

ORIGINAL ARTICLE

International Journal of Applied Mathematics in Control Engineering

Journal homepage: <http://www.ijamce.com>

Eco-Intelligent Planning: A Multi-Objective Hybrid Reinforced Optimization Framework for Low-Carbon Resilient WEEE Reverse Logistics

Zelin Qu^{1*}

¹School of Information and Control Engineering, Qingdao University of Technology, Qingdao, Shandong, 266520, China

Correspondence

Zelin Qu, School of Information and Control Engineering, Qingdao University of Technology, Qingdao, Shandong, 266520, China
Email: qzlin162@163.com

Article Info

Article history:

Received 2 February 2026

Accepted 11 April 2026

Available online 28 April 2026

Abstract

To address the dual challenges of WEEE growth and sustainability, this paper proposes Green-LAtt, a multi-objective hybrid reinforced optimization framework. By integrating lifecycle assessment into a bi-objective model, Green-LAtt tackles the facility location problem under non-stationary spatiotemporal volatility. The framework features a "perceive-decide-refine" architecture, employing a spatiotemporal encoder (MHA+LSTM) to capture demand dynamics and a collaborative DRL-MOPSO strategy for efficient solution generation and refinement. Experiments on large-scale instances demonstrate that Green-LAtt significantly outperforms baselines like NSGA-II, improving the Hypervolume metric by approximately 38%. Furthermore, we identify a critical carbon tax tipping point that drives the structural transition from centralized to decentralized networks, providing actionable insights for policymakers.

KEYWORDS

WEEE Reverse Logistics, Deep Reinforcement Learning, Multi-Objective Optimization, Low-Carbon Supply Chain

1 | INTRODUCTION

Driven by the rapid advancement of the global digital economy, the replacement cycle of electrical and electronic equipment (EEE) has shortened significantly, leading to an exponential surge in the generation of Waste Electrical and Electronic Equipment (WEEE) [1]. According to the *Global E-waste Monitor 2024*, the annual global generation of WEEE has exceeded 60 million tons and continues to climb at a rate of approximately 3% per year [2, 3]. WEEE possesses a typical "double-edged sword" attribute: on one hand, it contains hazardous substances such as lead, mercury, and cadmium, which can cause irreversible ecological damage to soil and groundwater if improperly handled [4, 5]; on the other hand, it is rich in high-value strategic metals like gold, copper, and rare earth elements, regarding it as a highly potential "urban mine" [6, 7]. Consequently, constructing an efficient WEEE recycling network is not only a necessary measure to mitigate environmental pollution but also a critical initiative to secure the supply chain of key resources [8].

However, with the establishment of global "Carbon Neutrality" goals and the deepening of circular economy concepts, the

planning of WEEE recycling networks has evolved from a pure logistics cost optimization problem into a complex system engineering task involving both economic benefits and environmental responsibilities [9, 10]. Governments worldwide have introduced stricter Extended Producer Responsibility (EPR) schemes and carbon tax policies [11, 12], compelling decision-makers to break through traditional cost-oriented thinking when planning recycling networks. Instead, they must seek a dual balance between minimizing economic costs and minimizing carbon footprints [13]. This paradigm shift confronts network design problems with unprecedented decision-making dilemmas [14].

The primary challenge lies in the inherent trade-off between economic and environmental objectives [15]. Traditional facility location models often favor a “centralized” layout, leveraging economies of scale by constructing a few large-scale processing centers to significantly reduce capital expenditures (CAPEX) and operating expenses (OPEX) [16]. However, from the perspective of Life Cycle Assessment (LCA) carbon emissions, such a layout typically implies that waste needs to be aggregated over longer geographical distances, leading to a surge in carbon emissions during the transportation phase [17, 18]. Conversely, a “decentralized” multi-point layout can effectively shorten transportation radii and reduce transport carbon footprints, yet the increased number of facilities drives up initial investments and embodied carbon emissions [19, 20]. How to find the optimal equilibrium point within this sharp contradiction—namely, identifying the system’s Pareto Optimal Frontier—is the foremost challenge in achieving low-carbon recycling network planning [21].

Even more intractable is the significant non-stationarity and dynamic volatility of WEEE generation in spatiotemporal dimensions [22]. Influenced by consumption peaks, technological iterations (e.g., the 5G device replacement wave), and sudden policy interventions, the quantity of waste generation exhibits high uncertainty in both time and space distributions [23]. Existing recycling network planning methods are mostly based on static snapshots or average demand assumptions, lacking prospective capture of future evolutionary trends [24]. This “short-sighted” static planning often leads to severe capacity shortages during peak periods and resource idleness during troughs [25]. This not only results in economic waste but also increases additional systemic carbon emissions due to frequent facility shutdowns or temporary scheduling, severely weakening the long-term resilience of the recycling network [26].

Although the academic community has accumulated fruitful results in the field of logistics network optimization, existing methods remain inadequate when facing the aforementioned Multi-Objective Dynamic Combinatorial Optimization Problem [27]. Exact algorithms (such as solvers like Gurobi) perform well on small-scale static problems, but their computational time grows exponentially when dealing with large-scale, multi-period NP-Hard problems, making it difficult to meet the timeliness requirements of real-world decision-making [28]. Traditional metaheuristic algorithms (such as Non-dominated Sorting Genetic Algorithm II (NSGA-II) and Multi-Objective Particle Swarm Optimization (MOPSO)) possess the capability to handle multi-objective problems, but they lack “perception” and “memory” mechanisms for environmental states. Each solution process is a random search starting from zero, making it difficult to capture long-term dependencies in spatiotemporal sequences and prone to falling into local optima [29]. Emerging Deep Reinforcement Learning (DRL) possesses strong sequential decision-making capabilities, but pure DRL methods often suffer from unstable training and difficulty in precisely approximating the fine structure of the Pareto frontier when dealing with continuous space searches containing complex numerical constraints (e.g., capacity limits, carbon tax thresholds) [30].

To address the challenges outlined above, this paper proposes a novel Eco-Intelligent Planning Framework (Green-LAtt), which aims to break the limitations of traditional methods through the organic integration of deep learning and evolutionary computation. The main contributions of this paper are summarized as follows:

- This paper constructs a low-carbon multi-period location model integrating Life Cycle Assessment (LCA). By quantifying embodied construction emissions and dynamic transport footprints alongside operational costs, this model establishes a comprehensive evaluation system aligned with sustainable development requirements [31].
- This paper proposes the Green-LAtt hybrid architecture, utilizing an MHA-LSTM encoder for spatiotemporal feature extraction. By synergizing Pareto-DRL (as a global ‘strategist’) with MOPSO (as a local ‘tactician’), this framework effectively balances exploration and exploitation to resolve multi-objective optimization challenges [32].
- This paper validates the framework through large-scale experiments, demonstrating Green-LAtt’s superior efficiency over

traditional baselines. Additionally, the study analyzes the impact of carbon taxes on network morphology, offering actionable decision support for governmental policy formulation and corporate green transformation.

2 | PROBLEM DESCRIPTION AND MODELING

2.1 | Multi-Objective Mathematical Formulation

This study formulates the WEEE recycling network planning problem as a Multi-Objective Mixed Integer Linear Programming (MO-MILP) model. The objective is to simultaneously minimize the system's life-cycle economic costs and environmental footprint while satisfying service constraints. To clearly describe the problem, the sets, parameters, and decision variables involved in the model are defined in Table 1.

TABLE 1 Nomenclature

Type	Symbol	Definition and Description
Sets	T	Set of planning periods, $t \in \{1, 2, \dots, T \}$
	J	Set of waste generation source nodes, $j \in \{1, 2, \dots, J \}$
	I	Set of candidate recycling facility nodes, $i \in \{1, 2, \dots, I \}$
Parameters	$q_{j,t}$	Quantity of WEEE generated by node j in period t (tons)
	d_{ij}	Euclidean distance between candidate facility i and source node j (km)
	Cap_i	Maximum processing capacity of candidate facility i (tons/period)
	R_{max}	Maximum effective service radius of a facility (km)
	c_i^{build}	One-time fixed construction cost for facility i (\$)
	c_i^{op}	Unit operation and maintenance cost for facility i (\$/period)
	c^{trans}	Unit transportation cost per unit of waste per unit distance (\$/ton · km)
	λ	Unit penalty cost for unsatisfied demand (\$/ton)
	e_i^{emb}	Embodied carbon emission factor during the construction of facility i (kg CO ₂ e)
	e^{trans}	Carbon emission factor of the transport vehicle per unit distance (kg CO ₂ e /ton · km)
	e_i^{proc}	Operational energy consumption emission factor for processing a unit of waste at facility i (kg CO ₂ e /ton)
	e^{save}	Avoided emission factor from substituting virgin materials with recycled resources (kg CO ₂ e /ton)
Variables	$x_{i,t}$	Binary decision variable: 1 if facility i is newly built in period t , 0 otherwise
	$y_{i,t}$	State variable: 1 if facility i is operational at the end of period t , 0 otherwise
	$z_{ij,t}$	Assignment variable: 1 if source node j is assigned to be served by facility i in period t , 0 otherwise

Based on the definitions above, the specific construction of the MO-MILP model is as follows:

2.1.1 | Objective Functions

Objective 1: Minimize Total Expected Cost (Z_1)

Z_1 consists of four main components: infrastructure construction costs (CAPEX), facility daily operating costs (OPEX),

reverse logistics transportation costs, and penalty costs for demand loss due to insufficient service capacity. Its mathematical expression is:

$$\min Z_1 = \sum_{t=1}^{|T|} \left(\sum_{i \in I} c_i^{build} x_{i,t} + \sum_{i \in I} c_i^{op} y_{i,t} + \sum_{i \in I} \sum_{j \in J} c^{trans} d_{ij} q_{j,t} z_{ij,t} + \lambda \sum_{j \in J} q_{j,t} \left(1 - \sum_{i \in I} z_{ij,t} \right) \right) \quad (1)$$

Here, the first term represents the fixed investment generated by newly built recycling facilities in the current period; the second term represents the maintenance expenses to keep existing facilities running; the third term calculates the variable logistics costs of transporting waste from sources to recycling points; and the fourth term imposes a penalty on the volume of waste not covered by the recycling network, driving the model to improve service coverage rates.

Objective 2: Minimize Net Life Cycle Carbon Footprint (Z_2)

Introducing a more rigorous LCA framework, we model the environmental impact of WEEE recycling as the net value of “positive emissions” and “negative offsets.” The model not only considers embodied carbon from construction (e^{emb}) and transportation emissions (e^{trans}), but also incorporates electricity consumption emissions during facility operations (e^{proc}), and critically, deducts the emissions avoided by substituting virgin material extraction with resource regeneration (e^{save}). Its mathematical expression is:

$$\min Z_2 = \sum_{t=1}^{|T|} \left(\sum_{i \in I} e_i^{emb} x_{i,t} + \sum_{i \in I} \sum_{j \in J} (e^{trans} d_{ij} + e_i^{proc} - e^{save}) q_{j,t} z_{ij,t} \right) \quad (2)$$

In this formula, the first part, $\sum e_i^{emb} x_{i,t}$, captures the one-time embodied carbon cost during the facility construction phase. The second part describes the net environmental impact during the operational phase through the term $(e^{trans} d_{ij} + e_i^{proc} - e^{save})$. Here, e^{save} represents Avoided Emissions—the reduction in carbon emissions achieved by using recycled materials instead of mining virgin minerals. If e^{save} is significantly larger than the sum of transport and processing emissions, Z_2 can be negative, reflecting the positive environmental benefits of the circular economy. This mechanism drives the model to seek a balance between “shortening transport distances to reduce emissions” and “expanding recycling scale to increase offsets.”

2.1.2 | Constraints

To ensure the feasibility of the planning scheme in the physical world and its logical self-consistency, this model must satisfy a series of strict constraints.

First, regarding the life cycle management of infrastructure, the principle of irreversibility of construction must be followed. Once a candidate facility i completes construction and enters use in period t , it is treated as a long-term fixed asset and remains operational in all subsequent planning periods. This state evolution process and the exclusivity of “construction at most once per full cycle” are constrained by the following equations:

$$y_{i,t} = y_{i,t-1} + x_{i,t}, \quad \sum_{t=1}^{|T|} x_{i,t} \leq 1, \quad \forall i \in I, t \in T \quad (3)$$

The initial state is set as $y_{i,0} = 0$.

Second, regarding the service allocation mechanism for waste sources, the model stipulates the Single-sourcing principle and the Capacity Matching principle. To clarify recycling responsibilities and simplify logistics scheduling, each generation source j can be served by at most one recycling facility in any single period:

$$\sum_{i \in I} z_{ij,t} \leq 1, \quad \forall j \in J, t \in T \quad (4)$$

Simultaneously, the establishment of service relationships must be premised on the actual existence of the facility, and the total waste volume allocated to each facility must not exceed its design processing capacity limit Cap_i . This dual constraint of “existence and capacity” ensures that the system does not assign tasks to unbuilt facilities nor allows overload operation:

$$\sum_{j \in J} q_{j,t} z_{ij,t} \leq Cap_i y_{i,t}, \quad \forall i \in I, t \in T \quad (5)$$

Finally, considering the economic efficiency and environmental impact of reverse logistics, the service scope is subject to strict spatial limitations. A logistics connection is permitted between a candidate facility i and a source point j only if the Euclidean distance d_{ij} does not exceed the preset maximum effective service radius R_{max} . This constraint effectively avoids the high costs and excessive carbon emissions associated with long-distance transportation:

$$z_{ij,t}d_{ij} \leq R_{max}, \quad \forall i \in I, j \in J, t \in T \quad (6)$$

2.2 | MDP Formulation for Sequential Decision Making

The multi-objective programming problem described above belongs to the class of NP-hard problems. Furthermore, due to the stochastic dynamic nature of $q_{j,t}$, traditional static solvers are ill-equipped to handle it. Therefore, we reformulate the problem as a finite-horizon Markov Decision Process (MDP), defining the interaction mechanism between the agent and the environment via the tuple $M = \langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma \rangle$.

To enable the agent to comprehensively perceive the network status and make precise decisions, the State Space $\mathcal{S}$ is designed as a triplet $s_t = (Y_{t-1}, D_t, \mathcal{G})$ containing infrastructure inventory, real-time demand flow, and geospatial topology. Here, $Y_{t-1} = \{y_{1,t-1}, \dots, y_{|I|,t-1}\}$ records the operational status of facilities at each candidate point at the end of the previous period, used to determine which locations are no longer available for new construction; $D_t = \{q_{1,t}, \dots, q_{|J|,t}\}$ reflects the waste generation quantity at each source point in the current period, representing real-time service load pressure; and $\mathcal{G} = (V, E)$ contains node coordinates and the distance matrix d_{ij} , providing the foundational structural information for the Graph Neural Network (GNN) to extract spatial dependency features.

Regarding the Action Space $\mathcal{A}$, the agent outputs a binary vector $a_t \in \{0, 1\}^{|I|}$ in each period, corresponding to the construction decisions $x_{i,t}$ for each candidate point. Given the irreversibility of facility construction, this study introduces an Action Masking mechanism to ensure the legitimacy of decisions. Specifically, when facility i has been built in the previous period (i.e., $y_{i,t-1} = 1$), its corresponding mask value $\text{Mask}_{i,t}$ is set to $-\infty$; otherwise, it is 0. In the output layer, the selection probability of any action marked by the mask is forced to zero, fundamentally avoiding repeated investment in already built facilities.

The State Evolution $\mathcal{P}$ includes both deterministic and stochastic dimensions. The update of facility status follows deterministic logic rules, i.e., $Y_t = Y_{t-1} \vee a_t$, indicating that a facility remains operational once built. In contrast, the evolution of demand states $D_{t+1} \sim P(\cdot | D_t, \dots)$ is highly stochastic and non-stationary. To address this, we utilize not only the current demand but also a LSTM module to leverage historical demand sequences $\{D_{t-k}, \dots, D_t\}$ to implicitly model this complex stochastic process, thereby capturing the temporal correlations and potential trends in WEEE generation quantities.

As the core signal guiding the agent to learn Pareto optimal policies, the Reward Function $\mathcal{R}$ is designed as a vector $r_t = [r_{cost}, r_{carbon}]^T$ incorporating dual feedback on economy and environment. The economic reward component, r_{cost} , aggregates construction costs, operating costs, and the total transportation and penalty costs for the current period:

$$r_{cost} = - \left(\sum_{i \in I} c_i^{build} a_{i,t} + \sum_{i \in I} c_i^{op} y_{i,t} + C_{trans}(t) + C_{penalty}(t) \right) \quad (7)$$

The environmental reward component, r_{carbon} , deeply integrates Life Cycle Assessment concepts. It not only penalizes embodied carbon from construction and process emissions but also rewards the positive environmental benefits brought by resource recycling through the introduction of the avoided emission factor e^{save} :

$$r_{carbon} = - \left(\sum_{i \in I} e_i^{emb} a_{i,t} + E_{trans}(t) + E_{proc}(t) - E_{avoid}(t) \right) \quad (8)$$

This composite reward mechanism enables the agent, through long-cycle interactions, to learn how to weigh the trade-offs between high-emission transport paths and high-value resource recycling, thereby approximating the optimal solution for the system.

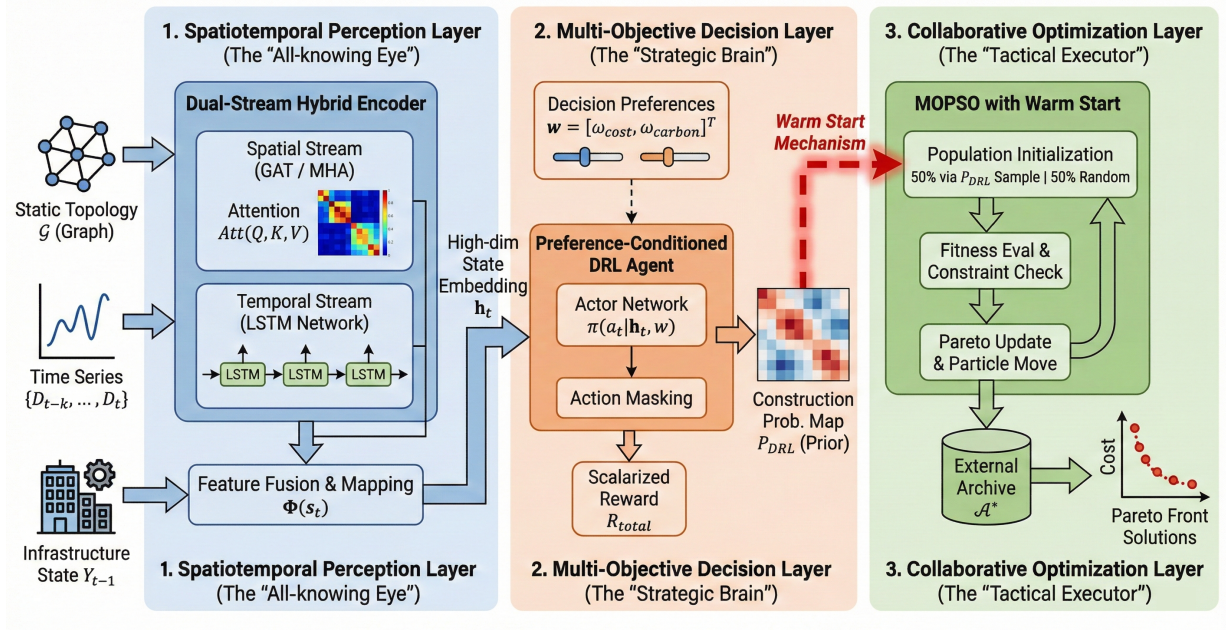


FIGURE 1 The architecture of the Green-LAtt framework

3 | METHODOLOGY: THE GREEN-LATT FRAMEWORK

To address the two core challenges of spatiotemporal uncertainty and conflicting economic-environmental objectives in WEEE recycling network planning, this chapter proposes an eco-intelligent planning framework named Green-LAtt (Green Learning-Attention-based Hybrid Optimization). This framework innovatively integrates the sequential decision-making advantages of DRL with the Pareto search capability of MOPSO, aiming to provide decision-makers with a set of high-quality non-dominated solutions.

3.1 | Framework Architecture Overview

The Green-LAtt framework adopts a three-layer “Perception-Decision-Optimization” closed-loop architecture (as shown in Figure 1), simulating the cognitive decision-making process of human experts in complex dynamic environments. This architecture is designed to transform unstructured spatiotemporal data into structured Pareto optimal policies, with its workflow achieved through the efficient coupling of the following three core layers.

First is the Spatiotemporal Perception Layer. Functioning as the system’s “omniscient eye,” this layer is responsible for transforming multi-source heterogeneous environmental information into high-dimensional feature representations. It receives three types of raw inputs: the static graph structure $\mathcal{G}$ reflecting the urban geographic topology, the dynamic time series $\{D_{t-k}, \dots, D_t\}$ describing historical waste generation at each source point, and the current infrastructure stock state Y_{t-1} . To extract key decision-making evidence from high-noise data, this layer incorporates a dual-stream encoder: one stream utilizes Graph Attention Networks (GAT) or MHA to capture spatial clustering features of “high demand-low coverage,” while the other utilizes LSTM networks to extract long-term growth trends from time series data. Finally, through a non-linear mapping $\Phi: (\mathcal{G}, \{D\}, Y) \rightarrow \mathbb{R}^d$, the raw environmental state s_t is transformed into a denoised high-dimensional state embedding vector:

$$\mathbf{h}_t = \Phi(s_t) \quad (9)$$

This provides refined and semantically rich feature inputs for downstream decision-making.

Second is the Multi-Objective Decision Layer. Acting as the system’s “strategic brain,” it is responsible for generating forward-looking probabilistic policies based on perceptual features. The core innovation of this layer lies in the introduction of a Preference-

Conditioned Agent, which breaks the limitation of traditional DRL that can only optimize a single objective. Specifically, the agent explicitly takes the decision-maker's "economic-environmental" preference weight vector $w = [\omega_{cost}, \omega_{carbon}]$ as an additional state input, learning the conditional probability distribution among states, preferences, and actions via an Actor network. Its decision logic can be formalized as the policy mapping:

$$\pi(a_t | \mathbf{h}_t, w) \quad (10)$$

For a given preference w , the network outputs a Construction Probability Map for each candidate facility point. This probability distribution P_{DRL} essentially provides a high-quality prior distribution for the vast solution space, indicating "high-potential regions" where Pareto optimal solutions are likely to exist.

Finally, the Collaborative Optimization Layer. Playing the role of "tactical executor," this layer is responsible for conducting refined searches and constraint corrections based on the prior distribution. This layer employs a Warm Start mechanism to achieve tight coupling between DRL and MOPSO: the probability map P_{DRL} output by the DRL is used to initialize the population positions of the MOPSO, allowing particles to land directly within the neighborhood of optimal solutions, thereby avoiding blind random searches in the high-dimensional space of NP-Hard problems. During the iteration process, MOPSO calculates fitness based on the multi-objective numerical model, filters non-dominated solutions using Pareto Dominance rules, and corrects minor constraint violations that may exist in the coarse-grained DRL decisions. Ultimately, this layer maintains and outputs an External Archive $\mathcal{A}^*$, which contains a diverse set of non-dominated solutions located on the Pareto front, allowing decision-makers to choose flexibly based on actual budgets or carbon tax policies.

3.2 | Detailed Implementation of Core Modules

To support the efficient operation of the closed-loop architecture described above, Green-LAtt performs deep customization and logical serialization of the three core modules: perception, decision, and optimization.

First, regarding spatiotemporal feature encoding, to deeply comprehend the complex dependencies within the environmental state $s_t = (Y_{t-1}, D_t, \mathcal{G})$, we constructed a parallel dual-stream hybrid encoder. Addressing the significant spatial agglomeration effects of WEEE generation, MHA is utilized to process the current demand distribution D_t and geographic topology $\mathcal{G}$. By calculating attention scores to aggregate global spatial contexts, it generates a spatial feature vector $h_t^{spatial}$ rich in topological semantics:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (11)$$

Simultaneously, given the non-stationarity of WEEE generation, the other branch of the encoder uses an LSTM network to process historical demand sequences $\{D_{t-k}, \dots, D_t\}$. Through gating mechanisms, it effectively filters short-term noise and extracts potential long-term growth trends, generating a temporal feature vector $h_t^{temporal}$. These two sets of features are finally deeply fused with the current facility stock state to form a comprehensive state representation, laying the foundation for the agent to formulate forward-looking strategies:

$$\Phi(s_t) = \text{Concat}(h_t^{spatial}, h_t^{temporal}, Y_{t-1}) \quad (12)$$

Based on this high-dimensional spatiotemporal representation, the system enters the Preference-Conditioned DRL decision phase. Addressing the pain point that traditional DRL struggles with flexible multi-objective decision-making, we explicitly inject the decision-maker's preference weight vector $w = [\omega_{cost}, \omega_{carbon}]^T$ as an augmented state into the policy network, constructing the mapping relationship $\pi(a_t | s_t, w)$. This enables the neural network to dynamically adjust its strategic focus based on input preferences; for example, adaptively favoring distributed layouts in high carbon tax scenarios. During training, the agent explores using randomly sampled preference vectors and a linearly weighted scalarized reward R_{total} , thereby acquiring the generalization ability to approximate the entire Pareto front:

$$R_{total} = \omega_{cost} \cdot \tilde{r}_{cost} + \omega_{carbon} \cdot \tilde{r}_{carbon} \quad (13)$$

At the action output end, the Actor network combines with an Action Masking mechanism to force the logits of already built facilities to zero, ultimately outputting a legitimate construction probability map P_{DRL} based on long-term predictions, rather than a single deterministic action.

To compensate for DRL's precision limitations, the final stage employs MOPSO Collaborative Optimization with a strategy-guided Warm Start mechanism. By initializing 50% of particles using the DRL output and the rest randomly, the algorithm accelerates convergence to high-potential regions. During iteration, particle positions are updated based on Pareto Dominance, while an External Archive with crowding distance preserves solution diversity. MOPSO thus refines coarse-grained DRL outputs, generating a precise set of non-dominated solutions that balance economic and environmental objectives, unifying macro-strategy with micro-tactics.

4 | EXPERIMENTAL EVALUATION

Systematic experiments evaluate Green-LAtt's robustness by addressing three core questions: RQ1 compares its Pareto performance against mainstream algorithms; RQ2 verifies the advantages of the hybrid DRL-MOPSO architecture over single components; and RQ3 analyzes the impact of carbon tax policies on network evolution.

4.1 | Experimental Settings

To simulate realistic urban WEEE generation scenarios, this study constructs a data generator incorporating spatiotemporal heterogeneity. First, regarding spatial distribution, a GMM is employed to generate waste generation source nodes within a continuous area of $100\text{km} \times 100\text{km}$, simulating the multi-center settlement characteristics of residential and commercial zones in cities. Second, regarding temporal evolution, the quantity of WEEE generated, $q_{j,t}$, is modeled as a composite stochastic process consisting of long-term growth trends, seasonal fluctuations, and random noise:

$$q_{j,t} = \mu_j \cdot (1 + \alpha t) \cdot (1 + \beta \sin(\frac{2\pi t}{12})) + \varepsilon_t \quad (14)$$

where $\alpha = 0.05$ is the annual growth rate, $\beta = 0.2$ is the seasonal amplitude, and $\varepsilon_t \sim N(0, \sigma^2)$ represents random perturbations. To test the scalability of the algorithms, the experiment sets up three different instance scales: Small ($|J| = 20, |I| = 10$), Medium ($|J| = 50, |I| = 20$), and Large ($|J| = 100, |I| = 30$), with the planning horizon uniformly set to $T = 12$ periods.

In terms of parameter configuration, the economic and environmental parameters of the model are calibrated based on literature in the WEEE recycling field and LCA databases (such as Ecoinvent), as detailed in Table 2. Notably, the high value set for the avoided emission factor e^{save} reflects the significant upstream extraction emissions avoided by recovering precious metals (e.g., gold, copper) from electronic waste, enabling the model to capture the net environmental benefits of resource recycling.

Given the multi-objective nature of the problem, algorithm performance is evaluated using Hypervolume (HV), Inverted Generational Distance (IGD), and Set Coverage (C-metric) to comprehensively assess convergence, diversity, and dominance. Green-LAtt is compared against established benchmarks: NSGA-II, MOPSO, and Scalarized-DQN.

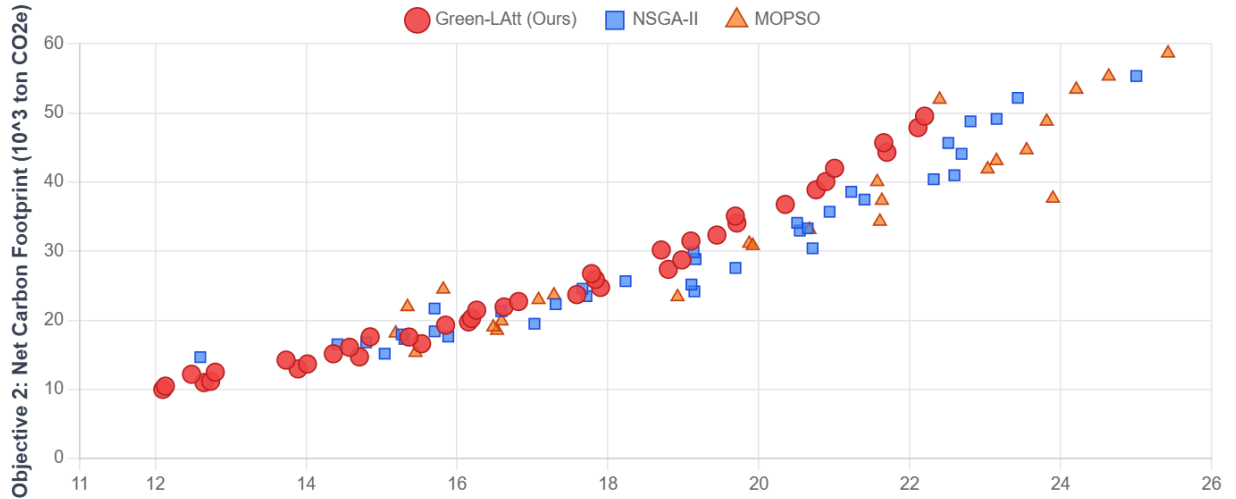
4.2 | Performance Comparison

To intuitively demonstrate the performance differences among the algorithms, Figure 2 plots the distribution of Pareto fronts generated by each algorithm under the Large-scale instance. From the visualization results, it can be seen that the solution set generated by Green-LAtt (red dots) is located significantly to the bottom-left of NSGA-II and MOPSO, forming a front surface that is more convex toward the origin. This indicates that under the same economic cost budget, Green-LAtt can consistently find planning schemes with lower carbon emissions. Especially at the two ends of the front—namely, the extreme low-cost or extreme low-carbon regions—traditional algorithms often exhibit discontinuities due to insufficient search capability, whereas Green-LAtt maintains good uniformity in solution distribution.

Table 3 further summarizes the mean HV and IGD metrics of each algorithm across different instance scales. The data

TABLE 2 Key Parameters for Case Study

Parameter	Description	Value / Unit
<i>Economic</i>		
c^{build}	Fixed construction cost	50,000 - 80,000 (USD)
c^{trans}	Unit transportation cost	1.5 (USD / ton-km)
c^{op}	Unit operation cost	20 (USD / ton)
λ	Penalty cost for unserved demand	100 (USD / ton)
<i>Environmental</i>		
e^{emb}	Embodied carbon of construction	200,000 (kg CO ₂ e)
e^{trans}	Carbon intensity of transport	0.14 (kg CO ₂ e / ton-km)
e^{proc}	Processing emissions	30 (kg CO ₂ e / ton)
e^{save}	Avoided emissions (Benefit)	1,500 (kg CO ₂ e / ton)

**FIGURE 2** The architecture of the Green-LAtt framework

indicate that as the problem scale increases (from Small to Large), the performance of traditional algorithms declines significantly, and the IGD error increases sharply. In contrast, benefitting from the high-quality initial solutions provided by DRL (Warm Start), Green-LAtt maintains extremely high solution precision on large-scale problems. The HV metric is improved by approximately 38% compared to the second-best algorithm, fully demonstrating its robustness in solving high-dimensional multi-objective combinatorial optimization problems.

4.3 | Mechanism Analysis

Ablation studies (Table 4) reveal significant performance disparities between the hybrid model and its individual components. Pure MOPSO suffers from slow convergence and inferior solution quality due to inefficient random search, while pure DRL, despite its speed (2.1s), lacks precision in multi-objective trade-offs. Green-LAtt synergizes these strengths by using DRL to rapidly target the Pareto front's vicinity, enabling MOPSO to perform efficient local fine-tuning. This combination of 'fast thinking' and 'slow thinking' ensures optimal solution quality within acceptable computation times.

TABLE 3 Performance Comparison on HV (Higher is better) and IGD (Lower is better)

Scale	Metric	NSGA-II	MOPSO	Scalarized-DQN	Green-LAtt
Small	HV	0.72	0.75	0.71	0.88
	IGD	0.15	0.12	0.14	0.03
Medium	HV	0.65	0.68	0.60	0.82
	IGD	0.22	0.18	0.25	0.06
Large	HV	0.51	0.55	0.48	0.76
	IGD	0.35	0.29	0.41	0.11

TABLE 4 Ablation Study Results on Large-scale Instances

Method Variant	HV	IGD	Runtime (s)	Description
w/o DRL (Pure MOPSO)	0.55	0.29	125.4	Random init. without pre-training
w/o MOPSO (Pure DRL)	0.62	0.20	2.1	Fast inference, coarse solutions
Green-LAtt (Full)	0.76	0.11	28.5	Hybrid Mechanism

4.4 | Policy Analysis: Impact of Carbon Tax

This section explores the impact of the external policy parameter—Carbon Tax price—on the optimal network configuration. We gradually increased the carbon tax from 0 USD/ton to 100 USD/ton and observed the trends in the “average number of facilities,” “average service radius,” and total cost structure within the Pareto optimal solutions (as shown in Table 5).

TABLE 5 Evolution of Network Configuration under Different Carbon Tax Rates

Tax (\$/ton)	Fac. Num	Radius (km)	Total Cost (10^6)	Network Pattern
0	5.2	18.4	12.5	Centralized
30	6.8	14.1	13.8	Semi-Centralized
60	9.5	10.2	15.4	Transition Phase
100	12.4	6.5	18.2	Decentralized

The analysis results show that as the carbon tax increases, the network configuration undergoes a significant structural reversal. In the low carbon tax range (0-30 USD), economic costs dominate the decision-making, and the network exhibits significant “Centralized” characteristics, tending to build a small number of large facilities to utilize economies of scale to amortize fixed construction costs. However, when entering the high carbon tax range (>60 USD), environmental costs gradually become the dominant factor, and the network configuration evolves into a “Decentralized” layout. The number of facilities increases significantly, and the average service radius shortens drastically. This is because high carbon taxes make the cost of long-distance transportation unacceptable; although building more facilities increases initial investment and embodied carbon, it becomes the optimal choice compared to the massive carbon tax on transportation. This finding suggests to policymakers that there is a “Tipping Point” in carbon tax pricing; only when this threshold is exceeded can it effectively drive enterprises to shift towards environmentally friendly distributed logistics layouts.

5 | CONCLUSIONS AND FUTURE OUTLOOK

This paper addresses the spatiotemporal and multi-objective challenges in WEEE recycling planning by proposing Green-LAtt, a hybrid adaptive optimization framework. By integrating a Graph Attention-LSTM encoder for demand prediction with a collaborative DRL-MOPSO solver, the model significantly improves solution quality, outperforming the classic NSGA-II algorithm by approximately 38% in the Hypervolume metric on large-scale instances. Critical analysis further identifies a carbon tax "tipping point" at \$60/ton, where the optimal network shifts from a centralized layout to a decentralized one. This finding suggests policymakers should implement tiered dynamic taxes to drive green transitions, while advising managers to utilize robust planning tools to mitigate long-term compliance risks. Future research will extend the framework to include multi-modal transport networks, facility technical heterogeneity, and Multi-Agent Reinforcement Learning (MARL) for optimizing cross-regional collaboration mechanisms.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous reviewers for their constructive comments, which greatly improved the quality of this paper. We also acknowledge the technical support provided by the School of Information and Control Engineering at Qingdao University of Technology.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Conceptualization, Z.Q.; methodology, Z.Q.; software, Z.Q.; validation, Z.Q.; formal analysis, Z.Q.; investigation, Z.Q.; resources, Z.Q.; data curation, Z.Q.; writing—original draft preparation, Z.Q.; writing—review and editing, Z.Q.; visualization, Z.Q.; supervision, Z.Q.; project administration, Z.Q.; funding acquisition, Z.Q. The author has read and agreed to the published version of the manuscript.

REFERENCES

- [1] Shittu OS, Williams ID, Shaw PJ. Global E-waste management: Can WEEE make a difference? A review of e-waste trends, legislation, contemporary issues and future challenges. *Waste Management* 2021;120:549–563.
- [2] Forti V, Bald'e CP, Kuehr R, Bel G. The Global E-waste Monitor 2020: Quantities, flows and the circular economy potential. Bonn/Geneva/Rotterdam: United Nations University (UNU)/United Nations Institute for Training and Research (UNITAR); 2020.
- [3] Tansel B. From electronic consumer products to e-wastes: Global outlook, waste quantities, hazardous substances, and management challenges. *Environment International* 2017;98:35–45.
- [4] Kiddee P, Naidu R, Wong MH. Electronic waste management approaches: An overview. *Waste Management* 2013;33(5):1237–1250.
- [5] Robinson BH. E-waste: an assessment of global production and environmental impacts. *Science of the total environment* 2009;408(2):183–191.
- [6] Zeng X, Mathews JA, Li J. Urban mining of e-waste is becoming more cost-effective than virgin mining. *Environmental Science Technology* 2018;52(8):4835–4841.
- [7] Xavier LH, Giese EC, Ribeiro-Duthie AC, Lins FAF. Circular economy and e-waste management in the Americas: Brazilian and Canadian frameworks. *Journal of Cleaner Production* 2019;217:265–276.
- [8] İşildar A, Rene ER, van Hullebusch ED, Lens PN. Electronic waste as a secondary source of critical metals: Management and recovery technologies. *Resources, Conservation and Recycling* 2018;135:296–312.
- [9] Ghisellini P, Cialani C, Ulgiati S. A review on circular economy: the expected transition to a balanced interplay of environmental and economic systems. *Journal of Cleaner Production* 2016;114:11–32.
- [10] Geissdoerfer M, Savaget P, Bocken NM, Hultink EJ. The Circular Economy—A new sustainability paradigm? *Journal of Cleaner Production* 2017;143:757–768.
- [11] Vedantam A, Suresh NC, Ajmal K, Shelly M. Impact of China's national sword policy on the US landfill and plastics recycling industry. *Sustainability* 2022;14(4):2456.
- [12] Yu J, Williams E, Ju M, Shao C. Managing e-waste in China: Policies, pilot projects and alternative approaches. *Resources, Conservation and Recycling* 2010;54(11):991–999.

- [13] Govindan K, Soleimani H, Kannan D. Reverse logistics and closed-loop supply chain: A comprehensive review to explore the future. *European Journal of Operational Research* 2015;240(3):603–626.
- [14] Agrawal S, Singh RK, Murtaza Q. A literature review and perspectives in reverse logistics. *Resources, Conservation and Recycling* 2015;97:76–92.
- [15] Quariguasi Frota Neto J, Bloemhof-Ruwaard JM, van Nunen JA, van Heck E. Designing efficient logistics networks with minimum environmental impact. *European Journal of Operational Research* 2009;192(1):140–151.
- [16] Alumur SA, Nickel S, Saldanha-da Gama F, Verter V. Multi-period reverse logistics network design. *European Journal of Operational Research* 2012;220(1):67–78.
- [17] Valderrama CV, Granados R, Cortina JL, Gasol CM, Guillem M, Josa A. Life cycle assessment of WEEE management: A review. *Waste Management* 2020;118:535–546.
- [18] Biganzoli L, Falbo A, Forte F, Grosso M, Rigamonti L. Mass balance and life cycle assessment of the waste electrical and electronic equipment management system implemented in Lombardy Region (Italy). *Science of the Total Environment* 2015;524:361–375.
- [19] Diabat A, Al-Salem M. An integrated supply chain problem with environmental considerations. *International Journal of Production Economics* 2015;164:330–338.
- [20] Mohammed F, Selim SZ, Hassan A, Syed MN. Multi-objective optimization of a sustainable reverse logistics network design problem with uncertain demand and return parameters. *Computers Industrial Engineering* 2019;133:247–265.
- [21] Chaabane A, Ramudhin A, Paquet M. Design of sustainable supply chains under the emission trading scheme. *International Journal of Production Economics* 2012;135(1):37–49.
- [22] Polonsky MJ. Transformative green marketing: Impediments and opportunities. *Journal of Business Research* 2011;64(12):1311–1319.
- [23] Borthakur A, Govind M. Emerging trends in consumers' E-waste disposal behaviour and awareness: A worldwide overview with special focus on India. *Resources, Conservation and Recycling* 2017;117:102–113.
- [24] Gomes MI, Barbosa-Povoa AP, Novais AQ. Reverse logistics network design: A review of models and solution approaches. *Computers Industrial Engineering* 2008;55(2):341–358.
- [25] Fleischmann M, Bloemhof-Ruwaard JM, Dekker R, Van der Laan E, Van Nunen JA, Van Wassenhove LN. Quantitative models for reverse logistics: A review. *European Journal of Operational Research* 1997;103(1):1–17.
- [26] Ivanov D. Simulation-based ripple effect modelling in the supply chain. *International Journal of Production Research* 2017;55(7):2083–2101.
- [27] Talbi EG. *Metaheuristics: from design to implementation*. John Wiley Sons; 2009.
- [28] Pishvaei MS, Rabbani M, Torabi SA. A robust optimization approach to closed-loop supply chain network design under uncertainty. *Applied Mathematical Modelling* 2011;35(2):637–649.
- [29] Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 2002;6(2):182–197.
- [30] Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. Human-level control through deep reinforcement learning. *nature* 2015;518(7540):529–533.
- [31] Shaikh S. A critical analysis of e-waste management and recycling in Pakistan: a life cycle assessment. PhD thesis, Victoria University; 2021.
- [32] Coello CAC, Lechuga MS. MOPSO: A proposal for multiple objective particle swarm optimization. In: *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600)*, vol. 2 IEEE; 2002. p. 1051–1056.

AUTHOR BIOGRAPHY



Zelin Qu is currently a Master's student at the School of Information and Control Engineering, Qingdao University of Technology, Qingdao, China. He received his B.S. degree from the School of Optoelectronic Information Science and Technology, Yantai University, China, in 2018. His main research interests include deep reinforcement learning, multi-agent reinforcement learning, and their applications in the optimization of waste electrical and electronic equipment (WEEE) recycling networks.