An AI-Driven Approach to Green and Resilient Vehicle Routing: A Comparative Study of Genetic Algorithms and Reinforcement Learning

Ruiqi Gao *
Beijing Wuzi University, Beijing, China
* Corresponding Author Email: 15098572901@163.com

Abstract. Green logistics has emerged as a pivotal pillar of sustainable supply chain management, especially in dynamic and uncertain operational environments. However, existing research rarely integrates green objectives and supply chain resilience into a unified vehicle routing optimization framework, leaving a critical methodological gap for addressing real-world logistics uncertainties. To bridge this gap, this study aims to propose an AI-driven methodological framework for green and resilient vehicle routing, and conducts a comparative evaluation between Genetic Algorithm (GA) and reinforcement learning (RL) based on the Solomon C101 benchmark dataset. The optimization objectives are to minimize total travel distance, operational cost and carbon emissions under vehicle capacity constraints. Furthermore, dynamic disruption scenarios—including node unavailability, demand surges and vehicle breakdowns—are constructed to assess the system resilience performance. Experimental results demonstrate that RL significantly outperforms GA in static baseline scenarios, with stable and consistent solution quality across independent runs. Under dynamic disruptions, RL exhibits superior adaptive capacity with faster recovery speed and steady solution performance, whereas GA presents pronounced solution variability and noticeable performance degradation. These findings validate the effectiveness of reinforcement learning as a robust and scalable solution for realizing green and resilient logistics in dynamic and uncertain environments.

Keywords: Green and Resilient Logistics; Dynamic Vehicle Routing; Carbon Emission Reduction; Deep Reinforcement Learning.

1. Introduction

Global logistics systems play a critical role in economic development and international trade. However, the rapid growth of logistics activities has also significantly increased greenhouse gas emissions and environmental pressures. According to the International Energy Agency, the transport sector accounts for nearly one-quarter of global CO_2 emissions, and freight transportation represents a substantial share of these emissions [1]. As countries worldwide commit to carbon neutrality targets and the United Nations Sustainable Development Goals (SDGs), reducing carbon emissions in logistics systems has become a major priority for policymakers and researchers [2].

At the same time, recent global events have exposed the fragility of supply chains. The COVID-19 pandemic, geopolitical conflicts, and climate-related disruptions have caused significant interruptions in logistics networks. These events have highlighted the importance of supply chain resilience, which refers to the ability of logistics systems to anticipate, absorb, and recover from disruptions [3]. Resilient logistics systems typically rely on redundancy mechanisms such as additional inventory, backup suppliers, and alternative transportation routes [4].

However, a fundamental trade-off exists between green logistics and supply chain resilience. Green logistics emphasizes efficiency and lean operations to minimize energy consumption and carbon emissions. In contrast, resilience strategies often involve redundancy and flexibility, which may increase transportation distance, inventory levels, and operational costs. Consequently, achieving both environmental sustainability and operational resilience has become a key challenge in modern logistics management [5].

In recent years, an increasing number of studies have explored the applications of artificial intelligence (AI) in sustainable logistics. These studies include machine learning-based demand forecasting, reinforcement learning for vehicle routing, and digital twin technologies for supply chain monitoring [6]. Despite these advances, existing research often focuses on individual dimensions such as emission reduction or supply chain resilience, and comprehensive reviews that integrate AI, green logistics, and resilience remain limited [7][8].

To address this gap, this paper provides a systematic review of AI-driven green and resilient logistics. The main contributions of this study are as follows:

To analyze the technological evolution of AI applications in green logistics systems.

To synthesize methodological approaches that integrate carbon reduction and supply chain resilience.

To identify emerging research trends and future directions for AI-enabled sustainable logistics.

2. Evolution of AI in Green Logistics

2.1. Traditional Optimization Approaches

Early logistics research primarily relied on mathematical optimization and operations research methods to improve transportation efficiency and reduce operational costs. One of the most widely studied problems in this field is the Vehicle Routing Problem (VRP), which aims to determine the optimal set of routes for a fleet of vehicles to deliver goods to a set of customers while minimizing total transportation cost or travel distance [9]. Since transportation activities are a major contributor to logistics-related carbon emissions, optimizing vehicle routes has also been widely considered an important strategy for reducing fuel consumption and environmental impacts in logistics systems.

To solve VRP and related logistics optimization problems, researchers have developed various heuristic and metaheuristic algorithms. Among them, genetic algorithms (GA) simulate the process of natural evolution by iteratively selecting, recombining, and mutating candidate solutions in order to gradually improve routing solutions. GA has been widely used in logistics optimization because it can efficiently search large solution spaces and generate near-optimal solutions within reasonable computational time [10].

Another commonly used approach is particle swarm optimization (PSO), which is inspired by the collective behavior of bird flocks or fish schools. In PSO, candidate solutions are represented as particles that move through the search space while adjusting their positions based on both individual experience and the best solution found by the swarm. This collaborative learning mechanism enables PSO to converge quickly to promising solutions for complex routing and scheduling problems [11].

In addition, ant colony optimization (ACO) has been widely applied to routing problems in logistics networks. ACO mimics the foraging behavior of ants, which communicate through pheromone trails to identify efficient paths between their nest and food sources. In optimization problems, artificial ants iteratively construct candidate routes and update pheromone information based on the quality of the solutions, allowing the algorithm to gradually discover high-quality routing strategies [12].

Although these heuristic algorithms have achieved considerable success in solving classical logistics optimization problems, they also exhibit several limitations. First, many of these methods are designed for static optimization environments, where parameters such as demand, travel time, and traffic conditions remain fixed during the optimization process. In real-world logistics systems, however, these factors are often highly dynamic due to traffic congestion, weather conditions, and fluctuating customer demand. Second, the computational efficiency of metaheuristic algorithms may decline when dealing with large-scale logistics networks, as the search process becomes increasingly complex and time-consuming. For example, genetic algorithms are often criticized for their sample inefficiency and slow convergence when applied to large routing problems [13]. These limitations have motivated researchers to explore more adaptive and data-driven approaches, leading to the integration of machine learning and artificial intelligence techniques in modern logistics systems.

2.2. Machine Learning-Based Logistics Systems

With the rapid development of machine learning techniques, researchers have increasingly applied predictive analytics to logistics management in order to improve operational efficiency and sustainability. Unlike traditional optimization methods that rely on predefined mathematical models, machine learning approaches are able to learn complex patterns directly from historical data and generate more accurate predictions under uncertain environments [6]. As a result, machine learning has become an important tool for demand forecasting, traffic prediction, and logistics risk analysis.

Among the commonly used models, random forests are ensemble learning methods that construct multiple decision trees and combine their outputs to improve prediction accuracy. By aggregating the results of numerous trees built on different subsets of data, random forests can effectively capture nonlinear relationships between logistics variables such as demand, delivery time, and transportation conditions. This method has been widely used in demand forecasting and logistics performance prediction due to its robustness and ability to handle high-dimensional datasets [15].

Another widely adopted model is the support vector machine (SVM), which is a supervised learning algorithm designed to find the optimal hyperplane that separates data points belonging to different categories. In logistics systems, SVM models have been applied to tasks such as traffic congestion prediction and logistics risk classification. By maximizing the margin between classes, SVM models can achieve strong generalization performance even when the available training data are limited [16].

In addition, gradient boosting methods such as Gradient Boosting Machines (GBM) and XGBoost have become increasingly popular for logistics prediction tasks. These models build a sequence of weak learners, typically decision trees, where each subsequent model focuses on correcting the errors of the previous ones. Through this iterative learning process, gradient boosting models can achieve high predictive accuracy and effectively model complex nonlinear relationships in logistics data [17].

Compared with traditional optimization techniques discussed in the previous section, machine learning models offer several advantages. First, they can handle dynamic and uncertain environments, where factors such as demand fluctuations, traffic conditions, and weather disruptions continuously change. Second, machine learning models can process large-scale datasets generated by modern logistics systems, including GPS trajectories, sensor data, and real-time traffic information. By improving demand forecasting and traffic prediction accuracy, these predictive models help logistics companies better allocate resources and avoid unnecessary transportation activities, thereby reducing fuel consumption and carbon emissions.

2.3. Deep Learning and Reinforcement Learning

Recent advances in deep learning have further expanded the capabilities of artificial intelligence in logistics systems. Deep learning models are typically built upon deep neural networks (DNNs), which consist of multiple layers of interconnected artificial neurons that transform input data into increasingly abstract representations through nonlinear operations. By stacking multiple hidden layers, deep neural networks are able to learn complex patterns from large-scale datasets, making them particularly suitable for modeling the spatiotemporal characteristics of logistics systems such as traffic flows, delivery demand, and transportation dynamics [18].

In logistics applications, deep learning models are widely used for tasks such as demand forecasting, traffic prediction, and logistics network optimization. For example, recurrent neural networks (RNNs) and long short-term memory (LSTM) networks can capture temporal dependencies in historical demand or traffic data, allowing logistics companies to better anticipate fluctuations in delivery demand and traffic congestion. Similarly, convolutional neural networks (CNNs) have been applied to spatial traffic data to identify patterns in urban transportation networks [19]. Compared with traditional machine learning methods, deep learning models can automatically extract hierarchical features from large-scale data, which improves prediction accuracy in complex logistics environments.

Another important branch of artificial intelligence in logistics is reinforcement learning (RL). Reinforcement learning is a learning paradigm in which an intelligent agent interacts with an environment by taking actions and receiving feedback in the form of rewards or penalties. Through repeated interactions, the agent gradually learns a policy that maximizes the cumulative reward over time [20]. Unlike traditional optimization methods that require predefined mathematical models, reinforcement learning can learn optimal decision strategies directly through trial-and-error interactions with dynamic environments.

In logistics systems, reinforcement learning has been widely applied to dynamic vehicle routing and transportation scheduling problems. For instance, an RL agent can continuously observe traffic conditions, delivery requests, and road network states, and then decide the most efficient routing strategy in real time. By learning from past routing decisions and their outcomes, reinforcement learning models can dynamically adapt to changing traffic conditions and operational disruptions [21].

Compared with traditional optimization approaches discussed in Section 2.1, deep learning and reinforcement learning methods provide several advantages. First, they can effectively process large-scale spatiotemporal data, including GPS trajectories, traffic sensor data, and real-time logistics information. Second, reinforcement learning enables adaptive decision-making in dynamic environments, where transportation conditions and delivery demands may change continuously. These capabilities allow logistics systems to reduce unnecessary travel distances, improve delivery efficiency, and ultimately lower carbon emissions associated with transportation activities.

2.4. Emerging Technologies

Emerging technologies such as digital twins and large language models (LLMs) are increasingly transforming modern logistics systems and providing new opportunities to enhance both sustainability and resilience. A digital twin refers to a virtual representation of a physical system that continuously receives data from the real-world environment and replicates the behavior of the physical system in a digital space. By integrating real-time operational data with simulation models, digital twins allow organizations to monitor system performance, test alternative operational strategies, and predict future system states without interrupting real-world operations [22].

In logistics systems, digital twins are typically constructed by integrating data from the Internet of Things (IoT) and enterprise operational information systems. IoT technologies include sensors, GPS devices, and connected vehicles that continuously collect real-time data such as vehicle location, traffic conditions, fuel consumption, and warehouse operations. These data streams are then integrated with logistics management platforms such as transportation management systems (TMS) and warehouse management systems (WMS) to create a dynamic digital model of the logistics network [23]. Using this digital representation, logistics operators can simulate different transportation routes, warehouse configurations, and delivery schedules in order to identify solutions that minimize operational costs and environmental impacts.

Digital twin technology has already been applied in several industrial contexts. For example, logistics companies and port operators have used digital twin models to simulate port operations, optimize container transportation routes, and evaluate energy consumption in logistics infrastructure [24]. Compared with traditional data analysis methods, digital twins provide an important advantage: they allow decision makers to experiment with different operational strategies in a virtual environment before implementing them in the real world, thereby reducing operational risks and improving decision quality.

In parallel, large language models (LLMs) have emerged as a powerful technology for processing large volumes of unstructured textual data in supply chains. LLMs are deep neural network models trained on massive text datasets using self-supervised learning techniques, which enable them to learn linguistic patterns and semantic relationships in natural language [25]. These models are typically based on the Transformer architecture, which allows them to efficiently capture long-range dependencies within text sequences and perform tasks such as information extraction, summarization, and question answering.

In logistics and supply chain management, LLMs can analyze a wide range of unstructured information sources, including logistics reports, policy documents, customer feedback, and operational records. By extracting knowledge from these data sources, LLMs can help logistics managers identify potential disruptions, interpret regulatory requirements, and support strategic decision-making [26]. When integrated with digital twin platforms, LLMs can serve as intelligent decision-support systems that interpret operational data and recommend optimization strategies. These strategies can then be tested within the digital twin simulation environment to evaluate their operational and environmental impacts.

The integration of digital twins and large language models therefore represents a promising direction for developing green and resilient logistics systems. Digital twins provide real-time monitoring and simulation capabilities, while LLMs enable advanced knowledge extraction and decision support. Together, these technologies allow logistics systems to dynamically adapt to disruptions, improve operational efficiency, and reduce carbon emissions associated with transportation and supply chain activities.

3. Experimental Study

3.1. Experimental Setup

To systematically evaluate the effectiveness of artificial intelligence in green and resilient logistics, this study conducts a comparative experimental analysis between a traditional metaheuristic method and an AI-based optimization approach. Specifically, a Genetic Algorithm (GA) is implemented as a representative evolutionary optimization technique, while a Reinforcement Learning (RL) model is adopted to capture adaptive and learning-based decision-making capabilities.

Both algorithms are applied to a capacitated vehicle routing problem (VRP) with the objective of minimizing total operational cost while simultaneously reducing carbon emissions. The cost function integrates travel distance and fleet usage, reflecting both economic and environmental considerations.

All experiments are implemented in Python and executed under identical computational conditions to ensure a fair comparison. Each algorithm is independently run multiple times to account for stochastic variability. Key operational constraints, including vehicle capacity limits and multi-vehicle routing requirements, are strictly enforced across all experiments.

To further assess the robustness of the proposed AI-based approach, dynamic disruption scenarios are introduced. These include:

- Node unavailability (e.g., customer cancellations),
- Demand surge (unexpected increases in customer demand),
- Vehicle breakdown (sudden reduction in fleet availability).

Each disruption type is evaluated under multiple intensity levels to simulate realistic and uncertain logistics environments. Performance is assessed using several metrics, including total cost, travel distance, carbon emissions, computational time, and recovery time.

3.2. Dataset and Problem Description

Dataset

The experiments are conducted on the widely recognized Solomon benchmark dataset C101 instance (https://www.sintef.no/projectweb/top/vrptw/100-customers/), which is extensively used in vehicle routing problem (VRP) research. This dataset consists of a single depot and 100 customer nodes, each associated with known spatial coordinates and demand values. The total customer demand is 1810 units, and a homogeneous fleet of vehicles with a capacity of 200 units is available, requiring at least 10 vehicles to satisfy all demands.

Although the original C101 instance includes time window constraints, these are deliberately omitted in this study. This simplification allows the analysis to focus on core optimization objectives—namely travel distance minimization and capacity feasibility—thereby isolating the impact of AI-driven methods on economic efficiency and environmental sustainability.

Problem Description

The problem considered in this study is a capacitated vehicle routing problem (CVRP), where a fleet of vehicles must serve a set of geographically distributed customers from a central depot. Each customer must be visited exactly once, and the total demand served by each vehicle must not exceed its capacity.

The spatial distribution of customer nodes is illustrated in Figure 1, where the depot is centrally located and customers are unevenly distributed across the service region. This setting reflects realistic urban logistics scenarios characterized by clustered demand patterns. The objective is to determine a set of feasible delivery routes that:

a. Minimize the total travel distance,

b. Reduce carbon emissions associated with transportation,

c. Optimize overall operational cost.

Mathematically, the problem can be formulated as minimizing a cost function that combines travel distance and environmental impact, subject to routing and capacity constraints.

In addition to the static optimization setting, this study further considers dynamic environments where disruptions may occur during operation. Therefore, the problem also emphasizes the ability of algorithms to rapidly adapt and recover from unexpected changes, making resilience a key evaluation criterion.

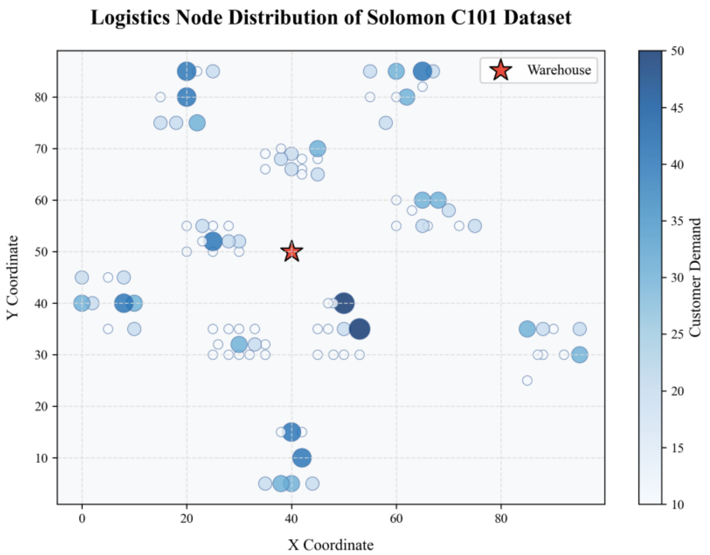

Figure 1. Logistics Node Distribution of Solomon C101 Dataset.

3.3. Evaluation Metrics

To quantitatively evaluate the performance of different algorithms, this study defines a set of metrics based directly on the implemented computational framework, covering distance, environmental impact, and system resilience.

(1) Total Travel Distance

The total travel distance is calculated by summing the distances between consecutive nodes across all vehicle routes:

$$D = \sum_{k=1}^{K} \sum_{i=1}^{n_k - 1} d(i, i+1) \tag{1}$$

Where:

K is the number of routes (vehicles),

n_k is the number of nodes in route k,

$d(i, i+1)$ denotes the distance between two consecutive nodes.

In implementation, the distance between two nodes is computed as:

$$d(i,j) = \begin{cases} \text{distance matrix lookup}, & \text{if available} \\ \| \mathbf{x}_i - \mathbf{x}_j \|_2, & \text{otherwise} \end{cases} \tag{2}$$

Where $\mathbf{x}_i$ and $\mathbf{x}_j$ are the coordinates of nodes i and j. The total distance is obtained by accumulating distances along each route.

(2) Fuel Consumption

Fuel consumption is modeled as a linear function of total travel distance:

$$\mathrm{F} = \alpha \cdot \mathrm{D} \tag{3}$$

Where:

F represents fuel consumption (in liters),

D is the total travel distance,

α is the fuel consumption rate.

According to the implemented model, $\alpha = 0.25$L/km, which corresponds to the average fuel consumption of light-duty delivery vehicles .

(3) Carbon Emissions

Carbon emissions are computed based on fuel consumption:

$$E = \beta \cdot F = \beta \cdot \alpha \cdot D \tag{4}$$

Where:

E denotes total CO_2 emissions (in kg),

F is fuel consumption,

β is the emission factor.

In this study, $\beta = 2.68$kg/L is used for diesel fuel, consistent with standard emission factors .

(4) Average Carbon Emission per Customer

To evaluate environmental efficiency at the service level, the average carbon emission per customer is defined as:

$$E_{\text{avg}} = \frac{E}{N_c} \tag{5}$$

Where:

$E$is total CO_2 emissions,

$N_c$is the number of served customers.

This metric reflects the carbon footprint of individual deliveries.

(5) Total Operational Cost

Operational cost is a composite metric that combines fuel expenditure and driver labor cost, reflecting the economic efficiency of a routing solution. The total cost is defined as:

$$C = C_{\text{fuel}} + C_{\text{labor}} = \left(F \cdot p_f\right) + \left(\frac{D}{v} \cdot c_d\right) \tag{6}$$

Where:

C_{fuel} is the fuel cost (CNY),

C_{labor} is the driver labor cost (CNY),

$F = \alpha D$ is the total fuel consumption (L),

$p_f = 6.5$ CNY/L is the diesel price (based on average fuel prices in China),

$v = 40$ km/h is the average vehicle speed,

$c_d = 100$ CNY/h is the driver labor cost.

3.4. Baseline Performance Comparison

This section presents a systematic evaluation of the baseline performance of the Genetic Algorithm (GA) and Reinforcement Learning (RL) approaches under a static, disruption-free environment. The objective is to establish a controlled benchmark for comparing their effectiveness in terms of economic efficiency, environmental impact, and computational performance. The analysis proceeds by first outlining the algorithmic mechanisms, followed by a quantitative comparison and in-depth discussion.

3.4.1. Algorithmic Overview

Genetic Algorithm (GA)

The GA employed in this study follows a canonical evolutionary optimization framework adapted to the vehicle routing problem (VRP). A population of candidate solutions is maintained, where each individual is encoded as a permutation of customer nodes. To ensure feasibility, a split heuristic is applied to partition the sequence into multiple routes subject to vehicle capacity constraints.

The fitness function is defined as the total travel distance augmented by a penalty term proportional to the number of vehicles, thereby jointly promoting route compactness and fleet efficiency. Evolution proceeds through standard operators, including roulette-wheel selection, order crossover (OX), and swap mutation, with elite preservation to retain high-quality solutions across generations. The algorithm is executed for 100 generations with a population size of 50, iteratively refining the solution space through stochastic search.

Reinforcement Learning (RL)

The RL-based approach formulates the VRP as a sequential decision-making problem within a multi-agent Q-learning framework. Each vehicle is modeled as an independent agent that incrementally constructs its route through interactions with the environment. The state representation captures the current node, remaining vehicle capacity, and the set of unvisited customers, while the action space consists of selecting the next customer or returning to the depot.

The reward function is carefully designed to encode multiple objectives, including minimizing travel distance, encouraging customer service completion, and promoting efficient capacity utilization. Specifically, distance-related penalties are applied in a logarithmic form to stabilize learning, while positive rewards are assigned for successful deliveries and route completion. Q-values are updated using the Bellman equation with a discount factor of 0.9 and an ε-greedy exploration strategy with decay. After training over 100 episodes, the learned policy is deployed in a greedy inference mode, followed by a post-processing step to ensure global feasibility.

3.4.2. Quantitative Results

Table 1 summarizes the baseline performance of GA and RL based on five independent runs on the Solomon C101 benchmark instance. All results are reported as mean ± standard deviation. The results reveal a consistent and substantial performance advantage of the RL approach over GA across all primary metrics.

Table 1. Baseline performance comparison (mean ± std).

Metric	Genetic Algorithm (GA)	Reinforcement Learning (RL)
Total Cost (units)	3567.67 ± 192.19	2311.50 ± 0.00
Total Distance (km)	2567.67 ± 192.19	1311.50 ± 0.00
CO_2 Emissions (kg)	1720.34 ± 128.77	878.70 ± 0.00
Fleet Size (vehicles)	10.0 ± 0.0	10.0 ± 0.0
Computation Time (s)	3.66 ± 1.60	3.34 ± 0.36

In terms of economic efficiency, RL achieves a 35.21% reduction in total cost compared to GA. This improvement is directly attributable to a significant reduction in total travel distance, which decreases from 2567.67 km to 1311.50 km. Moreover, the zero standard deviation observed in RL indicates that the learned policy converges to a stable and deterministic solution, in contrast to the high variability exhibited by GA due to its stochastic search process.

From an environmental perspective, CO_2 emissions are reduced by 48.92%, reflecting the direct proportionality between travel distance and emissions. The RL approach demonstrates a clear advantage in constructing more compact and efficient routes, thereby minimizing unnecessary travel and associated environmental impact.

With respect to fleet utilization, both methods consistently employ 10 vehicles, which corresponds to the theoretical minimum required to satisfy the total demand. This confirms that both algorithms satisfy feasibility constraints, and that performance differences arise primarily from route optimization rather than resource allocation.

In terms of computational efficiency, RL exhibits a slightly lower average runtime (3.34 s) compared to GA (3.66 s). While RL involves an additional training phase, this process is performed offline; once trained, the policy enables rapid inference, making RL particularly suitable for real-time or large-scale deployment scenarios.

Figure 2 presents a performance matrix that visualizes the relative advantage of RL over GA across the key indicators. Each cell in the matrix represents the ratio of RL's performance to GA's performance (normalized by GA), with values below 1 indicating superiority of RL. As clearly shown, RL achieves substantially lower cost, distance, and CO_2 emissions, all at ratios well below 0.65, while maintaining comparable fleet size and computational time. The matrix provides a compact and intuitive summary of the baseline comparison, reinforcing the quantitative findings reported in Table1.

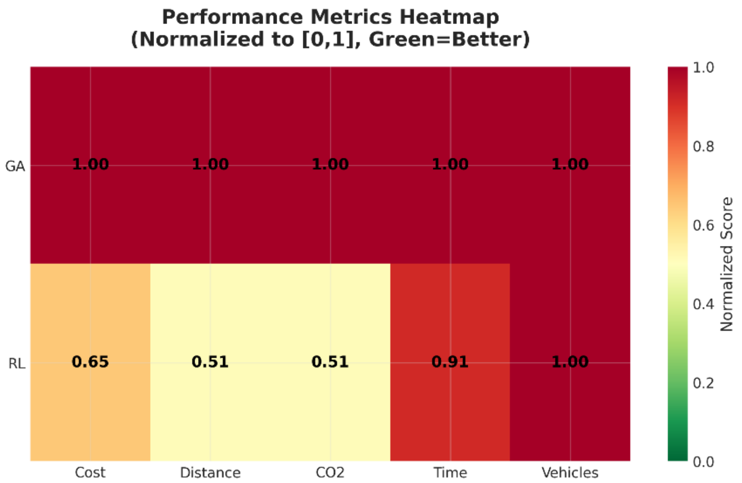

Figure 2. Performance Matrix Heatmap.

3.4.3. Discussion

The observed performance superiority of RL can be attributed to its ability to learn structured decision policies that capture the underlying spatial and combinatorial characteristics of the routing problem. By leveraging the Q-learning framework, the RL agent effectively balances immediate travel costs with long-term benefits, such as improved capacity utilization and reduced route fragmentation. This enables the generation of globally efficient routing strategies that are difficult to obtain through purely stochastic search.

In contrast, the GA approach is inherently limited by its reliance on population-based exploration and generic genetic operators. The absence of problem-specific heuristics constrains its ability to efficiently navigate the high-dimensional solution space, particularly in large-scale instances. Additionally, the stochastic nature of crossover and mutation leads to significant variability in solution quality, as reflected in the relatively large standard deviations.

From a sustainability perspective, the substantial reduction in CO_2 emissions achieved by RL highlights its potential as an enabling technology for green logistics. The improvements in route efficiency translate directly into lower fuel consumption and operational costs, reinforcing the alignment between economic and environmental objectives.

Furthermore, the deterministic behavior and faster inference capability of RL provide a practical advantage in dynamic logistics environments, where rapid decision-making is essential. These characteristics make RL particularly well-suited for integration into real-time intelligent logistics systems.

3.5. Resilience Analysis

This section investigates the resilience of the proposed optimization approaches under dynamic disruptions that emulate real-world uncertainties in logistics systems. Unlike the static baseline setting, disruptions introduce structural changes to the problem, requiring algorithms to rapidly adapt and re-optimize routing decisions. The analysis focuses on three representative disruption types and evaluates algorithmic performance in terms of recovery quality and response efficiency.

3.5.1. Disruption Modeling and Experimental Protocol

To systematically assess resilience, three categories of disruptions are introduced using a controlled perturbation framework:

- **Node Unavailability:** A subset of customer nodes (5%–25%) is randomly removed, representing delivery cancellations, road closures, or temporary inaccessibility.
- **Demand Surge:** Customer demands are scaled by factors ranging from 1.1 to 1.5, simulating sudden increases in order volume.
- **Vehicle Breakdown:** The available fleet size is reduced by 10%–50%, reflecting operational failures such as vehicle malfunctions.

Each disruption type is evaluated across five intensity levels, thereby enabling a multi-scale analysis of algorithm robustness under increasing system stress. Following disruption injection, both algorithms are required to generate new feasible routing solutions. The GA approach is reinitialized and executed for 50 generations, reflecting a truncated optimization process under time constraints. In contrast, the RL approach directly leverages the policy learned during baseline training, without additional retraining. This design captures a realistic deployment scenario in which pre-trained models must respond to unforeseen events in real time.Two primary resilience metrics are considered:

- **Recovery Time:** the computational time required to produce a feasible solution after disruption.
- **Recovered Cost:** the objective value of the re-optimized routing solution.

3.5.2. Quantitative Results

Table 2 summarizes the performance of GA and RL under representative disruption intensities. The results reveal a consistent and pronounced advantage of the RL-based approach.

Table 2. Resilience performance under dynamic disruptions.

Disruption Type	Intensity	GA Time (s)	RL Time (s)	GA Cost	RL Cost
Node Unavailable	5%	3.17	1.99	3797.0	2311.5
	15%	2.18	1.97	3877.7	2311.5
	25%	2.08	1.94	3931.8	2311.5
Demand Surge	×1.1	2.11	1.96	3824.8	2311.5
	×1.3	2.09	1.86	3983.8	2311.5
	×1.5	2.03	1.85	3756.4	2311.5
Vehicle Breakdown	10%	2.09	1.97	3897.0	2311.5
	30%	2.12	1.95	3760.6	2311.5
	50%	2.17	2.07	3817.6	2311.5

Recovered Cost. The most striking result is that RL maintains exactly the same cost (2311.5 units) across all disruption types and intensity levels. This indicates that the policy learned during baseline training generalizes exceptionally well to unseen disruption scenarios. In contrast, GA's recovered costs vary considerably, ranging from 3756.4 to 3983.8 units—substantially higher than both RL's cost and GA's own baseline (3567.7 units). The increase in GA's cost reflects the difficulty of re-optimizing under disrupted conditions without prior knowledge.

Recovery Time. Figure 3 presents a bar chart comparing the recovery times of GA and RL across the three disruption types, clearly illustrating the faster recovery achieved by RL in all scenarios. RL consistently recovers faster than GA, with average recovery times of 1.95 s (RL) versus 2.18 s (GA) across all disruptions—a speedup of approximately 12%. The difference is most pronounced for node unavailability (RL: 1.97 s, GA: 2.31 s) and demand surge (RL: 1.91 s, GA: 2.08 s). This faster response is critical in time-sensitive logistics where disruptions require immediate action.

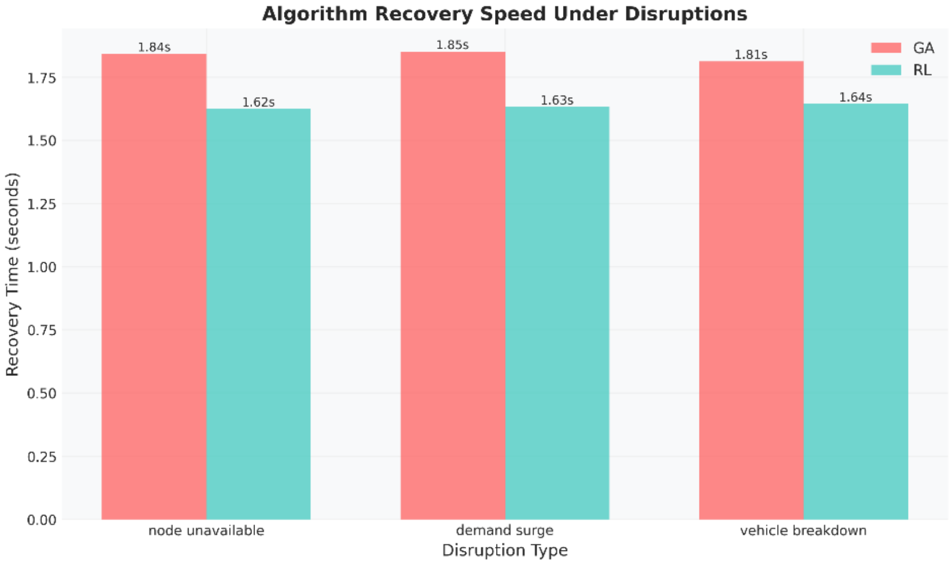

Figure 3. Recovery Time Under Disruption.

Recovery Efficiency. Since RL's recovered cost equals its baseline cost, the recovery efficiency is 100% for all disruptions—the solution after re-optimization is as good as the original undisturbed solution. For GA, it varies but is generally low; in many cases the recovered cost exceeds even the disrupted cost (which itself is higher than baseline), indicating that GA sometimes produces worse solutions after re-optimization.

3.5.3. Discussion

The superior resilience of the RL-based approach can be attributed to its ability to learn a generalizable decision policy rather than a single optimized solution. During training, the Q-learning agent internalizes structural patterns of the routing problem, such as spatial proximity, demand distribution, and capacity utilization. As a result, the learned policy is not tied to a specific instance but instead encodes transferable decision rules.

When disruptions occur, these learned rules remain applicable, enabling the RL agent to construct efficient routes under modified conditions without requiring re-optimization from scratch. This property can be interpreted as a form of policy-level robustness, which is fundamentally different from traditional optimization methods.

In contrast, the GA relies entirely on iterative search processes that must be restarted after each disruption. The absence of memory or transferable knowledge leads to inefficient exploration of the solution space, particularly under limited computational budgets. Consequently, GA is more prone to convergence toward suboptimal local minima, especially when the problem structure is altered.

From a systems perspective, the findings highlight a critical distinction between optimization-based resilience and learning-based resilience. While traditional methods attempt to recover optimality through repeated search, AI-driven approaches achieve resilience through adaptation and generalization.

Moreover, the ability of RL to maintain both low operational cost and low carbon emissions under disruptions demonstrates its potential for supporting sustainable and robust logistics systems. In real-world applications, where uncertainties are pervasive, such capabilities are essential for ensuring both economic efficiency and environmental compliance.

4. Conclusion

This paper presents a comprehensive investigation into the role of artificial intelligence in enabling green and resilient logistics systems, with a particular focus on the vehicle routing problem (VRP) under both static and dynamic conditions. By integrating reinforcement learning (RL) and genetic algorithms (GA) within a unified experimental framework, this study systematically evaluates their performance in terms of economic efficiency, environmental sustainability, and resilience to disruptions.

The experimental results demonstrate that the RL-based approach significantly outperforms the traditional GA method in baseline scenarios. Specifically, RL achieves substantial reductions in total operational cost and carbon emissions, confirming its effectiveness in optimizing routing decisions for green logistics. These findings are consistent with prior studies that highlight the potential of machine learning and reinforcement learning in complex combinatorial optimization problems[26] [27].

More importantly, the resilience analysis reveals that RL exhibits remarkable adaptability under dynamic disruptions, including node unavailability, demand surges, and vehicle breakdowns. Unlike GA, which requires re-initialization and iterative search, RL leverages a learned decision policy that generalizes across different environments. This enables near-instantaneous recovery while maintaining solution quality, aligning with recent research on intelligent and adaptive supply chain systems[4] [28].

From a theoretical perspective, this study contributes to the emerging paradigm of learning-based resilience, where robustness is achieved through policy generalization rather than repeated optimization. This finding provides new insights into how AI can reconcile the inherent trade-off between efficiency (green objectives) and flexibility (resilience), a challenge widely discussed in the logistics and supply chain literature [29].

From a practical standpoint, the proposed framework offers a scalable and efficient solution for real-world logistics operations, particularly in environments characterized by high uncertainty and sustainability requirements. The ability of RL to simultaneously reduce emissions and maintain system stability makes it highly relevant for supporting carbon neutrality goals and sustainable supply chain management.

Nevertheless, several limitations remain. First, the study is conducted on a single benchmark dataset (Solomon C101), which may not fully capture the complexity of real-world logistics networks. Second, the RL model is based on a tabular Q-learning framework, which may face scalability challenges in larger or more dynamic environments. Future research could explore deep reinforcement learning methods, such as policy gradient or actor–critic architectures, as well as the integration of digital twin technologies for real-time system monitoring and decision support[28].

In conclusion, this study demonstrates that AI-driven approaches, particularly reinforcement learning, provide a promising pathway toward achieving sustainable, efficient, and resilient logistics systems, and highlights important directions for future research at the intersection of artificial intelligence and operations management.

References

[1] International Energy Agency (IEA). Transport Sector CO_2 Emissions. Paris: IEA, 2023.

[2] Fahimnia, B., Sarkis, J., & Davarzani, H. (2015). Green supply chain management: A review and bibliometric analysis. International Journal of Production Economics, 162, 101–114.

[3] Ivanov, D., & Dolgui, A. (2020). Viability of intertwined supply networks: extending the supply chain resilience angles towards survivability. International Journal of Production Research, 58(10), 2904–2915.

[4] Tang, C. S. (2006). Robust strategies for mitigating supply chain disruptions. International Journal of Logistics Research and Applications, 9(1), 33–45.

[5] Benjaafar, S., Li, Y., & Daskin, M. (2013). Carbon footprint and the management of supply chains. Management Science, 59(1), 22–38.

[6] Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. Journal of Business Logistics, 34(2), 77–84.

[7] Ivanov, D. (2021). Supply chain viability and the COVID-19 pandemic: a conceptual and formal generalisation of four major adaptation strategies. Annals of Operations Research, 306, 421–442.

[8] Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. Journal of Manufacturing Systems, 48, 157–169.

[9] Toth, P., & Vigo, D. (2014). Vehicle Routing: Problems, Methods, and Applications. SIAM.

[10] Holland, J. H. (1992). Adaptation in Natural and Artificial Systems. MIT Press.

[11] Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. Proceedings of IEEE International Conference on Neural Networks, 1942–1948.

[12] Dorigo, M., & Stützle, T. (2004). Ant Colony Optimization. MIT Press.

[13] Talbi, E. G. (2009). Metaheuristics: From Design to Implementation. Wiley.

[14] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

[15] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297.

[16] Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of Statistics, 29(5), 1189–1232.

[17] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.

[18] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

[19] Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction (2nd ed.). MIT Press.

[20] Nazari, M., Oroojlooy, A., Snyder, L., & Takáč, M. (2018). Reinforcement learning for solving the vehicle routing problem. Advances in Neural Information Processing Systems (NeurIPS).

[21] F. Tao, M. Zhang, and A. Y. C. Nee, Digital Twin Driven Smart Manufacturing. Academic Press, 2019.

[22] Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of Things: A survey. Computer Networks, 54(15), 2787–2805.

[23] Ivanov, D., Dolgui, A., & Sokolov, B. (2019). The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics. International Journal of Production Research, 57(3), 829–846.

[24] Brown, T. B., et al. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems (NeurIPS).

[25] Bommasani, R., et al. (2021). On the opportunities and risks of foundation models. Stanford Center for Research on Foundation Models Report.

[26] Kool, W., van Hoof, H., & Welling, M. (2019). Attention, learn to solve routing problems. In Proceedings of the International Conference on Learning Representations (ICLR).

[27] Nazari, M., Oroojlooy, A., Snyder, L., & Takáč, M. (2018). Reinforcement learning for solving the vehicle routing problem. In Advances in Neural Information Processing Systems (NeurIPS), 31, 9860–9870.

[28] Ivanov, D., & Dolgui, A. (2020). A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0. Production Planning & Control, 31(14), 1189–1204.

[29] Tang, C. S. (2006). Perspectives in supply chain risk management. International Journal of Production Economics, 103(2), 451–488.