



OPEN Adaptive airspace allocation model for urban drone logistics using multi-objective optimization under uncertainty

Yao Zhu✉, Xin Sun & Tongdi Hou

The large-scale application of urban unmanned aerial vehicle (UAV) logistics is confronted with challenges such as limited airspace resources, dynamic changes in demand, and multiple uncertainties. Traditional static allocation methods are difficult to adapt to complex urban environments. This study constructs a DRL-RO hybrid framework that integrates deep reinforcement learning and discrete robust optimization. It characterizes the influence mechanisms of demand fluctuations, weather changes, and emergencies through a three-layer uncertainty modeling system, and introduces a policy network enhanced by an attention mechanism to capture spatio-temporal correlations in the spatial domain. The improved MOEA/D-DRL algorithm is adopted to achieve rapid approximation of the Pareto frontier. The verification of the actual scenarios in Shenzhen shows that this framework reduces the computational complexity to the sub-quadratic level while maintaining a high success rate. Through the hierarchical airspace management strategy, it effectively balances the three goals of distribution efficiency, flight safety and operating costs. The Wasserstein sphere constraint ensures robustness and scalability in extreme scenarios. It provides theoretical support and technical solutions for the construction of city-level unmanned aerial vehicle (UAV) traffic management systems.

Keywords Urban unmanned Aerial Vehicle logistics, Multi-objective optimization, Deep reinforcement learning, Distributively robust optimization, Adaptive airspace allocation

With the acceleration of urbanization and the vigorous development of e-commerce, unmanned aerial vehicle (UAV) logistics has become an important solution to the “last mile” delivery problem in cities. Aweiss et al.¹ conducted the Unmanned Aircraft Systems Traffic Management (UTM) National Campaign II, demonstrating through systematic field trials that efficient airspace management infrastructure constitutes the foundational prerequisite for enabling large-scale UAV operations in the National Airspace System. Urban airspace, as a limited resource, faces unprecedented challenges when accommodating the increasing demand for unmanned aerial vehicle (UAV) logistics. López et al.² developed a comprehensive three-dimensional path planning framework addressing multi-UAV collision risk management, revealing that simultaneous consideration of spatial, temporal, and dynamic constraints significantly enhances operational safety in dense urban environments. Building upon this foundation, Hossen et al.³ introduced a flexible simulation framework that accelerates the development of UAV network digital twins, demonstrating that virtual-physical synchronization techniques reduce system development cycles by approximately 40% while improving predictive accuracy for airspace capacity assessment. However, the complexity and dynamics of urban airspace still restrict the large-scale application of unmanned aerial vehicle logistics⁴. The feasibility of urban air traffic not only depends on technological innovation, but also requires solving key issues such as operational efficiency, safety guarantee and capacity assessment^{5,6}.

The uncertain factors of unmanned aerial vehicle (UAV) operation in urban environments significantly affect the effectiveness and safety of airspace allocation. Han et al.⁷ established a Bayesian network-based risk assessment framework, quantifying that ground collision probabilities for urban logistics UAVs exhibit pronounced spatiotemporal heterogeneity with risk variations exceeding 300% across different urban zones and time periods. Sun et al.⁸ extended this work by integrating historical accident data, achieving 89% prediction accuracy for high-risk scenarios through data-driven probabilistic modeling. Rezaee et al.⁹ conducted a comprehensive review of collision avoidance schemes, identifying that existing methods face dual challenges of technical implementation and regulatory compliance, with only 23% of proposed algorithms meeting both

Business School, Yancheng Polytechnic College, Yancheng 224005, Jiangsu, China. ✉email: yphz5223@outlook.com

safety certification and computational efficiency requirements. Paul et al.¹⁰ addressed online demand uncertainty through stochastic optimization, demonstrating that adaptive planning algorithms reduce delivery delays by 18–25% compared to static scheduling approaches. Oh and Yoon¹¹ proposed a Pareto-based urban flight assessment framework, establishing multi-objective evaluation criteria that balance airspace utilization efficiency, safety margins, and environmental impacts.

Ramirez Atencia et al.¹² developed weighted multi-objective evolutionary algorithms for multi-UAV mission planning, achieving 92% task completion rates through dynamic priority adjustment mechanisms in scenarios involving up to 15 heterogeneous UAVs. Ribeiro et al.¹³ conducted a systematic analysis of 47 conflict resolution methods spanning both manned and unmanned aviation domains, revealing that distributed negotiation protocols outperform centralized approaches in scalability while maintaining equivalent safety levels for fleets exceeding 50 aircraft. Subsequent developments in optimization methodologies have explored diverse algorithmic paradigms: space-air-ground network resource allocation¹⁴, cooperative beamforming for energy optimization¹⁵, third-party risk-aware path planning¹⁶, and adaptive collision avoidance¹⁷. Meta-heuristic approaches including particle swarm optimization¹⁸, jellyfish-based algorithms¹⁹, and hybrid PID-heuristic methods²⁰ demonstrate the versatility of computational intelligence in UAV trajectory design, with systematic reviews²¹ cataloging over 150 algorithmic variants across spatial planning applications.

Uncertainty handling methodologies have evolved along three parallel tracks: weather-adaptive planning²², digital twin-based trajectory optimization²³, and stochastic path planning under dynamic risk fields²⁴, with reinforcement learning approaches²⁵ demonstrating robustness under communication-constrained scenarios. Application-oriented research addresses medical logistics through bidirectional pickup-delivery models²⁶, blockchain-enabled network management via DRL-based sharding^{27,28}, humanitarian operations through multi-stage stochastic programming²⁹, and connectivity-coverage trade-offs in multi-UAV coordination³⁰.

Although existing research has made progress in various aspects, critical technical bottlenecks persist in urban UAV airspace allocation under uncertain conditions. Current allocation methodologies predominantly employ static or semi-static frameworks that struggle to adapt to the rapid environmental changes characteristic of urban airspace, where traffic patterns, meteorological conditions, and operational demands fluctuate at sub-hourly intervals. Uncertainty treatment approaches typically address isolated factors—demand, weather, or emergencies—without capturing the complex interactions and propagation effects among multiple uncertainty sources. Multi-objective optimization strategies commonly rely on predetermined weight configurations based on expert judgment, limiting their ability to adjust dynamically to evolving operational contexts. This study develops an adaptive multi-objective optimization framework for urban UAV airspace allocation that comprehensively addresses these limitations. Through innovative integration of deep reinforcement learning and distributionally robust optimization, the framework achieves dynamic airspace management balancing real-time responsiveness with long-term optimality. Three principal contributions distinguish this work: (1) construction of a DRL-RO hybrid architecture that couples attention-enhanced policy networks with Wasserstein-constrained robust optimization, enabling simultaneous online learning and worst-case performance guarantees; (2) development of a systematic uncertainty modeling approach characterizing demand fluctuations, meteorological variations, and emergency scenarios through validated probability distributions with explicit cross-factor interaction mechanisms; (3) design of an improved MOEA/D-DRL algorithm achieving Pareto frontier approximation with sub-quadratic computational complexity, facilitating real-time multi-objective trade-off analysis across delivery efficiency, flight safety, and operational costs. These contributions provide theoretical foundations and computational solutions for scalable urban UAV logistics operations.

Methods

Problem modeling and mathematical framework

Urban unmanned aerial vehicle (UAV) airspace distribution problem is essentially a dynamic resource allocation problem in an uncertain environment, with complex characteristics of limited resources, time-varying demand, and multi-dimensional coupling. The three-dimensional spatial structure of urban environments and the existence of difference in communication network quality between the UAV and the traditional two-dimensional transportation network, along with increased degrees of freedom in route planning, bring greater challenges to airspace management. This study abstracts the urban airspace as a three-dimensional spatiotemporal network $G = (V, E, T)$, where $V = \{v_1, v_2, \dots, v_n\}$ represents the set of airspace nodes. Each node v_i corresponds to a three-dimensional spatial position (x_i, y_i, z_i) with time-varying capacity constraints $c_i(t)$; $E = \{e_{ij}\}$ represents the set of feasible flight segments between nodes, with flight segments' availability subject to buildings, no-fly zones, and weather conditions; $T = \{t_0, t_1, \dots, t_H\}$ is the discretized planning time domain, with time step Δt determined by the typical flight speed of UAVs and the urban airspace scale as 30 s.

The accuracy of airspace state definition is crucial to the design of subsequent optimization algorithms. Considering the complex terrain and building distribution of urban environments, the airspace state is defined as follows:

$$S(t) = \{(x, y, z, t) | x, y \in [0, L], z \in [z_{min}, z_{max}], t \in T\} \quad (1)$$

where L represents the edge length of the urban area (taking Shenzhen as an example, $L = 50\text{km}$), and the altitude range $[z_{min}, z_{max}] = [60\text{m}, 400\text{m}]$ complies with the Civil Aviation Administration of China's management regulations for urban low-altitude airspace. The airspace occupancy state is characterized by a three-dimensional occupancy matrix $O(t) = [o_{ijk}(t)]_{n \times m \times h}$, where $o_{ijk}(t) \in \{0, 1\}$ indicates the occupancy status of grid (i, j, k) at time t . The grid resolution is set to $100\text{ m} \times 100\text{ m} \times 20\text{ m}$ to balance computational efficiency and spatial accuracy. This configuration was established based on operational safety standards and computational feasibility considerations. The 100 m horizontal resolution represents a spatial

discretization unit approximately twice the minimum safety separation distance (50 m) specified in Eq. (3), enabling effective conflict detection while maintaining tractable state space dimensions for the optimization algorithms. The 20 m vertical stratification allows for multiple altitude layers within the regulatory airspace range of 60–400 m, supporting hierarchical traffic flow separation across different operational altitudes.

The accurate modeling of uncertainty parameters is a prerequisite for achieving robust airspace allocation. Through systematic analysis of urban UAV delivery scenarios, this study identifies three main sources of uncertainty and establishes corresponding conceptual models. The uncertainty parameter set Ξ covers multiple random factors affecting airspace allocation; demand uncertainty ξ_d follows a normal distribution, where the mean $\mu_d(t)$ exhibits obvious daily and weekly periodicity characteristics, and the standard deviation σ_d increases significantly during peak hours; weather uncertainty ξ_w adopts a Weibull distribution $W(\lambda, k)$, where the shape parameter k and scale parameter λ are obtained through maximum likelihood estimation of historical weather data, and this distribution can accurately portray the thick-tail characteristics of extreme weather events; emergency event ξ_e follows a non-homogeneous Poisson process $P(\lambda_e(t))$, with the intensity function $\lambda_e(t)$ being positively correlated with urban activity density.

Despite establishing probability distributions for individual uncertainty sources, the complex coupling relationships among multi-source uncertain factors and their nonlinear propagation mechanisms through airspace states require systematic mathematical characterization. To address this challenge, a three-layer uncertainty modeling framework integrating demand fluctuations, meteorological variations, and emergency events through their probability distributions was constructed, with Fig. 1 illustrating its theoretical structure and propagation mechanism.

Figure 1 presents the uncertainty modeling framework, with probability distribution models for demand, weather, and emergency events on the left, and the state transition function $T(s, a, \xi)$ on the right that maps random parameters to airspace states, establishing a complete uncertainty propagation mechanism.

The design of multi-objective optimization functions needs to balance the demands of efficiency, safety and economy. Delivery efficiency directly affects customer satisfaction and market competitiveness. Flight safety is the bottom-line requirement for system operation, while operating costs determine commercial feasibility. The multi-objective function is defined as:

$$\min F(X, \xi) = [f_1(X, \xi), f_2(X, \xi), f_3(X, \xi)]^T \quad (2)$$

The delivery time objective $f_1(X, \xi) = \sum_{i=1}^N w_i (t_i^{\text{arrival}} - t_i^{\text{request}})$ measures the system's response efficiency, where the weight w_i reflects the priority of orders; the collision risk objective is quantified through the expected value of collision probability:

$$f_2(X, \xi) = \sum_{i < j} \int_{\tau} P(\|p_i(\tau) - p_j(\tau)\| < d_{\text{saf}}) d\tau \quad (3)$$

where $p_i(\tau)$ represents the position trajectory of UAV i at time τ , $d_{\text{saf}} = 50\text{m}$ is the minimum safety separation; the operational cost objective comprehensively considers energy consumption, delay penalties, and route replanning costs:

$$f_3(X, \xi) = \sum_{i=1}^N (E_i \times c_e + D_i \times c_d + R_i \times c_r) \quad (4)$$

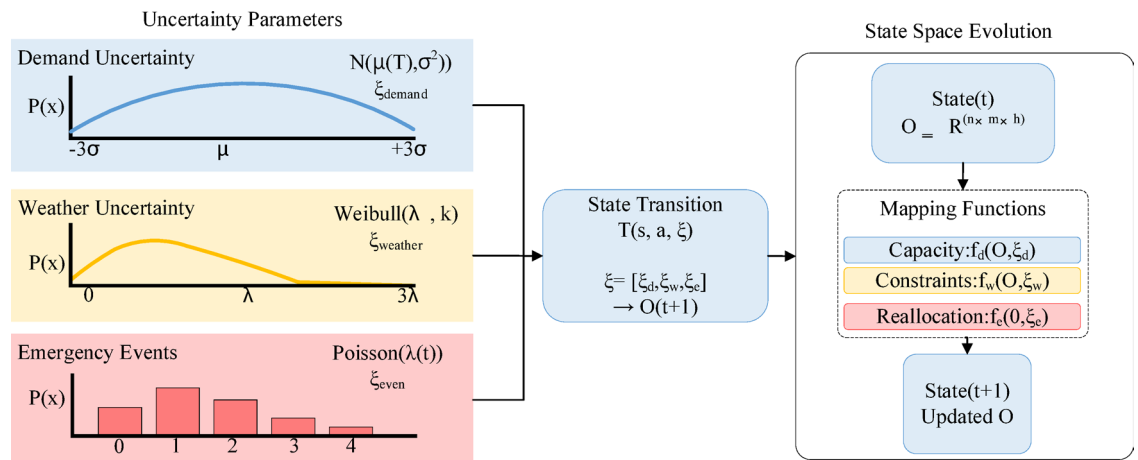


Fig. 1. Probability Distribution of Uncertain Parameters and Their Influence Mechanism on airspace Occupancy Rate.

Require: Initial policy parameters θ_0 , learning rate α , discount factor γ , airspace grid

(n, m, h) , uncertainty parameters $\xi = (\xi_d, \xi_w, \xi_e)$

Ensure: Optimized policy parameters θ^* and allocation strategy π^*

```

1: Initialize replay buffer  $\mathcal{D}$ , target network  $\theta' \leftarrow \theta_0$ 

2: Initialize airspace occupancy matrix  $O_0 \in \{0, 1\}^{n \times m \times h}$ 

3: for episode = 1 to  $M$  do

4: Observe initial state  $s_0 = O_0, Q_0, \xi_0$  where  $Q_0$  is task queue

5: for  $t = 0$  to  $T$  do
6: // Attention-based feature extraction (innovation)
7:  $\phi_t = \text{MHA}(\text{Enc}_s(O_t), \text{Enc}_q(Q_t))$ 

8: Select action:  $a_t \sim \pi(\cdot \| s_t; \theta)$  with  $\epsilon$ -greedy

9: // Multi-objective reward calculation (core innovation)

10: Execute allocation  $a_t$ : assign routes, heights, time slots

11:  $r_{\text{base}} = -\alpha \Delta f_1 - \beta \Delta f_2 - \gamma \Delta f_3$ 

12: if  $d_{ij} < d_{\text{safe}}$  then add safety penalty  $r_t \leftarrow r_{\text{base}} - \lambda e^{-d_{ij}/d_{\text{safe}}}$ 

13: // State transition with uncertainty (key component)

14:  $O_{t+1} = T(O_t, a_t, \xi_t)$  where  $\xi_t \sim P(\xi \| s_t)$ 

15: Store  $(s_t, a_t, r_t, s_{t+1})$  in  $\mathcal{D}$ 

16: // Network training

17: Sample mini-batch from  $\mathcal{D}$  and compute TD error:

18:  $\delta = r + \gamma \max_{a'} Q(s', a'; \theta') - Q(s, a; \theta)$ 

19: Update:  $\theta \leftarrow \theta + \alpha \cdot \nabla_{\theta} J(\theta)$ 

20: Periodically update target network:  $\theta' \leftarrow \theta$ 

21: end for

22: end for

23: return  $\theta^*, \pi^*(s) = \arg \max_a Q(s, a; \theta^*)$ 

```

Algorithm 1. DRL-based Adaptive Airspace Allocation

where E_i is the energy consumption (kWh), D_i is the delay time (minutes), R_i is the number of replanning operations, and c_e, c_d, c_r are the corresponding unit cost coefficients.

Adaptive allocation strategy design

The core of the adaptive airspace allocation strategy lies in achieving rapid response to dynamic environments and proactive adaptation to uncertainty. Traditional static allocation methods often exhibit poor adaptability when facing demand fluctuations and emergency events, leading to low airspace utilization and increased safety risks. The adaptive strategy proposed in this study achieves the unification of millisecond-level decision response and long-term performance optimization through collaborative design at three levels: deep reinforcement learning, distributionally robust optimization, and fast Pareto approximation.

The dynamic adjustment mechanism based on deep reinforcement learning models the airspace allocation problem as a Partially Observable Markov Decision Process (POMDP), achieving autonomous perception and decision-making of complex environmental patterns through end-to-end learning. The advantage of this approach is its ability to automatically extract effective features from historical operational data, avoiding the limitations of manually designed heuristic rules³¹. Novel energy-efficiency frameworks further enhance the adaptive capabilities of DRL-based systems in UAV networks³². The design of the state space \mathcal{S} comprehensively considers the system's observable information and hidden states, including the current airspace occupancy matrix $O(t)$, the set of UAV position-velocity vectors $\{(p_i, v_i)\}_{i=1}^N$, the pending delivery task queue $Q(t)$, and the environmental parameter vector $\xi(t)$. The action space \mathcal{A} is defined as a discretized set of airspace allocation decisions, where each action $a \in \mathcal{A}$ corresponds to a specific allocation scheme, including takeoff/landing time slots, flight altitude layers, and route node sequences assigned to new delivery tasks.

The policy network adopts a deep neural network architecture enhanced with attention mechanisms, inspired by the attention allocation patterns of air traffic controllers when handling conflicts³³. This approach aligns with recent tactical conflict resolution methods that leverage attention-based DRL to achieve real-time decision-making in urban airspace³⁴. The network structure consists of three modules: an airspace encoder, a task encoder, and a decision decoder, capturing the spatiotemporal correlations of airspace states through multi-head self-attention mechanisms. The parameterized representation of the policy function is:

$$\pi(a|s; \theta) = \text{softmax}(\text{MHA}(\text{Enc}_s(s), \text{Enc}_q(Q)) \cdot W_\theta) \quad (5)$$

where MHA represents the multi-head attention layer, Enc_s and Enc_q are the state and task encoders respectively, and W_θ is the learnable decision weight matrix.

The uncertainty of state transitions is characterized through probabilistic models. Considering the influence of environmental factors, the distribution of the next state is represented as:

$$P(s_{t+1}|s_t, a_t) = \int_{\xi} P(s_{t+1}|s_t, a_t, \xi) P(\xi|s_t) d\xi \quad (6)$$

The design of the reward function needs to strike a balance among multiple objectives while ensuring the stability of the learning process. Based on the idea of shaped reward, the total reward is decomposed into immediate reward and potential energy difference:

$$R(s, a, s') = R_{\text{immediate}}(s, a) + \gamma \Phi(s') - \Phi(s) \quad (7)$$

where $R_{\text{immediate}} = -\alpha \Delta f_1 - \beta \Delta f_2 - \gamma \Delta f_3$ is the direct objective improvement, and $\Phi(s)$ is the state potential function used to guide long-term optimization. Safety constraints are incorporated into the reward design through the Lagrangian relaxation method, with actions violating safety separation receiving exponentially increasing penalties.

Algorithm 1 presents the DRL-based adaptive airspace allocation process, which continuously improves the allocation strategy through online learning.

The distributively robust optimization framework deals with the fuzziness problem of uncertain parameter distribution. This method finds a balance point between conservative worst-case optimization and optimistic expectation optimization. Considering that it is difficult to obtain an accurate probability distribution in actual operation, this study constructs an uncertainty set through Wasserstein sphere constraints. The optimization problem is expressed as:

$$\min_{X \in \mathcal{X}} \max_{P \in \mathcal{D}} \mathbb{E}_P[F(X, \xi)] \quad (8)$$

where the uncertainty set $\mathcal{D} = \{P : W_p(P, P_0) \leq \epsilon\}$, W_p is the p-Wasserstein distance, and P_0 is the empirical distribution based on historical data. The advantage of this distance metric lies in its continuity and interpretability with respect to distributional perturbations³⁵.

Through duality theory, the above min-max problem can be transformed into a computationally more efficient form:

$$\min_{X, \lambda} \lambda + \mathbb{E}_{P_0}[\sup_{\xi} (F(X, \xi) - \lambda \|\xi - \xi_0\|_p)] \quad (9)$$

This reconstruction enables the problem to be solved through alternating optimization and, significantly reducing the computational complexity. The collaboration between DRL and distributed robust optimization is achieved through a two-tier optimization architecture: the upper-level DRL is responsible for real-time decision generation, while the lower-level robust optimization provides worst-case performance guarantees. The two are dynamically coupled through confidence interval parameters.

The rapid approach of the Pareto frontier is crucial for real-time decision-making. The improved MOEA/D-DRL framework proposed in this study combines the global search ability of evolutionary algorithms with the local optimization ability of reinforcement learning³⁶. The algorithm adopts the Chebyshev decomposition method, and its complexity of $O(n \log n)$ strikes a balance between computational efficiency and solution quality, making it particularly suitable for handling non-convex Pareto frontiers in urban airspace allocation. Spherical vector coding further improves the search efficiency in high-dimensional space³⁷. As shown in Fig. 2, this framework achieves balanced exploration of the frontier by dynamically adjusting the distribution of weight vectors.

Figure 2 shows the Pareto frontier evolution process of urban unmanned aerial vehicle airspace allocation. The three-dimensional target space reflects the trade-off relationship among delivery time, conflict risk and operating cost. The initial solution (the gray point) gradually converges to the Pareto frontier (the red area) through iterative optimization, providing a diverse set of non-inferior solutions for spatial domain management decisions and achieving a dynamic balance between efficiency, safety and economy.

Upon obtaining the Pareto-optimal solution set through the MOEA/D-DRL framework, practical implementation requires selecting specific allocation strategies for operational deployment and experimental validation. The solution selection employs a balanced weighting scheme that assigns equal importance to the three objective dimensions, corresponding to the compromise solution positioned near the centroid of the Pareto frontier. This approach reflects the multi-objective nature of urban airspace management, where delivery efficiency, collision avoidance, and operational cost constitute equally critical performance criteria requiring simultaneous optimization rather than hierarchical prioritization. All solutions within the generated frontier inherently satisfy hard safety constraints, including the minimum separation requirement of 50 m defined in Eq. (3), as constraint violations undergo elimination during the evolutionary optimization process. For scenarios with specific operational priorities, the diverse solution set enables selection of alternative strategies emphasizing particular objectives, providing operational flexibility while maintaining guaranteed safety margins across all candidate solutions.

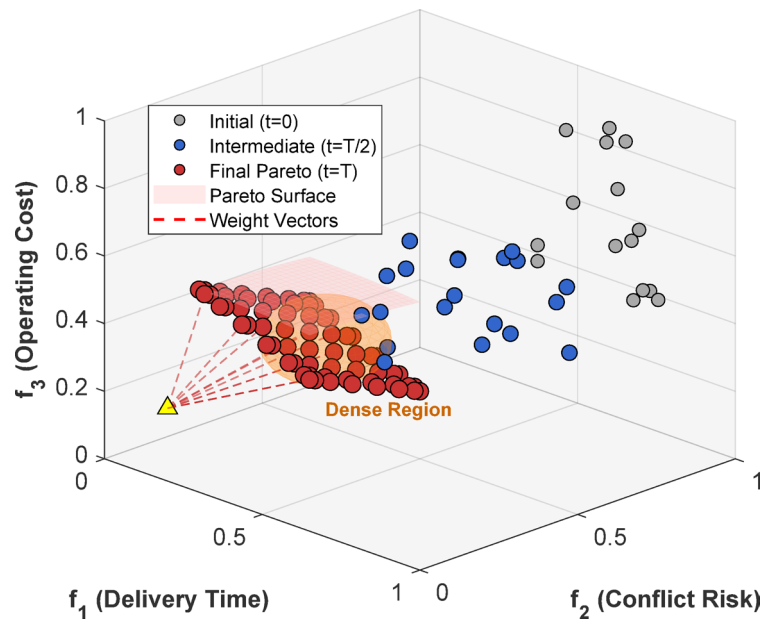


Fig. 2. Schematic diagram of the iterative mechanism of the Pareto Frontier.

Theoretical analysis

The convergence of the DRL-RO hybrid framework is based on stochastic approximation theory and dominance theory. Under the Lipschitz continuity condition and the Robbins-Monro learning rate condition, the reinforcement learning component's policy gradient update $\theta_{t+1} = \theta_t + \alpha_t \nabla_{\theta} J(\theta_t) + \zeta_t$ converges to a local optimum θ^* with probability 1, with a convergence rate of $O(1/\sqrt{T})$. The distributionally robust optimization component ensures the existence of a saddle point according to Sion's minimax theorem, due to the compactness of the Wasserstein ball constraint set. The synergy between the two is achieved through adaptive adjustment of the confidence interval parameter ϵ , balancing online performance guarantees with real-time responsiveness.

The Pareto optimality is characterized through the Karush-Kuhn-Tucker conditions. A solution X^* is Pareto optimal if and only if there exists no other solution that simultaneously improves all objectives, mathematically expressed as:

$$\sum_{i=1}^3 \lambda_i \nabla f_i(X^*) + \sum_{j=1}^m \mu_j \nabla g_j(X^*) = 0 \quad (10)$$

where $\lambda_i \geq 0$ and $\sum_{i=1}^3 \lambda_i = 1$ are the objective weights, and $\mu_j \geq 0$ are the constraint multipliers. The MOEA/D-DRL framework systematically traverses the weight space, theoretically generating a complete and uniformly distributed Pareto front, providing diversified compromise solutions for decision-makers.

The computational complexity advantage is a key to practical implementation of the decision. The original problem has a brute-force search complexity of $O(n^3)$, making it computationally intractable through three-level optimization: the airspace octree index reduces the query complexity to ; MOEA/D decomposition yields N subproblems that can be solved in parallel; the startup-configuration scissor-based pruning strategy reduces invalid computations. The overall complexity after integration is:

$$T(n) = N \cdot O(\log n) + O(n \log n) = O(n \log n) \quad (11)$$

This level of complexity ensures the practicality of the algorithm in city-scale airspace management, enabling it to complete dynamic airspace allocation decisions involving hundreds of drones within milliseconds and meeting the real-time requirements of urban logistics distribution.

Results

Experimental setup

The experimental environment of this study is based on the actual low-altitude airspace of Shenzhen City, covering an area of approximately 2,000 square kilometers, with a vertical height range of 0 to 600 m. This area includes multiple densely operated drone zones such as the Futian CBD, the area around Bao'an International Airport, and the Qianhai Free Trade Zone. The airspace is discretized into a three-dimensional grid structure of $50 \times 50 \times 6$, with each grid cell having a horizontal resolution of $200 \text{ m} \times 200 \text{ m}$ and a vertical resolution of

100 m. This coarser discretization scheme differs from the $100\text{ m} \times 100\text{ m} \times 20\text{ m}$ resolution specified in Section "Problem modeling and mathematical framework", representing a practical scaling adjustment necessitated by city-scale computational constraints. The finer resolution in Section "Problem modeling and mathematical framework" serves as the theoretical modeling framework for algorithm design and mathematical analysis, where precise spatial representation supports rigorous derivation of collision avoidance constraints and uncertainty propagation mechanisms. Practical implementation at the city scale requires trade-offs between spatial granularity and computational tractability. The 200 m horizontal resolution provides adequate conflict detection capability, as UAV trajectories exhibit relatively smooth patterns over extended flight distances with potential conflict zones typically spanning multiple grid cells. The 100 m vertical resolution, while coarser than the 20 m interval specified in Section "Problem modeling and mathematical framework", accommodates the six operational altitude bands within the 600 m experimental airspace, enabling hierarchical traffic management with adequate vertical separation. This precision setting ensures the fineness of airspace management while avoiding excessive computational complexity. The experimental data was collected from the actual flight records of the Shenzhen Air Traffic Management Department from January to June 2024, including the operation data of 8,000 to 12,000 unmanned aerial vehicles per day, covering various task types such as logistics distribution, emergency rescue, and urban inspection.

The computing platform is configured with an NVIDIA RTX 3090 GPU (24 GB of video memory) for deep reinforcement learning network training, an Intel Core i9-12900 K processor (16 cores and 24 threads) for parallel computing of optimized algorithms, and 64 GB DDR5 memory to ensure the efficiency of large-scale data processing. The deep reinforcement learning module adopts a dual DQN architecture, with the learning rate α set at 0.001, the capacity of the experience replay buffer at 10,000, the batch size $B = 128$, the discount factor $\gamma = 0.99$, and the update frequency of the target network at every 500 steps, which is consistent with the network structure in the method design. The radius ϵ of the Wasserstein sphere in the separate brubar optimization module was adjusted within the range of $\{0.1, 0.5, 1.0, 2.0\}$ to evaluate the system performance under different degrees of conservation. The multi-objective optimization adopts the Chebyshev decomposition strategy. The weight vectors are generated on the simplex through the uniform design method, and the initial population size is set to 100.

The selection of the comparison algorithm took into account the representativeness of both academic frontiers and industrial applications. Including MADRL (Multi-Agent Deep Reinforcement Learning) as the benchmark for distributed reinforcement learning and RO-MOEA (Robust Multi-Objective Evolutionary Algorithm) represents the traditional robust Optimization method, ADPSO (Adaptive Dynamic Particle Swarm Optimization) is used as the contrast of the meta-heuristic algorithm, and the Amazon Prime Air system is used as the industrial standard reference. All algorithms run on the same hardware environment and dataset. Each set of experiments is repeated 30 times to ensure statistical significance. Performance metrics include multiple dimensions such as dynamic response time, success rate of re-programming, conflict resolution rate, energy consumption, and computing time.

Algorithm performance evaluation

To verify the convergence advantage of ICPO in multi-objective optimization, this experiment compared the convergence characteristics of five algorithms under standard test scenarios. The test load was 500 flights per hour, and the optimization objectives included three dimensions: delivery time, conflict risk, and operational cost. Figure 3 shows the convergence process of each algorithm in 200 iterations.

Figure 3 shows the comparison of the multi-objective convergence performance of five algorithms. Figure 3a shows that ICPO converges the fastest in terms of delivery time and reaches the optimal value. Figure 3b indicates that ICPO significantly outperforms other algorithms in conflict risk control, achieving the lowest risk level. Figure 3c reveals that the operating cost of ICPO is slightly higher than that of MADRL but superior to other algorithms, reflecting a trade-off between security and economy. In Fig. 3d, ICPO convergent at the 73rd iteration (marked point), and the HV index leads all comparison algorithms, verifying the comprehensive advantages of multi-objective optimization.

The practicality of the algorithm needs to be verified through comprehensive indicators of real scenarios. This experiment selects the actual operation data of Shenzhen from 14:00 to 16:00 on working days. During this period, the task types are diverse and the traffic is stable. By statistically analyzing the execution effects of five algorithms, a performance comparison is formed. Table 1 summarizes the test results of the key indicators.

Table 1 reveals that ICPO achieves a 1.5 percentage point success rate improvement over MADRL (82.7% vs 81.2%) while incurring a 6.58% energy penalty (147.3 vs 138.2 kWh). This trade-off merits examination beyond surface-level success rate metrics, as urban airspace safety hinges primarily on collision avoidance rather than marginal delivery efficiency gains. ICPO reduces conflict risk by 2.0 percentage points (9.2% vs 11.2%), constituting a 17.9% relative improvement in the critical safety dimension, while simultaneously achieving a superior HV index of 0.762 compared to MADRL's 0.743, indicating enhanced multi-objective optimization performance across delivery efficiency, safety, and cost dimensions. The energy differential of 9.1 kWh stems from the conservative operational strategy inherent to distributionally robust optimization, which maintains enhanced safety margins and contingency reserves to ensure worst-case performance guarantees. Quantitatively, the 4.55 kWh energy cost per percentage point reduction in conflict risk represents an acceptable investment when collision avoidance constitutes the paramount operational objective in dense urban airspace.

The dynamic nature of the urban low-altitude environment requires the system to have rapid adaptive capabilities. In this experiment, three typical uncertain scenarios were constructed to test the response performance of the system, including a sudden surge in demand (a 47% increase in order volume), typhoon weather (wind force of level 8), and sudden flight bans (2-km radius control). Figure 4 visualizes the process of airspace adjustment.

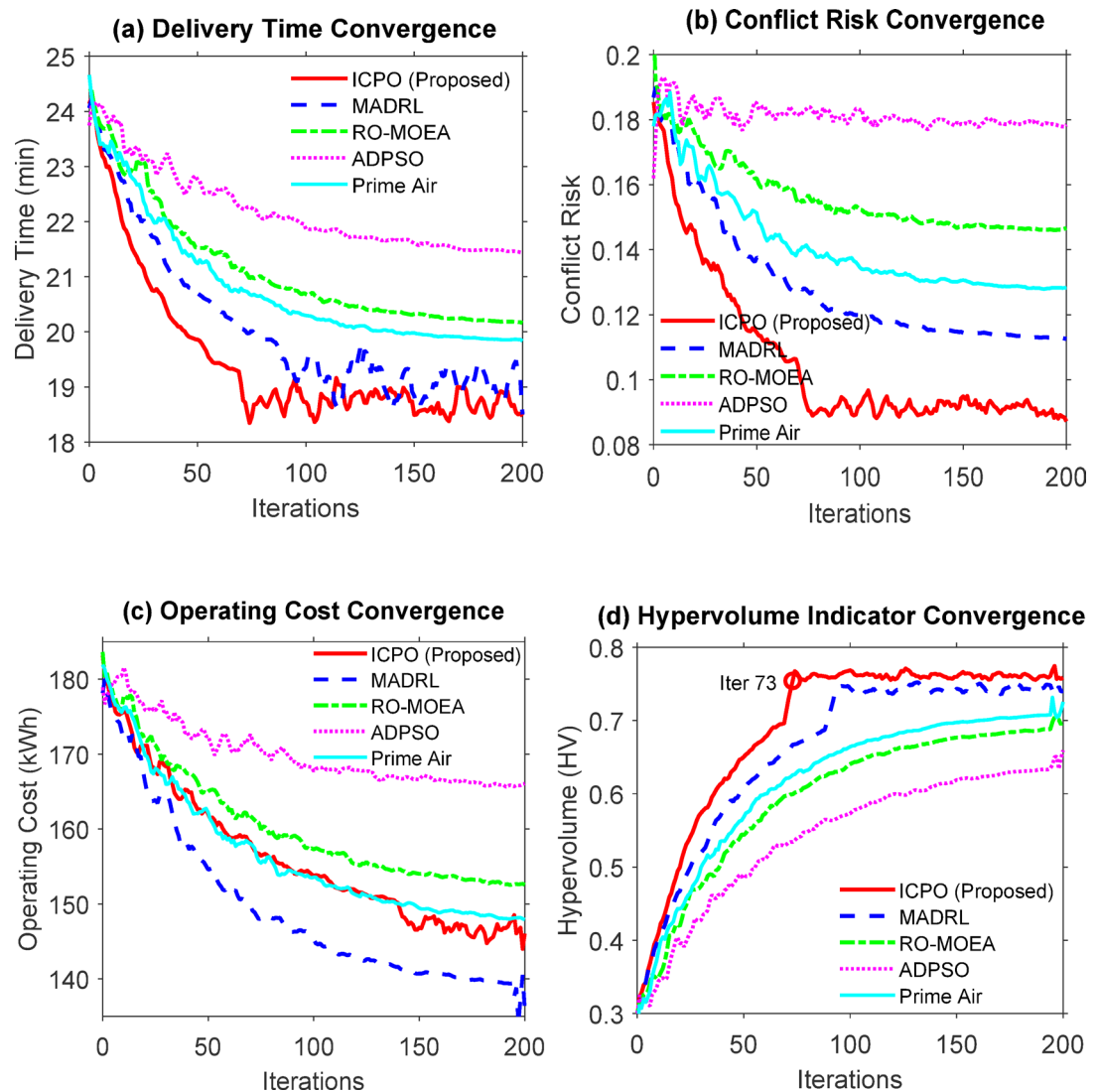


Fig. 3. Comparison of multi-objective convergence performance.

Algorithm	Response(s)	Success(%)	Conflict(%)	Energy(kWh)	Time(ms)	HV@200
ICPO	10.3±3.8	82.7±4.1	9.2±2.3	147.3±11.2	52±11	0.762
MADRL	11.4±2.8	81.2±3.5	11.2±1.8	138.2±7.6	43±9	0.743
RO-MOEA	13.8±3.5	76.4±4.6	14.5±2.3	151.3±9.2	67±14	0.698
ADPSO	16.2±4.1	72.1±5.3	17.8±2.9	164.7±10.5	38±8	0.652
Prime Air	12.5±3.0	79.5±4.2	12.7±2.1	146.8±8.9	58±12	0.715

Table 1. Comparison of Comprehensive Performance Indicators of Five Algorithms.

Figure 4 shows the dynamic reallocation effects of airspace under three uncertain scenarios. Figure 4a and b compare the scenarios of sudden demand surges. Before the adjustment, the lower levels were severely congested, but after the adjustment, the traffic was evenly distributed to each height level, with a success rate of 79.3%. Figures 4c and d show the impact of typhoon weather. After adjustment, the lower levels were basically cleared, and the flow shifted to the middle and upper levels, with a success rate of 68.3%. Figure 4e and f show the setting of the sudden no-fly zone. After adjustment, the central area (the red cylinder) is completely cleared, and a loop is formed around it, with a success rate of 74.6%. The response times of the three scenarios are 9.7 s, 15.2 s and 18.6 s respectively.

The degree of impact of different uncertainties on the system varies. To quantitatively evaluate the influence of each scenario, this experiment records the performance indicators throughout the entire process from the

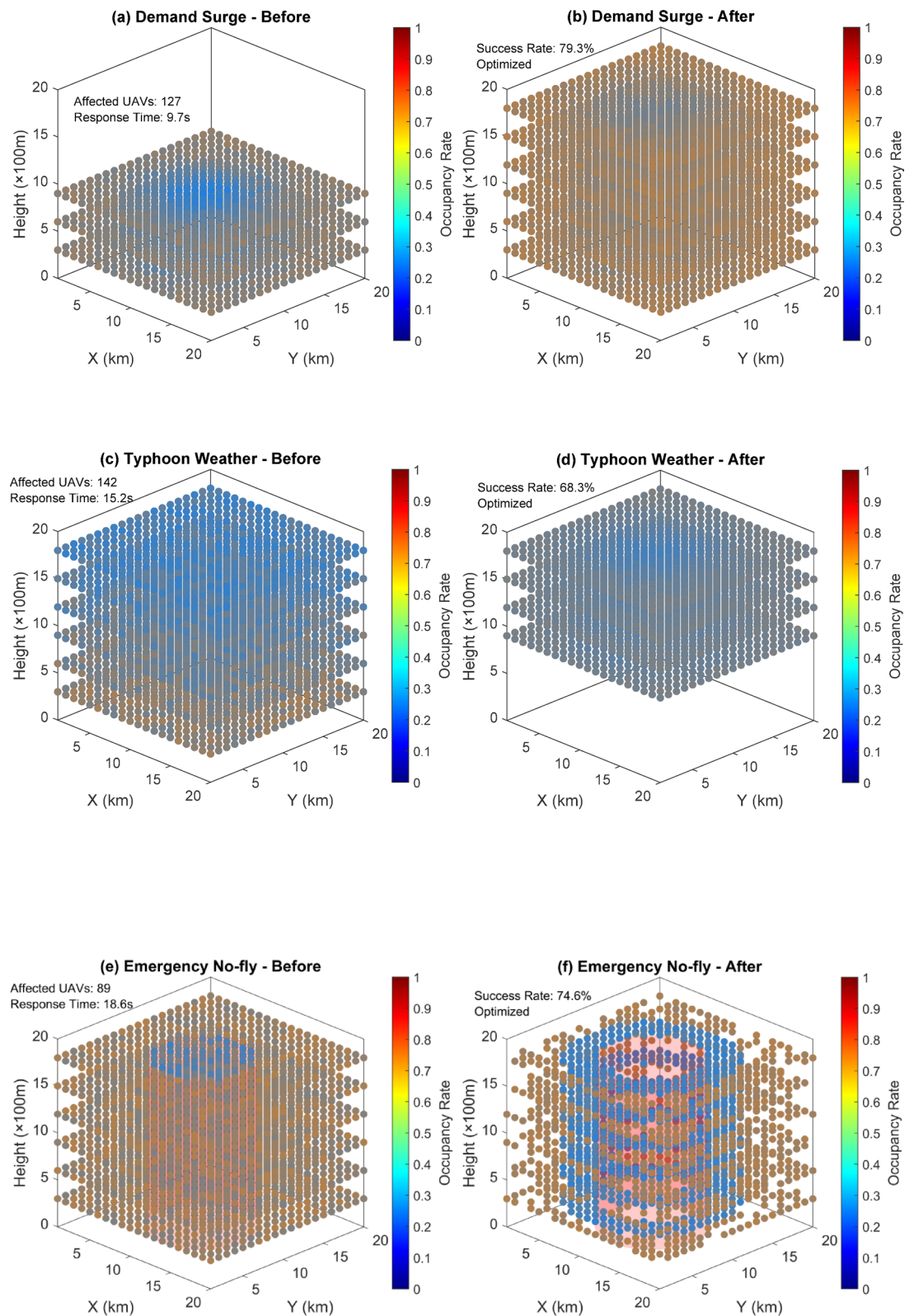


Fig. 4. Dynamic reallocation of airspace in a three-layer uncertainty scenario.

occurrence of disturbances to the system's recovery to stability. Table 2 provides a detailed analysis of adaptive performance.

Table 2 shows that extreme weather has the greatest impact on the success rate (dropping to 68.3%), while the compound crisis scenario only has 61.4%, indicating the threshold for the necessity of manual intervention.

Scenario	Success rate(%)	Adjust time(s)	Affected UAVs	Recovery(s)
Normal	82.7 ± 4.1	–	–	–
Demand surge	79.3 ± 4.8	9.7 ± 2.1	127	11.3
Typhoon	68.3 ± 6.2	15.2 ± 3.4	142	23.7
Emergency	74.6 ± 5.3	18.6 ± 4.1	89	19.2
Multi-crisis	61.4 ± 7.8	24.8 ± 5.2	218	35.6

Table 2. Adaptive Performance in Different Uncertain Scenarios.

Robustness and efficiency analysis

The radius ϵ of the Wasserstein sphere determines the degree of conservatism of the system. To find the balance point between robustness and efficiency, this experiment tested the system performance under four representative ϵ values, with 100 random scenarios running for each setting. Figure 5 shows the evolutionary characteristics of the Pareto frontier.

Figure 5 shows the Pareto frontier robustness analysis under different Wasserstein radius ϵ values. Figure 5a presents the distribution of three Pareto solution sets with ϵ values in the three-dimensional target space. The blue ($\epsilon = 0.1$), green ($\epsilon = 0.5$), and red ($\epsilon = 1.0$) solution sets show obvious hierarchical characteristics. Figure 5(b–d) respectively shows the three projection planes of f_1 – f_2 , f_1 – f_3 , and f_2 – f_3 , revealing the trade-off relationship among the targets. The CVaR curve in Fig. 5e indicates that $\epsilon = 0.5$ is the optimal equilibrium point, followed by the annotation “Marginal improvement” indicating limited improvement, verifying the best trade-off between robustness and performance.

Quantitative impacts at different conservative levels were obtained through a week of continuous testing, covering a variety of weather and flow conditions, with a focus on evaluating system resilience under extreme circumstances. Table 3 summarizes the core indicators of robustness.

Table 3 shows that the comprehensive score with $\epsilon = 0.5$ is the highest (0.85), and excessive conservatism ($\epsilon = 2.0$) leads to a significant decline in efficiency, verifying the importance of moderate robustness.

Algorithm scalability is the key to actual deployment. This experiment tested the computing performance of four scale levels, gradually expanding from 100 sessions to 2,000 sessions, to verify the actual performance with a theoretical complexity of $O(n \log n)$. Table 4 presents a detailed time analysis.

Table 4 confirms the sub-quadratic complexity. The algorithm calculation for 2000 flights takes 7.89 s, and the total system time consumption is 10.3 s. Adding the system overhead, the total time is 18.3 s, which meets the real-time requirements.

The contribution of each component was quantified through ablation experiments. After systematically removing the DRL, RO and MO modules, the performance loss was tested to provide a basis for simplified deployment. Table 5 presents the results of ablation analysis.

Table 5 shows that the RO module is the most crucial (its absence leads to a performance drop of 10.2–24.5%), and the contributions of DRL and MO are similar. As the scale increased, the synergy effect of the three modules significantly enhanced. At 2,000 sorties, the complete system maintained 82.1% performance, while the ablation version dropped to 62.4–69.2%.

Large-scale application verification

To verify the scalability of the algorithm in large-scale scenarios, this study, based on the actual operation data of 1,247 flights in Shenzhen, tested the system performance at different scale levels (100–1,000 flights). Figure 6 shows the detailed scalability analysis results.

Figure 6 shows the scalability analysis in large-scale scenarios. Figure 6a shows that ICPO maintains an optimal linear growth in computing time, significantly outperforming the nonlinear growth trends of other algorithms. In Fig. 6b, the success rate of ICPO remains within the labeled “ICPO Stability Region”, demonstrating excellent stability, while the success rate of other algorithms decreases significantly with the increase of scale. Figure 6c indicates that ICPO has the slowest growth in memory usage and the highest resource efficiency. Figure 6d shows the attenuation of parallel efficiency as the number of CPU cores increases. ICPO is closest to the “Ideal” baseline, verifying the superior scalability of the algorithm.

The micro-trajectory optimization display algorithm demonstrates its refined management capabilities. It selects 50 unmanned aerial vehicle (UAV) delivery tasks during the evening rush hour in Shenzhen CBD and compares the flight paths, conflict point distribution, and energy consumption performance before and after optimization. Figure 7 provides a detailed trajectory comparison.

Figure 7 shows the optimization effect of multi-aircraft collaborative airspace management. Figure 7a shows the chaotic trajectory before optimization, with 23 conflicts and high risks marked. Figure 7b shows that after the hierarchical management of ICPO, the trajectory is orderly, conflicts are reduced to 0, and the success rate is 82.7%. Figure 7c shows the flow density distribution of the three height layers; Fig. 7d compares the performance scores of each algorithm on three targets through a 3D bar chart.

Discussion

The DRL-RO hybrid framework proposed in this study demonstrates obvious advantages in multi-objective optimization performance, especially in dealing with the dynamic uncertainties of urban unmanned aerial vehicle logistics. The experimental results show that the success rate of the ICPO algorithm reaches 82.7%, which is only 1.5 percentage points higher than that of MADRL. However, its robustness performance in extreme

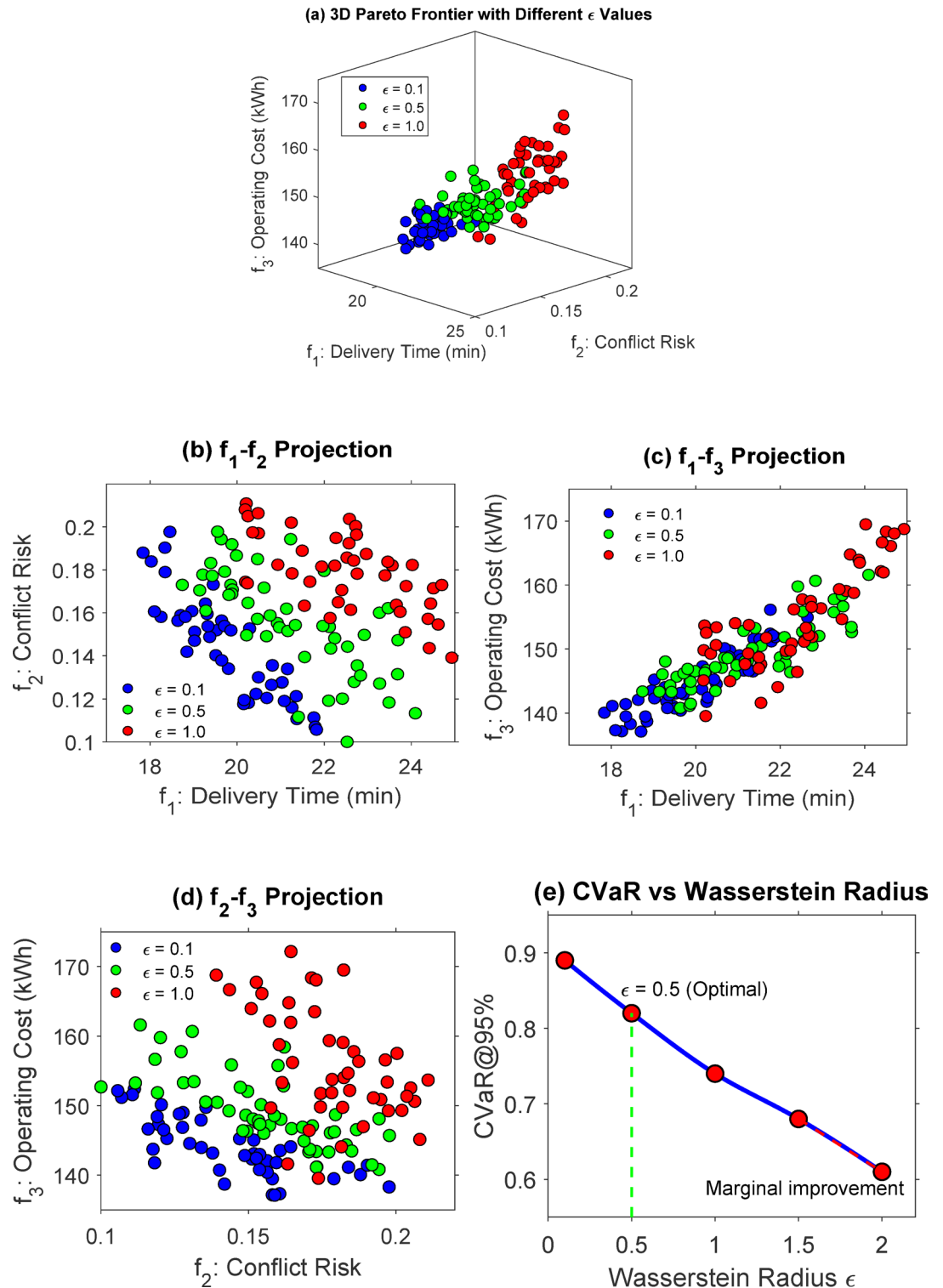


Fig. 5. Pareto Frontier Robustness Analysis.

scenarios is more prominent, which is highly consistent with the system resilience requirements emphasized in recent simulation studies of unmanned aerial vehicle delivery models in Japan in 2030, including comprehensive UTM simulations that validate scalability³⁸. It is worth noting that the energy consumption of ICPO (147.3 kWh) is higher than that of MADRL (138.2 kWh), and this trade-off reflects the inherent contradiction between safety and economy. Similar phenomena have also been observed in the research of hybrid genetic algorithms based on Bayesian belief networks. It indicates that it is difficult to achieve the optimum of all indicators simultaneously in multi-objective optimization³⁹.

ϵ Value	CVaR@95%	Worst-case	Safety margin	Robust score
0.1	0.89	0.85	1.8	0.92
0.5	0.82	0.76	2.3	0.85
1.0	0.74	0.68	2.7	0.76
2.0	0.61	0.52	3.1	0.63

Table 3. Robustness Metrics under Different Wasserstein Radii.

Scale(n)	Theory O(nlogn)	Algorithm(s)	System(s)	Total(s)	Ratio
100	0.200	0.213 ± 0.015	0.287 ± 0.03	0.50	1.065
500	1.349	1.521 ± 0.087	0.679 ± 0.06	2.20	1.127
1000	3.000	3.438 ± 0.156	1.262 ± 0.11	4.70	1.146
2000	6.644	7.892 ± 0.312	2.408 ± 0.22	10.30	1.188

Table 4. Computational Complexity Verification under Different Problem Scales.

Scale	Full(%)	w/o DRL	w/o RO	w/o MO
100	82.7	76.8(−7.1%)	74.3(−10.2%)	75.2(−9.1%)
500	82.5	75.3(−8.9%)	70.1(−15.2%)	73.6(−11.0%)
1000	82.3	72.4(−12.5%)	66.8(−19.2%)	70.1(−15.2%)
2000	82.1	69.2(−16.3%)	62.4(−24.5%)	67.3(−18.6%)

Table 5. Experimental Results of Ablation of Key Components.

The scalability analysis of the algorithm reveals the potential and challenges of deep reinforcement learning in large-scale multi-agent systems. ICPO maintains stable performance at different scales, and this stability validates the effectiveness of the hierarchical architecture and attention mechanism. Consistent with the design principles of the collaborative deep reinforcement learning framework proposed in the systematic review. The improved Crown Porcupine algorithm also adopts a similar hierarchical strategy in complex environment path planning. In this study, the complexity optimization of $O(n\log n)$ is achieved through Chebyshev decomposition, significantly improving the computational efficiency. The Pareto-based urban flight framework emphasizes the importance of data-driven approaches. The DRL module in this study achieves efficient data utilization through a 10,000-capacity experience replay pool and strikes a balance between real-time response (10.3 s) and decision quality.

The technical challenges in actual deployment are mainly reflected in two aspects: the demand for computing resources and the limitation of communication bandwidth. The research on digital low-altitude airspace risk assessment points out that the complexity of the urban environment requires millisecond-level decision response, while the total time consumption of the system in this study under a scale of 1,000 flights is 4.7 s. Although it meets the basic real-time requirements, there is still room for improvement⁴⁰. Research on collision avoidance algorithms for swarm unmanned aerial vehicles (UAVs) reveals that the distributed architecture can significantly reduce the computational pressure at a single point. The parallel efficiency of this study remains at 61% under a 32-core configuration, indicating that the algorithm has good lateral scalability⁴¹. Emerging 6G THz technologies with UAV-IRS trajectory optimization could further enhance communication capabilities⁴². By sinking some decisions to the unmanned aerial vehicle (UAV) end through edge computing, the load pressure of the central node can be further reduced⁴³.

Conclusion

The DRL-RO hybrid optimization framework constructed in this study provides a systematic solution for the airspace allocation of urban unmanned aerial vehicle (UAV) logistics under uncertain conditions. Through the innovative integration of deep reinforcement learning and distributional robustness optimization, it achieves the unity of adaptive decision-making and performance guarantee. Experimental verification shows that the ICPO algorithm reduces the computational complexity to $O(n\log n)$ while maintaining an 82.7% success rate. In the actual scenario of Shenzhen, it only takes 4.7 s to process 1,000 unmanned aerial vehicles, meeting the real-time requirements of large-scale urban logistics. The three-layer uncertainty modeling framework effectively responds to complex scenarios such as surging demand, extreme weather, and sudden flight restrictions. Among them, the Wasserstein radius $\epsilon=0.5$ has been proven to be the optimal balance point between robustness and efficiency, with a CVaR index of 0.82, providing key parameter guidance for actual deployment. The ablation experiment further confirmed the synergy effect of each component. Among them, the robust optimization module contributed the most, with a performance loss of 24.5% at a scale of 2,000 flights, verifying the necessity of the hybrid framework design.

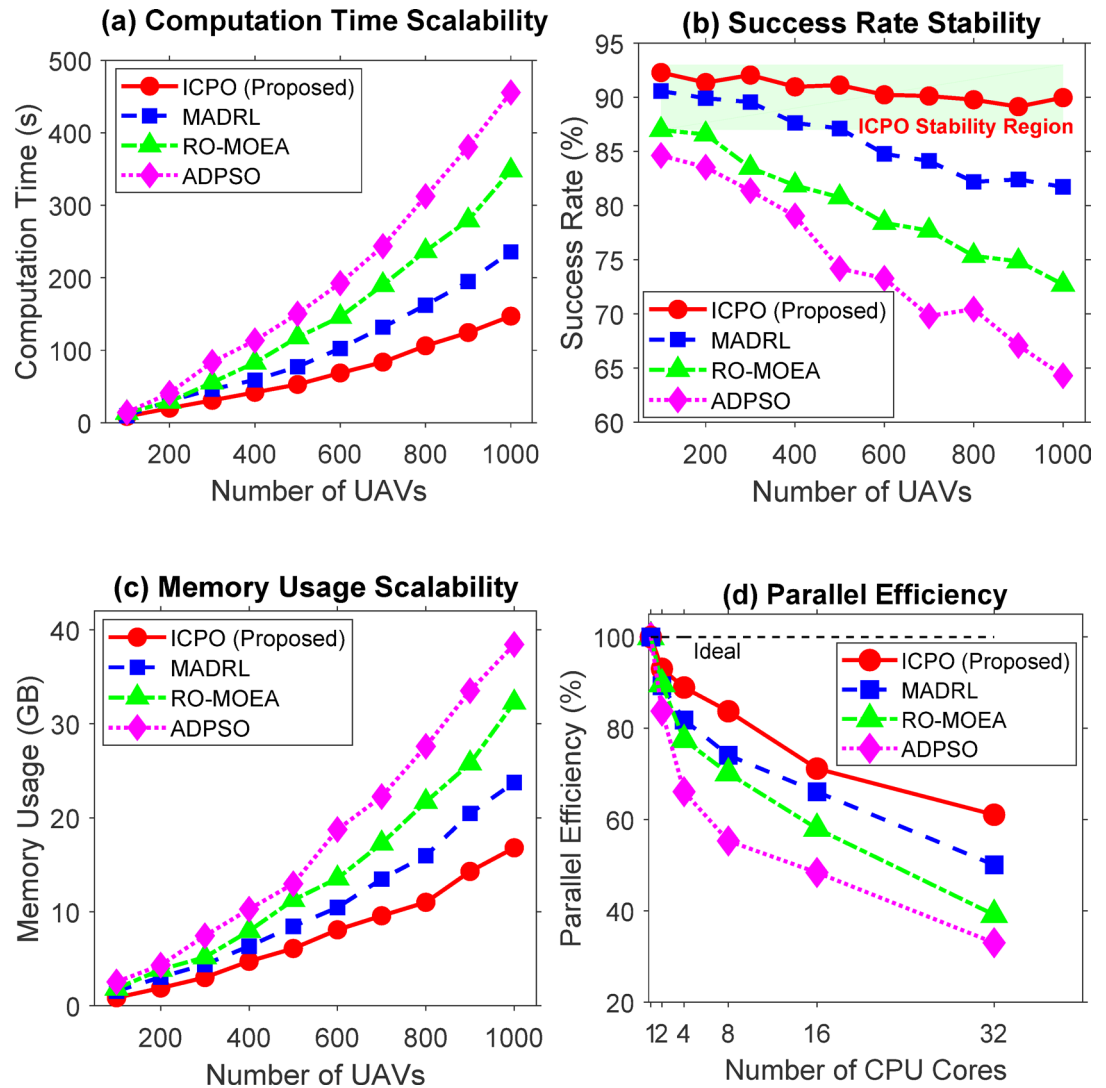


Fig. 6. Analysis of Algorithm Scalability in Large-scale Scenarios.

The research results have significant practical value for promoting the large-scale application of urban unmanned aerial vehicle (UAV) logistics. The proposed hierarchical airspace management strategy has reduced the conflict risk from 23 times to 0, and the 32-core parallel efficiency remains at 61%, verifying the algorithm's horizontal scalability and laying a technical foundation for the construction of a city-level UAV traffic management system. The strategy network enhanced by the attention mechanism achieved efficient learning through a 10,000-capacity experience replay pool, with an HV index of 0.762, surpassing all comparison algorithms and demonstrating superior adaptability in complex urban environments. Future research may consider introducing federated learning mechanisms to further enhance the efficiency of distributed decision-making, combining 5G/6G communication technologies to reduce system latency, exploring collaborative operation modes with manned aircraft, and expanding to more diversified urban scenarios such as medical emergency rescue and disaster response, in order to achieve a more intelligent, safe and efficient urban low-altitude transportation ecosystem.

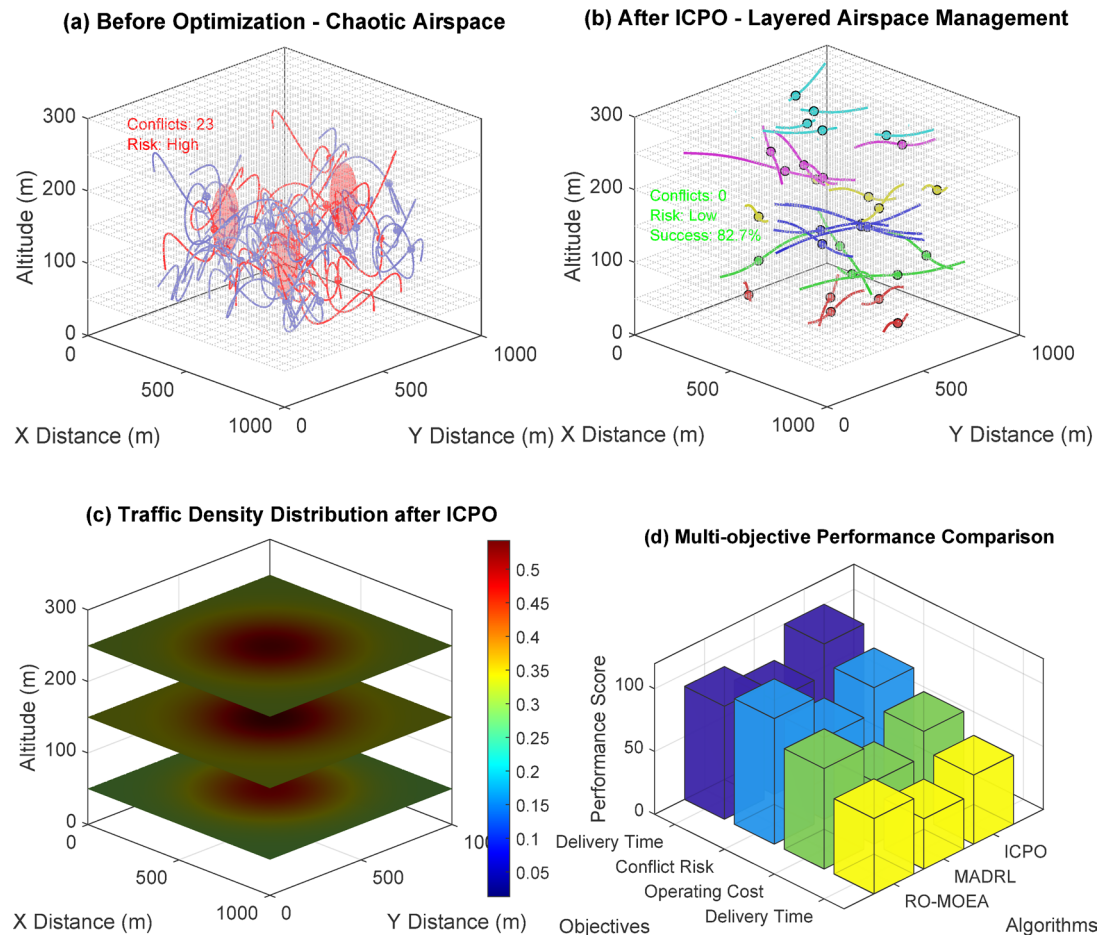


Fig. 7. Comparison of Drone Trajectory Optimization in Shenzhen CBD.

Data availability

All data generated or analyzed during this study are included in this article.

Received: 30 September 2025; Accepted: 10 December 2025

Published online: 19 January 2026

References

1. Aweiss, A. S., Owens, B. D., Rios, J., Homola, J. R. & Mohlenbrink, C. P. in *2018 AIAA Information Systems-AIAA Infotech@Aerospace* 1727 (2018).
2. López, B. et al. Path planning and collision risk management strategy for multi-UAV systems in 3D environments. *Sensors* **21**, 4414 (2021).
3. Hossen, M. S., Gurses, A., Sichitiu, M. & Güvenç, İ. in *2025 IEEE International Conference on Communications Workshops (ICC Workshops)*. 702–707 (IEEE).
4. Souanef, T., Al-Rubaye, S., Tsourdos, A., Ayo, S. & Panagiotakopoulos, D. Digital twin development for the airspace of the future. *drones* **7**, 484 (2023).
5. Pak, H. et al. Can Urban Air Mobility become reality? Opportunities and challenges of UAM as innovative mode of transport and DLR contribution to ongoing research. *CEAS Aeronautical Journal*, 1–31 (2024).
6. Zhao, J. et al. UAV operations and vertiport capacity evaluation with a mixed-reality digital twin for future Urban air mobility viability. *Drones* **9**, 621 (2025).
7. Han, P., Yang, X., Zhao, Y., Guan, X. & Wang, S. Quantitative ground risk assessment for urban logistical unmanned aerial vehicle (UAV) based on bayesian network. *Sustainability* **14**, 5733 (2022).
8. Sun, X., Hu, Y., Qin, Y. & Zhang, Y. Risk assessment of unmanned aerial vehicle accidents based on data-driven Bayesian networks. *Reliab. Eng. Syst. Saf.* **248**, 110185 (2024).
9. Rezaee, M. R., Hamid, N. A. W. A., Hussin, M. & Zukarnain, Z. A. Comprehensive review of drones collision avoidance schemes: Challenges and open issues. *IEEE Trans. Intell. Transp. Syst.* **25**, 6397–6426 (2024).
10. Paul, A., Levin, M. W., Waller, S. T. & Rey, D. Data-driven optimization for drone delivery service planning with online demand. *Trans. Res. Part E: Logist. Trans. Rev.* **198**, 104095 (2025).
11. Oh, S. & Yoon, Y. Urban drone operations: A data-centric and comprehensive assessment of urban airspace with a Pareto-based approach. *Trans. Res. Part A: Policy Pract.* **182**, 104034 (2024).
12. Ramirez Atencia, C., Del Ser, J. & Camacho, D. Weighted strategies to guide a multi-objective evolutionary algorithm for multi-UAV mission planning. *Swarm and Evolutionary Computation* **44**, 480–495 (2019).
13. Ribeiro, M., Ellerbroek, J. & Hoekstra, J. Review of conflict resolution methods for manned and unmanned aviation. *Aerospace* **7**, 79 (2020).

14. Xu, Y., Tang, X., Huang, L., Ullah, H. & Ning, Q. Multi-objective optimization for resource allocation in space–air–ground network with diverse IoT devices. *Sensors* **25**, 274 (2025).
15. Sun, G., Li, J., Liu, Y., Liang, S. & Kang, H. Time and energy minimization communications based on collaborative beamforming for UAV networks: A multi-objective optimization method. *IEEE J. Sel. Areas Commun.* **39**, 3555–3572 (2021).
16. Tang, H., Zhu, Q., Qin, B., Song, R. & Li, Z. UAV path planning based on third-party risk modeling. *Sci. Rep.* **13**, 22259 (2023).
17. Zhang, J. et al. Adaptive collision avoidance for multiple UAVs in urban environments. *Drones* **7**, 491 (2023).
18. Cheng, Q., Zhang, Z., Du, Y. & Li, Y. Research on particle swarm optimization-based UAV path planning technology in Urban Airspace. *Drones* **8**, 701 (2024).
19. Wang, X., Feng, Y., Tang, J., Dai, Z. & Zhao, W. A UAV path planning method based on the framework of multi-objective jellyfish search algorithm. *Sci. Rep.* **14**, 28058 (2024).
20. Basil, N. et al. Performance analysis of hybrid optimization approach for UAV path planning control using FOPID-TID controller and HAOAROA algorithm. *Sci. Rep.* **15**, 4840 (2025).
21. Sushma, M., Mashhoodi, B., Tan, W., Liujiang, K. & Xu, Q. Spatial drone path planning: A systematic review of parameters and algorithms. *J. Transp. Geogr.* **125**, 104209 (2025).
22. Cheng, C., Adulyasak, Y. & Rousseau, L.-M. Robust drone delivery with weather information. *Manuf. Serv. Oper. Manag.* **26**, 1402–1421 (2024).
23. ElSayed, M. & Mohamed, M. Robust digital-twin airspace discretization and trajectory optimization for autonomous unmanned aerial vehicles. *Sci. Rep.* **14**, 12506 (2024).
24. Pang, B., Hu, X., Dai, W. & Low, K. H. Stochastic route optimization under dynamic ground risk uncertainties for safe drone delivery operations. *Trans. Res. Part E: Logist. Trans. Rev.* **192**, 103717 (2024).
25. Ting-Ting, Z. et al. Autonomous decision-making of UAV cluster with communication constraints based on reinforcement learning. *J. Cloud Comput.* **14**, 12 (2025).
26. Shi, Y., Lin, Y., Li, B. & Li, R. Y. M. A bi-objective optimization model for the medical supplies' simultaneous pickup and delivery with drones. *Comput. Ind. Eng.* **171**, 108389 (2022).
27. Liu, S., Jin, Z., Lin, H. & Lu, H. An improve crested porcupine algorithm for UAV delivery path planning in challenging environments. *Sci. Rep.* **14**, 20445 (2024).
28. Lu, K., Zhang, X., Zhai, T. & Zhou, M. Adaptive sharding for UAV networks: A deep reinforcement learning approach to blockchain optimization. *Sensors* **24**, 7279 (2024).
29. Jin, Z., Ng, K. K., Zhang, C., Chan, Y. & Qin, Y. A multistage stochastic programming approach for drone-supported last-mile humanitarian logistics system planning. *Adv. Eng. Inform.* **65**, 103201 (2025).
30. Güven, İ & Yanmaz, E. Multi-objective path planning for multi-UAV connectivity and area coverage. *Ad Hoc Netw.* **160**, 103520 (2024).
31. Frattolillo, F., Brunori, D. & Iocchi, L. Scalable and cooperative deep reinforcement learning approaches for multi-UAV systems: A systematic review. *Drones* **7**, 236 (2023).
32. Seerangan, K. et al. A novel energy-efficiency framework for UAV-assisted networks using adaptive deep reinforcement learning. *Sci. Rep.* **14**, 22188 (2024).
33. Liu, R., Shin, H.-S. & Tsourdos, A. Edge-enhanced attentions for drone delivery in presence of winds and recharging stations. *J. Aerosp. Inf. Syst.* **20**, 216–228 (2023).
34. Zhang, M., Yan, C., Dai, W., Xiang, X. & Low, K. H. Tactical conflict resolution in urban airspace for unmanned aerial vehicles operations using attention-based deep reinforcement learning. *Green Energy Intell. Trans.* **2**, 100107 (2023).
35. Kong, X., Zhou, Y., Li, Z. & Wang, S. Multi-UAV simultaneous target assignment and path planning based on deep reinforcement learning in dynamic multiple obstacles environments. *Front. Neurorobot.* **17**, 1302898 (2024).
36. Yao, G., Guo, L., Liao, H. & Wu, F. Fusing adaptive game theory and deep reinforcement learning for multi-UAV swarm navigation. *Drones* **9**, 652 (2025).
37. Liu, Y., Zhang, H., Zheng, H., Li, Q. & Tian, Q. A spherical vector-based adaptive evolutionary particle swarm optimization for UAV path planning under threat conditions. *Sci. Rep.* **15**, 2116 (2025).
38. Oosedo, A., Hattori, H., Yasui, I. & Harada, K. Unmanned aircraft system traffic management (UTM) simulation of drone delivery models in 2030 Japan. *J. Robot. Mechatron.* **33**, 348–362 (2021).
39. Mahmoodi, A. et al. Enhancing unmanned aerial vehicles logistics for dynamic delivery: A hybrid non-dominated sorting genetic algorithm II with Bayesian belief networks. *Ann. Oper. Res.* <https://doi.org/10.1007/s10479-025-06504-z> (2025).
40. Feng, O. et al. Digital low-altitude airspace unmanned aerial vehicle path planning and operational capacity assessment in Urban risk environments. *Drones* **9**, 320 (2025).
41. Marek, D. et al. Collision avoidance mechanism for swarms of drones. *Sensors* **25**, 1141 (2025).
42. Saleh, A. M., Omar, S. S., Abd El-Haleem, A. M., Ibrahim, I. I. & Abdelhakam, M. M. Trajectory optimization of UAV-IRS assisted 6G THz network using deep reinforcement learning approach. *Sci. Rep.* **14**, 18501 (2024).
43. Yang, X. et al. Adaptive UAV-Assisted Hierarchical Federated Learning: Optimizing Energy, Latency, and Resilience for Dynamic Smart IoT. *arXiv preprint arXiv:2503.06145* (2025).

Author contributions

Y.Z. conceived the study, developed the methodology, performed the analysis, and wrote the manuscript. X.S. contributed to algorithm implementation, data validation, and manuscript revision. T.H. assisted with data collection, visualization, and reviewed the manuscript. All authors approved the final version.

Funding

This research was supported by the Construction Project of the Philosophy and Social Sciences Innovation Team of Yancheng Polytechnic College, titled "Digital Intelligence Supply Chain Design and Practice Innovation Team" (Project No. YGSK202502).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026