Risk-aware UAV-UGV Rendezvous with Chance-Constrained Markov Decision Process

Guangyao Shi[1], Nare Karapetyan[1], Ahmad Bilal Asghar[1], Jean-Paul Reddinger[2],
James Dotterweich[2], James Humann[2], Pratap Tokekar[1]

Abstract— We study a chance-constrained variant of the cooperative aerial-ground vehicle routing problem, in which an Unmanned Aerial Vehicle (UAV) with limited battery capacity and an Unmanned Ground Vehicle (UGV) that can also act as a mobile recharging station need to jointly accomplish a mission such as monitoring a set of points. Due to the limited battery capacity of the UAV, two vehicles sometimes have to deviate from their task to rendezvous and recharge the UAV. Unlike prior work that has focused on the deterministic case, we address the challenge of stochastic energy consumption of the UAV. We are interested in finding the optimal policy that decides when and where to rendezvous such that the expected travel time of the UAV is minimized and the probability of running out of charge is less than a user-defined tolerance. We formulate this problem as a Chance Constrained Markov Decision Process (CCMDP), which can be equivalently transformed into a Constrained Markov Decision Process (CMDP). To the best knowledge of the authors, this is the first CMDP-based formulation for the UAV-UGV routing problems under power consumption uncertainty. We adopt a Linear Programming (LP) based approach to solve the problem optimally. We demonstrate the effectiveness of our formulation in the context of an Intelligence Surveillance and Reconnaissance (ISR) mission.

I. Introduction

Unmanned Aerial Vehicles (UAVs) are increasingly being sought in applications such as surveillance, environmental monitoring, and agriculture due to their ability to monitor large areas in a short period of time. One bottleneck in practice that limits their application is the limited battery capacity, especially for multi-rotor UAVs. One way to overcome this bottleneck is to use a team of aerial and ground vehicles for such tasks, in which the UGV can work as a mobile recharging station and will recharge the UAV during long-range operations. The key to achieving such cooperation on the decision-making level is to design efficient routing algorithms that can tell robots *which task node to visit next*, and *when and where the UAV should be recharged*. Moreover, the rate of battery discharge of a UAV is stochastic in the real world. The routing algorithm should be able to deal with such uncertainties, e.g., trade off task performance with failure risks.

In this paper, we consider the cooperative routing problem with a team of a single UGV that can work as a mobile
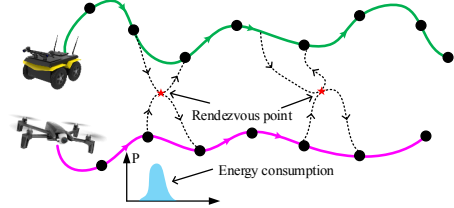

Fig. 1. An illustrative example of the rendezvous problem considered in this paper. When the UAV and UGV are executing tasks, they need to decide when and where to rendezvous to replenish the battery of the UAV while minimizing the travel time of the UAV and satisfying the risk constraint induced by stochastic energy consumption. When they need to rendezvous, they will deviate from their task and meet at a chosen rendezvous location.

charger and a single energy-constrained UAV, in which the UAV and UGV need to complete a task by visiting task nodes distributed throughout the task area. The UGV can only move on the road network but the UAV can directly fly between any pair of nodes (assuming it has enough charge). Given the task nodes to visit and the stochastic energy consumption model of the UAV, we are interested in finding a routing strategy for the UAV and the UGV such that the expected time to finish the task is minimized and the probability of running out of charge is less than a user-defined tolerance. Such problems can be formulated within the stochastic programming (SP) framework [1]. However, since we need to consider not only routes but also recharge decisions and chance constraints, SP-based formulation would involve too many variables, rendering the formulation only solvable for very small instances.

In this paper, we propose to find the routing strategy in two decoupled phases. In the first phase, a higher-level planner finds deterministic routes for the UAV and the UGV without considering stochasticity in energy consumption based on the task requirement. In the second phase, a risk-aware planner will refine the routes generated from the previous phase to find when and where to rendezvous to satisfy the chance constraint while minimizing the time to finish the UAV task[1]. Our focus in this paper is mainly on the second phase. We formulate our risk-aware refinement as a Markov Decision Process (CMDP), in which the chance constraint is modeled

[1]We focus on minimizing the time taken for the UAV task instead of the total time which would be the maximum of the UAV and UGV travel times. If the UGV task takes longer than the UAV, then once the UAV's task is done, the UGV simply executes the remaining portion of its route. Therefore, optimizing the UAV time is appropriate. Furthermore, In many applications [2], the UGV's task is to simply act as a mobile recharging station

as the secondary cost in the constraint. To the best of our knowledge, this is the first CMDP-based formulation for the UAV-UGV routing problems under energy consumption uncertainty. We use Linear Programming (LP) to find the optimal stationary policy. We validate our formulation and the solution in an Intelligence Surveillance and Reconnaissance (ISR) mission.

The main contributions of this paper are:

- We show how to formulate the stochastic cooperative UAV and UGV rendezvous problem with energy constraint as a CMDP and use LP to solve the problem optimally in polynomial time.
- Demonstrate the effectiveness of the formulation and the solution using realistic models that are obtained using field data.

The rest of this paper is organized as follows. We first give a brief overview of the related work in Section II. Then we discuss the general problem definition along with its mathematical formulation as a Chance-Constrained Markov Decision Process (CCMDP) in Section III. Next in Section IV, we discuss the solution to this problem as an LP instance. Finally, we present the results from numerical simulations in Section V with conclusions in Section VI.

II. RELATED WORK

The routing of energy-constrained UAVs with stationary recharging stations or assistive UGVs has been studied extensively [3]–[6]. Even with deterministic environmental changes or stationary conditions, this problem can be reduced to the Traveling Salesman Problem (TSP), [7], [8] making it an NP-hard problem.

The cooperative UAV and UGV routing problem has been studied from different perspectives and thus received various formulations in the literature. It is most commonly formulated as a type of vehicle routing problem. Manyam et al. [9] use a team of one UAV and one UGV with communication constraints to cooperatively visit targets. Along with an exact solver to Mixed Integer Linear Programming (MILP) formulation, they also provide heuristic reduction to the generalized traveling salesman problem (GTSP). Maini et al. [10] present a two-fold strategy: first, they identify feasible rendezvous points, then they formulate a MILP to find the optimal routes for the UAV and UGV. Thayer et al. [11] present a solution to the Stochastic Orienteering Problem, where the objective is to maximize the sum of rewards associated with each visited node while constrained by the maximum budget over edges with stochastic cost.

Murray and Chu [12] introduced the flying sidekick TSP (FSTSP) for parcel delivery systems, which was later adopted in last-mile delivery applications using drones [13], [14]. In literature, the term multi-echelon scheme is often used for systems where delivery consists of multiple layers. Specifically, the two-echelon vehicle routing problem (2E-VRP) is concerned with finding minimal cost routes to deliver packages with trucks/UGVs and drones [4], [15]. An important differentiation from the original vehicle routing problem is the synchronization of UAV and UGV tasks.

Learning-based approaches have been used to address cooperative UAV and UGV routing problems. Ermugan et al. [16] also propose a two-phase approach. First, they find a route for UAVs without taking into account the energy constraints. Then, the planner learns to insert into the route recharging stations and replans a new TSP route. Reinforcement learning has also proven to be a possible approach to solving this problem [17].

In our previous work, we studied cooperative planning with a single UGV and an energy-constrained UAV as well [2], [7], [18]. Our proposed approach in [18] demonstrated how to maximize the number of sites visited in a single charge in conjunction with the ability to land a UAV on top of a UGV to be transported to the next take-off site. We extended this in [2] to allow the UAV to also be recharged while either being transported or stationary on the UGV. We extended the latter to the area coverage path planning problem by formulating it as a GTSP [7]. Here, we extend this body of work by introducing the stochasticity of the UAV's energy consumption and by assuming that the UGV has its own required set of tasks to be carried out. To the best of our knowledge and based on the presented literature review, none of the works takes into account the stochastic nature of energy consumption.

III. PROBLEM FORMULATION

The cooperative routing problem studied in this paper involves one UAV and one UGV. The UAV and the UGV are executing tasks, which are given by some task planners as shown in Section V-A. The UAV needs to visit a sequence of task nodes in order to finish the task, but its battery may not be enough for it to finish the task in a single flight without recharging. Also, the energy consumption of the UAV is stochastic. The UAV needs to decide *when* and *where* it should rendezvous with the UGV to replenish the battery while minimizing the total travel time to finish the task. When the UAV decides to rendezvous with the UGV to replenish power, both the UAV and the UGV will take a detour (as defined in Sec. III-D) from their respective tasks and go back to their tasks after recharging.

At a high level, the problem studied in this paper is stated below.

Problem 1 (Risk-aware UAV-UGV rendezvous). *Given a route of nodes for the UAV $\mathcal{T}_A$, a route of nodes for the UGV $\mathcal{T}_G$, and the stochastic energy consumption model and battery capacity of the UAV, find a policy for the UAV to decide when and where to rendezvous with the UGV for recharging such that the total travel time is minimized and the probability of running out of charge during flight is less than a given tolerance.*

Next, we will present the setup and the assumptions that we use in this paper. Then we will present our CCMDP based formulation and show how to transform a CCMDP into a CMDP.

A. Environment and Task Model

Our problem considers a two-dimensional Euclidean space, which consists of a road network graph $G = (V_r, E)$ and a set of task points V_t for the UAV to visit.

The UGV has to move on the road network and its task is specified as a sequence of nodes of the road network. UAV's task is specified by some task planners using nodes in V_t. More details on the task planner will be discussed in the Section V-A. Both UAV and UGV should follow the task specification to visit the task nodes in order and they will deviate from the task route to rendezvous when necessary.

B. Vehicle Motion Model

The UGV will move at a fixed speed v_g when it transits between two nodes in the road network. When the UAV transits between two nodes, it will fly with either the best endurance speed, v_{be}, or the best range speed, v_{br}. The best endurance speed is the speed at which the energy consumption rate is minimized. At this speed, the propellers of the multirotor operate more efficiently than in hover, and the UAV is capable of the greatest flight duration. By contrast, when a UAV flies at the best range speed, it minimizes the derivative of energy consumption rate with respect to velocity. This flight speed results in a lower flight duration than operation at v_{be}, but will allow a greater range to be traveled per unit of energy. For a no-wind condition, the velocity of the best range is always better than the velocity of best endurance.

C. Recharging and Stochastic Energy Consumption Model

We assume that it takes constant time T to finish the recharging process, which includes the landing/take-off time and battery-swapping time.

In this paper, we only consider the power consumption when a UAV traverses the route with the assumption that the power needed for computation, takeoff, and landing has been reserved by the power management system. As described in the transition model, the UAV will fly at a fixed speed when it transits between two nodes in the environment. However, given that constant speed, the energy consumption is stochastic considering the disturbances in the environment.

Given the distance l between two task nodes and the flying speed v, the energy consumption can be computed as

$$e_{l,v} = \int_{t=0}^{\frac{l}{v}} P_{\Theta}(v)dt, \tag{1}$$

where $P_{\Theta}(v)$ is the power consumption of the UAV when it flies at a speed v, and Θ is a vector of parameters for stochastic variables.

D. Rendezvous Model

When the UAV reaches a node in $\mathcal{T}_A$ and decides to rendezvous with the UGV, the UAV and UGV will deviate from their task temporarily to finish the rendezvous process. There are two steps in the rendezvous process. In the first step, the UAV and UGV will meet at a rendezvous point as shown in Fig. 2 and in the second step, they will go to
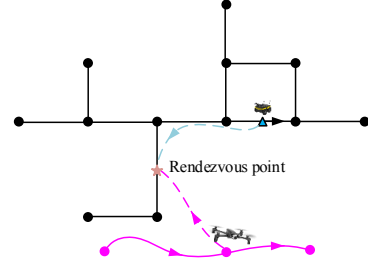

Fig. 2. First step in the rendezvous process. The UGV (blue triangle) needs to deviate from its task node to rendezvous with the UAV at the rendezvous point (pink star). The rendezvous paths are in dashed lines.

the next task node in $\mathcal{T}_a$ and $\mathcal{T}_g$ respectively. We want to optimize the time consumed in these two steps to find the optimal rendezvous points.

Let $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}_+$ be the distance metric between two points in the Euclidean space. We use $d_G : V_r \times V_r \to \mathbb{R}_+$ to denote the length of the shortest path between two nodes in the road network. We use $\mathcal{T}_a(k)$ to denote the position of the UAV when it decides to rendezvous at the kth node in its task route and $\mathcal{T}_a(k+1)$ to denote the next position to visit for UAV after the rendezvous. With a slight abuse of notation, we use $\mathcal{T}_g(k)$ to denote the position of the UGV in the road network when the UAV decides to rendezvous at the kth node in its task route. With the above notations, the problem to find the rendezvous point can be stated below.

Problem 2 (Where to rendezvous). *Given the positions of UAV ($\mathcal{T}_a(k)$) and UGV ($\mathcal{T}_g(k)$) at the beginning of the rendezvous process, UAV's next position to go $\mathcal{T}_a(k+1)$, UAV's flight speed v_a, UGV's transition speed v_g, and the road network G, we want to find a rendezvous point $p_r \in G$ such that the time consumed in the rendezvous process is minimized. Mathematically,*

$$\min_{\Delta \geq 0,\ p_r \in G} \Delta + \frac{d(p_r,\ \mathcal{T}_a(k+1))}{v_a} \tag{2}$$

$$s.t.\ \Delta = \max\left(\frac{d_G(\mathcal{T}_g(k),\ p_r)}{v_g},\ \frac{d(\mathcal{T}_a(k),\ p_r)}{v_a}\right). \tag{3}$$

In the first step of the rendezvous process, if the UAV or UGV reaches the rendezvous first, it has to wait for the other vehicle. Therefore, the time consumed in the first step is decided by the vehicle that reaches the rendezvous point later than the other. We encode this fact in the optimization problem by introducing the variable Δ, which describes the maximum time needed for both UAV and UGV to reach the rendezvous point. The time consumed in the second step of the rendezvous process is the time needed for the UAV to fly back to its next task node.

Problem 2 can be solved by iterating over the nodes in the road network as what we do in the case study. But such a method will increase the time to extract transition information for the CMDP. A more efficient way to solve Problem 2 is left for our future work.
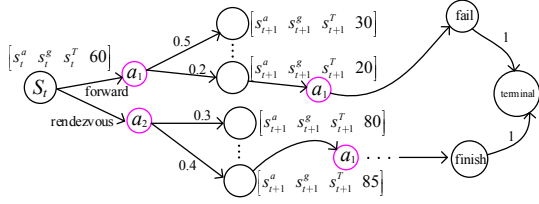
Fig. 3. State transition graph in CMDP.

E. Chance-Constrained Markov Decision Process

One natural choice to model the sequential decision-making problems described in Problem 1 is to use MDP. In this section, we first show how to formulate Problem 1 as a CCMDP and then show how to transform a CCMDP into a CMDP in the following section.

The rigorous definition of an MDP can be found in [19]. Here we define the MDP from the perspective of the application. The MDP corresponding to Problem 1 is defined as a tuple $\mathcal{M} = (S, A, T, C, s_0)$, where

- $S = \mathcal{T}_a \times S_g \times \mathcal{T}_g \times \mathcal{B} \cup \{s_{ob}, s_l\}$ is the state space of the problem, where $\mathcal{T}_a$ here is used as an un-ordered set, which describes all possible positions of UAV in a task route; S_g is the set of positions of UGV and this information is needed when we compute the rendezvous points; $\mathcal{T}_g$ here is used as an un-ordered set, which describes the task nodes UGV will visit. $\mathcal{T}_g$ is included in the state space to inform the MDP about the next node the UGV needs to visit after a rendezvous. Without this information, the system will be non-Markovian; $\mathcal{B}$ is a discretized variable for describing the state of the charge of the UAV; s_{ob} is one failure state representing the out-of-charge state and the UAV will transit to this state whenever it cannot finish its task route; s_l is added as an absorbing state and UAV will transit to this state when it either finishes UAV's route or runs into a failure state. One illustrative example of state transitions is given in Fig. 3.
- A is the action space of the UAV. If the UAV has not finished its route and is not in a failure state, there are four actions for the UAV to choose: 1) v_{be}: move to the next node in $\mathcal{T}_a$ with the best endurance velocity. 2) v_{br}: move to the next node in $\mathcal{T}_a$ with the best range velocity. 3) v_{be_be}: rendezvous with the best endurance velocity. 4) v_{br_br}: rendezvous with the best range velocity. When UAV is in a failure state or has finished its route, there is only one action that makes the system transit to the terminal state s_l.
- $T_a(s, s', a) = P(s' \mid s, a)$ is the transition function, which depends on the stochastic energy consumption model. When UAV chooses to move forward to its next task node, its battery state at the destination node is a random variable that depends on the current battery state and Equation (1). Since we have discretized the battery charging levels at each node, the probability of reaching the destination node with a given battery charge can be calculated using Equation (1). When it cannot reach the next task node, i.e., with non-zero probability, it will run out of charge,

it transits to the failure state s_{ob}. When UAV chooses to rendezvous, a rendezvous point is first computed by solving the Problem 2. Then the distribution of battery remaining when it reaches the rendezvous point can be computed based on Equation (1). The non-positive portion of the distribution corresponds to the failure probability. After recharging, UAV will transit to its next task node starting with a full battery. When the UAV transits to the failure state or it finishes the task route, it will transit to the terminal state s_l with probability 1 as shown in Fig. 3. In the terminal state s_l, the system will loop over this state.

- $C(s, s', a)$ is the cost function for the UAV. We define it as the time needed to transit between two states. If the UAV chooses to move to the next node, the cost will be time consumed during that transition. If a UAV chooses to rendezvous, the cost will be the sum of the time consumed in two steps of the rendezvous process. When the state transits to the failure state or to the terminal state, it takes zero cost.
- s_0 is the initial state of the system.

Definition 1 (Risk). Let π be a policy, the risk of the policy given initial state s_0 is defined as

$$\rho^\pi(s_0) = \mathbf{P}(\exists\ t\ s_t = s_{ob} \mid s_0). \tag{4}$$

We seek the optimal policy π^* that satisfies

$$\pi^* = \underset{\pi}{\operatorname{argmin}} \mathbb{E}\left[\sum_{i=0}^{\infty} C(s_i, \pi(s_i))\right] \tag{5}$$

$$\text{s.t.}\ \ \rho^\pi(s_0) \le \delta, \tag{6}$$

where δ is the user-specified risk tolerance.

F. Constrained Markov Decision Process

We can transform a CCMDP into a CMDP by introducing a new cost function $\overline{C}: S \times S \times A \to \{0, 1\}$ [20]. As shown in Fig. 3, when the system transits from a non-failure state to the failure state s_{ob}, it will incur a cost of one and other transitions will incur zero cost. The new cost function $\overline{C}$ is defined as

$$\overline{C}(s, a, s') = \begin{cases} 1 & \text{if } s \neq s_{ob} \text{ and } s' = s_{ob} \\ 0 & \text{else.} \end{cases} \tag{7}$$

As shown in [20] [Proposition 4.1], the risk can be defined using the new cost function $\overline{C}$ as

$$\rho^\pi(s_0) = \mathbb{E}\left[\sum_{i=0}^{\infty} \overline{C}(S_i, \pi(S_i)) \mid s_0\right]. \tag{8}$$

As a result, the CCMDP problem can be formulated as

$$\pi^* = \underset{\pi}{\operatorname{argmin}} \mathbb{E}\left[\sum_{i=0}^{\infty} C(s_i, \pi(s_i))\right] \tag{9}$$

$$\text{s.t.}\ \ \mathbb{E}\left[\sum_{i=0}^{\infty} \overline{C}(S_i, \pi(S_i)) \mid s_0\right] \le \delta. \tag{10}$$

IV. SOLUTIONS TO CMDP

A CMDP can be solved using Linear Programming (LP) [21], [22]. The decision variables y in LP are the occupancy measure for each state-action pair and are defined as

$$y(s, a) = \sum_t \Pr(S_t = s, A_t = a). \quad (11)$$

The LP is formulated as:

$$\min_{y(s,a), \forall s, a} \sum_{(s,a) \in S \times A} y(s, a) C(s, a) \quad (12)$$

$$\text{s.t.} \sum_{s,a} y(s, a) \overline{C}(s, a) \leq \delta \quad (13)$$

$$\sum_{a'} y(s', a') = \mathbb{I}(s', s_0) + \sum_{s,a} y(s, a) \Pr(s' \mid s, a) \quad (14)$$
$$\forall s' \in S \setminus \{s_l\}$$

$$y(s, a) \geq 0 \ \forall s, a, \quad (15)$$

where $\mathbb{I}(s', s_0)$ is a Dirac delta function that returns 1 when $s' = s_0$ and 0 otherwise. This LP corresponds to the dual linear program for MDPs [21] with one extra cost constraint (13), which enforces that the cost of entering the failure state be lower than the predefined risk tolerance. Constraint (14) is a flow conservation constraint to define valid occupancy measures and is defined by the initial state and the transition probability (see [21], ch. 8 for details). The last constraint (15) is added to guarantee that $y(s, a)$ is non-negative.

If LP admits a solution, we can construct the policy from the occupancy measures by normalizing them:

$$\pi^*(s, a) = \frac{y(s, a)}{\sum_{a'} y(s, a')} \ \forall (s, a) \in S \times A, \quad (16)$$

where $\pi^*(s, a)$ is the probability of taking action a in the state s in the optimal stationary randomized policy. If Eq. (16) has a zero denominator, which suggests that state s is not reachable from s_0, the policy for (s, a) can be defined arbitrarily.

An illustrative example to explain the policy extracted from the solution of LP is given in Fig. 4. When the UAV reaches s_k^a, it has 20% battery remaining. It can choose to rendezvous to replenish itself or move forward to its next task node s_{k+1}^a. If the UAV chooses to rendezvous, the distribution of the battery remaining at the rendezvous point is shown in Fig. 4.(a). The distribution corresponding to forward action is given in Fig. 4.(b). If we set $\delta = 0.01$, the policy constructed from LP results in rendezvous with a probability of 0.997 and forward with 0.003. Otherwise, if the UAV were to choose to move forward, it would have less than 10% battery charge remaining with a probability of 0.2, which is not enough to ensure a rendezvous or forward action at s_{k+1}^a, suggesting the failure probability is at least 0.2. By contrast, if we set the $\delta = 0.5$, the policy becomes that rendezvous with a probability of 0.504 and moves forward with 0.496. One reason for such a policy is that with a probability of 0.8 the UAV will have 10% battery remaining when it moves to s_{k+1}^a, where a rendezvous process with a
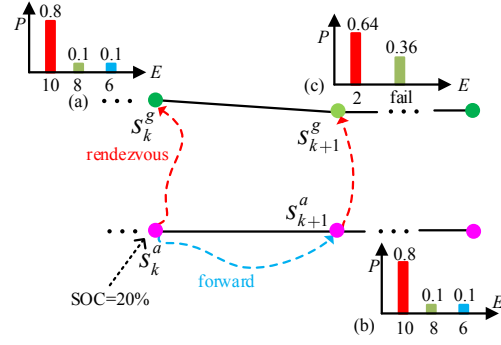

Fig. 4. An illustrative example to explain the policy constructed with occupancy measures from LP. When the UAV is at state s_k^a with State of Charge (SOC) equal 20%, it has two actions: move forward (blue dashed line) to its next task node s_{k+1}^a or rendezvous with the UGV (red dashed line) to recharge itself. The distribution of SOC after taking rendezvous is shown in (a) and that for forward action is shown in (b). (c) shows the distribution of SOC if the UAV first chooses forward action at s_k^a and then chooses rendezvous s_{k+1}^a.

shorter travel distance may be possible with a lower failure probability than the risk tolerance as shown in Fig.4.(c).

V. EXPERIMENTS

In this section, we first present a qualitative example to show what the input and output look like for our problem. Next, we study how system parameters (different risk tolerances) influence the rendezvous behaviors between the UAV and the UGV. Then, we present quantitative results for the ISR application that motivates our research. Specifically, we will use Monte Carlo (MC) simulations to evaluate 1) the satisfaction of the risk constraint for the policy constructed from LP; 2) the effectiveness of the policy in minimizing the expected task duration; 3) the risk tolerance-task duration Pareto curves. Moreover, the running time of LP for CMDP is empirically evaluated. All experiments are conducted using Python 3.8 on a PC with the i9-8950HK processor. LP is solved using Gurobi 9.5.0.

A. Task Route Planner

The task routes $\mathcal{T}_a$ and $\mathcal{T}_g$ used in Problem 1 can be either generated jointly by some existing task planners [2], [9] or can be generated by separately by different task planners. In our case study, the task for the UGV is to persistently monitor nodes A, B, and C (blue squares in Fig. 5). The task nodes for the UAV are red dots in Fig. 5 and the task route (from node 0 to 18 and back to 0) is generated by a planner for Traveling Salesman Problem (TSP).

B. System Models

TABLE I
COEFFICIENTS FOR STOCHASTIC ENERGY CONSUMPTION MODEL

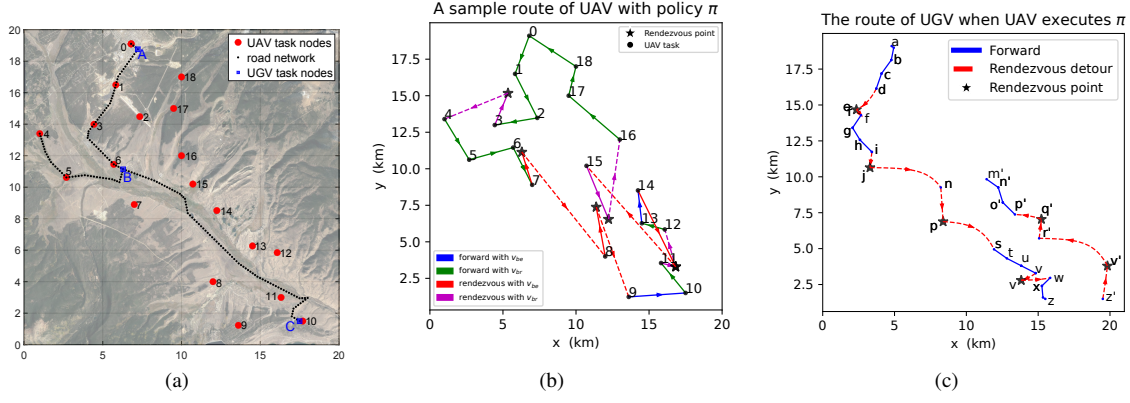|	b_0	b_1	b_2	b_3	b_4	b_5
Value	-88.77	3.53	-0.42	0.043	107.5	-2.74

Fig. 5. A qualitative example to illustrate how UAV and UGV rendezvous with each other under the policy π that is obtained by solving the CMDP. The risk tolerance is set to be $\delta = 0.1$ in this case study. (a) The input of the risk-aware rendezvous problem. (b) One sample route of UAV when it executes the policy π. (c) The route of UGV corresponds to the route of UAV. Lower case letters with or without prime denote the same road node that is visited at different times

The UAV task and UGV tasks are from our ongoing project on intelligence, surveillance, and reconnaissance (ISR) as shown in Fig. 5a. In this project, we are interested in the case where $\delta = 0.1$. UAV has about 240 KJ energy and its best range speed and best endurance speed are 14 m/s and 9.8 m/s respectively. UGV moves at 4.5 m/s. The rendezvous process will take 300 seconds.

We consider two sources of stochasticity in the energy consumption model of UAVs: weight and wind velocity contribution to longitudinal steady airspeed. The deterministic energy consumption model of the UAV is a polynomial fit constructed from analytical aircraft modeling data, given as

$$P(\boldsymbol{v}_\infty) = b_0 + b_1 \boldsymbol{v}_\infty + b_2 \boldsymbol{v}_\infty^2 + b_3 \boldsymbol{v}_\infty^3 + b_4 \boldsymbol{w} + b_5 \boldsymbol{v}_\infty \boldsymbol{w}, \quad (17)$$

where b_0 to b_5 are coefficients, and their experimental values are listed in Table I. Figure 6 shows the agreement between the polynomial regression fit model and the analytical data that it was derived from.

Weight is randomly selected following a normal distribution with a mean of 2.3 kg and a standard deviation of 0.05 kg, $\boldsymbol{w} \sim \mathcal{N}(\mu_{\boldsymbol{w}}, \sigma_{\boldsymbol{w}}^2)$. Vehicle airspeed, v_∞, is the sum of the vehicle ground speed, $\bar{v}_g$, and the component of the wind velocity that is parallel to the vehicle ground speed, ignoring sideslip angle and lateral wind components.

$$v_\infty = |\bar{v}_g + \cos(-\psi)\boldsymbol{\xi}_{a,b}| \quad (18)$$

The longitudinal wind speed contribution is derived from two random parameters; wind speed, and wind direction. Wind speed is modeled using the Weibull probability distribution model of wind speed distribution, $\boldsymbol{\xi}_{a,b}$, with a characteristic velocity $a = 1.5$ m/s and a shape parameter $b = 3$. This is representative of a fairly mild steady wind near ground level. Wind direction ψ is the heading direction of the wind, and is uniformly randomly selected on a range of $[0, 360)$ degrees.

C. Simulation Results

An illustrative example of the input and the output of the problem considered is shown in Fig. 5. The input of
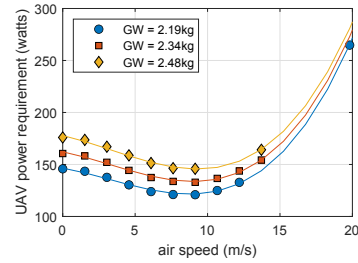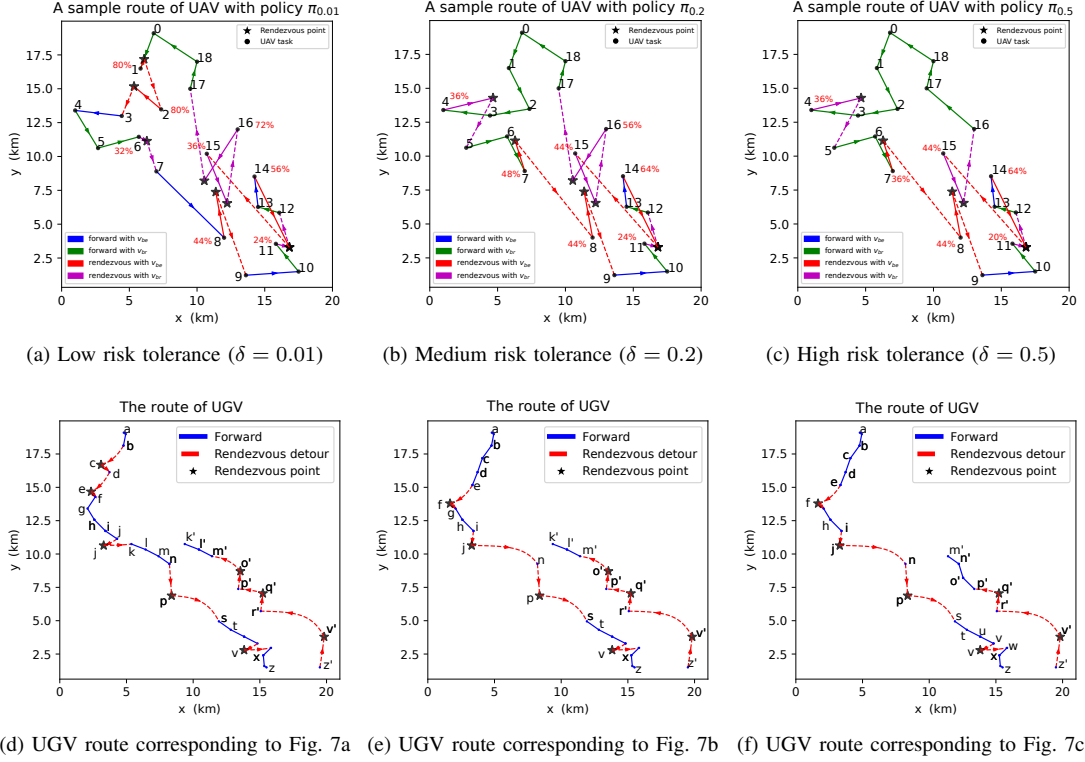

Fig. 6. Comparison of analytical data used to derive the polynomial regression fit model of UAV power requirement at three weights and across 11 airspeeds.

the problem is shown in Fig. 5a, which consists of UAV task nodes (red dots), UGV task nodes (blue square), and road network (black nodes). Fig. 5b and 5c shows one sample route of UAV and UGV respectively when the system executes the policy computed by LP. UAV's route starts from node 0. When the UAV reaches node 4, it will choose to rendezvous with UGV using the best range speed in a rendezvous point, which is denoted as a star, and then go to its next task node 5. Similarly, the UAV will rendezvous with the UGV when it reaches nodes 7, 8, 11, 14, and 15. The corresponding route of UGV is presented in Fig. 5c.

Next, we show how different risk tolerances influence rendezvous behaviors under our CMDP formulation. In these experiments, we set the risk tolerance δ to be 0.01, 0.2, and 0.5. Results shown in Fig. 7 includes sample routes for the UAV and the UGV and statistical data of the policies. Fig. 7a, 7b, and 7c are sample routes for the UAV when it executes the policy. Fig. 7d, 7e, and 7f are corresponding routes of the UGV. The rendezvous point is denoted as a star. The SOC is annotated in red text close to the task node at which the UAV decides to rendezvous. Some statistical data are summarized in table 7g. In general, we observe that when the risk tolerance is set to be small, the UAV tends to rendezvous more often, and the average route travel time is higher. Here the average route travel time is computed by considering only

(a) Low risk tolerance ($\delta = 0.01$) (b) Medium risk tolerance ($\delta = 0.2$) (c) High risk tolerance ($\delta = 0.5$)

(d) UGV route corresponding to Fig. 7a (e) UGV route corresponding to Fig. 7b (f) UGV route corresponding to Fig. 7c

UAV data	$\delta = 0.01$	$\delta = 0.2$	$\delta = 0.5$
Empirical failure rate	0.00975	0.201	0.497
route travel time overhead	160.2%	139.2 %	138.9%

(g) Quantitative results.

Fig. 7. How different risk thresholds influence the rendezvous behaviors. UAV route time with the best range speed is 4354 s and the route distance is 61.0 km. (a) UAV is very risk-averse to failures with a risk threshold equal to 0.01. (b) UAV is less risk-averse to failures with a risk threshold equal to 0.2. (c) UAV is neutral to the failures with a risk threshold equal to 0.5.

TABLE II
EMPIRICAL EVALUATION OF FAILURE PROBABILITY $\delta = 0.1$.

N_{MC} (# of MC trials)	500	1000	3000	5000
Failure rate	0.108	0.105	0.099	0.097

Fig. 8. Results comparisons for the CMDP policy and the greedy policies with $\delta = 0.1$. The dashed black line represents the task duration if the UAV moves with the best range speed without considering battery limitation.

trials in which the UAV finishes its task route. By contrast, as the risk tolerance is relaxed to a larger value, the average route travel time will decrease, which comes at the cost of a high failure probability.

We also conducted several quantitative experiments to validate our formulation. The first experiment is to use MC simulation to check whether the failure probability is upper bounded by the set risk tolerance of 0.1. We use FR to denote the empirical failure rate. As can be seen in Table II, as MC increases, the empirical failure rate is close to and below the theoretical PF 0.1. In the following experiments, we will use $N_{MC} = 2000$ for simulation.

To validate that the policy constructed from LP can minimize the expected travel distance. We compare our
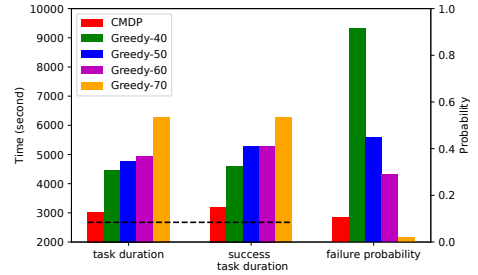
policy with a greedy baseline. The greedy policy is set as: always flies with the best range speed and chooses to rendezvous when state-of-charge drops below a set value. What we observe in experiments is that when the route of a UAV is long, for example, there are more than 15 nodes, the probability of finishing the route is close to zero for the greedy baseline no matter what threshold we set. To
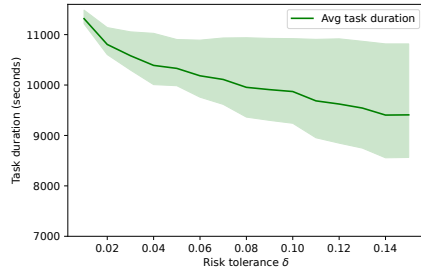
Fig. 9. Risk tolerance vs task duration Pareto Curve.

have a more informative comparison, we use only nodes 0 to 11 for the task route in the following experiments. We consider four-set values 40%, 50%, 60%, and 70%, and the corresponding policies are denoted as *Greedy-40*, *Greedy-50*, *Greedy-60*, and *Greedy-70*. As shown in Fig.8, our policy can guarantee success probability above the set value of 0.9 and the expected travel time of UAV is shorter compared to the baseline. Though the baseline can achieve a higher success probability in some cases as shown in Fig.8, its expected task duration is still longer than our policy.

The empirical Pareto curve for risk tolerance and the task duration is shown in Fig. 9. The green curve is the mean value and the shaded area is formed using one standard deviation from the mean. When the risk level is set to be a higher value, the UAV will tend to make more risky decisions, leading to a lower travel time at the cost of a higher failure probability.

The running time for the proposed routing problem consists of three parts. The first part is devoted to extracting transition information for LP. The second part is about constructing an LP model with Gurobi and the last part is for solving the LP. In our case study, there are about 54000 states and it takes about 6 min to extract transition information, 9 minutes to create an LP model, and about 1 second to solve the LP.

VI. Conclusion

In this paper, we study a variant of the cooperative aerial-ground routing problem with energy chance constraint on the UAV. We formulate the problem as a CMDP and use LP to find the optimal policy for the CMDP. We validate our formulation and the solution in one ISR application. In future work, one direction we will explore is to extend one UAV and one UGV routing problems to multiple UAVs and UGVs and consider the distributed solution for such types of risk-aware routing problems.

References

[1] B. Du, D. Sun, S. G. Manyam, and D. W. Casbeer, "Cooperative air-ground vehicle routing using chance-constrained optimization," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 392–397.

[2] K. Yu, A. K. Budhiraja, and P. Tokekar, "Algorithms for routing of unmanned aerial vehicles with mobile recharging stations," in *2018 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2018, pp. 5720–5725.

[3] A. Otto, N. Agatz, J. Campbell, B. Golden, and E. Pesch, "Optimization approaches for civil applications of unmanned aerial vehicles (uavs) or aerial drones: A survey," *Networks*, vol. 72, no. 4, pp. 411–458, 2018.

[4] H. Li, J. Chen, F. Wang, and M. Bai, "Ground-vehicle and unmanned-aerial-vehicle routing problems from two-echelon scheme perspective: A review," *European Journal of Operational Research*, vol. 294, no. 3, pp. 1078–1095, 2021.

[5] S. Choudhury, K. Solovey, M. J. Kochenderfer, and M. Pavone, "Efficient large-scale multi-drone delivery using transit networks," *Journal of Artificial Intelligence Research*, vol. 70, pp. 757–788, 2021.

[6] P. Maini, K. Yu, P. Sujit, and P. Tokekar, "Persistent monitoring with refueling on a terrain using a team of aerial and ground robots," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 8493–8498.

[7] K. Yu, J. M. O'Kane, and P. Tokekar, "Coverage of an environment using energy-constrained unmanned aerial vehicles," in *2019 international conference on robotics and automation (ICRA)*. IEEE, 2019, pp. 3259–3265.

[8] N. Mathew, S. L. Smith, and S. L. Waslander, "Planning paths for package delivery in heterogeneous multirobot teams," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 4, pp. 1298–1308, 2015.

[9] S. G. Manyam, K. Sundar, and D. W. Casbeer, "Cooperative routing for an air–ground vehicle team—exact algorithm, transformation method, and heuristics," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 1, pp. 537–547, 2019.

[10] P. Maini, K. Sundar, M. Singh, S. Rathinam, and P. Sujit, "Cooperative aerial–ground vehicle route planning with fuel constraints for coverage applications," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 6, pp. 3016–3028, 2019.

[11] T. C. Thayer and S. Carpin, "An adaptive method for the stochastic orienteering problem," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 4185–4192, 2021.

[12] C. C. Murray and A. G. Chu, "The flying sidekick traveling salesman problem: Optimization of drone-assisted parcel delivery," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 86–109, 2015.

[13] N. Agatz, P. Bouman, and M. Schmidt, "Optimization approaches for the traveling salesman problem with drone," *Transportation Science*, vol. 52, no. 4, pp. 965–981, 2018.

[14] Q. M. Ha, Y. Deville, Q. D. Pham, and M. H. Hà, "On the min-cost traveling salesman problem with drone," *Transportation Research Part C: Emerging Technologies*, vol. 86, pp. 597–621, 2018.

[15] Y. Liu, Z. Luo, Z. Liu, J. Shi, and G. Cheng, "Cooperative routing problem for ground vehicle and unmanned aerial vehicle: The application on intelligence, surveillance, and reconnaissance missions," *IEEE Access*, vol. 7, pp. 63 504–63 518, 2019.

[16] U. Ermağan, B. Yıldız, and F. S. Salman, "A learning based algorithm for drone routing," *Computers & Operations Research*, vol. 137, p. 105524, 2022.

[17] G. Wu, M. Fan, J. Shi, and Y. Feng, "Reinforcement learning based truck-and-drone coordinated delivery," *IEEE Transactions on Artificial Intelligence*, 2021.

[18] P. Tokekar, J. Vander Hook, D. Mulla, and V. Isler, "Sensor planning for a symbiotic UAV and UGV system for precision agriculture," *IEEE Transactions on Robotics*, 2016.

[19] D. Bertsekas, *Dynamic programming and optimal control: Volume I*. Athena scientific, 2012, vol. 1.

[20] P. Geibel and F. Wysotzki, "Risk-sensitive reinforcement learning applied to control under constraints," *Journal of Artificial Intelligence Research*, vol. 24, pp. 81–108, 2005.

[21] E. Altman, *Constrained Markov decision processes: stochastic modeling*. Routledge, 1999.

[22] S. Thiébaux, B. Williams *et al.*, "Rao*: An algorithm for chance-constrained pomdp's," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30, no. 1, 2016.