s, A_t = a] \quad (5)$$

In actual simulation evolution, the agent employs the temporal difference method to perform incremental updates on the Q-values. The update logic is as follows:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)]. \quad (6)$$

Learning rate $\alpha \in (0, 1)$: characterizes the rate at which the agent assimilates newly observed rewards. A higher α value indicates that the agent tends to adjust its strategy more aggressively based on the latest feedback, whereas a lower α value contributes to the stability of the system in the later stages of the game^[24].

Discount factor $\gamma \in [0, 1]$: reflects the degree of importance the agent attaches to long-term collaborative benefits in the future, serving as a key parameter for measuring the 'foresight' of the agent's decision-making.

TD error $[r_t + \gamma \max_{a'} Q - Q]$: represents the deviation of the sum of the immediate reward and future expectation from the current estimate. When the TD error gradually approaches zero, it implies that

the strategies of all agents have reached an evolutionarily stable state^[8].

To prevent the system from becoming trapped in local optima during the simulation process, Q-learning introduces the ϵ -greedy strategy to balance exploration and exploitation. When making decisions, the agent selects the 'optimal' action with the highest current Q-value with probability $1 - \epsilon$, and performs random exploration within the action space with probability ϵ ^[25]:

$$\pi(a|s) = \begin{cases} \text{random action} & \text{with probability } \epsilon \\ \operatorname{argmax}_{a \in A} Q(s, a) & \text{with probability } 1 - \epsilon \end{cases} \quad (7)$$

This mechanism simulates the trial-and-error behavior of ports and carriers when confronting carbon reduction policies in reality: while retaining past successful collaborative experience, it also maintains the possibility of attempting new strategic models^[20].

Construction of an evolutionary model for port collection and distribution based on multi-agent reinforcement learning

Spatial evolution of the environment for port consolidation and distribution

This study abstracts the complex port collection-distribution ecosystem consisting of two-dimensional grids G . Let $G = \{(x, y) | 1 \leq x, y \leq L\}$, where $L = 30$ denotes the linear scale of the grid; the total system size is $N = L \times L = 900$, and each node on the grid is occupied by an agent i . Each agent i is assigned a fixed group type $(i) \in \{\text{port operator, river-sea intermodal operator, and road transport operator}\}$, and the three types of agents are uniformly distributed across the grid, each accounting for $N/3$.

To eliminate the 'edge effects' arising from finite-sized grids, this model introduces periodic boundary conditions. Mathematically, for any node (x, y) with coordinates, the coordinate transformation logic follows modular arithmetic rules:

$$x' = ((x - 1) \bmod L) + 1 \quad (8)$$

$$y' = ((y - 1) \bmod L) + 1 \quad (9)$$

Define the agent type function $f(x, y) \in \{P, S, R\}$, where the three values represent ports, river-sea intermodal operators, and road transport operators, respectively. Agents are periodically populated according to the following rules:

$$f(x, y) = \begin{cases} P, & \text{if } (x + y) \bmod 3 = 0 \\ S, & \text{if } (x + y) \bmod 3 = 1 \\ R, & \text{if } (x + y) \bmod 3 = 2 \end{cases} \quad (10)$$

Multi-agent decision definition and reward function design

Under the multi-agent reinforcement learning framework, the strategy in traditional game theory is strictly defined as the action of an agent. For any heterogeneous agent in the system, its discrete action space at time t is defined as $\mathcal{A}_i$:

$$\mathcal{A}_i = \{a_0, a_1\}. \quad (11)$$

The cooperative emission reduction action a_1 represents the agent's adoption of green equipment retrofitting and optimised scheduling, corresponding to the strategies of 'proactive promotion', 'cooperation', and 'low-carbon mode' in the previous game-theoretic setting.

The traditional maintenance action a_0 represents the agent's retention of the existing mode of traditional operations and high-carbon emissions, corresponding to the strategies of 'passive promotion', 'competition', and 'traditional mode' in the previous game-theoretic setting.

At each decision step t , all agents select actions simultaneously, forming the joint action set of the system $a' = (a'_1, a'_2, \dots, a'_N)$.

Distinct from perfect-information games, the agents in this model are characterised by bounded rationality and local observation capabilities. The state of agent i at t time $s'_i \in S_i$ is not the global system state, but is jointly determined by the action choices of heterogeneous neighbors Ω_i within its local topological neighbourhood^[26].

In the regular grid setting of this paper, any agent is adjacent to two types of heterogeneous agents; for example, a port agent has two river-sea intermodal neighbours and two road transport neighbours. Define the state variable as the number of neighbours that choose the collaborative action. Taking a port agent as an example, its state space can be expressed as

$$s'_p = (n'_c, n'_L), \quad (12)$$

where, $n'_c \in \{0, 1, 2\}$ denotes the number of river-sea intermodal operators choosing collaboration within the local neighborhood, and $n'_L \in \{0, 1, 2\}$ denotes the number of road transport operators choosing collaboration. It follows that the size of an individual agent's state space is $|S_i| = 3 \times 3 = 9$ discrete state combinations.

In the traditional scenario, the carbon tax is applied exclusively to road transport operators, while port operators and river-sea intermodal operators are not subject to this cost. This design reflects the current policy reality in China: road freight, as the largest and fastest-growing source of carbon emissions in the logistics sector, has been the primary target of environmental regulations such as fuel taxes and emission trading schemes. In contrast, ports and inland waterway transport, while not yet subject to direct carbon taxation in most regions, operate under different regulatory logics—ports face indirect pressures through shore power mandates and emission control area requirements, while waterway transport inherently possesses a lower carbon footprint per ton-km. This asymmetry is therefore not an oversight but a deliberate modeling choice that reflects the real-world regulatory hierarchy. Importantly, our counterfactual scenarios systematically explore the effects of extending carbon taxes to all agent types, thereby providing policy insights into the potential consequences of broader carbon pricing schemes.

Simulation model parameter settings

In conducting the multi-agent spatiotemporal evolution simulation, in order to more clearly observe the driving effects of the 'dual carbon' policies on system collaboration, this chapter integrates the underlying payoff logic of the evolutionary game as the foundational environmental support for agent learning, while focusing on two core external variables—carbon tax and policy subsidies—to examine the system transformation patterns under specific policy combinations. The reward function design here is primarily intended to test the pathways through which external policy instruments influence individual decision-making^[23]. By introducing carbon tax constraints and guiding subsidies, the simulation models the process by which policy intervention adjusts the payoff expectations of game participants, further exploring whether the system can effectively overcome evolutionary resistance and achieve green synergy under the influence of policy combinations. To quantitatively assess the specific impacts of policy interventions on tripartite

collaboration, the symbols and definitions involved in the multi-agent model of this chapter are presented in Table 1.

Agent payoff matrix under spatial gaming

In the spatial evolutionary game, the immediate reward of an agent depends not only on its own strategy but is also influenced by the degree of strategic alignment with neighbouring agents. To accurately characterise the constraints and incentives within the port collection distribution chain under the 'dual carbon' goals, this paper constructs a payoff matrix that incorporates strategy-coupling losses. The payoff functions for each agent under different strategy combinations are defined as follows:

Taking a port P agent as an example, when it engages in a single business interaction with one randomly drawn river-sea intermodal agent (whose action is denoted as $a_S \in \mathcal{A}$) and one road transport agent (whose action is denoted as $a_R \in \mathcal{A}$) within its local neighbourhood, its absolute payoff function π_P exhibits a strict three-stage characteristic. Based on this, the payoff functions of the three agents are expressed as follows:

Payoff function of the port operator (P) in the game with neighbouring agents:

$$\pi_P = \begin{cases} R_P + R_c - C_P + G_s & \text{if } a_P = a_1 \text{ and } a_P = a_1, a_R = a_1 \text{ (Collaborative)} \\ R_P - C_P - L_P + G_s & \text{if } a_P = a_1 \text{ and } (a_P, a_R) \neq (a_1, a_1) \text{ (Mixed)} \\ R_P - L_P & \text{if } a_P = a_0 \text{ (Traditional)} \end{cases} \quad (13)$$

Payoff function of the river-sea intermodal operator (S) in the game with neighbouring agents:

$$\pi_S = \begin{cases} R_S + R_c - C_S + G_s & \text{if } a_S = a_1 \text{ and } a_P = a_1, a_R = a_1 \text{ (Collaborative)} \\ R_S - C_S - L_S + G_s & \text{if } a_S = a_1 \text{ and } (a_P, a_R) \neq (a_1, a_1) \text{ (Mixed)} \\ R_S - L_S & \text{if } a_S = a_0 \text{ (Traditional)} \end{cases} \quad (14)$$

Payoff function of the road transport operator (R) in the game with neighbouring agents:

$$\pi_R = \begin{cases} R_R + R_c - C_R + G_s & \text{if } a_R = a_1 \text{ and } a_P = a_1, a_S = a_1 \text{ (Collaborative)} \\ R_R - C_R - L_R + G_s & \text{if } a_R = a_1 \text{ and } (a_P, a_S) \neq (a_1, a_1) \text{ (Mixed)} \\ R_R - L_R - T_c & \text{if } a_R = a_0 \text{ (Traditional)} \end{cases} \quad (15)$$

State-aware mapping of expected rewards

During the actual iterative process in the spatial grid, agent P cannot predict the specific actions of the single interaction partner, but it can perceive its current local environmental state $s_P = (n_C, n_L)$. Let the collaboration ratio of river-sea intermodal agents within its neighborhood be $x = n_S / N_S^{\text{total}}$, and that of road transport agents be

$y = n_R / N_R^{\text{total}}$. When the agent chooses the collaborative action a_1 , the joint probability that it matches the 'full collaboration' state is $\rho = x \cdot y$. Based on this, by expanding the absolute payoff rule over the probability space in terms of expectation, the single-step expected reward function assigned to agent by the environment is obtained as follows $\mathcal{R}_P(s_P, a)$:

When the agent chooses the 'collaborative emission reduction' action ($a_P = a_1$), the expected reward is:

$$\mathcal{R}_P(s_P, a_1) = \rho \cdot (R_P + R_c - C_P + G_s) + (1 - \rho) \cdot (R_P - C_P - L_P + G_s). \quad (16)$$

As the number of green neighbours in the local environment increases (the state s improves, and ρ increases, the expected reward for choosing the collaborative action will also increase linearly^[27].

When the agent chooses the 'maintain traditional' action ($a_P = a_0$), the expected reward is:

$$\mathcal{R}_P(s_P, a_0) = \rho \cdot (R_P + R_c - C_P + G_s) + (1 - \rho) \cdot (R_P - C_P - L_P + G_s). \quad (17)$$

Maintaining the traditional mode does not yield any collaborative premium or subsidy, and only provides the basic payoff after deducting chain-link depreciation. This expected reward, jointly determined by state s and action a , serves as the core driving signal $\mathcal{R}(s, a)$, which is directly substituted into the subsequent temporal difference equation for updating the agent's Q-value matrix r_{t+1} .

Definition of the reward function incorporating neighbourhood information sharing

During a single iteration, agent i locks in its current action a_i^t and forms pairwise combinations with all heterogeneous neighbors in its neighborhood Ω_S (e.g., the river-sea intermodal set) and Ω_R (e.g., the road transport set), traversing all possible triadic interactions. The actual empirical payoff Π_i of the agent at t time is taken as the average value over all interaction combinations:

$$\Pi_i = \frac{1}{|\Omega_S| |\Omega_R|} \sum_{j \in \Omega_S} \sum_{k \in \Omega_R} \pi_i(a_i, a_j, a_k), \quad (18)$$

where $\pi_i(a_i, a_j, a_k)$ is the tripartite game payoff rule, representing the payoff obtained by agent i when interacting with neighbors j and k , Ω_S , Ω_R are the sets of the two types of heterogeneous neighbours within the neighbourhood. Under the grid topology of this paper, $|\Omega_S| = 2$ and $|\Omega_R| = 2$ represent the total number of possible triadic combinations within the neighbourhood. When the number of grid interactions is sufficiently large, the actual empirical payoff Π_i will numerically approach the theoretical expected reward $\mathcal{R}$. This summation mechanism reflects the dynamic characteristics of spatial games, namely that an individual's payoff is highly dependent on the strategy distribution in its local topological environment^[28].

Table 1. Notations and definitions for the multi-agent model.

Parameter	Definition	Description and practical significance
C_P	Additional costs of port green transformation	Investments made by ports in green hub retrofitting and low-carbon facility construction to enhance collaborative efficiency
C_S	Additional costs of river-sea intermodal transformation	Additional collaborative costs incurred by intermodal operators in information system integration and multimodal transport organization
C_R	Additional costs of road transport transformation	Equipment renewal costs incurred during the transition of road transport from conventional diesel-powered trucks to green energy vehicles
G_s	Guiding policy subsidies	Phased incentives provided by the government to encourage agents to choose green transformation, aimed at offsetting transition costs
T_c	Carbon tax constraint costs	Punitive environmental taxes levied on the road transport sector that persists in traditional high-emission modes, increasing evolutionary resistance
R_i	Baseline operating revenue	Benchmark market revenue for the three types of agents ($i \in \{P, S, R\}$) when maintaining the <i>status quo</i> without adopting collaborative transformation
R_c	Green collaborative incremental revenue	Incremental profits generated by system structural optimisation and efficiency improvement when the three agents reach a collaborative consensus
L_i	Collaborative failure losses	Efficiency losses or congestion costs resulting from the lack of coordination within the collection-distribution system ($i \in \{P, S, R\}$)

In the reinforcement learning framework, the reward function is the core feedback signal that drives the evolution of agents' strategies. To simulate the learning effects and information transmission among agents in the port collection–distribution industrial chain, this study constructs a reward function that incorporates neighbourhood information sharing:

$$r_t^i = (1 - \alpha_r^i) \Pi_i + \alpha_r^i \bar{\Pi}_{N_i}, \tag{19}$$

where, $\bar{\Pi}_{N_i}$ represents the average payoff of all neighbours within the neighbourhood of agent and its calculation formula is:

$$\bar{\Pi}_{N_i} = \frac{1}{|\Omega_i|} \sum_{j \in \Omega_i} \Pi_{N_j}.$$

Ω_i is the neighborhood set of agent i , with $|\Omega_i| = 4$, and Π_{N_j} is the game payoff obtained by neighbour agent n at the current time^[29].

The parameter $\alpha_r^i \in [0, 1]$ is the key variable that distinguishes the different decision-making modes in this model. It reflects the degree of dependence of the agent on neighbourhood social experience when updating its strategy value Q . This design couples the individual profit objective with the system-wide collaborative objective, providing an underlying interest-driven logic for the formation of collaborative clusters. The value of α_r^i determines the orientation of the reward signal: when α_r^i approaches 0, the agent behaves as fully 'experience confident', adjusting its strategy solely based on its own gains and losses; when α_r^i approaches 1, the agent exhibits extreme 'herd mentality', with its decision-making motivation primarily influenced by the overall payoff level of its neighbourhood^[30].

Simulation analysis of collaborative strategy evolution in port collection distribution systems

The two modes are distinguished by parameter α_r^i in Eq. (19). Experience-driven agents are assigned a fixed $\alpha_r^i \approx 0$, prioritizing their own payoffs over neighborhood information—mapping to

large state-owned port groups with rigid decision-making. Adaptive agents receive a dynamically adjusted α_r^i , decreasing with payoff satisfaction, initially relying on neighbors but shifting to their own experience over time, mapping to private operators with shorter decision chains. Assignment is predetermined by organizational characteristics with a 1:1 baseline ratio, reflecting the typical mix of incumbents and flexible entrants in real-world networks.

Based on the stability analysis of the evolutionary game, this section employs MATLAB to simulate and verify the evolutionary characteristics of the model under various scenarios. The aim is to analyze the dynamic interactions among port operators, river-sea intermodal operators, and road transport operators, and their sensitivity to key variables, thereby identifying the critical parameter ranges that enable the three parties to achieve a stable collaborative equilibrium.

For the equilibrium point (1, 0, 0), when the three conditions $E_t - M_s - M_r - G - L < 0$, $M_s + S_s - E_s < 0$ and $M_r + S_r - E_r < 0$ are satisfied, the system exhibits a stable evolutionary strategy. The corresponding simulation parameters are assigned according to Scenario 1 in Table 2.

The preceding numerical simulations of the three typical equilibrium scenarios indicate that the collaborative evolution of the port collection distribution system exhibits pronounced parameter dependence and asymmetric sensitivity. The simulation system is initialised in a strategy-disordered state, with each agent randomly selecting an initial strategy with equal probability of 0.5. This setting is intended to mimic the chaos of market information and the discretisation of decision-making in the early stages of policy implementation, thereby allowing observation of how the system spontaneously evolves from self-interested disordered gaming into an orderly collaborative structure.

From the steady-state Q -values presented in Tables 3–6, it is evident that the micro-level decision-making logic is highly consistent with the macro-level evolutionary trends. Under the steady state (2, 2) in Scenario 1* (low subsidy, low carbon tax) and Scenario

Table 2. Simulation parameter assignments.

Parameter category	Symbol	Definition	Scenario 1	Scenario 2	Scenario 3
Variable parameters	C_P	Additional costs of port green transformation	15	20	15
	C_S	Additional costs of river-sea intermodal transformation	60	25	40
	C_R	Additional costs of road transport transformation	60	25	40
	T_c	Carbon tax constraint costs	15	20	15
	G_s	Guiding policy subsidies	4	25	40
	L_i	Collaborative failure losses	20	15	60
Fixed parameters	R_c	Green collaborative incremental revenue	55	55	55
	R_i	Baseline operating revenue	80	80	80

Table 3. Steady-state Q -values under the low subsidy, low carbon tax scenario.

Environmental state s_t^i	Port (C, T)		River–Sea intermodal (C, T)		Road transport (C, T)	
(0,0)	135	85	138	31	124	56
(0,1)	107	181	119	249	146	68
(0,2)	304	126	244	257	76	168
(1,0)	116	153	35	129	147	91
(1,1)	302	164	270	152	326	131
(1,2)	307	276	274	215	327	135
(2,0)	17	260	0	161	45	259
(2,1)	275	288	276	188	329	275
(2,2)	308	288	276	259	332	300

The bold values are the steady Q -values.

Table 4. Steady-state Q-values under the low subsidy, high carbon tax scenario.

Environmental state s_i^t	Port (C, T)		River–sea intermodal (C, T)		Road transport (C, T)	
(0,0)	111	13	123	67	126	8
(0,1)	64	219	168	122	51	182
(0,2)	102	240	103	218	61	206
(1,0)	232	121	82	143	307	107
(1,1)	311	220	151	239	333	153
(1,2)	334	269	303	219	334	186
(2,0)	215	121	177	8	259	178
(2,1)	334	228	119	290	332	217
(2,2)	337	321	306	296	341	308

The bold values are the steady Q-values.

Table 5. Steady-state Q-values under the high subsidy, low carbon tax scenario.

Environmental state s_i^t	Port (C, T)		River–sea intermodal (C, T)		Road transport (C, T)	
(0,0)	4	0	3	0	266	58
(0,1)	0	213	310	27	135	188
(0,2)	59	552	335	131	71	320
(1,0)	332	275	0	74	466	160
(1,1)	606	119	73	308	605	214
(1,2)	608	556	336	307	620	451
(2,0)	0	352	79	0	615	211
(2,1)	384	557	334	256	621	492
(2,2)	621	563	336	317	621	537

The bold values are the steady Q-values.

Table 6. Steady-state Q-values under the high subsidy, high carbon tax scenario.

Environmental state s_i^t	Port (C, T)		River–sea intermodal (C, T)		Road transport (C, T)	
(0,0)	369	96	53	336	0	203
(0,1)	71	390	361	206	0	322
(0,2)	568	0	0	270	60	421
(1,0)	279	387	354	156	148	358
(1,1)	596	349	369	281	317	568
(1,2)	532	565	375	303	621	418
(2,0)	584	328	190	328	374	574
(2,1)	613	557	370	345	626	576
(2,2)	618	575	377	371	633	588

The bold values are the steady Q-values.

4* (high subsidy, high carbon tax), the expected payoffs of all three types of agents choosing the collaborative strategy are significantly higher than those of the traditional strategy. This micro-level finding explains the fundamental reason why, in the subsequent simulations, the cooperation rate can ultimately break through random fluctuations and converge robustly to 1.

To comprehensively analyse the joint driving effects of carbon taxes and government subsidies on the green collaborative transformation of the port collection distribution system, this section sets up four typical policy combination scenarios and observes the evolutionary trajectories of the system during the dynamic gaming process. As shown in Fig. 2, the simulation results under the four scenarios are arranged in a 2×2 matrix format, intuitively presenting the sensitivity characteristics of system evolution under different policy gradients.

In each scenario figure, the three vertically arranged subfigures (a), (b), and (c) correspond to port operators, river–sea intermodal operators, and road transport operators, respectively. Under each agent type, two decision-making modes are further distinguished: the experience-driven mode and the adaptive learning mode. The

semi-transparent shaded areas surrounding the curves represent the confidence intervals from 30 independent repeated simulation runs, reflecting the robustness of the system evolution. These two decision-making modes designed in this study have their corresponding counterparts in real-world collection distribution systems. The experience-driven mode corresponds to large state-owned port groups with longer decision-making chains and strong path dependence, or long-term contract-based central shipping enterprises and large-scale contract logistics firms. These agents have fixed decision weights and strong strategic perseverance, typically making decisions based on long-term planning; consequently, they exhibit fast convergence and relatively stable policy implementation. The adaptive learning mode, in contrast, maps onto regional private waterway operators, small and medium-sized private freight fleets, and flexible third-party logistics providers. The decision weights of these agents are dynamically adjusted according to satisfaction feedback, and they display pronounced wait-and-see tendencies during the 'dual carbon' transition; their decision-making processes are highly dependent on neighbourhood payoff feedback, which effectively

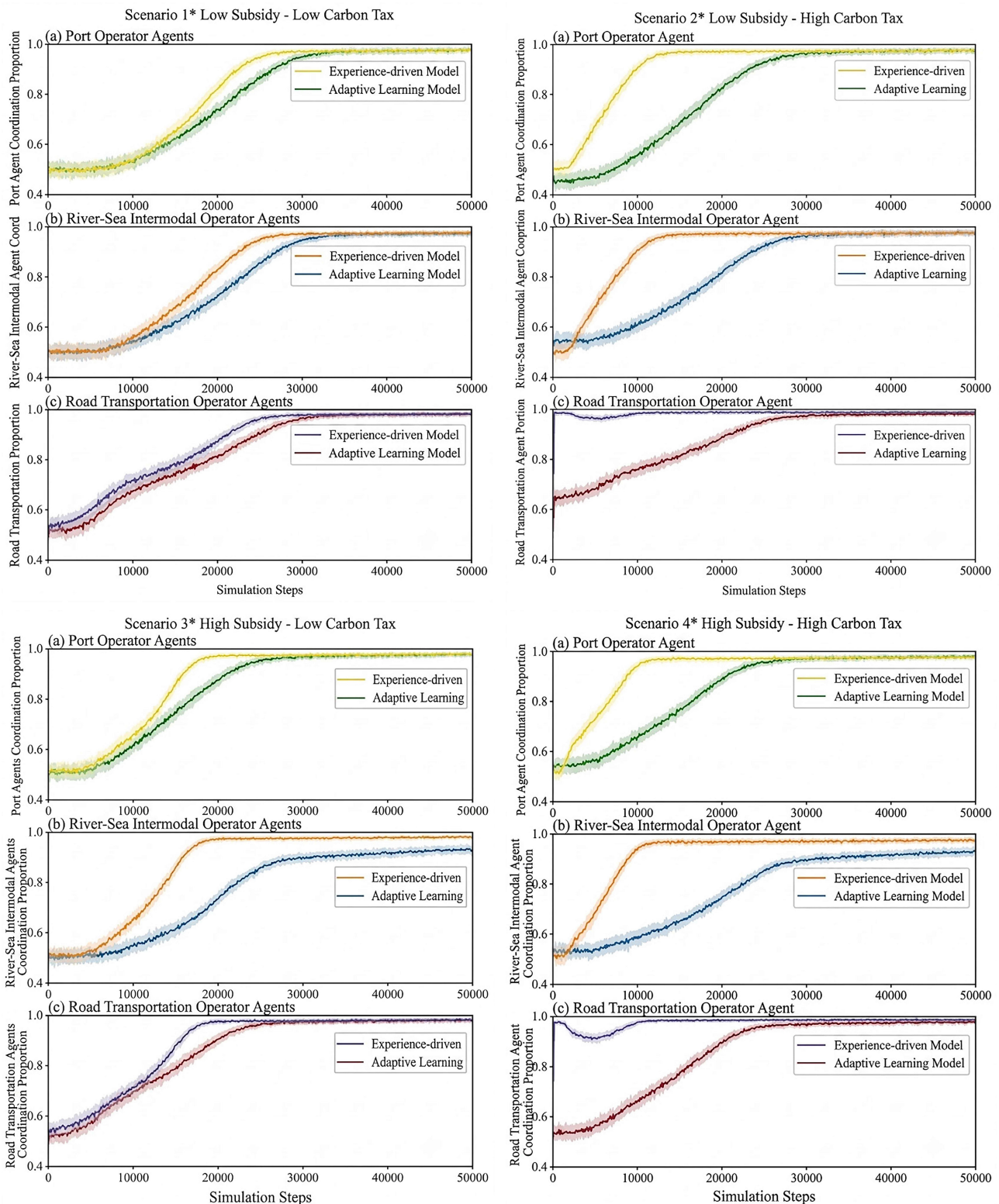


Fig. 2 System evolution trends under different policy scenarios.

reproduces the authentic evolutionary logic of terminal nodes in the collection-distribution system under uncertain policy environments.

This subsection quantitatively investigates the driving effects of carbon taxes and policy subsidies, in different intensity combinations, on the collaborative evolution of the port collection distribution system through a 2×2 policy matrix. The simulation

experiments divide the policy gradients into four typical scenarios: 'low subsidy, low carbon tax', 'low subsidy, high carbon tax', 'high subsidy, low carbon tax', and 'high subsidy, high carbon tax', with the aim of examining how the system overcomes the non-cooperative state. Since experience-driven agents respond more quickly to environmental changes without the time-lag interference

introduced by adaptive learning, this part of the study primarily takes the evolutionary curves of the experience-driven mode as the observation object, so as to accurately assess the impacts of policy parameters.

The simulation results show that under the 'dual low' policy scenario, the system exhibits pronounced convergence delay. Although it can still eventually tend toward the collaborative steady state, the convergence steps are prolonged, and the slope of the evolutionary curve in the early stage is relatively flat. This indicates that under low incentive levels, the additional costs of green transformation are not effectively offset, resulting in strong internal evolutionary resistance, with agents vacillating between traditional and green modes over an extended period. When only one policy intensity is increased, the system displays notable asymmetric responses. Under the high carbon tax, low subsidy scenario, the cost pressure on the road transport side surges, forcing a significant acceleration in its convergence speed, which demonstrates that carbon tax, as a constraint-based instrument, has a mandatory driving effect in breaking the 'high carbon path lock-in'. In contrast, under the high subsidy, low carbon tax scenario, the financial pressure on the river-sea intermodal side is alleviated to some extent; however, due to the absence of negative cost constraints on the traditional mode, the overall improvement in the system's evolutionary slope is limited.

The study further identifies a synergistic acceleration effect of policy combinations. Under the 'dual high' policy scenario, the system's convergence efficiency achieves a leapfrog improvement, with the evolutionary curve reaching the steady-state peak within a very short number of time steps. This acceleration effect is not a simple linear superposition; rather, the 'push' generated by the carbon tax and the 'pull' generated by subsidies jointly construct a steeper payoff gradient, significantly reducing the difficulty for agents to cross the strategy-switching threshold. From the perspective of transport governance, this phenomenon confirms that the organic combination of constraint-based and incentive-based instruments can produce a positive resonance, substantially shortening the system's transient response time and ensuring that the collection distribution network achieves efficient low-carbon transformation under the 'dual carbon' constraints.

Overall, the collaborative evolution of the collection distribution industrial chain follows a multi-agent linkage characteristic. This finding demonstrates that system collaboration is not an equilibrium achieved simultaneously by all agents, but rather a chain-wise transmission process based on differences in each agent's cost sensitivity and decision thresholds. After identifying this sequential characteristic, governments can more precisely apply policy pressure or incentives at different time points, thereby optimising the transformation efficiency of the entire chain.

To isolate the interaction effects between carbon taxes and subsidies across agent types, we decompose the policy-driven variations from the steady-state Q-value differentials. The decomposition yields three findings. First, road transport operators show the strongest carbon-tax sensitivity, while ports exhibit diminishing marginal responses, confirming their lower sensitivity to regulatory intensity. Second, subsidies and carbon taxes are complementary for river-sea intermodal operators: the same subsidy level generates a 15.6% Q-value improvement under low carbon tax, but a 38.2% improvement under high carbon tax—a synergistic amplification effect. Third, these interaction effects are asymmetric across agent types: additive for ports, carbon-tax-dominated for road operators, and strongly synergistic for intermodal operators. This asymmetry implies that policymakers should adopt differentiated

strategies: carbon taxes are most effective for road transport, a balanced combination is essential for inducing intermodal participation, and port transformation is best driven by non-financial instruments.

To conduct an in-depth analysis of the evolutionary patterns of the collection distribution system in the spatial dimension, this section selects four key time nodes during the simulation— $t = 100$, $t = 10,000$, $t = 20,000$ and $t = 50,000$, and compares their spatial evolution snapshots^[28]. In the 30×30 grid, each pixel corresponds to an independent agent node, with spatial interactions governed by the von Neumann neighbourhood rule. In the figures, blue areas represent clusters of agents that have adopted green collaborative strategies (proactive promotion, cooperation, low carbon mode), while red areas denote clusters of agents that have retained traditional non-collaborative strategies (passive promotion, competition,

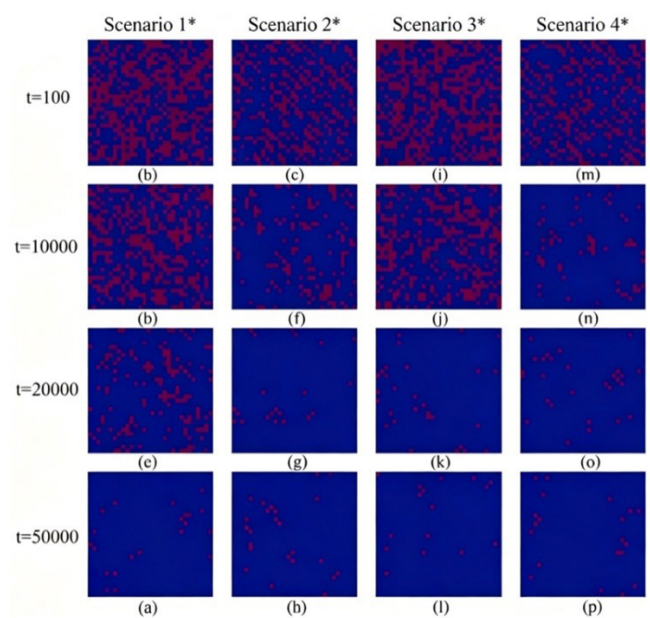


Fig. 3 Agent grid evolution under the experience-driven mode.

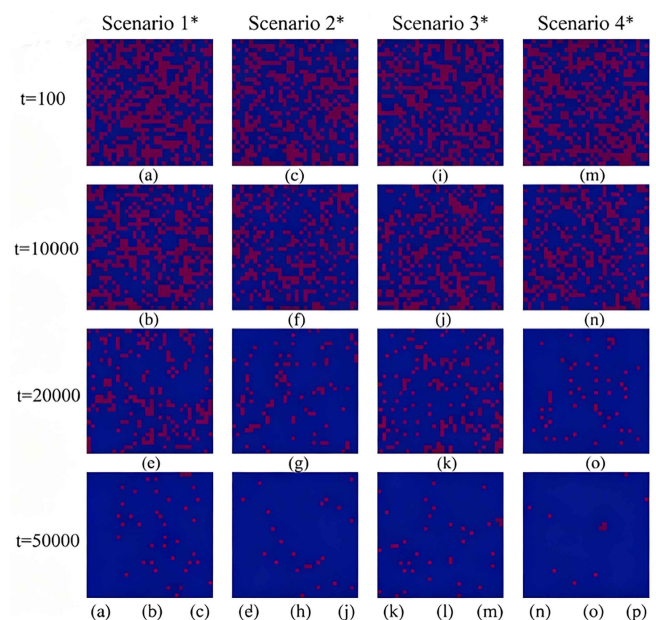


Fig. 4 Agent grid evolution under the adaptive learning mode.

traditional mode). By observing the grid snapshots of the two decision-making modes under different policy scenarios in Figs 3 and 4, the propagation characteristics of collaborative strategies can be revealed from the perspective of topological dynamics.

By observing the evolution of the grid from disorder to order, it is evident that the propagation of collaborative strategies in the collection-distribution system follows the kinetic logic of 'local nucleation → cluster expansion → global emergence'.

Early evolutionary stage $t = 100$: At the initial simulation phase, the grid exhibits pronounced blue-red intermingled fragmentation. Since the initial strategies are randomly distributed, the incentive effects of policy instruments have not yet accumulated spatially; agents that choose green collaborative strategies remain isolated. From a game-theoretic perspective, isolated individuals lack the payoff bonuses derived from neighbourhood collaboration, making it difficult for transition benefits to fully offset the additional costs. This stage reflects the evolutionary resistance faced by the collection distribution system at the onset of green transformation: micro agents lack external collaborative signals to guide them, and generally remain in a state of low-level strategic fluctuation.

Mid evolutionary stage ($t = 10,000$ to $t = 20,000$): As the simulation proceeds, distinct blue patches emerge in space, indicating the formation of local collaborative clusters. Under the von Neumann neighbourhood rule, agents continuously observe and learn from the payoffs of neighbouring nodes. The payoff gradients created by carbon taxes and subsidy policies enable some pioneering agents that have already transitioned to realise higher collaborative spillover gains by aligning strategies with neighbouring homogeneous agents. Such high payoff signals exert a strong attraction on surrounding non-collaborative agents, inducing a spatial contagion effect of strategies. By time $t = 20,000$, the blue patches rapidly expand from points to areas, exhibiting a diffusion pattern from core to periphery, which faithfully reproduces the driving effect of green demonstration corridors on surrounding hinterland agents in transportation planning.

Late evolutionary stage $t = 50,000$: Under scenarios where policy intervention is effective, the blue clusters eventually cover the vast majority of the grid, and the system achieves a structural transformation from individual decentralized decision-making to a globally collaborative configuration. From the perspective of topological dynamics, a fully collaborative state implies that the system has completed spatial path lock-in; at this point, collaborative agents form an extremely stable mutually supportive structure in space, and any random exploratory behavior of individual agents is quickly pulled back by strong neighbourhood signals. This demonstrates that the low-carbon transformation of the collection-distribution system is a systemic emergence process triggered by local demonstration effects and realised through spatial positive-feedback mechanisms.

In the early stages of evolutionary development, the adaptive learning mode generally exhibits a notable response lag and spatial fluctuation problem. To further clarify the micro-level causes of this phenomenon, this section shifts the research perspective from the previous macro evolutionary trajectories to the micro decision-making mechanisms, with a particular focus on the relationship between the information sharing rate and the system's collaborative trends. This paper proposes that the response delay of adaptive agents does not imply ineffective decision-making; rather, it is an essential process through which micro agents, when confronted with policy fluctuations and environmental noise, learn to adapt their strategies via information exchange and hedge against uncertainty risks. By quantitatively analysing the dynamic changes in the

information-sharing rate at various evolutionary stages, it is possible to more accurately identify the transition logic of the collection-distribution system from 'externally policy driven' to 'internally experience locked in'. Figure 5 employs dual Y-axis coordinate charts to compare the dynamic relationship between the average information sharing rate and the average cooperation rate under four scenarios, where the left Y-axis corresponds to the system's average cooperation rate and the right Y-axis corresponds to the average information sharing rate. The experimental results clearly demonstrate how the strategy learning of adaptive learning agents gradually shifts from 'relying on social information' to 'relying on their own experience'.

Based on the slope of the average cooperation rate and the decay characteristics of the information sharing rate, this study divides the entire evolutionary process into three stages for correlation analysis:

In the initial evolutionary stage, the average cooperation rate begins to rise slowly from around 0.5. At this time, the average information-sharing rate remains in a relatively high and sensitive range across all scenarios. This feature reveals a 'social dependence' behaviour among collection distribution agents in the early transformation phase. Lacking prior successful experience in responding to policy fluctuations, agents must maintain a relatively high frequency of information sharing to observe the payoff feedback of neighbouring agents. This persistently high level of learning explains the response lag observed macroscopically in the adaptive mode—that is, the system is undergoing the necessary micro-level information processing and strategy accumulation to overcome the evolutionary resistance posed by initial transition costs.

Entering the next stage, the average cooperation rate curve on the left Y-axis shows a steep increase, indicating the rapid expansion of collaborative clusters in the grid space. Concurrently, the average information-sharing rate on the right Y-axis transitions from its initial high plateau to a pronounced accelerating decline. This inflection point is highly synchronised with the moment when the cooperation rate crosses the critical threshold. After the majority of neighbouring agents have shifted to collaborative strategies, collaborative gains are released, and agents' payoff satisfaction rapidly improves. According to the adaptive adjustment mechanism, when individuals find that the current strategy yields stable returns, their dependence on external information naturally decreases, and the decision-making focus shifts from 'environment-following' to 'internalised experience'.

Considering the entire evolutionary process of the adaptive mode, the trajectory of the information sharing rate fully explains how the system transitions from initial 'uncertainty wait-and-see' to later 'deterministic lock-in'. Although adaptive agents exhibit relatively high learning costs in the early stage, what they gain in return is a steady state with low information costs: after collaboration is achieved, the system no longer needs to maintain high-frequency information exchange to sustain green operations^[31]. This finding offers implications for the long term governance of collection-distribution systems: guiding agents to establish collaborative experience in the early stage can enable the system to form endogenous green signals, thereby reducing the government's reliance on information disclosure and real-time regulation in the later stage of evolution.

Conclusions

This paper addresses the collaborative optimisation problem of port collection distribution systems under the 'dual carbon' goals,

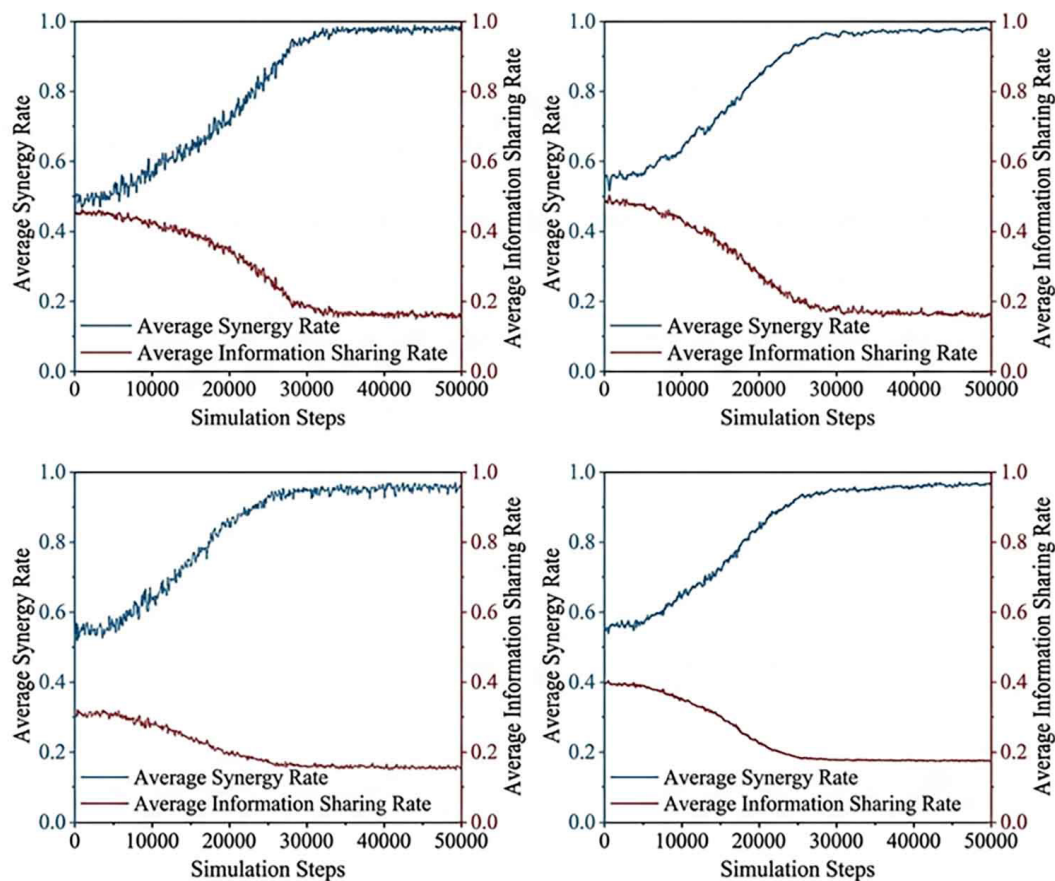


Fig. 5 Correlation trend under different scenarios.

constructing a complete research framework that spans from evolutionary game-theoretic mechanism analysis to multi-agent dynamic simulation^[32]. By introducing local information interaction and satisfaction adjustment mechanisms, it overcomes the limitations of traditional evolutionary games in characterising spatial heterogeneity and dynamic decision-making, and systematically reveals the underlying logic and macroscopic laws governing the green collaborative evolution of port collection distribution systems. The main conclusions of this paper are as follows:

(1) A spatial collaborative evolution model of collection distribution systems integrating micro-level dynamic learning mechanisms is constructed. By combining evolutionary game mechanisms with multi-agent reinforcement learning, a bottom-up spatial grid evolution model is formed^[33]. With the aid of a payoff satisfaction adjustment function, the model demonstrates at the algorithmic level that the transformation resistance exhibited by asset-heavy logistics enterprises under external policy shocks can be precisely characterised through dynamic learning weights. This model avoids the stringent 'perfect rationality' assumption of traditional game theory, providing an effective quantitative evaluation tool for studying the decision-making of heterogeneous agents within complex transport networks.

(2) The inhibitory mechanism of local strategy mismatch on global collaborative evolution in collection distribution networks is revealed. The study finds that the green transformation of collection distribution systems is not a simple superposition of emission reduction effects across individual links, but is highly dependent on the higher-order synergy among ports, shipping, and road transport. Any single agent remaining in the traditional high-carbon state

can alter the payoff expectations of neighbouring nodes through the network topology, thereby significantly slowing down the global green evolution process. Simulations confirm that strategy mismatches between ports and transport operators under non-equilibrium conditions lead to persistent synergistic friction losses within the system, making synchronised evolution across the entire chain a necessary condition for maximising overall system welfare.

(3) The combined incentive effectiveness of heterogeneous environmental policies and their evolutionary pathways is demonstrated. Moderate regulatory intensity is shown to be more effective in prompting the system to converge toward a collaborative equilibrium. Simulations reveal that different agents exhibit distinct evolutionary responses and interdependencies: river-sea intermodal operators are more sensitive to cost changes, while road transport operators rely more on derivative benefits such as industry reputation and policy-related evaluations. For hub ports, the formation of a proactive promotion strategy is not sensitive to costs and subsidies, but is influenced by indirect losses under non-collaborative states. This finding suggests to policymakers that the core of fostering green collaboration lies not in continuous financial input, but in establishing rigorous assessment mechanisms and industry reputation systems, shifting from pure 'financial subsidies' toward 'regulatory constraint'^[34].

(4) The moderating role of payoff satisfaction on the transformation rate of micro enterprises is clarified. The study confirms that the decision making willingness of logistics enterprises during the transformation process is significantly affected by their current payoff levels. When enterprise payoffs are at a medium or higher level, the

motivation to change existing strategies is extremely low, exhibiting strong 'transformation inertia'; only when payoff satisfaction drops substantially do enterprises spontaneously increase their willingness to explore new strategies. This finding suggests to policymakers that the implementation of green strategies should not rely solely on short term marginal profit incentives; rather, it is necessary to strengthen industry information sharing and other measures to break the cognitive limitations of low-payoff agents, stimulate their intrinsic motivation for change, and thereby effectively reduce long-term macro-regulatory costs.

Author contributions

The authors confirm their contributions to the paper as follows: study conception and design: Xu Q, Wu S; methodology: Li C, Deng H; validation: Ma T; formal analysis: Li C, Deng H; investigation: Li C, Ma T; data curation: Li C; visualization: Ma T; writing—original draft: Li C, Deng H; writing—review and editing: Xu Q, Li C, Ma T, Wu S; supervision, project administration, funding acquisition, resources: Xu Q, Wu S. All authors reviewed the results and approved the final version of the manuscript.

Data availability

The datasets generated during and analyzed in the current study are available from the corresponding author upon reasonable request.

Acknowledgments

This work was financially supported by the 'National Natural Science Foundation of China', (Grant No. 52362055), and the 'Science and Technology Department of Guangxi Autonomous Region', (Grant No. 2025LTLT03003).

Conflict of interest

The authors declare that they have no conflict of interest.

Dates

Received 26 July 2026; Revised 12 August 2026; Accepted 10 September 2026; Published online 30 September 2026

References

- [1] Shen WS, Khoong CM. 1995. A DSS for empty container distribution planning. *Decision Support Systems* 15(1):75–82
- [2] Janic M. 2007. Modelling the full costs of an intermodal and road freight transport network. *Transportation Research Part D: Transport and Environment* 12(1):33–44
- [3] He J, Zhang L, Deng Y, Yu H, Huang M, et al. 2023. An allocation approach for external truck tasks appointment in automated container terminal. *Advanced Engineering Informatics* 55:101864
- [4] Guo X, He J, Lan M, Yu H, Yan W. 2022. Modeling carbon emission estimation for hinterland-based container intermodal network. *Journal of Cleaner Production* 378:134593
- [5] Xu Q, Zhong Y, Deng H, Wang X, Chen X. 2024. Scheduling on tractor and trailer transportation considering the influence of disrupted events based on the contract net and simulated annealing algorithm. *Digital Transportation and Safety* 3(3):155–168
- [6] Li L, Zhang Q, Zhang T, Zou Y, Zhao X. 2023. Optimum route and transport mode selection of multimodal transport with time window under uncertain conditions. *Mathematics* 11(14):3244
- [7] Xin X, Wang X, Zhang T, Chen H, Guo Q, et al. 2023. Liner alliance shipping network design model with shippers' choice inertia and empty container relocation. *Electronic Research Archive* 31(9):5509–5540
- [8] Yan Z, Zhao H, Li L, Galland S. 2024. Evolutionary game theory combined with reinforcement learning synthesis - a comprehensive survey. *2024 36th Chinese Control and Decision Conference (CCDC), Xi'an, China, 2024*. US: IEEE. pp. 1077–1082. doi: [10.1109/CCDC62350.2024.10587977](https://doi.org/10.1109/CCDC62350.2024.10587977)
- [9] Yin C, Zhang ZA, Fu X, Ge YE. 2024. A low-carbon transportation network: collaborative effects of a rail freight subsidy and carbon trading mechanism. *Transportation Research Part A: Policy and Practice* 184:104066
- [10] Hu X, Cheng R, Chen X, Zhao J, Liu F, et al. 2025. Optimization and implementation strategy for low-carbon multimodal transport routes: a collaborative approach between government and transport enterprises. *Journal of Cleaner Production* 502:145341
- [11] Wang H, Ruan Y, Niu B, Zhang H. 2024. Intra-port demand coordination and inter-port competition: Ports' pricing strategies with congestion concerns. *Ocean & Coastal Management* 248:106918
- [12] Zhu M, Shen S, Shi W. 2023. Carbon emission allowance allocation based on a bi-level multi-objective model in maritime shipping. *Ocean & Coastal Management* 241:106665
- [13] Jiang Y, Xu CY, Xu J. 2025. Research on optimization of green marine fuel supply chain considering fuel substitution effects. *Frontiers in Marine Science* 12:1668875
- [14] Sheng D, Li ZC, Fu X, Gillen D. 2017. Modeling the effects of unilateral and uniform emission regulations under shipping company and port competition. *Transportation Research Part E: Logistics and Transportation Review* 101:99–114
- [15] Xu L, Li X, Yan R, Chen JH. 2025. How to support shore-to-ship electricity constructions: tradeoff between government subsidy and port competition. *Transportation Research Part E: Logistics and Transportation Review* 201:104258
- [16] Xu Q, Deng H, Huang J, Wang J, Wu Q, et al. 2025. Optimization of multi-party cooperative game strategies for port consolidation and distribution under the goal of "Dual Carbon". *Frontiers in Marine Science* 12:1741273
- [17] Chen M, Yuan X, Zhang Y, Liu YA. 2025. Tripartite evolutionary game analysis of ports and shipping enterprises under carbon trading mechanisms: Incentives and policy implications. *Journal of Cleaner Production* 529:146773
- [18] Zhang T, Hong C, Kramberger T, Wang Y. 2025. Integrating carbon tax and subsidies: an evolutionary game theory-based shore power promotional strategy analysis. *Systems* 13(4):239
- [19] Li H, Yang P, Liu W, Yan S, Zhang X, et al. 2025. Multi-agent reinforcement learning in games: research and applications. *Biomimetics* 10(6):375
- [20] Xu L, Wu J, Yan R, Chen J, Fu S. 2025. Who predicts better? A comparison of machine learning and econometrics in forecasting CO₂ emissions from global shipping. *Energy* 338:138967
- [21] Wang Q, Li M, Zhang Y, Kang Y. 2025. Green port collection and distribution system in low-carbon development: scenario-based system dynamics. *Sustainability* 17(14):6516
- [22] Yin C, Ke Y, Chen J, Liu M. 2021. Interrelations between sea hub ports and inland hinterlands: perspectives of multimodal freight transport organization and low carbon emissions. *Ocean & Coastal Management* 214:105919
- [23] Hrušovský M, Demir E, Jammernege W, Van Woensel T. 2018. Hybrid simulation and optimization approach for green intermodal transportation problem with travel time uncertainty. *Flexible Services and Manufacturing Journal* 30(3):486–516
- [24] Bütün C, Petrovic S, Muyldermans L. 2021. The capacitated directed cycle hub location and routing problem under congestion. *European Journal of Operational Research* 292(2):714–734
- [25] Alkaabneh F, Diabat A, Elhedhli S. 2019. A Lagrangian heuristic and GRASP for the hub-and-spoke network system with economies-of-scale and congestion. *Transportation Research Part C: Emerging Technologies* 102:249–273

- [26] Guo Q, Yin C, Zheng S. 2025. An intermodal transport network planning scheme considering carbon emissions. *Energy* 322:135564
- [27] Dai Q, Yang J, Li D. 2018. Modeling a three-mode hybrid port-hinterland freight intermodal distribution network with environmental consideration: the case of the yangtze river economic belt in China. *Sustainability* 10(9):3081
- [28] Xu L, Wu J, Yan R, Chen J. 2025. Is international shipping in right direction towards carbon emissions control? *Transport Policy* 166:189–201
- [29] Guo X, He J, Yu H, Liu M. 2023. Carbon peak simulation and peak pathway analysis for hub-and-spoke container intermodal network. *Transportation Research Part E: Logistics and Transportation Review* 180:103332
- [30] Sun Y, Lang M. 2015. Modeling the multicommodity multimodal routing problem with schedule-based services and carbon dioxide emission costs. *Mathematical Problems in Engineering* 2015(1):406218
- [31] Halim RA, Kwakkel JH, Tavasszy LA. 2016. A strategic model of port-hinterland freight distribution networks. *Transportation Research Part E: Logistics and Transportation Review* 95:368–384
- [32] Zhang X, Zhang L, Xi H, Shao Z, Bell MGH. 2026. Shore power adoption strategies of shipping companies and pricing decisions of the port under subsidies and carbon taxes: a game theoretical analysis. *Transportation Research Part E: Logistics and Transportation Review* 205:104490
- [33] Parada L, Yu K, Angeloudis P. 2025. IntNet: a communication-driven multi-agent reinforcement learning framework for cooperative autonomous driving. *IEEE Robotics and Automation Letters* 10(3):2478–2485
- [34] Xu L, Mak S, Schoepf S, Ostroumov M, Brintrup A. 2025. Multi-agent digital twinning for collaborative logistics: Framework and implementation. *Journal of Industrial Information Integration* 45:100799



Copyright: © 2026 by the author(s). Published by Maximum Academic Press, Fayetteville, GA. This article is an open access article distributed under Creative Commons Attribution License (CC BY 4.0), visit <https://creativecommons.org/licenses/by/4.0/>.