

Maximizing Seaweed Growth on Autonomous Farms: A Dynamic Programming Approach for Underpowered Systems Operating in Uncertain Ocean Currents

Matthias Killer , Marius Wiggert , Hanna Krasowski , *Member, IEEE*, Manan Doshi, Pierre F.J. Lermusiaux, *Member, IEEE*, and Claire J. Tomlin , *Fellow, IEEE*

Abstract—Seaweed biomass presents a substantial opportunity for climate mitigation, yet to realize its potential, farming must be expanded to the vast open oceans. However, in the open ocean neither anchored farming nor floating farms with powerful engines are economically viable. Thus, a potential solution are farms that operate by *going with the flow*, utilizing minimal propulsion to strategically leverage beneficial ocean currents. In this work, we focus on low-power autonomous seaweed farms and design controllers that maximize seaweed growth by taking advantage of ocean currents. We first introduce a Dynamic Programming (DP) formulation to solve for the growth-optimal value function when the true currents are known. However, in reality only short-term imperfect forecasts with increasing uncertainty are available. Hence, we present three additional extensions. Firstly, we use frequent replanning to mitigate forecast errors. Second, to optimize for long-term growth, we extend the value function beyond the forecast horizon by estimating the expected future growth based on seasonal average currents. Lastly, we introduce a discounted finite-time DP formulation to account for the increasing uncertainty in future ocean current estimates. We empirically evaluate our approach with 30-day simulations of farms in realistic ocean conditions. Our method achieves 95.8% of the best possible growth using only 5-day forecasts. This demonstrates that low-power propulsion is a promising method to operate autonomous seaweed farms in real-world conditions.

Index Terms—Marine robotics, underactuated robots, planning under uncertainty, optimization and optimal control.

Received 2 April 2025; accepted 10 August 2025. Date of publication 1 September 2025; date of current version 10 September 2025. This article was recommended for publication by Associate Editor K. Chatzilygeroudis and Editor O. Stasse upon evaluation of the reviewers' comments. This work was supported in part by C3.ai Digital Transformation Institute, in part by the IFI Program of the German Academic Exchange Service (DAAD), and in part by the DARPA Assured Autonomy program. (Matthias Killer and Marius Wiggert contributed equally to this work.) (Corresponding authors: Matthias Killer; Marius Wiggert.)

Matthias Killer and Hanna Krasowski are with the Department of EECS, University of California, Berkeley, CA 94720 USA, and also with the Technical University of Munich, 80333 München, Germany (e-mail: matthias.killer@tum.de).

Marius Wiggert and Claire J. Tomlin are with the Department of EECS, University of California, Berkeley, CA 94720 USA (e-mail: mariuswiggert@berkeley.edu).

Manan Doshi and Pierre F.J. Lermusiaux are with the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Digital Object Identifier 10.1109/LRA.2025.3604727

I. INTRODUCTION

RECENT research has shown promising applications of seaweed biomass for climate mitigation. It can be used as human food, as cattle feed that reduces methane emissions [1], for biofuel and plastic [2], and for carbon capture i.e. when the biomass is sunk to the ocean floor, it removes carbon dioxide from the atmosphere [3]. To deliver on this promise, production must scale by expanding seaweed farming from labor-intensive operations near shore to automated solutions utilizing the vast expanse of the open oceans [4]. But conventional farming becomes economically infeasible in deeper waters as anchoring costs increase with depth [5].

A promising solution could be non-tethered, autonomous seaweed farms that roam the oceans while growing seaweed [6], [7]. These floating farms need to be able to control their position to prevent stranding, colliding with ships, or drifting to nutrient-depleted waters. While they could be steered with powerful ship engines, the power and, hence, energy costs are prohibitively high due to the drag force F_{Drag} . Vessels in water require power $P = F_{Drag} \cdot v = C_d A_C \cdot v^3$ where C_d is the drag coefficient A_C the cross-sectional area, both of which are large for such farms [8]. As P is cubic in the relative velocity v , low propulsion implies low energy usage. Recent studies [9], [10] demonstrated that an autonomous vessel can navigate reliably to nearby targets by *going with the flow*, using its limited propulsion ($0.1 \frac{m}{s}$) strategically to nudge itself into ocean currents ($[0 - 2 \frac{m}{s}]$) that drift towards its destination. These studies have been extended to reduce the risk of stranding by incorporating obstacles [11] and to multi-agent fleets of vessels that navigate while staying connected in a local communication network [12].

In this letter, we use this low-power steering paradigm for operating seaweed farms. In contrast to [9], [10], which solves navigating to a target within a 5-day forecast horizon, our objective is to maximize seaweed growth along the trajectory of the farms over longer periods beyond the forecast horizon. For an autonomous vessel operating approach, there are four key challenges that we need to address. First, the currents are non-linear and time-varying. Second, in realistic settings, only coarse uncertain forecasts are available [13], [14], [15]. Third, the farm itself is *underpowered* by which we mean that

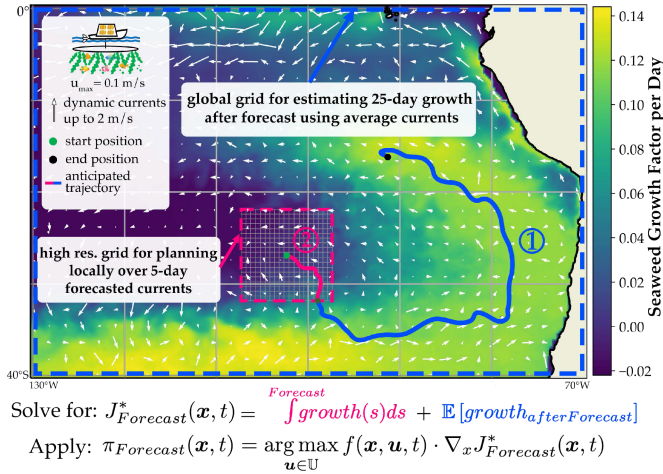


Fig. 1. Our method maximizes long-term growth on autonomous seaweed farm that operate by harnessing ocean currents. We solve for the value function $J_{Forecast}^*$ that is long-term growth-optimal under the forecast with dynamic programming. We first compute the expected 25-day growth after the forecast based on historical average currents (1) and then use it to regularly solve for the value function over the next 5 days using daily current forecasts (2). Applying the induced policy $\pi_{Forecast}$ as feedback controller ensures high growth despite imperfect short-term forecasts.

its propulsion is smaller than the surrounding currents, so it cannot easily compensate for forecast errors. Lastly, we want to maximize seaweed growth over weeks but forecasts from the leading providers are only 5-10 days long [14], [15] and the uncertainty for long-time predictions is high [16]. As summarized in Fig. 1, we are tackling long-term horizon optimization of a state-dependent running cost with an underpowered agent in non-linear time-varying dynamics under uncertainty that increases over time. The long-term dependency of seaweed growth means the objective cannot easily be decomposed into multiple short-term objectives. While we showcase the method specifically on autonomous seaweed farms, many other applications could use environmental dynamics while navigating over long timeframes in the ocean currents or winds, e.g., for low-powered sensor systems (balloons, sailboats, or zeppelins for sensing or building mesh networks for communication) [17].

A. Related Work

Various approaches for time- and energy-optimal path planning exist for non-linear, time-varying dynamics like ocean currents [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31]. In the context of navigating within *known currents or flows*, researchers have derived Hamilton-Jacobi (HJ) reachability equations for exact solutions [18], [19], non-linear programming [20], [32], evolutionary algorithms [21], and graph-based search methods [22], [24], [33]. However, the last three techniques are prone to discretization errors and the non-convex nature of the problem, can lead to infeasibility or solvers getting stuck in local minima. In contrast, DP based on the HJ equations can solve the exact continuous-time control problem.

There is limited research that focuses on maximizing seaweed growth. In [34], the authors maximize seaweed harvesting using

autonomous vessels in varied settings. They use a 3D HJ reachability framework in which the harvesting state is augmented into the third dimension. They optimize harvesting from stationary seaweed farms and assume the currents to be known making it not applicable in realistic settings, additionally the value function is not suitable for a closed-loop control policy.

For managing current uncertainty, previous work optimizes the expectation or a risk function over a stochastic solution of probabilistic ocean flows [26]. However, this is not yet suitable for operational settings as it demands a principled uncertainty distribution for flows but most operational forecasts are deterministic. At the same time, robust control techniques, which aim to maximize the objective even in the face of worst-case disturbances, are not suitable when considering realistic error bounds, as the forecast error often equals or exceeds our low propulsion capabilities. Thus, to mitigate forecast inaccuracies, frequent replanning in a Model Predictive Control (MPC) fashion has been proposed using either non-linear programming [35], [36] or employing the HJ value function as feedback policy [9], which offers the benefits of being both fast and optimal. Another approach is to use Reinforcement Learning (RL) to learn how to best operate stratospheric balloons despite wind forecast uncertainty [31], [37]. While they ran operational experiments over more than 30 days their objective of staying above a certain area is relatively short-term, rendering RL appropriate. However, the applicability of RL for long-term objectives, similar to ours, remains uncertain. To address the increasing complexity associated with long-time horizons, problems are frequently divided into multiple subproblems using graph-based methods or hierarchical RL [38], [39]. These approaches are appropriate for combinatorial optimization problems, where dividing and conquering in subtasks is effective. However, this is not suitable for our problem involving continuous state space and long-time dependencies. An approach to handle growing uncertainty of the currents over time is to discount future rewards which is common in RL settings [40] but not in maritime control.

B. Overview of Method & Contributions

In this letter, we make five main contributions towards controllers that optimize seaweed growth on autonomous seaweed farms over long periods.

First, we formulate maximizing seaweed growth on an autonomous farm as an optimization problem that can be solved exactly with DP in the 2D spatial state of the system (Section III-A). Compared to prior work using HJ Reachability in 3D [34] to model seaweed growth on stationary farms, our formulation leads to two advantages: significant reduction of computational complexity (Section III) and the value function can be used as feedback policy. This allows for frequent replanning in the MPC spirit which is critical to mitigate forecast uncertainty and can be used for multiple farms [12]. Second, we extend the value function beyond the forecast horizon, which leads to a feedback policy that optimizes for long-term optimal growth (Section III-C). For that, we estimate the expected growth using historical average currents over a coarse grid and then initialize the DP over the forecast horizon with these values

(Fig. 1, Section III-C). Third, to account for the growing uncertainty of the ocean current estimates, we introduce finite-time discounting into the DP formulation, which to the best of the authors knowledge has not been applied to the context of handling uncertainty in flow field forecast. (Section III-D). Fourth, we are the first to run extensive empirical simulations of autonomous seaweed farms in realistic current settings over 30 days. We first investigate how different propulsion of the farms would affect the best achievable seaweed growth with known currents. We then evaluate how close different configurations of our method can get to the best achievable growth when only daily, 5-day forecasts are available (Section IV). Lastly, we open-source our code, which contains extensive features to simulate, visualize, and study controllers for 2D vessels operating by harnessing uncertain ocean currents.

The remainder of the article is structured as follows: in Section II we define the problem. Section III details the four components of our method. Section IV contains the performance evaluation of our methods and baselines, and we conclude with Section VI and outline future work.

II. PROBLEM STATEMENT

A. System Dynamics

We consider an autonomous seaweed farm as surface vessel on the ocean with the spatial state $\mathbf{x} \in \mathbb{R}^2$. Let the control input be denoted by $\mathbf{u}$ from a bounded set $\mathbb{U} \in \mathbb{R}^{n_u}$ where n_u is the dimensionality of the control. Then, the spatial dynamics of the system at time t can be modelled by the first order Ordinary Differential Equation (ODE):

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}, t) = v(\mathbf{x}, t) + g(\mathbf{x}, \mathbf{u}, t), \quad t \in [0, T] \quad (1)$$

where the movement of the vessel depends on the drift due to the time-varying, non-linear flow field $v(\mathbf{x}, t) \rightarrow \mathbb{R}^2$ and its control $\mathbf{u}$. We choose a first-order model where the drift and control directly influence the state, disregarding inertial effects from motor acceleration and drag forces. This is justified by the fact that high-drag seaweed farms attain equilibrium velocity within a few minutes, a timescale considerably shorter than our 30-day planning horizon.

While our method is generally applicable, we focus on *underpowered* settings in the sense that most of the time $\max \|g(\mathbf{x}, \mathbf{u}, t)\|_2 \ll \|v(\mathbf{x}, t)\|_2$. We denote the spatial trajectory induced by this ODE with ξ . For a vessel starting at the initial state $\mathbf{x}_0$ at time t_0 with control sequence $\mathbf{u}(\cdot)$, we denote the state at time t by $\xi_{t_0, \mathbf{x}_0}^{\mathbf{u}(\cdot)}(t) \in \mathbb{R}^2$. The system dynamics ((1)) are assumed to be continuous, bounded, and Lipschitz continuous in $\mathbf{x}, \mathbf{u}$ [10].

Additionally, we assume the farm has seaweed mass m which evolves according an exponential growth ODE:

$$\dot{m} = m \cdot \Psi(\mathbf{x}, t), \quad t \in [0, T] \quad (2)$$

where Ψ is the growth factor per time unit, e.g. $20 \frac{\%}{\text{day}}$, which depends on nutrients, incoming solar radiation, and water temperature at the spatial state $\mathbf{x}$ and time t .

B. Problem Setting

The objective of the seaweed farm starting from $\mathbf{x}_0$ at t_0 with seaweed mass $m(t_0)$ is to maximize the seaweed mass at the final time T . This implies optimizing the growth along its trajectory $\xi_{t_0, \mathbf{x}_0}^{\mathbf{u}(\cdot)}$.

$$\max_{\mathbf{u}(\cdot)} m(T) = m(t_0) + \max_{\mathbf{u}(\cdot)} \int_{t_0}^T m(s) \cdot \underbrace{\Psi(\xi_{t_0, \mathbf{x}_0}^{\mathbf{u}(\cdot)}(s), s)}_{\text{growth factor}} ds \quad (3)$$

If the currents v are known, our method (Section III) is guaranteed to find the optimal value function from which the optimal control $\mathbf{u}^*(\cdot)$ and trajectory can be obtained. However, in realistic scenarios only inaccurate, short-term forecasts $\hat{v}_{FC}$ are available at regular intervals. These differ from the true flow v by the forecast error $\delta(\mathbf{x}, t)$. Our goal is then to determine a feedback policy $\pi(\mathbf{x}, t)$ that results in a high expected seaweed mass $\mathbb{E}[m(T)]$. Hence, in our experiments (Section IV) we evaluate our method empirically over a set of missions $(\mathbf{x}_0, t) \sim \mathbb{M}$ and a realistic distribution of true and forecasted ocean currents $v, \hat{v}_{FC} \sim \mathbb{V}$.

III. METHOD

Our method consists of a core DP formulation that optimizes seaweed growth when the currents are known and three extensions to get a feedback policy π that performs well over long-time horizons when only limited forecasts are available. We first introduce the core DP formulation to obtain the growth-optimal value function (Section III-B). Then, we demonstrate using the value function as feedback policy π , which is equivalent to replanning at every time step (Section III-C). This leads to reliable performance even if the value function was computed with inaccurate forecasts. Next, we extend the feedback policy by estimating the growth beyond the forecast horizon (Section III-D and introduce a finite-time discount factor III-D). Lastly, we describe the control algorithm variations developed and discuss computational aspects (Section III-E and III-F).

A. Maximizing Seaweed Mass With Known Dynamics

We use continuous-time optimal control where the value function $J(\mathbf{x}, \mathbf{u}(\cdot), t)$ of a trajectory ξ is based on a state and time-dependent reward R and a terminal reward R_T :

$$J(\mathbf{x}, \mathbf{u}(\cdot), t) = \int_t^T R(\xi_{t, \mathbf{x}}^{\mathbf{u}(\cdot)}(s), s) ds + R_T(\xi_{t, \mathbf{x}}^{\mathbf{u}(\cdot)}(T), T).$$

Let $J^*(\mathbf{x}, t) = \max_{\mathbf{u}(\cdot)} J(\mathbf{x}, \mathbf{u}(\cdot), t)$ be the optimal value function. Using DP we can derive the corresponding Hamilton-Jacobi Partial Differential Equation (PDE) [41]:

$$-\frac{\partial J^*(\mathbf{x}, t)}{\partial t} = \max_{\mathbf{u}} [\nabla_{\mathbf{x}} J^*(\mathbf{x}, t) \cdot f(\mathbf{x}, \mathbf{u}, t) + R(\mathbf{x}, t)] \quad (4)$$

$$J^*(\mathbf{x}, T) = R_T(\mathbf{x}, T). \quad (5)$$

Computationally, this PDE is solved in the state space discretized into N grid points along each dimension d [42]. At each step backward in time we need to compute the gradient

$\nabla_x J^*(\mathbf{x}, t)$ which is $O(d)$ for each grid point, so the complexity scales exponentially with the state dimension as $O(dN^d)$, which is called the curse of dimensionality [41].

Next, we define the reward R and terminal reward R_T to maximize $m(T)$. One approach is to solve the PDE in an augmented state space $x_{aug} = (\mathbf{x}, m)^\top \in \mathbb{R}^3$. If we set $R_T = 0$ and define the reward as $R = m \cdot \Psi(\mathbf{x}, t)$, the value function is our objective (3). However, as the computational complexity of solving for J^* scales exponentially with the state dimension, we want a reward R that does not depend on the augmented state m . For that, we introduce the variable $\eta = \ln(m)$ with the new dynamics $\dot{\eta} = \frac{\dot{m}}{m} = \Psi(\mathbf{x}, t)$. As $\eta(m)$ is strictly increasing in m , the control $\mathbf{u}^*(\cdot)$ that maximizes $\eta(T)$ is equivalent to $\mathbf{u}^*(\cdot)$ maximizing $m(T)$. We can then reformulate (3) to $\eta(T)$:

$$\max_{\mathbf{u}(\cdot)} \eta(T) = \eta(t_0) + \max_{\mathbf{u}(\cdot)} \int_{t_0}^T \Psi(\xi_{t,x}^{\mathbf{u}(\cdot)}(s), s) ds. \quad (6)$$

By setting the reward to $R = \Psi(\mathbf{x}, t)$ the optimal value function captures this optimization without requiring m :

$$J^*(\mathbf{x}, t) = \max_{\mathbf{u}(\cdot)} \int_t^T \Psi(\xi_{t,x}^{\mathbf{u}(\cdot)}(s), s) ds. \quad (7)$$

We then solve the HJ PDE for the growth-optimal $J^*(\mathbf{x}, t)$ in the spatial state $\mathbf{x}$ and obtain $\mathbf{u}^*(\cdot)$ and the trajectory $\xi_{t_0,x_0}^{\mathbf{u}^*(\cdot)}$ that maximize $m(T)$ at $\frac{2}{3}N$ the computational cost (Section III-F). This formulation can be applied more generally to optimize the state of exponential growth or decay ODEs. We can convert the value of $J^*(\mathbf{x}_0, t_0)$ to the final seaweed mass of the optimal trajectory starting at $\mathbf{x}_0, t_0$ with $m(t_0)$:

$$m(T) = m(t_0) \cdot e^{\int_{t_0}^T \Psi(\xi_{t_0,x_0}^{\mathbf{u}^*(\cdot)}(s), s) ds} = m(t_0) \cdot e^{J^*(\mathbf{x}_0, t_0)}.$$

B. Feedback Policy Based on Regular Forecasts

The value function J^* from Section III-A allows us to compute the optimal control $\mathbf{u}^*(\mathbf{x}, t)$ for all $\mathbf{x}, t$ and hence a feedback policy $\pi(\mathbf{x}, t)$ for the vessel or multiple vessels in the same region [12]. This policy is the optimizer of the Hamiltonian (right side (4)):

$$\pi(\mathbf{x}, t) = \arg \max_{\mathbf{u} \in \mathcal{U}} f(\mathbf{x}, \mathbf{u}, t) \cdot \nabla_x J^*(\mathbf{x}, t), \quad (8)$$

which can often be computed analytically depending on $g(\mathbf{x}, \mathbf{u}, t)$. For our experiment, we take thrust and direction θ as control variables: $g(\mathbf{u}(t)) = u_{thrust}(t)\hat{\mathbf{e}}_{\theta(t)}$ with limited scalar thrust $u_{thrust} \leq u_{max}$ and $\hat{\mathbf{e}}_{\theta}$ as the 2D unit vector with angle θ . The policy is then full thrust along $\nabla_x J^*(\mathbf{x}, t)$: $\pi(\mathbf{x}, t) = \frac{\nabla_x J^*(\mathbf{x}, t)}{\|\nabla_x J^*(\mathbf{x}, t)\|_2}$. While π is optimal if J^* is based on the true currents v , it can also be applied when imperfect forecasts $\hat{v}_{FC}$ were used to compute the value function $J_{\hat{v}_{FC}}^*(\mathbf{x}, t)$. In that case, an agent at state $\mathbf{x}$ executing $\pi_{\hat{v}_{FC}}(\mathbf{x}, t)$ will find itself at a different state $\mathbf{x}'$ than anticipated as v differs from $\hat{v}_{FC}$. But the control that would be growth optimal under $\hat{v}_{FC}$ can again be computed with $\pi_{\hat{v}_{FC}}(\mathbf{x}', t + \Delta t)$. Applying $\pi_{\hat{v}_{FC}}$ closed-loop is hence equivalent to full-time horizon re-planning with $\hat{v}_{FC}$ at each time step. This notion of re-planning at every time step at the low cost of a 2D gradient computation (Section III-F) ensures

good performance despite forecast errors [9]. $J_{\hat{v}_{FC}}^*(\mathbf{x}, t)$ can be updated daily as new forecasts arrive.

C. Reasoning Beyond the Forecast Horizon

As the growth cycles of seaweed typically spans months, our aim is to maximize the seaweed mass at an *extended* future time T_{ext} after the final time of the 5-day forecast T_{FC} . A principled way to reason beyond the planning horizon is to estimate the expected growth our seaweed farm will experience from the state $\xi_{t,x}^{\mathbf{u}(\cdot)}(T_{FC})$ onward and add this as terminal reward R_T to (7).

$$J_{\hat{v}_{FC}, ext}^*(\mathbf{x}, t) = J_{\hat{v}_{FC}, T_{FC}}^*(\mathbf{x}, t) + \mathbb{E} \left[J_{T_{ext}}^*(\xi_{t,x}^{\mathbf{u}(\cdot)}(T_{FC}), T_{FC}) \right] \\ J_{\hat{v}_{FC}, T_{FC}}^*(\mathbf{x}, t) = \max_{\mathbf{u}(\cdot)} \int_t^{T_{FC}} \Psi(\xi_{t,x}^{\mathbf{u}(\cdot)}(s), s) ds \quad (9)$$

where $J_{\hat{v}_{FC}, T_{FC}}^*(\mathbf{x}, t)$ is the growth a vessel starting from $\mathbf{x}$ at t will achieve at T_{FC} and $\mathbb{E}[J_{T_{ext}}^*(\xi_{t,x}^{\mathbf{u}(\cdot)}(T_{FC}), T_{FC})]$ estimates the additional growth from T_{FC} to T_{ext} . The expectation is over the uncertain future ocean currents.

We propose to estimate $\mathbb{E}[J_{T_{ext}}^*]$ by computing the value function $J_{\bar{v}, T_{ext}}^*$ based on monthly average currents $\bar{v}$ for the region using Section III-A. To compute $J_{\hat{v}_{FC}, ext}^*$ we then solve (4) with $R_T(\mathbf{x}, T_{FC}) = J_{\bar{v}, T_{ext}}^*(\mathbf{x}, T_{FC})$.

D. Finite-Time Discounting to Mitigate Uncertainty

As the oceans are a chaotic system, the uncertainty of the forecasted ocean currents increases over time. We can incorporate this increasing uncertainty in the value function by using the finite-time discounted optimal control formulation:

$$J^\tau(\mathbf{x}, \mathbf{u}(\cdot), t) = \int_t^T e^{\frac{-(s-t)}{\tau}} R(\xi_{t,x}^{\mathbf{u}(\cdot)}(s), s) ds \\ + R_T(\xi_{t,x}^{\mathbf{u}(\cdot)}(T), T),$$

where τ is the discount factor. Note that in contrast to discrete time dynamics, where discount factors range from 0 to 1, our application of τ conforms to the conventional interpretation in continuous dynamic programming [40]: τ can assume values significantly greater than 1 and for higher τ future rewards are discounted less. We derive the corresponding HJ PDE by following the steps in [40] and in place of (4) we obtain:

$$\frac{\partial J^{*,\tau}(\mathbf{x}, t)}{\partial t} = - \max_{\mathbf{u}} [\nabla_x J^{*,\tau} \cdot f(\mathbf{x}, \mathbf{u}, t) \\ + R(\mathbf{x}, t)] + \frac{J^{*,\tau}(\mathbf{x}, t)}{\tau}$$

E. Control Algorithm Variations

All variations of our method are feedback policies π derived from a value function (Section III-B). The four variations differ only in how the value function is computed. When the true currents v are known we compute J^* (7) for optimal control. When only forecasts $\hat{v}_{FC}$ are available, we calculate the $J_{\hat{v}_{FC}}^*$ for planning horizons up to the end of the forecasts T_{FC} and update it as new forecasts become available (Section III-B). Thirdly, to optimize for growth until $T_{ext} > T_{FC}$ we calculate an extended

Algorithm 1: Discounted HJ Closed-Loop Control.

Input: Forecast Flow(s) $\hat{v}_{FC}$, $t = 0$, $\mathbf{x}(t) = \mathbf{x}_0$, average Flows $\bar{v}$, discount τ , plan until T_{ext}

- 1 Compute $J_{\bar{v}, T_{\text{ext}}}^{*, \tau}$ using $\bar{v}$ (Sec. III-C and III-D);
- 2 **while** $t \leq T$ **do**
- 3 **if** new forecast $\hat{v}_{FC}$ available **then**
- 4 Compute $J_{\hat{v}_{FC}, \text{ext}}^{*, \tau}$ (Sec. III-C and III-D);
- 5 $\mathbf{u}_t = \pi_{\hat{v}_{FC}, \text{ext}}^{*, \tau}(\mathbf{x}_t, t)$; using $J_{\hat{v}_{FC}, \text{ext}}^{*, \tau}$ (Sec. III-B)
- 6 $\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \int_t^{t+\Delta t} f(\mathbf{u}_s, \mathbf{x}(s), s) ds$;
- 7 $t \leftarrow t + \Delta t$;

value function $J_{\hat{v}_{FC}, \text{ext}}^{*}$ (Section III-C) using average currents ($\hat{v}_{FC}$ then $\bar{v}$). Lastly, we can discount future rewards with $J_{\bar{v}, T_{\text{ext}}}^{*, \tau}$ (Section III-D) in any of the above value functions. In Algorithm 1 we detail the discounted, long-term version as it contains all components.

F. Computational Considerations

To illustrate the computational advantage of our approach let's consider our realistic simulation experiments in Section IV. The computational complexity is $O(dN^d)$ so with a spatial discretization of $N = 120$ solving for J^* in $d = 3$ compared to $d = 2$ dimensions would be $\frac{2}{3}N = 180$ times as expensive. We only need to solve the 2D HJ PDE for $J_{\hat{v}_{FC}}^*(\mathbf{x}, t)$ once per day as new forecasts become available. Concretely, for a greedy $T=3$ day horizon and 1° spatial grid this takes $\approx 2.3 \frac{s}{\text{day}} \cdot T$ so 7s with a 2.7GHz Intel i7. From the value function, we obtain the control every 10 minutes with a cheap 2D gradient computation $\nabla_{\mathbf{x}} J^*(\mathbf{x}, t)$ which takes 1ms and is $O(d)$. That means ≈ 7.14 s per operational day.

In contrast, using non-linear programming MPC, we would need to solve an optimization problem 144 times per day. We implemented a multiple-shooting MPC approach with Euler dynamics and CasADi [43]. With the same settings as above, $T=3$ days and 10 mins time-steps the IPOPT solver [44] took a minimum of 4s for mostly linear ocean current flows, to 10s for more non-linear flow fields, resulting in a total of at least 576s per operational day.

In addition to these savings for a single farm, $J_{\hat{v}_{FC}}^*(\mathbf{x}, t)$ can be used for hundreds of farms in the same region [12], whereas MPC would need to be run for each farms separately. Considering control resilience, HJ guarantees convergence and optimality [41] whereas non-linear MPC sometimes does not find a feasible solution.

One limitation of planning with ocean current forecasts provided as matrices is that it requires significant RAM due to the interpolation of currents at each time-step (e.g., the above MPC and HJ examples took 1.2 GB and 1.1 GB peak RAM, respectively). We use JAX to first compile a computational graph for the value function computation before solving the PDE. This adaption yields a significant speed-up over the Matlab-based helperOC. Nevertheless, it took 60 GB of RAM for 30 d planning of $J_{\hat{v}_{FC}, \text{ext}}^*(\mathbf{x}, t)$, which limited our simulation horizon in Section IV. This can be optimized further e.g. by using GPU acceleration and moving the interpolation outside of the PDE solving.

IV. EXPERIMENTS

As our system is underpowered (Section II), it is impossible to prove robustness of our method against potentially adversarial currents [11]. Hence, we evaluate our method empirically by simulating the operation of an autonomous seaweed farm in realistic ocean currents and growth conditions. We will open-source the code for our simulator and controllers for others to replicate results and build on.¹ We run two main experiments. First, we investigate how varying the propulsion $\mathbf{u}_{max}$ impacts the best achievable seaweed growth under known currents v and compare it to the growth achieved by 30-day planning without discounting relying on daily, 5-day forecasts v and average currents $\bar{v}$ (Section IV-B1). Second, we fix the propulsion to $\mathbf{u}_{max}=0.1 \frac{m}{s}$ and evaluate how the planning horizon and discounting in our method affect growth and how close we can get to the best achievable growth while relying on daily forecasts $\hat{v}_{FC}$ and average currents $\bar{v}$ (Section IV-B2). The experimental setup for both is the same and will be explained next.

A. Experimental Setup

1) *Seaweed Growth Model:* The growth of macroalgae depends on the species, water temperature, solar irradiance, and dissolved nutrient concentrations, specifically nitrate (NO_3) and phosphate (PO_4) [34]. We use the model of the Net Growth Rate (NGR) of Wu et al. [45] and temperate species parameters from [46], [47]. In this model, the time-dependent NGR is determined by the growth rate r_{growth} and the respiration rate r_{resp} caused by metabolism as:

$$\dot{m}(t) = m(t) \cdot \text{NGR}(t) = m(t) \cdot (r_{\text{growth}}(t) - r_{\text{resp}}(t)). \quad (10)$$

Fig. 1 shows the NGR for our region at the apex of the sun's motion in January 2022.

2) *Realistic Ocean Forecast Simulation:* In realistic operations the vessel receives daily forecasts for replanning. In our simulations, we use Copernicus [15] hindcasts as true currents v and mimic daily 5-day forecasts $\hat{v}_{FC}$ by giving the planner access to a 5-day sliding time window of HYCOM [14] hindcasts. Aligned with previous work [9], we find that the forecast error δ with this setting is comparable to the evaluated forecast error of HYCOM [13] in key metrics. To estimate the expected growth beyond the forecast horizon of $\hat{v}_{FC}$ (Section III-C) we use $\frac{1}{6}$ th deg seasonal averages $\bar{v}$ of the ocean currents from Copernicus 2021.

3) *Large Scale Mission Generation:* We simulate operations in the southeast Pacific due to high nutrient densities. For a large representative set of missions $\mathbb{M}$, we sampled 1325 tuples $(\mathbf{x}_0, t_0, m(t_0)=100 \text{ kg})$, uniformly distributed in time between January and October 2022 and across the region of longitude $[-130, -70]^\circ\text{W}$ and latitude $[-40, 0]^\circ\text{S}$. This allows for varying current distributions. As our method is not aware of land obstacles, we had 290 missions where at least one of the farms stranded or left the simulation region. While stranding can be avoided by modifying the HJ PDE and initial value function without increasing the computational complexity as described

¹The code will be available in a github repository

TABLE I
COMPARED CONTROLLER SETTINGS

controller	planning horizon T_{ext}	discount τ
w/o discount (v)	30 days	-
floating	-	-
greedy 1 h ($\hat{v}_{FC}$)	1 h	-
greedy 5 days ($\hat{v}_{FC}$)	5 days	-
w/o discount ($\hat{v}_{FC}$ then $\bar{v}$)	30 days	-
w/ discount I ($\hat{v}_{FC}$ then $\bar{v}$)	30 days	1.296.000
w/ discount II ($\hat{v}_{FC}$ then $\bar{v}$)	30 days	1.728.000

in detail by Doering et al. [11], we consider only the remaining 1035 missions for our results as this letter focuses on long-term planning and not obstacle avoidance.

4) *Evaluated Controllers and Baselines*: We evaluate our method in different configurations categorized by: a) the ocean current data used by the controller for planning, either the true currents v or daily forecasts $\hat{v}_{FC}$ and average currents $\bar{v}$, and b) the controller's planning horizon T_{ext} over which it optimizes growth, either the entire 30-day period or more short term *greedy* (5-day and 1 h). We also examine the effect of using of a discounted value function. An overview of the configurations is provided in Table I.

The simplest baseline that we compare against is the seaweed growth on passively floating farms. We also consider the greedy HJ-based controllers as baselines representing all short-term controllers. That is because they are optimal under $\hat{v}_{FC}$, hence we would only expect MPC or another approach to be better in the unlikely case that their approximation errors would systematically improve performance.

For long-term ($T_{\text{ext}}=30$ days) controllers, we compute the growth-to-go after T_{FC} , i.e., $J_{\bar{v}, T_{\text{ext}}}^*(x, T_{FC})$, over the full area on a coarse $\frac{1}{6}^\circ$ grid, as illustrated in Fig. 1. The value function $J_{\bar{v}, T_{\text{ext}}}^*(x, t)$ used for the control policy is then computed daily on new forecasts using a smaller $\frac{1}{12}^\circ$ grid around the current farm's position (10° square).

5) *Evaluation Metrics*: Our objective is to maximize the seaweed mass at the end of each mission $m(T)$. Additionally, we compute the relative improvement in final seaweed mass by normalizing within each mission with the baseline final mass. We then present the average relative improvement across all missions which allows us to gauge how much more/less biomass a specific controller can grow above the baseline. This is important as the start x_0 of a mission is a major indicator of achievable growth as illustrated in Fig. 3. As baselines we use either passively floating or the best achievable growth based on the true currents v .

B. Experimental Results

1) *How Does Varying Propulsion Affect Growth?*: We vary the maximum propulsion u_{max} of the farm and evaluate how this impacts the best achievable seaweed growth under known currents v . Fig. 2 and Table II compare the final seaweed mass distributions for different propulsion levels, starting with passively floating. We observe that the average seaweed growth scales almost linearly with u_{max} , yielding between 15% and 12% more biomass per $0.1 \frac{m}{s}$ propulsion. We also compare how much growth our method w/o discount ($\hat{v}_{FC}$ then $\bar{v}$) achieves

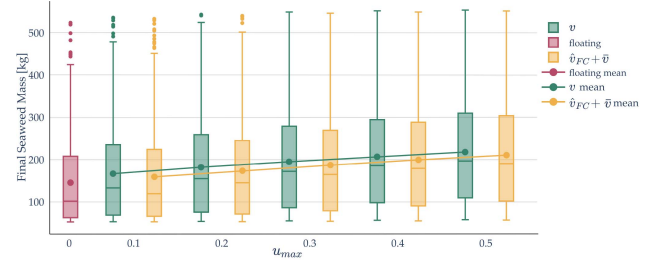


Fig. 2. The best achievable seaweed mass given v increases linearly with u_{max} . Operating with our long-term control method using forecasts $\hat{v}_{FC}$ and average currents $\bar{v}$ achieves $\approx 95\%$ of growth.

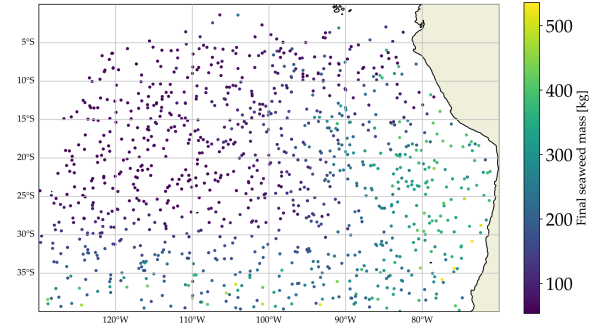


Fig. 3. We sample a diverse set of starts (x_0, t_0) for seaweed farms to empirically evaluate our controllers. The colored starts show the best achievable seaweed mass after 30 days using $u_{max} = 0.1 \frac{m}{s}$.

TABLE II
AVERAGE SEAWEED GROWTH FOR DIFFERENT PROPULSIONS u_{max}

u_{max}	planning input	rel. growth	final seaweed mass	95% CIs
0.0 $\frac{m}{s}$	(floating)	100%	145.29kg $\pm$ 100.30kg	[139.38, 151.60]kg
0.1 $\frac{m}{s}$	v	115.38%	166.45kg $\pm$ 109.67kg	[159.96, 173.41]kg
0.1 $\frac{m}{s}$	$\hat{v}_{FC}$ then $\bar{v}$	109.62%	159.29kg $\pm$ 107.46kg	[152.74, 165.84]kg
0.2 $\frac{m}{s}$	v	128.69%	182.04kg $\pm$ 115.11kg	[175.03, 189.06]kg
0.2 $\frac{m}{s}$	$\hat{v}_{FC}$ then $\bar{v}$	121.29%	173.72kg $\pm$ 112.94kg	[166.84, 180.60]kg
0.3 $\frac{m}{s}$	v	141.27%	194.98kg $\pm$ 117.39kg	[187.83, 202.13]kg
0.3 $\frac{m}{s}$	$\hat{v}_{FC}$ then $\bar{v}$	133.28%	187.01kg $\pm$ 116.60kg	[179.91, 194.11]kg
0.4 $\frac{m}{s}$	v	153.71%	206.96kg $\pm$ 118.34kg	[199.75, 214.17]kg
0.4 $\frac{m}{s}$	$\hat{v}_{FC}$ then $\bar{v}$	145.79%	199.50kg $\pm$ 118.09kg	[192.31, 206.70]kg
0.5 $\frac{m}{s}$	v	165.79%	218.10kg $\pm$ 118.59kg	[210.87, 225.32]kg
0.5 $\frac{m}{s}$	$\hat{v}_{FC}$ then $\bar{v}$	158.14%	210.78kg $\pm$ 117.72kg	[203.61, 217.95]kg

with varying propulsion. As expected this achieves slightly less biomass ($\approx 95\text{--}96\%$ of v) due to forecast errors for all propulsion levels. For higher u_{max} the gap is slightly smaller, possibly because the farm can better compensate for forecast errors. Nonetheless, even small propulsion of $u_{max}=0.1 \frac{m}{s}$ enables 9.6% more biomass than a passively floating farm. The start x_0 of a mission significantly influences 30-day growth, as shown in Fig. 3. High-growth missions are situated in the east and south of our region, aligning with nutrient-rich areas (see Fig. 1). This also explains the high variance in absolute seaweed mass.

2) *The Impact of Planning Horizon and Discounting*: As the energy consumption scales cubically with u_{max} , higher propulsion may be economically infeasible for real-world applications. Therefore, for this experiment we fix $u_{max}=0.1 \frac{m}{s}$. We investigate how different planning horizons and discounting affect performance when operating with forecasts $\hat{v}_{FC}$ and how close we can get to the best achievable growth. We evaluate two

TABLE III
SEAWEED GROWTH OF DIFFERENT CONTROLLERS FOR 1035 MISSIONS

controller	$u_{max}=0.1 \frac{m}{s}$	relative growth	final seaweed mass
w/o discount (v)		100%	168.45kg $\pm$ 109.67kg
floating (-)		88.20%	145.29kg $\pm$ 99.54kg
greedy 1 h ($\hat{v}_{FC}$)		92.24%	152.48kg $\pm$ 102.89kg
greedy 5 days ($\hat{v}_{FC}$)		95.19%	157.78kg $\pm$ 106.04kg
w/o discount ($\hat{v}_{FC}$ then $\bar{v}$)		95.61%	158.84kg $\pm$ 106.71kg
w/ discount I ($\hat{v}_{FC}$ then $\bar{v}$)		95.77%	159.16kg $\pm$ 106.62kg
w/ discount II ($\hat{v}_{FC}$ then $\bar{v}$)		95.77%	159.17kg$\pm$106.66kg

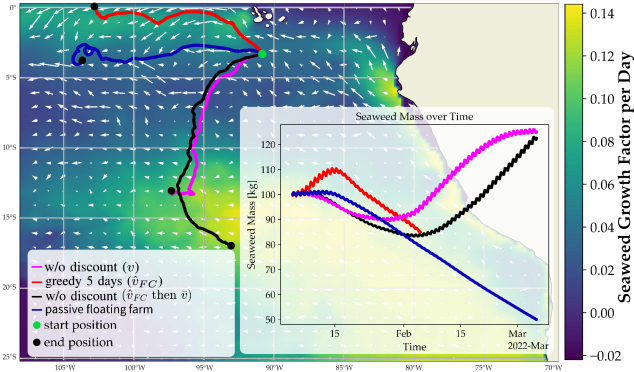


Fig. 4. 60-day Case Study: The greedy controller optimizes for 5-day growth, thereby navigating to the closest growth region. It fails to anticipate the strong currents that push it out of the region. The long-term controllers reach a more distant growth-rich area while incurring short-term losses.

greedy controllers that repeatedly optimize over short T_{ext} (1h and 5-days) and compare to 30-day time-horizon with different discounting settings (Table I).

Table III shows the results. As expected, both the greedy and long-term controllers outperform passively floating. Surprisingly, the performance of the 5-day greedy controller, is close to the 30-day controllers. Using the discounted formulation slightly improves the long-term controller, yielding on average 95.77% of the best achievable growth.

3) *Case Study of 60-Day Scenario:* We were intrigued that the 5-day controller did achieve almost the same seaweed growth as by planning over 30-days (Section IV-B2). Hence, we conducted a case study with planning and operating the farm over 60 days instead of 30 days and with $u_{max}=0.3 \frac{m}{s}$ (Fig. 4). As discussed in Section III-F, large scale experiments on planning horizons longer than 30 days go beyond our currently available computational resources. In the 60-day case study, we find that the greedy controller then aims for the nearest growth region, while the long-term controller properly balances short-term lower growth against the long-term gains of reaching a high-growth region. This leads to the greedy controller being driven out of the simulated region while the long-term controller achieves close to the best achievable growth (see sub-figure Fig. 4). Note that the zig-zag shapes of the lines are due to day-night cycles.

V. DISCUSSION

The experiments demonstrate that controllers using forecasts $\hat{v}_{FC}$ substantially outperform a passively floating farm. The 95%

confidence intervals of floating and our method w/o discounting (Table II) are not overlapping, floating [139.38, 151.60]kg and [152.74, 165.84]kg for $0.1 \frac{m}{s} \hat{v}_{FC}$ then $\bar{v}$. The myopic behavior of a greedy policy not only leads it to navigate toward low-growth regions in the vicinity but also fails to account for the possibility of being pushed out of good-growth regions by strong currents, as in our 60-day case study in Fig. 4. Therefore, we were surprised that our 5-day optimizing controller was nearly on par with our 30-day optimizing controllers (Section IV-B2).

We attribute this to several factors. First, the initial position determines most of the possible growth within the 30 days which can be observed in Fig. 3. Farms starting from suboptimal positions cannot reach and grow seaweed in more distant, high-nutrient regions. We believe that experiments over the full seaweed growth cycle of 60-90 days would yield more significant differences between the controllers as long-term high growth and avoiding low-growth regions become more important. Second, the growth map in our region exhibits a smooth gradient, which means that even greedy controllers might move toward globally optimal growth regions without planning for it. Third, in our experimental evaluation, we do not consider missions where any controller leaves the predefined region (Section IV-A3). This often occurs with greedy or floating controllers (Fig. 4); consequently, the performance increase with long-term controllers would be greater if we accounted for the filtered missions.

VI. CONCLUSION AND FUTURE WORK

In this work, we maximize seaweed growth on autonomous farms that are underpowered and operate by harnessing uncertain ocean currents. We introduced a 2D DP formulation to solve for the growth-optimal value function when the true currents are known. Next, we showed how the value function computed on forecasted currents can be used as feedback policy, which is equivalent to replanning on the forecast at every time step and hence mitigates forecast errors. As operational forecasts are only 5 days long, we extended our method to reason beyond the forecast horizon by estimating expected future growth based on seasonal average currents and using finite-time discounting DP PDE to account for increasing uncertainty in ocean currents. We conducted extensive empirical evaluations based on realistic ocean conditions over 30 days. Our method achieved 95.8% of the best achievable growth and 9.6% more growth than passively floating despite its low propulsion of $u_{max}=0.1 \frac{m}{s}$ and relying on daily 5-day forecasts. This demonstrates that low-power propulsion is a promising method to operate autonomous seaweed farms in real-world conditions.

A future direction is to learn the expected growth after the forecast horizon using experience and approximate value iteration [48] or a value network [49]. This could implicitly learn the distribution shift between $\hat{v}_{FC}$ and v . Another direction is to make the discount factor state-dependent based on the uncertainty of current predictions, which could be estimated historically or from forecast ensembles [50]. Lastly, we can conduct field tests with Phykos [6] to further validate our method in real-world ocean conditions.

REFERENCES

- [1] B. M. Roque et al., "Effect of the macroalgae *Asparagopsis taxiformis* on methane production and rumen microbiome assemblage," *Animal Microbiome*, vol. 1, pp. 1–14, 2019.
- [2] W. T. L. Yong, V. Y. Thien, R. Rupert, and K. F. Rodrigues, "Seaweed: A potential climate change solution," *Renewable Sustain. Energy Rev.*, vol. 159, 2022, Art. no. 112222.
- [3] "Seaweed farming as climate solution," *Project Drawdown*. Accessed: Aug. 2025. [Online]. Available: <https://drawdown.org/solutions/seaweed-farming>
- [4] R. M. Tullberg, H. P. Nguyen, and C. M. Wang, "Review of the status and developments in seaweed farming infrastructure," *J. Mar. Sci. Eng.*, vol. 10, no. 10, 2022, Art. no. 1447.
- [5] F. Ross, P. Tarbuck, and P. I. Macreadie, "Seaweed afforestation at large-scales exclusively for carbon sequestration: Critical assessment of risks, viability and the state of knowledge," *Front. Mar. Sci.*, vol. 9, 2022, Art. no. 1015612.
- [6] "Floating seaweed farms," Phykos PBC. (n.d.). [Online]. Available: <https://www.phykos.com>
- [7] M. T. Sherman et al., "SeaweedPaddock: Initial modeling and design for a Sargassum ranch," in *Proc. OCEANS MTS/IEEE*, 2018, pp. 1–6.
- [8] M. M. Doshi, M. S. Bhabra, and P. F. J. Lermusiaux, "Energy-time optimal path planning in dynamic flows: Theory and schemes," *Comput. Methods Appl. Mech. Eng.*, vol. 405, Feb. 2023, Art. no. 115865.
- [9] M. Wiggert, M. Doshi, P. F. Lermusiaux, and C. J. Tomlin, "Navigating underactuated agents by hitchhiking forecast flows," in *Proc. IEEE Conf. Decis. Control*, 2022, pp. 2417–2424.
- [10] M. Doshi, M. Bhabra, M. Wiggert, C. J. Tomlin, and P. F. Lermusiaux, "Hamilton-Jacobi multi-time reachability," in *Proc. IEEE Conf. Decis. Control*, 2022, pp. 2443–2450.
- [11] A. Doering, M. Wiggert, H. Krasowski, M. Doshi, P. F. Lermusiaux, and C. J. Tomlin, "Stranding risk for underactuated vessels in complex ocean currents: Analysis and controllers," in *Proc. IEEE Conf. Decis. Control*, 2023, pp. 7055–7060.
- [12] N. Hoischen, M. Wiggert, and C. J. Tomlin, "Safe connectivity maintenance of underactuated multi-agent networks in dynamic oceanic environments," in *Proc. Eur. Control Conf.*, 2024, pp. 1922–1929.
- [13] E. Metzger et al., "Validation test report for the global ocean forecast system 3.5-1/25 degree HYCOM/CICE with tides," Naval Res. Lab, Washington, DC, USA, Tech. Rep. NRL/MR/7323-19-9975, 2020. [Online]. Available: <https://apps.dtic.mil/sti/pdfs/AD1113389.pdf>
- [14] E. P. Chassignet et al., "US GODAE: Global ocean prediction with the hybrid coordinate ocean model (HYCOM)," *Oceanography*, vol. 22, no. 2, pp. 64–75, 2009.
- [15] EU Copernicus Marine Service Information, "Global ocean 1/12 physics analysis and forecast updated daily product [data set]," Accessed: Aug. 2025. [Online]. Available: <https://doi.org/10.48670/moi-00016>
- [16] N. Pinardi, P. F. J. Lermusiaux, K. H. Brink, and R. Preller, "The sea: The science of ocean prediction," *J. Mar. Res.*, vol. 75, no. 3, pp. 101–102, 2017.
- [17] P. F. J. Lermusiaux et al., "A future for intelligent autonomous ocean observing systems," *J. Mar. Res.*, vol. 75, no. 6, pp. 765–813, Nov. 2017.
- [18] T. Lolla, M. P. Ueckermann, K. Yiğit, P. J. Haley, and P. F. Lermusiaux, "Path planning in time dependent flow fields using level set methods," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2012, pp. 166–173.
- [19] D. N. Subramani et al., "Time-optimal path planning: Real-time sea exercises," in *Proc. Oceans '17 Aberdeen*, Aberdeen, U.K., Jun. 2017, pp. 1–10.
- [20] W. Zhang, T. Inanc, S. Ober-Blobaum, and J. E. Marsden, "Optimal trajectory generation for a glider in time-varying 2D ocean flows B-spline model," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2008, pp. 1083–1088.
- [21] M. -C. Tsou and H. -C. Cheng, "An ant colony algorithm for efficient ship routing," *Polish Maritime Res.*, vol. 20, pp. 28–38, 2013.
- [22] V. T. Huynh, M. Dunbabin, and R. N. Smith, "Predictive motion planning for AUVs subject to strong time-varying currents and forecasting uncertainties," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2015, pp. 1144–1151.
- [23] D. Kularatne, S. Bhattacharya, and M. A. Hsieh, "Going with the flow: A graph based approach to optimal path planning in general flows," *Auton. Robots*, vol. 42, no. 7, pp. 1369–1387, 2018.
- [24] G. Mannarini and L. Carelli, "VISIR-I. B: Ocean surface gravity waves and currents for energy-efficient navigation," *Geoscientific Model Develop.*, vol. 12, no. 8, pp. 3449–3480, 2019.
- [25] D. Kularatne, H. Hajieghrary, and M. A. Hsieh, "Optimal path planning in time-varying flows with forecasting uncertainties," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 4857–4864.
- [26] D. N. Subramani and P. F. Lermusiaux, "Risk-optimal path planning in stochastic dynamic environments," *Comput. Methods Appl. Mech. Eng.*, vol. 353, pp. 391–415, 2019.
- [27] A. A. Pereira, J. Binney, G. Hollinger, and G. S. Sukhatme, "Risk-aware path planning for autonomous underwater vehicles using predictive ocean models," *J. Field Robot.*, vol. 30, pp. 741–762, 2013.
- [28] G. A. Hollinger, A. A. Pereira, J. Binney, T. Somers, and G. S. Sukhatme, "Learning uncertainty in ocean current predictions for safe and reliable navigation of underwater vehicles," *J. Field Robot.*, vol. 33, no. 1, pp. 47–66, 2016.
- [29] P. Gunnarson, I. Mandralis, G. Novati, P. Koumoutsakos, and J. O. Dabiri, "Learning efficient navigation in vortical flow fields," *Nature Commun.*, vol. 12, no. 1, pp. 1–7, 2021.
- [30] W. H. Al-Sabban, L. F. Gonzalez, and R. N. Smith, "Wind-energy based path planning for unmanned aerial vehicles using Markov decision processes," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2013, pp. 784–789.
- [31] M. G. Bellemare et al., "Autonomous navigation of stratospheric balloons using reinforcement learning," *Nature*, vol. 588, no. 7836, pp. 77–82, 2020.
- [32] D. Jones and G. A. Hollinger, "Planning energy-efficient trajectories in strong disturbances," *IEEE Robot. Autom. Lett.*, vol. 2, no. 4, pp. 2080–2087, Oct. 2017.
- [33] M. Otte, W. Silva, and E. Frew, "Any-time path-planning: Time-varying wind field moving obstacles," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2016, pp. 2575–2582.
- [34] M. S. Bhabra et al., "Optimal harvesting with autonomous tow vessels for offshore macroalgae farming," in *Proc. Glob. OCEANS*, 2020, pp. 1–10.
- [35] N. K. Yilmaz, C. Evangelinos, P. F. J. Lermusiaux, and N. M. Patrikalakis, "Path planning of autonomous underwater vehicles for adaptive sampling using mixed integer linear programming," *IEEE J. Ocean. Eng.*, vol. 33, no. 4, pp. 522–537, Oct. 2008.
- [36] N. K. Yilmaz et al., "Path planning methods for adaptive sampling of environmental and acoustical ocean fields," in *Proc. OCEANS MTS/IEEE*, 2006, pp. 1–6.
- [37] A. Alexander et al., "Loon library: Lessons from building Loon's stratospheric communications service," Loon LLC, Mountain View, CA, USA, 2021. [Online]. Available: <https://storage.googleapis.com/x-prod.appspot.com/files/The%20Loon%20Library.pdf>
- [38] B. Eysenbach, R. R. Salakhutdinov, and S. Levine, "Search on the replay buffer: Bridging planning and reinforcement learning," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, vol. 32. [Online]. Available: <https://proceedings.neurips.cc/paper/2019/hash/5c48ff18e0a47baaf81d8b8ea51ee92-Abstract.html>
- [39] S. Pateria, B. Subagdjia, A. -H. Tan, and C. Quek, "Hierarchical reinforcement learning: A comprehensive survey," *ACM Comput. Surv.*, vol. 54, no. 5, pp. 1–35, 2021.
- [40] K. Doya, "Reinforcement learning in continuous time and space," *Neural Comput.*, vol. 12, no. 1, pp. 219–245, 2000.
- [41] S. Bansal, M. Chen, S. Herbert, and C. J. Tomlin, "Hamilton-Jacobi reachability: A brief overview and recent advances," in *Proc. IEEE Int. Conf. Dec. Control*, 2017, pp. 2242–2253.
- [42] I. M. Mitchell and J. A. Templeton, "A toolbox of Hamilton-Jacobi solvers for analysis of nondeterministic continuous and hybrid systems," in *Proc. Hybrid Syst.: Computation Control*, 2005, pp. 480–494.
- [43] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "CasADi: A software framework for nonlinear optimization and optimal control," *Math. Program. Computation*, vol. 11, pp. 1–36, 2019.
- [44] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Math. Program.*, vol. 106, pp. 25–57, 2006.
- [45] J. Wu, D. P. Keller, and A. Oschlies, "Carbon dioxide removal via macroalgae open-ocean mariculture and sinking: An Earth system modeling study," *Earth Syst. Dyn.*, vol. 14, no. 1, pp. 185–221, 2022.
- [46] I. Martins and J. Marques, "A model for the growth of opportunistic macroalgae (*Enteromorpha* sp.) in tidal estuaries," *Estuarine, Coastal Shelf Sci.*, vol. 55, no. 2, pp. 247–257, 2002.
- [47] J. Zhang, W. Wu, J. Ren, and F. Lin, "A model for the growth of mariculture kelp *Saccharina japonica* in Sanggou Bay, China," *Aquaculture Environ. Interact.*, vol. 8, pp. 273–283, 2016.
- [48] L. Beckenbach and S. Streif, "Approximate infinite-horizon predictive control," in *Proc. 61st IEEE Int. Conf. Decis. Control*, 2022, pp. 3711–3717.
- [49] D. Silver et al., "Mastering the game of go with deep neural networks and tree search," *Nature*, vol. 529, pp. 484–503, 2016.
- [50] D. N. Subramani, Q. J. Wei, and P. F. J. Lermusiaux, "Stochastic time-optimal path-planning in uncertain, strong, and dynamic flows," *Comput. Methods Appl. Mech. Eng.*, vol. 333, pp. 218–237, 2018.