ITM Web of Conferences **78**, 03028 (2025) https://doi.org/10.1051/itmconf/20257803028
CSEIT 2025

Optimization of Uav Obstacle Avoidance Based on Multi-Armed Bandit Algorithms

Jiashen He

College of Information and Computing Science, Jilin University, Changchun, 130012, China

Abstract. This paper presents a novel Unmanned Aerial Vehicle (UAV) path planning framework that integrates the UCB1-Tuned multi-armed bandit algorithm with precise numerical integration techniques (Runge-Kutta 4th order) to address navigation challenges in stochastic wind environments with obstacles. The proposed method discretizes the joint action space of wind direction and acceleration to enable efficient exploration and exploitation for adaptive control under uncertainty. The use of UCB1-Tuned balances reward means and variance to select optimal control actions without requiring Bayesian sampling, enhancing computational efficiency and stability. High-fidelity trajectory generation is ensured through accurate numerical integration respecting UAV dynamic constraints. Extensive simulations demonstrate that the UCB1-Tuned, combined with Runge-Kutta integration, achieves a success rate over 99% and reduces average energy consumption significantly compared to random baselines and Euler integration variants. The approach effectively learns favorable control policies by dynamically adapting to wind and obstacle conditions, ensuring safety and energy efficiency. This work offers a scalable decision-making framework that combines principled learning with model-based trajectory prediction, providing a promising direction for real-time autonomous UAV navigation in complex and uncertain environments.

1 Introduction

UAVs have become critical tools for diverse applications such as surveillance, delivery, and emergency response, demanding autonomous navigation capabilities in complex and uncertain environments. UAV path planning faces two core challenges: unpredictable environmental dynamics, including wind disturbances and moving obstacles, and stringent onboard computational limits requiring highly efficient and reliable decision-making algorithms.

Traditional path planning methods like Rapidly-exploring Random Trees (RRT*) and Model Predictive Control (MPC) provide theoretical guarantees but often incur high computational costs, causing delayed replanning unsuitable for real-time UAV operations. These methods typically involve iterative optimization that can require hundreds of milliseconds or more, exceeding the responsiveness needed for rapidly evolving scenarios. Conversely, purely learning-based approaches, particularly deep reinforcement learning, offer faster inference times but frequently suffer from lacking explicit physical or dynamic

hejs9922@mails.jlu.edu.cn

constraints enforcement. This deficiency leads to trajectories that may violate UAV kinematics or dynamics, threatening flight safety and limiting practical deployment.

To address the critical trade-off between computational efficiency and physical feasibility, this paper proposes a novel framework that integrates the UCB1-Tuned multi-armed bandit learning algorithm with reliable numerical integration methods, specifically the Runge-Kutta 4th Order (RK4) and Euler schemes. This dual strategy enables efficient exploration and exploitation of a discretized action space composed of wind directions and acceleration levels, quickly adapting to stochastic environmental conditions while ensuring trajectory feasibility through precise dynamical simulations.

This experimental result demonstrates the superiority of this combined approach. The UCB1-Tuned + RK4 method achieves an impressive success rate exceeding 99% while maintaining low average energy consumption (~9.3%), substantially outperforming a naive random policy that achieves only a 6% success rate with significantly higher energy demands (~36%). When compared with Euler integration, RK4 integration displays clear advantages: it produces more energy-efficient trajectories and offers enhanced stability and accuracy in simulating UAV dynamics, which directly translates to improved control performance and mission reliability. These findings underscore the practical value of adopting higher-order numerical integration techniques in autonomous UAV control tasks.

Additionally, the distribution of arm selections across trials confirms that UCB1-Tuned efficiently identifies and exploits more beneficial control actions, demonstrating robust learning capabilities in uncertain and dynamic environments. This empirical evidence supports the potential of bandit-based methods as powerful tools for online UAV path planning and control.

In summary, this work presents a computationally efficient and physically grounded UAV navigation framework that successfully balances safety, performance, and real-time adaptability. By blending principled bandit learning with model-based trajectory prediction, the approach not only advances state-of-the-art UAV control methodologies but also provides a scalable foundation for handling increasingly complex autonomous flight scenarios.

2 Related Work

2.1 Traditional Navigation Methods

Traditional navigation methods play a critical role in robotics, especially in stable environments. Among these methods, potential fields have gained significant attention for their simplicity and effectiveness in obstacle avoidance and goal-oriented navigation. Khatib introduced the potential field method in 1986, guiding robots toward desired targets by leveraging attractive and repulsive forces, effectively navigating obstacles in various scenarios [1]. Visual odometry (VO) algorithms, as demonstrated by Davison, allow robots to estimate their position and orientation by tracking changes in sequentially captured images. This technique has become essential in dynamic environments where Global Positioning System (GPS) signals might be unreliable or unavailable.

As the demand for UAVs to operate in wind-affected environments increases, physics-based motion primitives have emerged as a notable area of research. For instance, Wang et al. provided an overview of various wind effects on UAVs, stressing the necessity for robust navigation strategies that account for such uncertainties in real-world applications [2]. These studies emphasize the ongoing need for enhancing autonomous navigation capabilities for UAVs in complex environments.

2.2 Bandit Learning in Robotics

Bandit learning has rapidly evolved in the robotics field, particularly in applications involving non-stationary environments. The non-stationary Multi-Armed Bandit (MAB) framework allows robots to adapt their exploration strategies while navigating dynamic reward landscapes. Mohamed et al. demonstrated the effectiveness of MAB for gateway selection in dynamic networks, emphasizing the importance of adapting to changing conditions to optimize resource allocation [3]. Similarly, Lin et al. explored the application of MAB in UAV-assisted emergency communication systems, highlighting how these strategies can improve communication reliability during critical situations [4].

Contextual bandits further extend the applicability of MAB by incorporating environmental features into the decision-making process. Amrallah et al. applied dual energy-aware bandits for UAV trajectory optimization in post-disaster settings, illustrating the model's effectiveness in responding to environmental changes and allowing UAVs to autonomously adjust their paths based on real-time evaluations of the environment's reward structure [5].

2.3 Hybrid Approaches

Hybrid approaches that combine classical control methods with modern learning techniques are gaining traction in the robotics domain. These methods leverage the strengths of reinforcement learning (RL) and traditional control strategies to achieve superior results in complex tasks. Wang and Hong discussed the opportunities and challenges associated with applying reinforcement learning to building controls, asserting that integrating Reinforcement Learning (RL) with MPC could lead to more adaptive and efficient energy management systems [6].

To illustrate further, Arroyo et al. presented a learning approach using Reinforcement Learning Model Predictive Control (RL-MPC) for building energy management, demonstrating how combining learning and control can enhance performance in managing energy-efficient operations [7]. Hewing et al. also highlighted insights on safe learning in control systems, emphasizing the critical need for safety in reinforcement learning applications, particularly when utilized in dynamic environments [8]. Schwenzer et al. provided an engineering perspective on model predictive control, which can complement RL techniques to optimize control strategies efficiently in managing building energy and other complex systems [9].

2.4 Open Challenges

Despite significant advancements in navigation and learning methodologies, several open challenges remain. A prominent issue is how to bridge discrete decision-making processes with continuous control, particularly in high-dimensional spaces. Addressing this issue requires further research into developing advanced algorithms capable of operating under uncertainty and augmenting existing models with sophisticated perception and prediction capabilities.

Environmental uncertainties represent significant challenges for UAV operations, as they can greatly impact decision-making processes. Studies on wind field estimation are particularly relevant here. Larrabee et al. explored wind field estimation in UAV formation flight and highlighted the complexities involved in managing multiple UAVs under varying environmental conditions [10]. Additionally, Xiang et al. focused on wind field estimation through autonomous quadcopter avionics, providing insights into how UAVs can autonomously adjust their flight strategies based on real-time wind data [11]. Addressing these challenges is critical for advancing the reliability and effectiveness of UAVs in increasingly complex environments.

3 Methodology

Fig. 1 illustrates a real-world UAV obstacle avoidance situation, it shows a Vision mini-helicopter flying close to a power line to demonstrate its potential inspection capability [12].

Fig. 1 An UAV demonstrates its potential inspection capability [12]

To achieve the practical problem mentioned above. The Fig. 2 illustrates the UAV flight simulation process using a Multi-Armed Bandit approach. It begins with environment setup and initialization of the UCB1-Tuned algorithm. The UAV flight simulation then iteratively selects actions, updates its state, and checks constraints like collision and energy. Rewards are updated accordingly. Once the UAV reaches its target, the simulation ends, followed by calculating the overall success rate and energy consumption, ensuring optimization of flight performance.

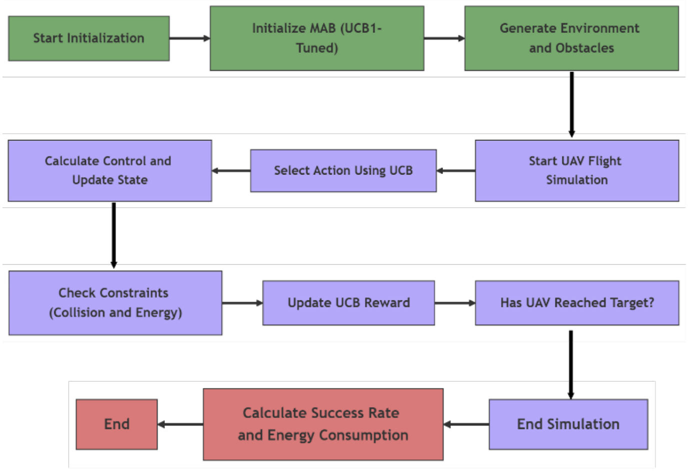

Fig. 2 Flowchart of the experiment (Picture credit: Original)

3.1 Experimental Setup

The experimental setup is designed to evaluate and compare the performance of different UAV flight control strategies under stochastic wind conditions and obstacle-rich environments. The key components of the setup include:

Simulation Environment: A three-dimensional space that includes randomly positioned spherical obstacles generated by sampling coordinates from a UAV flight log dataset and assigning each obstacle a random radius and altitude. Each simulation run contains 10 obstacles.

UAV Dynamics: The UAV motion is governed by a 6-dimensional state vector (position and velocity in 3D), updated at discrete time steps using either the RK4 or Euler integration method. External wind effects are modeled as directionally discretized vectors with speed uniformly sampled between 5 and 25 m/s.

Action Space: The control action space consists of a set of "arms" formed by combinations of discretized wind directions (8 directions) and acceleration magnitudes (3 levels: 0.3, 0.6, 1.0 m/s^2), yielding a total of 24 arms.

Bandit Algorithms: The UAV controller selects control actions using a multi-armed bandit approach. The UCB1-Tuned algorithm is used to balance exploration and exploitation by maintaining per-arm reward statistics and adapting confidence bounds based on empirical variance.

Metrics: Each experimental condition runs 800 UAV flights. Metrics evaluated include success rate (percentage of flights reaching the goal within a tolerance) and average energy consumption, which accounts for velocity and acceleration-related costs.

Comparisons: This paper compares UCB1-Tuned + RK4, Random action selection + RK4, and UCB1-Tuned + Euler integration methods to assess the influence of exploration strategy and numerical integration precision on control performance.

3.2 Two-Tier Architecture

3.2.1 Strategic Layer (Bandit)

The MAB problem considers N = 8 * 3=24 arms (combining 8 discrete wind directions with 3 acceleration levels), each with an unknown reward distribution. The goal is to maximize cumulative rewards by balancing exploration and exploitation through repeated arm selections.

Unlike Thompson Sampling, which is a Bayesian approach maintaining posterior Beta distributions for each arm, the experiment uses the UCB1-Tuned algorithm. This method estimates the mean reward and variance of each arm, and constructs an upper confidence bound for each arm to guide the selection in a deterministic manner.

UCB1-Tuned Algorithm Overview: For each arm k, maintain:

- n_k: the number of times arm k has been selected.
- $\hat{\mu}_k$: empirical mean reward of arm k.
- s_k^2: empirical variance estimate of arm k.

At each decision step t, select arm:

$$k^* = arg\,\max_k \left(\hat{\mu}_k + c * \sqrt{\frac{lnt}{n_k} min\left(\frac{1}{4}, s_k^2 + \sqrt{\frac{2lnt}{n_k}}\right)} \right) \tag{1}$$

where:

- k^*: selected arm at time step t.
- t: current total number of arm pulls (time step).
- n_k: number of times arm k has been selected so far.
- $\hat{\mu}_k$: empirical mean reward of arm k, representing the average observed reward.

- s_k^2: empirical variance of rewards of arm kk, representing reward variability.
- $c > 0$: exploration parameter that balances exploration and exploitation.
- $ln\ t$: natural logarithm of time step t, increasing exploration over time.
- $min\left(\frac{1}{4}, s_k^2 + \sqrt{\frac{2lnt}{n_k}}\right)$: a variance upper bound to stabilize the confidence bound.

Reward Function Design: The reward r_k depends on the expected travel time T_k and risk penalty R_k:

$$r_k = exp(-\lambda_T T_k + \lambda_R R_k) \tag{2}$$

where:

- r_k: reward for arm k, continuous and used as feedback.
- T_k: expected travel time associated with arm k.
- R_k: risk penalty associated with arm k.
- $\lambda_T, \lambda_R > 0$: weighting parameters controlling the influence of travel time and risk penalty, respectively.

The reward is continuous and used directly as feedback to update empirical means and variances for each arm in UCB1-Tuned.

Initialization and Update:

Step1: Initialize all arms with zero counts and zero mean and variance estimates.

Step2: At each iteration, select arm k^* according to the UCB1-Tuned formula.

Step3: Observe reward r_t.

Step4: Increment n_k^*, update $\hat{\mu}_k^*$ and $s_{k^*}^2$ incrementally using observed rewards.

This approach avoids sampling from distributions and relies on confidence bounds calculated from observed statistics to guide arm selection.

3.2.2 Tactical Layer (Numerical)

- UAV Dynamics Model (with Wind): State vector: $P = (x\ , y\ , z)$ is position and $V = (v_x, v_y, v_z)$ is velocity. The combined dynamics can be represented by the ODE:

$$\frac{dX}{dt} = f(X, u) \tag{3}$$

where the state vector is $X = (P, V)$. The function f captures the dynamics defined by:

$$\frac{dp}{dt} = v + w(t) \tag{4}$$

$$\frac{dv}{dt} = a(u, x) - d(v) \tag{5}$$

where $w(t)$ is the wind velocity vector. $a(u, x)$ is the control acceleration depending on control input u and position x. $d(v)$ is the drag or damping forces acting on velocity V.

- RK4 Integration: For ODE:

$$\frac{dx}{dt} = f(t, x) \tag{6}$$

Time step h update from t to $t + h$ via:

$$\begin{cases} k_1 = f(t, x_t) \\ k_2 = f\left(t + \frac{h}{2}, x_t + \frac{h}{2}k_1\right) \\ k_3 = f\left(t + \frac{h}{2}, x_t + \frac{h}{2}k_2\right) \\ k_4 = f(t + h, x_t + hk_3) \end{cases} \tag{7}$$

The state is updated as:

$$X_{t+h} = X_t + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{8}$$

RK4 updates x considering current dynamics with wind and control.

- Wind-Adaptive Control: Wind vector $w(t)$ varies with time and space from turbulence models. Control input adapts in real time:

$$u(t) = \pi\big(x(t), w(t)\big) \tag{9}$$

where π is the control policy compensating wind disturbances for a stable trajectory.

- Trajectory Propagation Workflow: First, Initialize x_0 at the start position. Then, Loop over simulation time steps with size h. At each step: Obtain current wind $w(t)$, calculate control $u(t)$ based on current state and wind, apply RK4 to find $x(t+h)$, check constraints. Repeat until the mission is complete or the termination condition is met.

3.3 Optimization Constraints

To ensure the safety and operational integrity during UAV missions, a minimum obstacle clearance of 2 meters is mandated. This clearance is crucial for avoiding collisions with dynamic objects, such as moving vehicles and pedestrians, as well as static obstacles like buildings and trees. Maintaining this safety buffer not only protects the UAV but also minimizes risks to individuals and property below. The system should continuously monitor the distance to obstacles in real-time using onboard sensors, and adjustments to the UAV's flight path should be made whenever the clearance falls below the required threshold.

Effective energy management is essential for UAV operations, especially for lengthy missions requiring sustained flight. A requirement of maintaining at least 20% energy reserve is enforced to ensure mission completion safely and effectively. This reserve allows the UAV to have enough power available for emergency maneuvers, unexpected diversions, or additional tasks that may arise during the operation. The UAV's onboard systems should be equipped with algorithms to predict energy consumption based on current flight conditions, including altitude, speed, and payload, allowing for optimal route planning within the energy reserve constraints.

For coordinated operations involving multiple UAVs, synchronization of the mission clock is paramount. All UAVs must maintain synchronized timing to execute tasks effectively, such as formation flying, coordinated payload delivery, or simultaneous data collection. This synchronization reduces the risk of timing discrepancies that could lead to collisions or inefficient operation. A centralized timekeeping system or an agreed-upon time source is recommended to calibrate each UAV’s internal clock upon mission initiation. Regular synchronization checks should be performed throughout the mission to ensure all UAVs remain aligned in timing.

3.4 Evaluation Metrics and Quantitative Formulas

3.4.1 Success Rate

Defined as the proportion of UAVs that successfully complete the mission within the target tolerance without violating constraints:

$$SuccessRate = \frac{N_{success}}{N_{total}} \tag{10}$$

$N_{success}$ is the number of UAVs that reached the goal within tolerance and respected all constraints. N_{total} is the total number of UAV simulations.

3.4.2 Average Energy Used

For each successful mission, energy consumption is calculated as the fraction of consumed energy from a full initial energy of a UAV:

$$\bar{E}_{used} = \frac{1}{N_{success}} \sum_{i=1}^{N_{success}} \left(1 - E_i^{left}\right) \tag{11}$$

E_i^{left} is the remaining energy of the i^{th} UAV at mission completion. Lower values indicate more energy-efficient missions.

3.4.3 Stepwise Reward

At each step t, a reward is assigned depending on progress toward the goal and constraint satisfaction. If all constraints are satisfied:

$$r_t = 1 - \frac{d_t}{d_{max}} \tag{12}$$

$$d_t = \ ||p_t - p_{goal}|| \tag{13}$$

d_t is the Euclidean distance between the current position p_t and the goal position p_{goal}. d_{max} is the initial distance between the start and goal. The cumulative reward reflects trajectory efficiency and safety. If constraints violated: $r_t = 0$.

3.4.4 Obstacle Clearance Constraint

The minimum distance between UAV and any obstacle must satisfy: $\min_{t} \min_{o \in O} ||p_t - o|| \geq r_o + d_{safe}$. p_t is the UAV position at time t, o is the obstacle center position, r_o is the obstacle radius,$d_{safe} = 2.0m$ is the minimum clearance distance, violation leads to zero reward and invalid trajectory in simulation.

3.4.5 Energy Reserve Constraint

Remaining energy at any time must satisfy: $E_t^{left} \geq E_{reserve} = 0.2$. If energy drops below this threshold, the simulation terminates early, marking failure.

4 Experimental Design and Results

This paper conducted a series of batch simulations comparing three different UAV control methods, each running 800 flights in a simulated obstacle environment. The following subsections describe the experimental design for each method, along with their quantitative results.

4.1 UCB1-Tuned with RK4 Integration

4.1.1 Experimental Setup

This method employs the UCB1-Tuned multi-armed bandit algorithm for intelligent action selection combined with the precise RK4 numerical integrator to simulate UAV dynamics. The 24-arm action space covers 8 discrete wind directions and 3 acceleration levels.

4.1.2 Results:

First, the experiment achieved a remarkably high success rate of 99.38%. Besides that, it maintained efficient energy usage with an average consumption of 9.28%, as shown in Fig. 3. Fig. 3 compares the average energy consumption of three UAV control methods. The UCB1-Tuned algorithm, combined with the RK4 integrator and Euler method, both demonstrate significantly lower energy usage compared to the Random method with RK4. Notably, UCB1-Tuned with RK4 achieves the lowest energy consumption, indicating its efficiency in optimizing UAV flight dynamics. This suggests that leveraging intelligent action selection with precise numerical integration can greatly reduce energy costs, improving the sustainability and effectiveness of UAV operations.

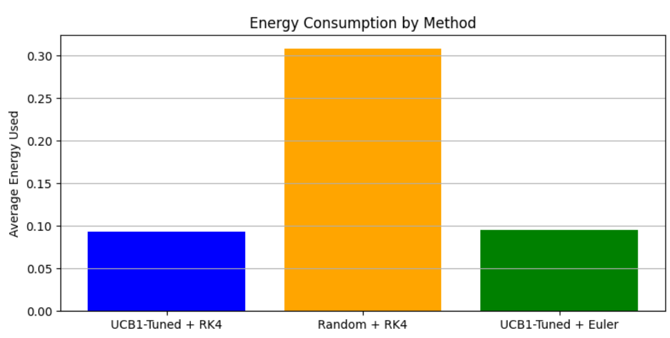

Fig. 3 Comparison of energy consumption among three methods (Picture credit: Original)

Furthermore, it demonstrated steady reward accumulation over steps, as shown in Fig. 4. This confirms the effectiveness of combining the UCB1-Tuned bandit algorithm with high-order integration for safe and energy-efficient flight control.

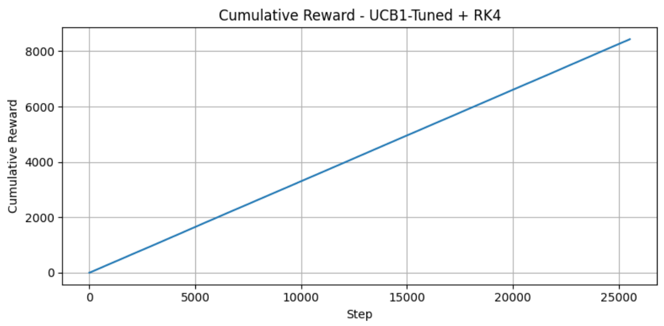

Fig. 4 The relationship between steps and cumulative reward (Picture credit: Original)

4.2 Random Action Selection with RK4 Integration

4.2.1 Experimental Setup

This baseline method uses random action selection within the same 24-arm action space combined with the RK4 integrator for state updates. It serves to highlight the benefit of intelligent action selection over naive sampling.

4.2.2 Results

First, it performed poorly with only a 6.00% success rate,

and it consumed substantially more energy, averaging 35.74% energy usage (also in Fig. 3). Also, it exhibited erratic cumulative reward growth, reflecting ineffective navigation.

The results emphasize the necessity of intelligent adaptive strategies for UAV operation under uncertain and dynamic conditions.

4.3 UCB1-Tuned with Euler Integration

4.3.1 Experimental Setup

This method tests the UCB1-Tuned bandit algorithm combined with the simpler Euler forward integration method for UAV dynamics.

4.3.2 Results

First, it achieved a success rate of 99.00%, a little bit lower compared to the RK4 variant. Apart from that, it used slightly higher average energy at 9.56% (also in Fig. 3). It also indicates that although Euler integration is less precise, it still supports effective control when combined with UCB1-Tuned. In total, it shows that RK4 provides marginal improvements in accuracy and energy efficiency.

4.4 Summary of Quantitative Results

Table 1 shows that the UCB1-Tuned + RK4 method achieved a remarkably high success rate (99.38%) with very efficient energy consumption (9.28% on average), demonstrating the effectiveness of the UCB1-Tuned bandit algorithm combined with the precise RK4 integrator.

Table 1 Key quantitative results

Method	Success Rate	Average Energy Used
UCB1-Tuned + RK4	99.38%	9.28%
Random + RK4	6.00%	35.74%
UCB1-Tuned + Euler	99.00%	9.56%

In contrast, the random policy performed poorly, with only a 6% success rate and substantially higher energy usage, highlighting the importance of intelligent action selection in UAV flight control under uncertainty.

The UCB1-Tuned, combined with the Euler integrator, also performed well, with a 99% success rate and slightly higher energy usage than RK4, indicating that the simpler integration method is effective, but RK4 provides marginal improvements in accuracy and efficiency.

The bandit arm selection counts show a strong preference for particular arms, indicating that UCB1-Tuned reliably identifies more rewarding wind direction and acceleration combinations.

5 Conclusion

This work demonstrates the effectiveness of using the UCB1-Tuned algorithm combined with accurate numerical integration (RK4) for controlling UAV flight under uncertain wind conditions and obstacle-rich environments. Compared to a random selection baseline, the UCB1-Tuned + RK4 approach achieves a significantly higher success rate (99.38%) with substantially lower average energy consumption (9.28%). The results also show that while the simpler Euler integration method paired with UCB1-Tuned is effective (99% success rate), RK4 integration yields slight improvements in energy efficiency and stability.

The adaptive nature of the UCB1-Tuned algorithm allows the controller to efficiently explore and exploit the discrete joint action space of wind directions and accelerations, rapidly focusing on favourable combinations. This approach is robust, scalable, and can be extended to other model-based control problems involving stochastic dynamics and safety constraints. Future research can extend this framework in several directions:

Firstly, incorporating continuous action spaces and leaning on function approximation methods (such as contextual bandits or reinforcement learning) for more complex and larger state-action domains.

Secondly, integrating more realistic UAV dynamics models with additional degrees of freedom and considering varying payload or battery models.

Thirdly, studying non-stationary environments with time-varying wind distributions and dynamic obstacles, possibly by using discounted or sliding-window variants of the bandit algorithm.

Apart from that, it can also combining UCB-based action selection with onboard sensing and real-time adaptation to enhance robustness in real flight scenarios.

Finally, exploring multi-agent cooperation where multiple UAVs coordinate flight paths using decentralized bandit approaches.

References

1. Khatib, O.: 'Real-time obstacle avoidance for manipulators and mobile robots', The International Journal of Robotics Research, 1986, 5, (1), pp. 90-98

2. Wang, B.H., Wang, D.B., Ali, Z.A., Ting Ting, B., Wang, H.: 'An overview of various kinds of wind effects on unmanned aerial vehicle', Measurement and Control, 2019, 52, (7-8), pp. 731-739
3. Mohamed, E.M., Hashima, S., Aldosary, A., Hatano, K., Abdelghany, M.A.: 'Gateway selection in millimeter wave UAV wireless networks using multi-player multi-armed bandit', Sensors, 2020, 20, (14), pp. 3947
4. Lin, Y., Wang, T., Wang, S.: 'UAV-assisted emergency communications: An extended multi-armed bandit perspective', IEEE Communications Letters, 2019, 23, (5), pp. 938-941
5. Amrallah, A., Mohamed, E.M., Tran, G.K., Sakaguchi, K.: 'UAV trajectory optimization in a post-disaster area using dual energy-aware bandits', Sensors, 2023, 23, (3), pp. 1402
6. Wang, Z., Hong, T.: 'Reinforcement learning for building controls: The opportunities and challenges', Applied Energy, 2020, 269, pp. 115036
7. Arroyo, J., Manna, C., Spiessens, F., Helsen, L.: 'Reinforced model predictive control (RL-MPC) for building energy management', Applied Energy, 2022, 309, pp. 118346
8. Hewing, L., Wabersich, K.P., Menner, M., Zeilinger, M.N.: 'Learning-based model predictive control: Toward safe learning in control', Annual Review of Control, Robotics, and Autonomous Systems, 2020, 3, (1), pp. 269-296
9. Schwenzer, M., Ay, M., Bergs, T., Abel, D.: 'Review on model predictive control: An engineering perspective', The International Journal of Advanced Manufacturing Technology, 2021, 117, (5), pp. 1327-1349
10. Larrabee, T., Chao, H., Rhudy, M., Gu, Y., Napolitano, M.R.: 'Wind field estimation in UAV formation flight'. Proc. 2014 American Control Conference, June 2014, pp. 5408-5413
11. Xiang, X., Wang, Z., Mo, Z., Chen, G., Pham, K., Blasch, E.: 'Wind field estimation through autonomous quadcopter avionics'. Proc. IEEE/AIAA 35th Digital Avionics Systems Conference (DASC), September 2016, pp. 1-6
12. Hrabar, S.: '3D path planning and stereo-based obstacle avoidance for rotorcraft UAVs', IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008, September, pp. 807-814. IEEE.