

Multi-vehicle Dynamic Water Surface Monitoring

František Nekovář^{id}, Jan Faigl^{id}, Martin Saska^{id}

Abstract—Repeated exploration of a water surface to detect objects of interest and their subsequent monitoring is important in search-and-rescue or ocean clean-up operations. Since the location of any detected object is dynamic, we propose to address the combined surface exploration and monitoring of the detected objects by modeling spatio-temporal reward states and coordinating a team of vehicles to collect the rewards. The model characterizes the dynamics of the water surface and enables the planner to predict future system states. The state reward value relevant to the particular water surface cell increases over time and is nullified by being in a sensor range of a vehicle. Thus, the proposed multi-vehicle planning approach is to minimize the collective value of the dynamic model reward states. The purpose is to address vehicles' motion constraints by using model predictive control on receding horizon and fully exploiting the utilized vehicles' motion capabilities. Based on the evaluation results, the approach indicates improvement in a solution to the kinematic orienteering problem and the team orienteering problem in the monitoring task compared to the existing solutions. The proposed approach has been experimentally verified, supporting its feasibility in real-world monitoring tasks.

I. INTRODUCTION

In this letter, we present a novel formulation of the model-based multi-vehicle planning, denoted *Incremental Motion Planning with Dynamic Reward* (IMP-DR), to address continual exploration and monitoring of water surface with objects of interest using dynamic spatio-temporal reward model. The studied problem belongs to a class of robotic scenarios in which repeated monitoring for state (or location) changes is needed upon encountering an object of interest during the initial exploration. Besides, repeated monitoring can identify new objects that become detectable in time.

The motivational scenario is to employ a fleet of *Unmanned Aerial Vehicles* (UAVs) in top-down visual monitoring of large water surfaces surrounding a central ship, where the water surfaces that are not static due to tidal and weather conditions. The problem is most closely related to the informative motion planning to maximize information gathering along the planned trajectory [1], the kinematic *Orienteering Problem* (OP) [2] to find vehicle's constrained reward-collecting trajectory and persistent monitoring with limited sensing range [3]. Hence, we

Manuscript received: February 23, 2023; Revised June 17, 2023; Accepted July 21, 2023.

This paper was recommended for publication by Editor Pauline Pounds upon evaluation of the Associate Editor and Reviewers' comments. This work was supported by the Czech Science Foundation (GAČR) under research projects No. 22-05762S and No. 22-24425S, the CTU grant No. SGS23/177/OHK3/3T/13, and the Technology Innovation Institute - Sole Proprietorship LLC, UAE, under the Research Project Contract No. TII/ARRC/2055/2021.

Authors are with the Czech Technical University, Faculty of Electrical Engineering, Technická 2, 166 27, Prague, Czech Republic, email: {nekovfra|faigl|saskam1}@fel.cvut.cz.



Figure 1. Experimental UAV performing a water surface monitoring mission.

model the information gained from exploration and monitoring as dynamic reward states similar to the discretized field in [3], where the states' position might change.

The underlying spatio-temporal reward model dynamics influence the expected information gain, and we thus formulate the combined information-collecting task as maximizing the collected reward on a receding horizon with minimizing the weighted sum of the reward states. While the value of any reward state steadily increases in time, it is nullified by its presence in the vehicle's sensor range. Besides, the information gain is constrained by limited sensor range, and movement planning needs to consider vehicle motion constraints.

We propose to address the studied problem using the *Model Predictive Control* (MPC) approach that allows us to consider dynamical models of the vehicles, reward states, and positions of objects being monitored. The dynamical models enable predicting future environment and vehicle states, with computational complexity exponentially increasing with prediction horizon length. Thus, information-gathering trajectories are computed on a limited control horizon, and the iterative closed-loop operation scheme of the MPC planner with the system model is used for the continual exploration and monitoring using multiple vehicles. The approach has been experimentally verified using a real vehicle shown in Fig. 1. A top-down snapshot from the deployment is depicted in Fig. 2 with overlaid reward states' values spaced on a symmetric grid. The main contributions are considered as follows.

- Novel IMP-DR problem formulation to coordinate a team of vehicles to search a priori unknown (water) surface and monitor detected dynamic objects continually.
- IMP-DR solution using MPC-based planning on receding horizon with a dynamic environment and vehicle models,

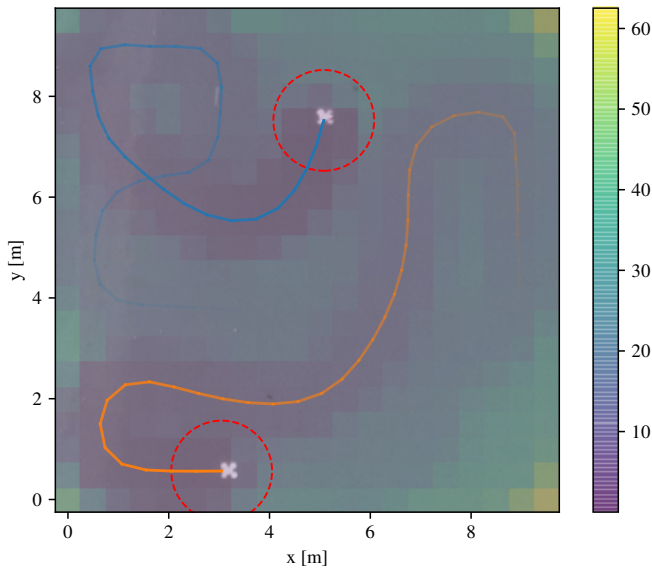


Figure 2. Reward value overlay during monitoring experiment using two UAVs. The reward collection range is shown as red circles around the vehicles.

predicting future states of the dynamic environment.

- Evaluation of the proposed solution and comparison with the most related (to the best of the authors' knowledge) approach to the *Team Orienteering Problem* (TOP) [4] adapted for the receding horizon planning with dynamic model and solved by the state-of-the-art meta-heuristic [5]. The proposed approach is also compared to the state-of-the-art kinematic OP methods [2] for a single vehicle showing improved results on evaluation scenarios.
- Experimental deployment of the proposed approach in real-world flight in a water surface monitoring scenario.

The remainder of the letter is organized as follows. Related work is overviewed in the following section. The addressed IMP-DR is formally introduced in Section III. The sensor and reward models are described in Section IV together with the vehicle model and its motion constraints. The proposed MPC-based solution is presented in Section V. Evaluation results and experimental deployment are reported in Section VI. Concluding remarks are summarized in Section VII.

II. RELATED WORK

Environment monitoring can be formulated as a problem of creating a quantifiable phenomena model using sampled data from the environment [6] with the regression about the phenomena state using regression models such as *Gaussian Processes* (GPs) [7] and sensor model, such as in [8]. Nevertheless, the studied continual monitoring task has to address vehicle motion constraints and information dynamics. Therefore, path planning is a part of the navigation toward the locations where the studied phenomena can be measured, such as the signal strength [9]. Thus, informative path planning [10] is to determine the most informative path improving the phenomena model with respect to motion constraints of information collecting vehicles [1], [11]. However, these approaches are suitable for modeling spatial phenomena in environment monitoring tasks, but they do not address the underlying

temporal dynamics that we call information dynamics and its interaction with the vehicles.

Persistent multi-vehicle monitoring of changing environments with limited vehicle sensing range and pre-determined closed vehicles' paths is presented in [3]. Continuous dynamic scalar field monitoring with value estimation based on individual filtered measurements is proposed in [12], where the authors employ Pontryagin's minimum principle for planning locally optimal, single-vehicle sensing trajectories minimizing the estimation uncertainty. However, single-order integrator vehicle dynamics is considered with unconstrained input.

In our addressed task, we explicitly model the environment information dynamics similarly to the accumulation function [3] in combination with sensor and vehicle dynamics as a set of *Discrete Algebraic Equations* (DAEs) used for moving-horizon prediction and control. The limited prediction horizon allows planning to visit only some locations to collect measurements. Hence, we need to select the most informative locations reachable, and the task can be formulated as a generalization of the *Orienteering Problem* (OP) [13].

Although not directly addressing the information dynamics, existing OP formulations can be utilized in receding horizon reward collection planning schemes. Multi-vehicle planning for search-and-rescue operations with grid-sampled rewards determined by satellite imaging of the area of interest is formulated in [5]. However, exact visits of the sampled grid positions are required, constraining the vehicle's movement and neglecting its dynamics. Non-zero sensor range and Dubins vehicle dynamics are addressed in [14]. The authors of [2] address the OP with time-optimal multi-rotor vehicle kinematic trajectory generation. The OP with multiple Dubins vehicles is studied in [15], providing background work on existing solutions to the planning problem with a limited travel budget.

Furthermore, we consider discretized environment representations inspiring, such as in exploring flooded areas [16], where the collection of the information reward is addressed by the uniform sampling of grid cells with travel budget-constrained vehicle-driving policy. The Correlated OP [17] formulates cyclic informative tours with the information gain correlated between neighboring nodes of the graph-based environment representation to estimate a scalar field. However, the sensor range is limited to the individual nodes, and the approaches do not account for the vehicle dynamics.

The dynamic position of the target objects is addressed in [18] by receding horizon planning and dynamic clustering to maximize the reward collected from moving targets with uncertain dynamics. In [19], the approach is generalized to address obstacles in cooperative planning that can also be based on the Fisher information matrix model, for example, used in multi-UAV target tracking [20].

In [21], the dynamics of water surface debris show additional effects to the local oscillations to be taken into account, such as water currents, sail effect, and locally most dominant Stokes drift [22]. Therefore, we model the target movement as a combination of local oscillations and time-dependent drift.

The existing receding horizon formulations address targets' and vehicles' position dynamics but not the reward values'

dynamics for continual monitoring. Therefore, we generalize the existing approaches to account for all the dynamics in the studied problem of combined water surface monitoring and tracking of the detected object of interest.

III. PROBLEM STATEMENT

The *Incremental Motion Planning with Dynamic Reward* (IMP-DR) problem is formulated as a multi-vehicle scenario in a two-dimensional environment represented by n_p sampled locations or predicted positions of objects of interest, further referred to as targets, each encoded by a position vector \mathbf{p}_i . Each \mathbf{p}_i represents coordinates $\mathbf{p}_i = [p_{i,x}, p_{i,y}]$, and targets form a discrete finite set $\mathcal{P} = \{\mathbf{p}_i \in \mathbb{R}^2 : 1 \leq i \leq n_p\}$. The number of targets n_p can evolve during the mission as new targets might be added by the object detection. Targets' coordinates can change with predictable dynamics, a model that is assumed to be known. However, the model parameterization can change during the monitoring based on the visual observations. Thus, the target set is a function of time $\mathcal{P}(t)$.

The presence of m vehicles is assumed at $\mathbf{q}_i = [q_{i,x}, q_{i,y}]$ forming the set $\mathcal{Q} = \{\mathbf{q}_i \in \mathbb{R}^2 : 1 \leq i \leq m\}$. Since the vehicles move in time, we denote $\mathcal{Q}(t)$. Besides, the vehicles are subject to dynamical constraints given by the maximum velocity magnitude v_{\max} and acceleration in the axes a_{\max} . Mutual vehicle collision avoidance and safety are addressed by the constraint on the minimum allowed distance between the vehicles d_{\min} at any time event t , such that $\|\mathbf{q}_i(t) - \mathbf{q}_j(t)\| \geq d_{\min}, \forall \mathbf{q}_i(t), \mathbf{q}_j(t) \in \mathcal{Q}(t), i \neq j$, and $i, j \in \{1, \dots, m\}$.

We seek to minimize the time elapsed between visits to the targets in \mathcal{P}_t on a planning horizon T_h while penalizing maximal value. It is addressed by defining the dynamical reward state, analogous to the information in the informative planning approaches. The reward states $r_i(t) \in \mathbb{R}_0^+ : 0 \leq i \leq n_p$ are modeled for each $\mathbf{p}_i \in \mathcal{P}(t)$. In the presence of any of the vehicles near the reward state position \mathbf{p}_i , its reward value r_i is nullified. Squared reward state value is used in optimization problems to penalize high values and improve solver performance. The function $f_r(\dot{r}, r, \mathcal{P}, \mathcal{Q})$ describes implicit reward dynamics with relation to the targets and positions of the vehicles. The variables of f_r are functions of time; the notation is omitted for clarity. We formulate the IMP-DR planning problem step as follows.

$$\begin{aligned} \min_{\mathcal{Q}} &= \sum_{i=1}^{n_p} \int_{t_0}^{t_0+T_h} r_i^2 dt \\ \text{s.t.} & \\ & 0 = f_r(\dot{r}, r, \mathcal{P}, \mathcal{Q}). \end{aligned} \quad (1)$$

We use the IMP-DR to perform monitoring tasks by iterative solution of (1), incrementing the initial time t_0 with the sampling period $T_s \leq T_h$ on each iteration and utilizing the current reward state feedback.

IV. DYNAMIC MODELS

The sensor model, reward model, and second-order multi-rotor UAV model used in the proposed monitoring solution are presented in this section.

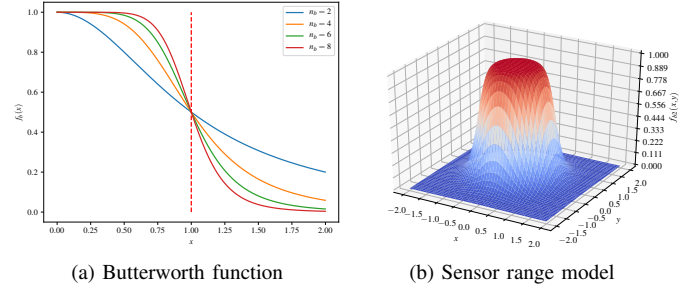


Figure 3. (a) The Butterworth function (2) plotted for different values of n_b . The red vertical dashed line denotes the cut-off range parameter $c_b = 1$. The function satisfies $f_b(c_b) = 0.5$; therefore, the value of c_b indicates an approximate sensor range. (b) The sensor range model is a function of two variables (3), plotted for $n_b = 8$ and $c_b = 1$.

A. Sensor Model

The UAV sensor range, modeled as a radius around the vehicle where the reward is collected, is depicted in Fig. 3. The proposed sensor range model is based on the continuous differentiable Butterworth function (2) employed to serve as a reward collection indicator function in the reward model.

The degree n_b of the Butterworth function influences the convergence rate in limits and function shape. The value of c_b specifies the effective sensor range as f_b approximates square shape with increasing n_b . It enables us to model the sensor coverage range as a function of the UAV's position (3).

$$f_b(x) = \frac{1}{1 + (\frac{x}{c_b})^{n_b}} \quad (2)$$

$$f_{b2}(x, y) = \frac{1}{1 + (\frac{\sqrt{x^2+y^2}}{c_b})^{n_b}} \quad (3)$$

Due to the asymptotic convergence in $\lim_{x \rightarrow \pm\infty} f_b(x) = 0$, the modeled reward collection is exhibited even for large distances $\gg c_b$; the model reward values are approximate. It positively influences solution convergence of interior-point methods, as the value of f_b , which directly affects the reward gain, is always non-zero for all reward pairs of state-vehicle (up to the limits of the numerical precision). On a UAV visit to the neighborhood of a rewarding target, the related reward state is approximately nullified due to $f_b(0) = 1$; the overall reward state error is assumed to converge to some finite value during the continual monitoring.

B. Reward Model

The dynamics of discrete reward state r_i , where $1 \leq i \leq n_p$, are modeled as a DAE (4), which is similar to the information accumulation function in [3].

$$\begin{aligned} r_{i,k+1} &= (r_i + T_s k_{\text{gain}}) \\ & (1 - \max(1, \sum_{j=1}^m f_{b2}(p_{i,x,k} - q_{j,x,k}, p_{i,y,k} - q_{j,y,k}))) \end{aligned} \quad (4)$$

The reward variables' dynamics is autonomous if no vehicle is present proportional to k_{gain} . On the presence of a vehicle in the reward state neighborhood indicated by (3), the reward state is nullified. We consider the sampling rate of the discrete

dynamical system T_s . Limiting the maximal influence of the multiple vehicles on the reward states is necessary for the model, as the reward state values are assumed to be non-negative.

Due to the approximate sensor range radius c_b , the results of the proposed monitoring approach are processed by the evaluation model (5). Discrete vehicle positions $q_{j,x,k}, p_{j,y,k}$ at sampled time-steps k are used as the inputs of an indicator function to determine if a vehicle passed through the reward-collecting range during continual monitoring

$$r_{i,k+1}^{\text{eval}} = \begin{cases} 0, & \text{if } d_{i,j} \leq n_e, \forall 1 \leq j \leq m, \\ r_{i,k}^{\text{eval}} + T_s k_{\text{gain}} & \text{otherwise,} \end{cases} \quad (5)$$

$$\text{where } d_{i,j} = \sqrt{(p_{i,x} - q_{j,x,k})^2 + (p_{i,y} - q_{j,y,k})^2}.$$

C. Vehicle Model

The vehicles are modeled as the second-order DAE systems of m multi-rotor UAV agents in (6a) to (6d) with the accelerations $a_{i,x}, a_{i,y}$ as the inputs. The model is subject to motion constraints expressed in (7a) to (7c).

$$q_{i,x,k+1} = v_{i,x,k} T_s + a_{i,x,k} \frac{T_s^2}{2} \quad \sqrt{v_{i,x}^2 + v_{i,y}^2} \leq v_{\max} \quad (7a)$$

$$(6a) \quad -a_{\max} \leq a_{i,x} \leq a_{\max} \quad (7b)$$

$$q_{i,y,k+1} = v_{i,y,k} T_s + a_{i,y,k} \frac{T_s^2}{2} \quad -a_{\max} \leq a_{i,y} \leq a_{\max} \quad (7c)$$

$$(6b) \quad \text{for } 1 \leq i \leq m.$$

$$v_{i,x,k+1} = a_{i,x,k} T_s \quad (6c)$$

$$v_{i,y,k+1} = a_{i,y,k} T_s \quad (6d)$$

for $1 \leq i \leq m$.

Collision avoidance puts minimal mutual distance constraints (8a) and (8b) on the agents as:

$$d_{i,j} \geq d_{\min}, \quad (8a)$$

$$d_{i,j} = \sqrt{(q_{i,x} - q_{j,x})^2 + (q_{i,y} - q_{j,y})^2}, \quad (8b)$$

for $1 \leq i, j \leq m, i \neq j$.

D. Water Surface Model

The water surface area is modeled as a finite grid of target points. The movement of dynamic debris on the surface is modeled as local periodic oscillations combined with a time-dependent drift, a simplified combination of the water currents' influence, sail effect, and Stokes drift. Precise modeling of underlying causes might improve the movement prediction, but it is considered out of this letter's scope. The model is expressed for each axis as simplified time-dependent function (9) parameterized by the amplitude A_p , angular velocity ω_s , and the drift velocity v_d .

$$x_d(t) = tv_d + A_p \sin(t\omega_p) \quad (9)$$

The function parameters are assumed to be known but might vary during the monitoring for each target. Then, parameter changes can be determined from the visual measurements during the adaptive monitoring.

V. PROPOSED MPC-BASED SOLUTION FOR IMP-DR

We propose to address the motivational problem of multi-vehicle dynamic water surface monitoring using the proposed IMP-DR formulation and planning solution method based on the MPC (10). In the combined adaptive exploration and monitoring problem, we seek to maximize the information gained from the specified targets. The information is quantified as a reward value associated with the targets. As the reward values continually increase, the planning goal is to minimize the values present in the reward states of the modeled environment.

We propose employing an MPC-based control technique to solve the multi-vehicle planning problem, as it is suitable for addressing vehicle constraints and time-evolving reward dynamics. The utilized Optimal Control Problem of the MPC on N_s steps, where $N_s = \frac{T_h}{T_s}$, is stated as

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{u}} &= \sum_{k=0}^{N_s-1} (f_l(\mathbf{x}_k) + \Delta \mathbf{u}_k^T k_R \Delta \mathbf{u}_k) + f_m(\mathbf{x}_{N_s}), \quad (10) \\ \text{s.t.} & \quad \mathbf{x}_{lb} \leq \mathbf{x}_k \leq \mathbf{x}_{ub}, \\ & \quad \mathbf{u}_{lb} \leq \mathbf{u}_k \leq \mathbf{u}_{ub}, \\ & \quad \mathbf{z}_{lb} \leq \mathbf{z}_k \leq \mathbf{z}_{ub}. \end{aligned}$$

The state vector \mathbf{x} (11) and input vector \mathbf{u} (12) are:

$$\mathbf{x} = [q_{1,x}, q_{1,y}, q_{2,x}, q_{2,y}, \dots, q_{m,x}, q_{m,y}, v_{1,x}, v_{1,y}, v_{2,x}, v_{2,y}, \dots, v_{m,x}, v_{m,y}, r_1, r_2, \dots, r_{n_p}]^T, \quad (11)$$

$$\mathbf{u} = [a_{1,x}, a_{1,y}, a_{2,x}, a_{2,y}, \dots, a_{m,x}, a_{m,y}]^T. \quad (12)$$

The functions $f_l(\mathbf{x}) = f_m(\mathbf{x}) = \sum_{k=1}^{n_p} r_k^2$ are the Lagrange and Meyer terms. Input penalty $k_R = 10^{-3}$ is used to penalize the system input changes $\Delta \mathbf{u}_k = \mathbf{u}_{k+1} - \mathbf{u}_k$. Including the penalty term prevents oscillatory behavior, and it was empirically observed to decrease solution times. The constant vectors \mathbf{x}_{ub} , \mathbf{x}_{lb} , \mathbf{u}_{ub} , and \mathbf{u}_{lb} enforce state and input constraints. Algebraic vector $\mathbf{z}(\mathbf{x}, \mathbf{u})$ and its bounding vectors \mathbf{z}_{ub} and \mathbf{z}_{lb} enforce minimum vehicle distance and velocity magnitude constraints.

VI. RESULTS

The proposed solution has been empirically examined in several scenarios to evaluate its properties. In Section VI-A, we present results on computational performance according to the problem size, number of vehicles, and planning horizon length. Furthermore, the introduced IMP-DR generalizes both the Kinematic and Team OP. Therefore, the proposed approach is utilized to address these problems as well. In particular, the proposed approach's performance is compared to the state-of-the-art Kinematic OP planner [2] and Team OP meta-heuristic planner [5] with the results reported in Section VI-B and Section VI-C, respectively. Finally, simulation results and reports on experimental field deployment on continual water surface exploration and monitoring tasks are presented in Section VI-D and Section VI-E, respectively.

A. Computational Evaluation

For the computational evaluation, the static targets are sampled on a symmetric two-dimensional grid of the width w_{grid} . The spacing of targets is 1 m on both axes if not otherwise specified. The reward increase rate is set to $k_{\text{gain}} = 1 \text{ s}^{-1}$. The initial reward state value is set to $r_0 = 10$, as the system input change is penalized in the solver cost function. The vehicle dynamics is limited by the maximum velocity vector magnitude $v_{\text{max}} = 1 \text{ m s}^{-1}$ and acceleration $a_{\text{max}} = 2 \text{ m s}^{-2}$ in each axis. The sensor model is set to $c_b = 0.25$ and $n_b = 8$. The solutions were determined at each planning step with the period of $T_s = 0.25 \text{ s}$. The number of sampling steps in the planner is set to $N_s = 20$ with the resulting planning horizon of 5 s. Monitoring was performed for 5 min, translating to $S = 1200$ planning steps.

The proposed approach is implemented using the `do-mpc` toolbox [23] with the `Ipopt` [24] non-linear optimization framework. The MA97 sparse linear system solver, part of the HSL collection [25], was utilized. The evaluation was performed using *Robot Operating System* (ROS) control pipeline of the MRS UAV System [26] and the AMD Ryzen 4750U processor running at the base clock of 1.7 GHz, accompanied with 32 GB RAM. The solution convergence tolerance of 10^{-8} was used as the stopping criterion during evaluation.

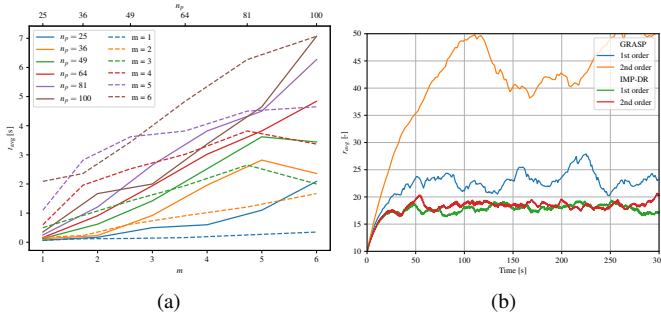


Figure 4. (a) Average computational time t_{avg} per planning step for the number of targets n_p and m vehicles. The time t_{avg} does not necessarily increase with n_p due to the utilization of multi-threading for larger problems by the MA97 solver; however, the overall tendency is linear with increasing problem size, as shown for $m = 6$. (b) Average reward state values r_{avg} in time of the GRASP-based TOP planner [5] and the proposed IMP-DR-based approach for the 1st order (velocity only) and the full 2nd order vehicle dynamics models. The monitoring problem is to minimize the system reward states by repeatedly visiting targets, nullifying the dynamic value.

The influence of the number of vehicles m and targets n_p on the computational performance is depicted in Fig. 4a and listed in Table I. Since the computational time differs at each planning step, the average t_{wall} planning time is reported as t_{avg} and the maximal planning time as t_{max} . The average reward state value $r_{\text{avg}} = \frac{1}{n_p} \sum_{i=1}^{n_p} r_i$ is continually stable around the equilibrium value r_{eq} , the maximum system state reward is reported in the column r_{max} in Table I. From the observations of r_{avg} , its value exhibits Lyapunov stability as it does not diverge from a neighborhood of r_{eq} for $t \rightarrow \infty$ that has been pragmatically examined for 300 s as an average of r_{avg} . It demonstrates the numerical stability of the proposed continual monitoring approach in indefinite operation. The average step computational time t_{avg} increases linearly with the number of targets n_p and polynomially with the number of vehicles m .

Table I
COMPUTATIONAL RESULTS FOR n_p TARGETS AND m VEHICLES

n_p	m	t_{avg} [s]	t_{max} [s]	r_{eq} [-]	r_{max} [-]
25	1	0.05	1.5	32.7	109.8
25	2	0.12	2.0	27.6	139.2
25	3	0.55	5.8	16.7	97.7
25	4	0.77	8.9	9.9	56.2
25	5	1.06	8.8	13.7	94.6
25	6	1.22	12.1	16.4	86.2
49	1	0.18	5.8	31.3	125.0
49	2	0.49	4.3	32.3	143.5
49	3	1.42	19.9	18.3	80.6
49	4	1.63	21.1	25.5	98.5
49	5	3.30	18.3	14.4	94.6
49	6	3.53	33.0	13.3	70.0
100	1	0.43	20.9	55.2	279.9
100	2	1.50	8.9	33.4	149.0
100	3	1.85	26.9	28.1	164.1
100	4	2.67	28.5	28.2	118.9
100	5	4.16	31.8	24.5	136.5
100	6	7.07	77.3	22.7	118.4

Table II
COMPUTATIONAL RESULTS FOR m VEHICLES AND HORIZON N_s

m	N_s	t_{avg} [s]	t_{max} [s]	r_{eq} [-]	r_{max} [-]
1	10	0.05	0.4	78.4	309.7
1	15	0.16	11.9	64.9	268.2
1	20	0.34	2.8	61.0	237.5
1	25	0.61	6.0	59.3	299.2
1	30	1.12	9.9	54.7	255.0
3	10	0.23	2.3	71.0	294.9
3	15	1.32	19.3	28.6	191.8
3	20	1.63	10.4	28.3	155.5
3	25	3.17	17.6	26.6	138.7
3	30	5.39	33.2	26.2	135.0
5	10	0.5	11.1	52.5	279.2
5	15	1.61	9.4	23.4	153.4
5	20	4.29	26.5	20.4	133.3
5	25	11.02	84.5	17.1	100.9
5	30	21.68	224.7	21.2	137.7

The influence of the planning horizon N_s on the computational performance with $w_{\text{grid}} = 10$ is depicted in Table II. While increasing N_s yields an exponential increase in t_{avg} , it does not necessarily lead to improved average reward equilibrium r_{eq} and the maximal reward r_{max} . Hence, lower values of N_s can be preferred.

Here, it is worth noting that the used `Ipopt` solver relies on the third-party code to solve sparse symmetric indefinite linear systems repeatedly, and the choice of the third-party solver influences the required computational time and the quality of the solution. A solution is needed in less than the sampling rate T_s for the ideal operation of the proposed monitoring approach. Although planning at each sampling step provides the best performance, it can be performed at a reduced rate to allow for a longer computational time. Besides, the solver can be terminated prematurely before its convergence to obtain an intermediate solution suitable for field deployment, which is used for the results presented in Section VI-E.

B. Kinematic Orienteering Problem (Kinematic OP)

The IMP-DR problem can be formulated as the Kinematic OP if the reward dynamics is non-existent and end-point constraints are introduced. The proposed MPC-based approach can provide an approximate solution to the single-vehicle

Kinematic OP formulated in [2]. Therefore, the proposed IMP-DR is compared with the state-of-the-art solutions to the Kinematic OP denoted KOP-1 and KOP-6^{lns} [2].

For the comparison, the OP is addressed by modeling the neighborhood size negligible to the overall target distances in the problem instance. The sampling of $T_s = 0.1$ s was used with the input penalties of $k_Q = 10^{-3}$. The cost function terms were modified as: $f_l(\mathbf{x}) = \sum_{k=1}^{n_p} r_k$, $f_m(\mathbf{x}) = 10^3((q_x - x_{\text{final}})^2 + (q_y - y_{\text{final}})^2)$ to apply a quadratic soft-constraint on the final vehicle position $q_{N_s, x}, q_{N_s, y}$. The sensor parameters were set to $c_b = 0.05$ and $n_b = 2$, approximating a spiking function at each target position. Dynamical constraints were $v_{\text{max}} = 3 \text{ m s}^{-1}$ and $a_{\text{max}} = 1.5 \text{ m s}^{-2}$.

Table III
RESULTS FOR THE KINEMATIC ORIENTEERING PROBLEM

C_{max} [s]	KOP-1	KOP-6 ^{lns}	MPC _{best}	t_{avg} [s]
10	95	80	115	1.7
15	180	165	200	2.8
20	250	250	260	8.7
25	325	330	370	11.4
30	390	390	450	15.9
35	430	435	450	22.8
40	450	450	450	45.7

Furthermore, since the I_{popt} solver provides a locally optimal solution, a potentially better solution can be obtained by perturbing the initial reward values in the dynamic model. Thus, a noise with $\mathcal{N}(0, 0.1)$ distribution was added to the initial reward values on each solution iteration, and 10 iterations were performed for each benchmark instance. The best-found results are presented in Table III under the column MPC_{best} and shown in Fig. 5. The average computational time of the proposed MPC-based solver is denoted t_{avg} .

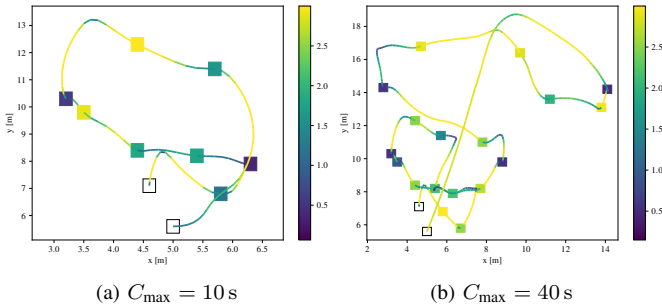


Figure 5. Velocity profiles of two solutions to the Kinematic OP on the Tsigilrides-2 problem instances with the travel budget C_{max} along the traveled path. The path color indicates the velocity for the scale depicted at the right of the plot in m s^{-1} . The solutions were obtained using the proposed MPC-based reward collection approach adjusted for solving the Kinematic OP.

The proposed solution method provided improved reward gain in all but one testing instance, where the travel budget C_{max} [s] is sufficiently large for the compared methods to catch up with the proposed approach performance. The trajectory deviation from the visited targets and final position does not exceed 0.02 m in all benchmark instances, which we consider negligible with regard to the problem scale. As the number of planning steps is fixed, the provided solutions are not time-optimal. They, however, satisfy the travel budget constraint.

In addition to better performance than the existing solvers in the collected reward, the proposed MPC-based approach can further address the reward dynamics, neighborhood, and multiple vehicles.

C. Team Orienteering Problem (Team OP)

A variant of the multi-vehicle monitoring scenario can be formulated as the Team OP, for which the proposed solver can also be utilized. Therefore, we compared its performance with the grid-based Team OP planner proposed in [5]. The *Greedy Random Adaptive Search Procedure* (GRASP) meta-heuristic planner [5] determines paths to maximize the collected reward gain from the sampled grid locations using multiple vehicles with the given travel budget. Hence, it was generalized for use in the receding horizon planning scenario by removing the end-point constraint and using individual vehicle starting locations. Besides, the GRASP cost function was modified to $f_l(\mathbf{x}) = \sum_{i=1}^{n_p} r_i^2$. It provides a trade-off between the computational requirements and solution quality compared to the exact *Mixed Integer Linear Programming* solution. The computational effectiveness is essential for the real-time operation of the proposed monitoring approach, where the Team OP is solved repeatedly on each sampling step. Therefore, we consider the heuristic approach in the comparison.

Grid width $w_{\text{grid}} = 10$ with $n_p = 100$ and $m = 3$ vehicles were used. The GRASP planner was limited to movement in the 4-neighborhood of a two-dimensional grid at 1 m s^{-1} . Neglecting the second-order dynamics and utilizing maximal vehicle velocity provides a *Lower Bound* (LB) on the GRASP-based monitoring performance. Assuming the vehicles have to accelerate and decelerate between the targets at 2 m s^{-2} provides an *Upper Bound* (UB) solution cost. As the second-order dynamics become negligible (due to large target distances), the UB converges to the LB.

Five GRASP planning iterations were performed on each planning step to obtain a quality solution. The planning was performed on the horizon of the 20 steps, i.e., 20 s for the LB and 40 s for the UB. The tuning of the proposed IMP-DR solver was the same as for the Kinematic OP reported in Section VI-B with the sampling rate $T_s = 0.1$ s, input penalties $k_Q = 10^{-3}$, neighborhood $c_b = 0.05$, $n_b = 2$, and $N_s = 20$ planning steps. Solutions were obtained addressing the first-order velocity dynamics and the second-order acceleration dynamics vehicle models. The benchmark was run for 300 s, and the average system state reward values r_{avg} provided by the evaluation model are plotted in Fig. 4b. On the target visit, vehicles did not exceed the neighborhood of 0.02 m as in the KOP benchmark.

The proposed IMP-DR planning approach demonstrates a lower average reward r_{avg} in continual monitoring over the GRASP-based TOP planner. The IMP-DR planning performance is similar for both the first- and acceleration-constrained second-order models. The results demonstrate the importance of the dynamic reward model in continual monitoring as it leads to lower r_{eq} . However, IMP-DR is impractical for planning multi-vehicle information-gathering tasks with large travel budgets, as the t_{avg} increases exponentially with planning horizon length.

D. Water Surface Monitoring and Flotsam Monitoring

The addressed motivational scenarios of repeated water surface exploration to detect objects of interest and dynamic object monitoring are presented. In the formulated exploration task, a symmetric target grid of $w_{\text{grid}} = 20$ with the target spacing of 0.5 m was used. The sensor function was set to $c_b = 0.5$ and $n_b = 4$, modeling the reward collection dependent on the distance. The MPC parameters were set to $T_s = 0.25$ s and $N_s = 20$. The evolution of the system state

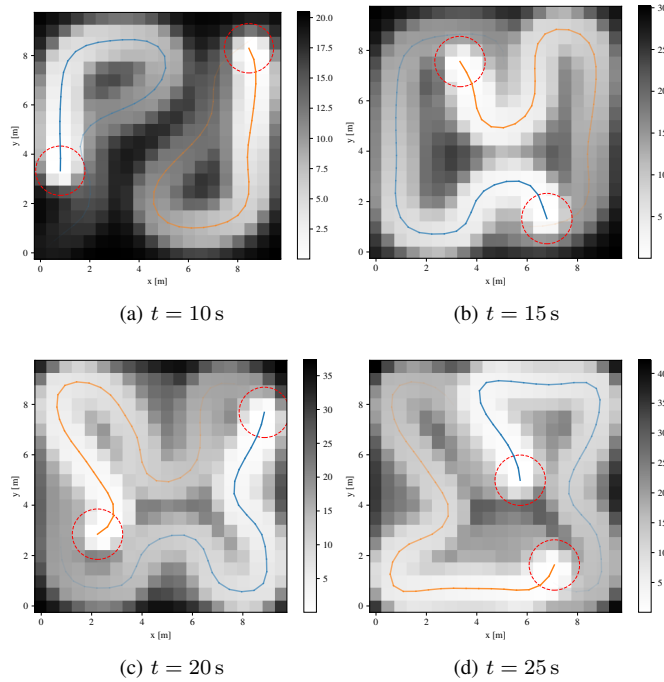


Figure 6. Water surface monitoring scenario on 20×20 large grid utilizing two vehicles. The color of the cells indicates the reward state value according to the scale depicted at the right part of the plot. A circle with the radius 1 m is shown in the red around each vehicle at the time instant t , approximating the maximal reward collection range. The paths traveled by vehicles are in blue and orange.

rewards at a particular time instant t is depicted in Fig. 6 together with the visualization of the vehicles' trajectories. The task was experimentally evaluated, and results are reported in Section VI-E.

Regarding the motivational scenario of monitoring debris or flotsam on a water surface with predictable movement dynamics, we consider the following problem setup. The targets are within a 3×4 large grid with the spacing of 2 m and the exact vehicle dynamic constraints as in the previous water surface monitoring scenario. The target dynamics in the x -axis are modeled as (9) with $A_p = 0.5$, $\omega_p = \frac{\pi}{10}$, and $v_d = 0$. Modeling target movement is possible by introducing time-varying parameters in the MPC planner. The parameters for the scenario instance are set to $T_s = 0.25$ s and $N_s = 40$, and the sensor function to $c_b = 0.5$ and $n_b = 8$.

The monitoring for the time horizon of 25 s is shown in Fig. 7. The vehicle positions are initiated from the opposite left-most targets and proceed to monitor moving targets continuously. Note that the locally optimal solutions provided by the utilized solver tend to degrade if vehicles are in close

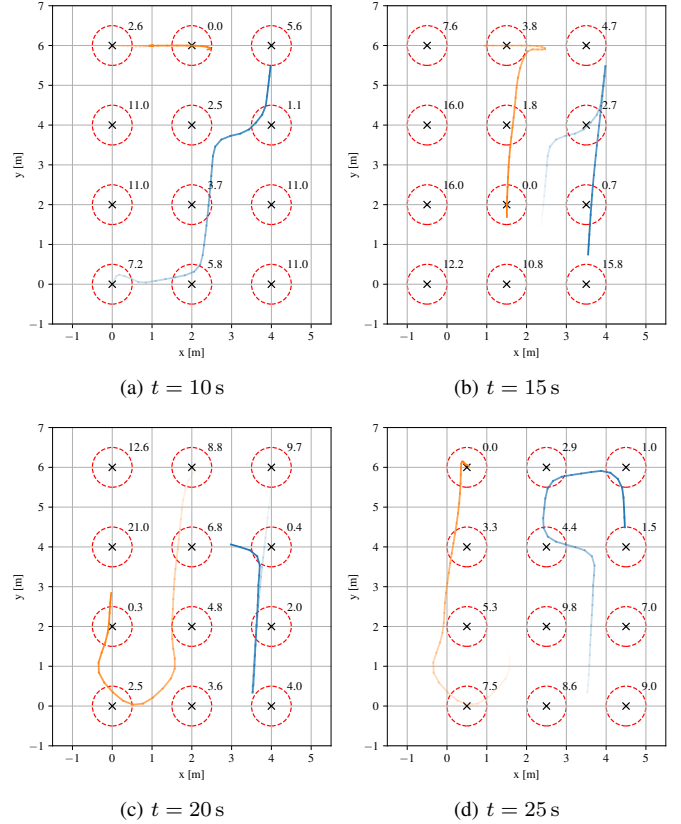


Figure 7. Flotsam monitoring scenario with dynamic target positions and two vehicles. The red (dashed) circles denote the reward collection radius of 0.5 m. Paths traveled by vehicles are shown in blue and orange.

proximity, as seen in Fig. 7a, where one of the vehicles is stuck in a feasible solution space.

E. Experimental Field Deployment

The feasibility of the proposed solution has been further examined in an experimental field deployment with limited computational resources. The proposed approach was experimentally tested with two UAV research platforms based on DJI F450, shown in Fig. 8a, performing a water surface exploration task as depicted in Fig. 2. The vehicles were equipped with GPS-based navigation and a top-down visual sensor. The connection between the vehicles and the planning computer was established over a Wi-Fi network using ROS. The used communication network introduced transport delays due to networking limits over long distances. Therefore, the parameters of the planning scheme from Section VI-A were modified to identical re-planning rate and horizon length of 5 s and the maximum solver processing time was limited to $t_{\text{wall}} \leq 3$ s.

After obtaining a solution, time-stamped command trajectories were passed to the UAV trajectory trackers using ROS messages for open-loop monitoring control. The onboard trackers adjusted to the trajectories on the ROS message arrival. The planner loop initiated from the UAV states 5 s in the future. The 2 s long window left for data transfer ensured the continuity of the flight trajectories. The vehicles operated over the surface of the Orlik dam, located on the Vltava river

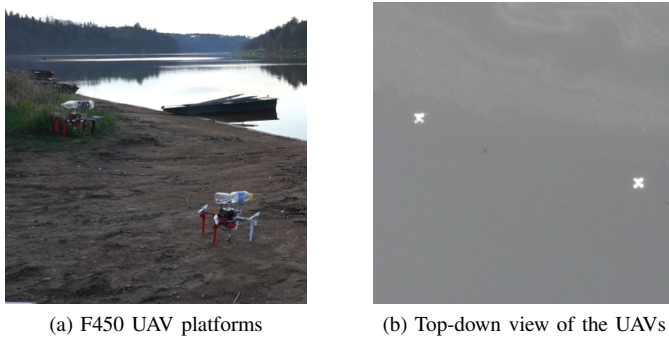


Figure 8. Snapshots from the real field deployment of the proposed approach. Supporting material and media from the experiment are available at [27].

in southern Bohemia. The vehicles are depicted in Fig. 8b. Compared to $r_{eq} = 18.9$ provided by the ideal planning rate identical to $T_s = 0.25$, the monitoring performance of adjusted real-world setup was $r_{eq} = 43.7$ during the 5 min flight. The performance degradation over an ideal configuration due to the limited computational time and resources can be mitigated using more powerful computational hardware. Nevertheless, the presented monitoring approach was shown to be feasible in a real-world deployment.

VII. CONCLUSION

We propose a novel formulation of the introduced IMP-DR model-based multi-vehicle monitoring approach. The proposed solution is based on the MPC on receding horizon evaluated on several problem instances. Based on the reported results, the performance of the proposed approach surpasses the state-of-the-art Kinematic OP solver in static reward collection and the grid-based Team OP solver in dynamic monitoring tasks. The proposed approach has been utilized in the combined dynamic water surface exploration and monitoring missions with multiple UAVs and limited sensor range. The results demonstrated the IMP-DR in theory, and the experimental deployment supports its practical viability in monitoring tasks. Future research is directed at formulating the planning-oriented model prediction to increase performance and address multi-rotor vehicle limited-thrust model and battery charge constraints in prolonged monitoring missions.

REFERENCES

- [1] G. A. Hollinger and G. S. Sukhatme, "Sampling-based robotic information gathering algorithms," *The International Journal of Robotics Research*, vol. 33, no. 9, pp. 1271–1287, 2014.
- [2] F. Meyer and K. Glock, "Kinematic orienteering problem with time-optimal trajectories for multirotor uavs," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 11402–11409, 2022.
- [3] S. L. Smith, M. Schwager, and D. Rus, "Persistent robotic tasks: Monitoring and sweeping in changing environments," *IEEE Transactions on Robotics*, vol. 28, no. 2, pp. 410–426, 2012.
- [4] I.-M. Chao, B. L. Golden, and E. A. Wasil, "The team orienteering problem," *European Journal of Operational Research*, vol. 88, no. 3, pp. 464–474, 1996.
- [5] C. B. Pedersen, K. G. Nielsen, K. Rosenkrands, A. E. Vasegaard, P. Nielsen, and M. El Yafrani, "A grasp-based approach for planning uav-assisted search and rescue missions," *Sensors*, vol. 22, no. 1, 2021.
- [6] W. Luo and K. Sycara, "Adaptive sampling and online learning in multi-robot sensor coverage with mixture of gaussian processes," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 6359–6364.
- [7] V. Tresp, "Mixtures of gaussian processes," in *Advances in Neural Information Processing Systems*, T. Leen, T. Dietterich, and V. Tresp, Eds., vol. 13, 2000.
- [8] B. Moon, S. Chatterjee, and S. Scherer, "Tigris: An informed sampling-based algorithm for informative path planning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 5760–5766.
- [9] Y. Shi, N. Wang, J. Zheng, Y. Zhang, S. Yi, W. Luo, and K. Sycara, "Adaptive informative sampling with environment partitioning for heterogeneous multi-robot systems," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 11718–11723.
- [10] T. Choi and G. Cielniak, "Adaptive selection of informative path planning strategies via reinforcement learning," in *European Conference on Mobile Robots (ECMR)*, 2021, pp. 1–6.
- [11] M. Ghaffari Jadidi, J. Valls Miro, and G. Dissanayake, "Sampling-based incremental information gathering with applications to robotic exploration and environmental monitoring," *The International Journal of Robotics Research*, vol. 38, no. 6, pp. 658–685, 2019.
- [12] X. Lan and M. Schwager, "A variational approach to trajectory planning for persistent monitoring of spatiotemporal fields," in *American Control Conference*. IEEE, 2014, pp. 5627–5632.
- [13] P. Vansteenwegen, W. Souffriau, and D. V. Oudheusden, "The orienteering problem: A survey," *European Journal of Operational Research*, vol. 209, no. 1, pp. 1–10, 2011.
- [14] R. Pěnička, J. Faigl, M. Saska, and P. Váňa, "Data collection planning with non-zero sensing distance for a budget and curvature constrained unmanned aerial vehicle," *Autonomous Robots*, vol. 43, no. 8, pp. 1937–1956, 2019.
- [15] D. Zahrádka, R. Pěnička, and M. Saska, "Route planning for teams of unmanned aerial vehicles using dubins vehicle model with budget constraint," in *International Conference on Modelling and Simulation for Autonomous Systems*, 2018, pp. 365–389.
- [16] A. Garg and S. S. Jha, "Directed explorations during flood disasters using multi-uav system," in *IEEE International Conference on Automation Science and Engineering (CASE)*, 2022, pp. 2154–2161.
- [17] J. Yu, M. Schwager, and D. Rus, "Correlated orienteering problem and its application to informative path planning for persistent monitoring tasks," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 342–349.
- [18] M. Khosravi, H. Khodadadi, A. G. Aghdam, and H. Rivaz, "Maximum reward collection problem: a cooperative receding horizon approach for dynamic clustering," in *Conference on Research in Adaptive and Convergent Systems*, 2015, pp. 38–43.
- [19] —, "A cooperative receding horizon controller for multi-target interception with obstacle avoidance," in *IEEE Conference on Decision and Control (CDC)*, 2016, pp. 93–98.
- [20] F. Koohifar, A. Kumbhar, and I. Guvenc, "Receding horizon multi-uav cooperative tracking of moving rf source," *IEEE Communications Letters*, vol. 21, no. 6, pp. 1433–1436, 2017.
- [21] J. V. Durgadoo, A. Biastoch, A. L. New, S. Rühls, A. J. Nurser, Y. Drillet, and J.-R. Bidlot, "Strategies for simulating the drift of marine debris," *Journal of operational Oceanography*, vol. 14, no. 1, pp. 1–12, 2021.
- [22] T. S. van den Bremer and Ø. Breivik, "Stokes drift," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 376, no. 2111, p. 20170104, 2018.
- [23] S. Lucia, A. Tătulea-Codrean, C. Schoppmeyer, and S. Engell, "Rapid development of modular and sustainable nonlinear model predictive control solutions," *Control Engineering Practice*, vol. 60, pp. 51–62, 2017.
- [24] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.
- [25] HSL, "A collection of fortran codes for large-scale scientific computation," 2013, [cited 15 Nov 2022]. [Online]. Available: <http://www.hsl.rl.ac.uk>
- [26] T. Bába, M. Petrлік, M. Vrba, V. Spurný, R. Penička, D. Heřt, and M. Saska, "The mrs uav system: Pushing the frontiers of reproducible research, real-world deployment, and education with autonomous unmanned aerial vehicles," *Journal of Intelligent & Robotic Systems*, vol. 102, no. 1, p. 26, 2021.
- [27] "Supporting Material for the Multi-vehicle Dynamic Water Surface Monitoring," 2023, [cited 17 Jun 2023]. [Online]. Available: <http://mrs.felk.cvut.cz/uavmon-ral>