Spatiotemporal Optimization for Vertical Path Planning of an Ocean Current Turbine

Arezoo Hasankhani, *Student Member, IEEE,* Yufei Tang, *Member, IEEE,* James VanZwieten, and Cornel Sultan

***Abstract*—This paper presents a novel spatiotemporal optimization approach for vertical path planning (i.e., waypoint optimization) to maximize the net output power of an ocean current turbine (OCT) under uncertain ocean velocities. To determine the net power, OCT power generation from hydrokinetic energy and the power consumption for controlling the depth are modeled. The stochastic behavior of ocean velocities is a function of spatial and temporal parameters, which is modeled through a Gaussian process (GP) approach. Two different algorithms, including model predictive control (MPC) as a model-based method and reinforcement learning (RL) as a learning-based method, are applied to solve the formulated spatiotemporal optimization problem with constraints. Comparative studies show that the MPC- and RL-based methods are computationally feasible to address vertical path planning, which are evaluated with a baseline A^* approach. Analysis of the robustness is further carried out under the inaccurate ocean velocity predictions. Results verify the efficiency of the presented methods in finding the optimal path to maximize the total power of an OCT system, where the total harnessed energy after 200 hours shows over 18% increase compared to the case without optimization.**

***Index Terms*—Ocean current turbine, vertical path planning, spatiotemporal optimization, model predictive control, reinforcement learning**

I. Introduction

MARINE hydrokinetic (MHK) energy is considered as one of the most promising renewable energy resources [1]. For example, high potential of electricity production exists in ocean currents of the Gulf Stream within 200 miles of the US coastline from Florida to North Carolina (i.e., 163 TWh/year), mostly located in the coastal areas with high population densities [2]. The main hurdles in developing MHK-based energy are high investment and maintenance costs (e.g., due to hostile operating environment) and the difficulty of integrating their produced electricity into the grid. There are many approaches to address the high cost, such as lowering the cost of turbine design through control co-design [3], [4] or optimizing the operational strategy [5]. This study focuses on maximizing the net output power of an ocean current turbine (OCT) through optimal vertical path planning.

This work was supported in part by the U.S. National Science Foundation under Grant Nos. ECCS-1809164, ECCS-1809404, CMMI-2145571, & OAC-2017597 and the U.S. Department of Energy under Grant No. DE-EE0008955.

A. Hasankhani and Y. Tang are with the Department of Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL 33431 USA. E-mail: {ahasankhani2019, tangy}@fau.edu.

J. VanZwieten is with the Department of Civil, Environmental, and Geomatics Engineering, Florida Atlantic University, Boca Raton, FL 33431, USA. (e-mail: jvanzwi@fau.edu).

C. Sultan is with the Department of Aerospace and Ocean Engineering, Virginia Tech, 460 Old Turner St., Blacksburg, VA 24061, USA. (e-mail: csultan@vt.edu).

Path planning is studied for many other applications, such as unmanned vehicle trajectory optimization [6]–[9]; however, the challenge for OCT path planning in this paper is unique - the turbine is tethered to the seafloor through a mooring system and treated as an "autonomous underwater vehicle (AUV)" but with the primary role of energy generation. Given this primary role, it is critical that the OCT maintains an accurate spatiotemporal estimate of the ocean current profile and navigate itself at or near the water depth with the most intensive ocean flow. It is noteworthy to mention that the trajectory can be planned either offline or online depending on the application, while the spatiotemporal uncertainties in the ocean environment create a need for real-time iterative path optimization.

Among many approaches to address real-time vertical path planning, extremum seeking, model predictive control (MPC), and reinforcement learning (RL) are gaining increasing attention in recent years due to their success in both academia and industry. These approaches can be justified with a classical path planning algorithm, e.g., A^* algorithm, that is extensively studied for the autonomous vehicles' path planning, particularly the AUV [10], [11]. The extremum seeking approach is applied to deal with the path planning for similar applications to ours, such as airborne wind energy systems [12]–[14] and AUVs [15]–[18]. To address the path planning for autonomous vehicles in a dynamic environment, MPC-based approaches are applied to determine the optimal path and operation mode of vehicles [19]–[21]. A spatial-based path planning method for autonomous vehicles is introduced in [22], verifying autonomous driving in an obstacle-free environment and in the presence of obstacles. For a specific application of the airborne wind energy system, MPC-based techniques are employed to set the optimal altitude [23], [24], in which the future wind speeds are predicted and the output power is optimized by changing the location of the system to access the optimal velocities.

On the other hand, the path planning of autonomous vehicles is applied through RL approaches considering safety, security, and communication issues [25]. An RL-based path planning of mobile robots with obstacles is proposed to avoid collision and determine the optimal path through identifying environmental spatiotemporal data [26], [27]. To realize the minimum power consumption for plug-in hybrid electric vehicles, path planning and energy management are addressed through an RL-based approach [28]. However, due to the error induced by the prediction modeling, spatiotemporal uncertainties, etc. [29], the feasibility and robustness of these methods applied to OCT vertical path planning are not investigated. To incorporate the

uncertainties into the control planning frameworks, a Gaussian process (GP) approach can be employed, such as applying the GP for modeling probabilistic dynamics in RL [30], [31], as well as using the numerical GP approach to model the uncertainty associated with noisy data for partial differential equations [32].

A vast literature is devoted to addressing an independent task of path planning for AUV applications. For example, an estimator is designed to predict the spatially dependent ocean velocity, and trajectory planning is further developed to determine the location of an AUV applied as an energy harvester [33]. Trajectory planning of AUVs is addressed to avoid collisions in the presence of dynamic obstacles [34] and to reduce expected cost [35]. Transoceanic gliders treated as AUVs are studied, aiming to a decreased battery consumption on long-duration missions [36]. Bayesian optimization is used to specify the configuration and location of an OCT array to maximize harvested energy [37]; however, the research is focused on the economic side and a more detailed and realistic model is desired.

There exists limited literature on path planning for OCTs operating in a turbulent environment with the absence of human intervention (e.g., remote areas and deep-sea). A predictive methodology is nominated in the literature to design the planner in the presence of an approach to model environment uncertainties (e.g., the GP) and give a fair prediction from the environment, such as the predictive approach used in similar applications of airborne wind turbines [23] and underwater vehicles [38]. However, the error between the predicted ocean environment and real current conditions should not be neglected, which is the main motivation to move towards the approach relying more on the real measured data from the oceanic environment instead of predicted data. The learning-based approaches (e.g., the RL) can be a suitable option to deal with path planning in the stochastic environment but are seldom studied in the literature. Given the limitations of the predictive approach in the previous works [23], [38], the RL-based approach is presented to limit the error from the prediction by directly learning from real data recorded from the oceanic environment. Hence, a scientific gap is identified in that the MPC approach and its dependency on the predicted ocean environment are unknown, which motivates us to apply the RL-based method. The performance of the learning-based method should be then validated through qualitative studies and performing robustness analyses.

This paper focuses on the vertical path planning of a buoyancy-controlled OCT. An analytical and closed-form expression is developed to define the net output power model of the OCT system, which is missing in the literature, enabling the substantial expansion of the optimization problem. The OCT system is treated as an intelligent agent, navigating itself at or near the depth with the highest current velocity in the uncertain ocean environment without human intervention. In this regard, our work targets maximization of the output power of the OCT through real-time vertical path planning, which distinguishes us from the studies of maximum power point tracking control in the literature [39], [40]. Furthermore, since the OCT varies its vertical location at each time step, our path optimization problem is different from the tidal turbine control and optimization [41]. Two approaches (MPC and RL) are proposed to solve the spatiotemporal vertical path planning, which are assessed through the comparative results. Moreover, the robustness of these two approaches to the prediction error is analyzed.

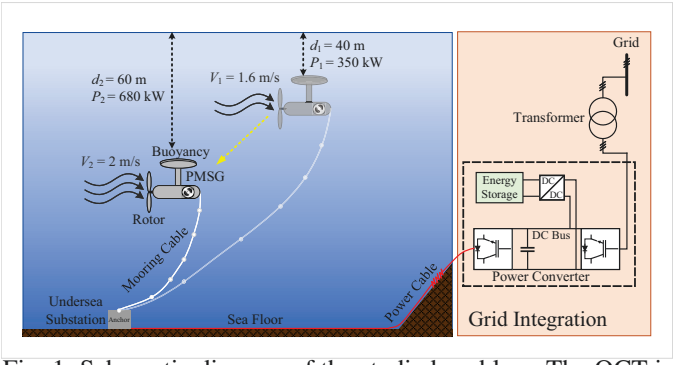


Fig. 1: Schematic diagram of the studied problem. The OCT is controlled spatiotemporally, where the optimal path is planned through changing the depth, e.g., move from $d_1 = -40$ m to $d_2 = -50$ m over one sampling time, resulting in a 300 kW power increase.

The contribution of this study is two-fold:

- This paper proposes a precise linear model to represent the net harvested power of a buoyancy-controlled OCT system, which includes directly generated power from ocean currents, consumed power for stabilizing the system at specified water depths, and consumed power to navigate to new optimal operating depths. The linear power model is derived based on the authors' previous work [42] on the nonlinear modeling of the OCT to lessen the computational complexity.
- This paper further formulates a novel spatiotemporal optimization problem for the OCT vertical path planning operating in an uncertain oceanic environment to maximize the total harvested power of the system. Under this problem formulation, an RL-based method is designed to explore the optimal control actions, and results are compared with an MPC-based strategy and a baseline A* approach.

The rest of this paper is organized as follows: Section II formulates the spatiotemporal problem. Section III describes the modeling of the ocean velocity as a function of time, location, and the power output model of an OCT system. Section IV presents our proposed solution methodology based on MPC and RL. Section V presents simulation results and discussions. Finally, section VI draws conclusions and perspectives for future research.

II. Problem Formulation

The schematic diagram of the studied OCT vertical path planning problem is shown in Fig. 1. In this paper, the ultimate objective is to maximize the OCT generated power through a spatiotemporal optimization approach. Based on the predicted ocean velocity, the optimal path should be determined at each sampling time.

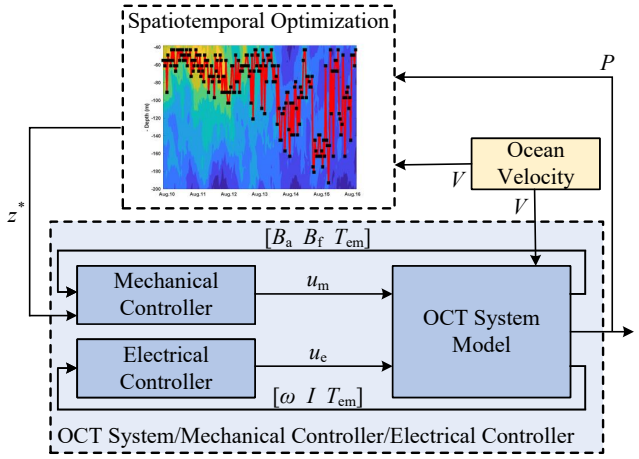


Fig. 2: Proposed spatiotemporal optimization for OCT vertical path planning. The system is controlled hierarchically, where the upper-level is designed to enable spatiotemporal optimization, and the lower-level is a dynamic tracking control. This paper focuses on the spatiotemporal optimization.

To address the complexity of this problem, a novel hierarchical spatiotemporal optimization and control framework is proposed, as shown in Fig. 2. The ocean velocity is first modeled with the GP method, and the forecasted velocity is used to calculate the output power of the OCT system. The OCT output power P is received by the spatiotemporal optimization at each time step in the upper-level, while the optimal water depth z^* as a set-point is determined accordingly for the lower-level controllers. The lower-level control tracks the prescribed set-points from the upper-level and adjust the turbine dynamics.

We focus on the upper-level spatiotemporal optimization in this paper, assuming that the lower-level control exists and follows the optimal path found through the spatiotemporal optimization. The major goal of the proposed approach is to plan the vertical path that maximizes the harnessed power from the OCT (i.e., minimize the difference between the rated power as the maximum output power of the system and the real output power), subject to the operational constraints of the OCT system and environmental model. The spatiotemporal optimization problem and constraints are formulated as follows:

$$\mathcal{OP}: \mathbf{u}^*(p) = \underset{\mathbf{u}(p)}{\arg\min} \sum_{i=p}^{p+T-1} [P_{\text{r}} - E(P_{\text{net}}(u(i|p), v_*(u(i|p), i|p))] \tag{1a}$$

subject to

$$v_*(u(i|p), i|p) \sim f(z,t) \tag{1b}$$

$$u^{\min} \leq u(i|p) \leq u^{\max} \tag{1c}$$

$$\dot{u}(i|p) \leq \dot{u}^{\max} \tag{1d}$$

where $\mathbf{u}(p) \triangleq [u(p|p), ..., u(p+T-1|p)]$ denotes the decision variables vector, the optimal finite-horizon decision variables sequence $\mathbf{u}^*(p) \triangleq [u^*(p|p), ..., u^*(p+T-1|p)]$ is computed as the solution of the optimization problem. The predicted velocity is denoted by $v_*(u(i|p), i|p)$, the rated power is denoted by P_{r}, the net output power is denoted by P_{net}, the expected power is denoted by $E(P_{\text{net}}(.))$, the minimum and maximum allowable decision variables are denoted by $u^{\min}$ and $u^{\max}$, and the maximum limit on the slew rate is denoted by $\dot{u}^{\max}$. Here, constraint (1b) shows a model to predict the ocean velocity over the prediction horizon; also, constraints (1c) and (1d) limit decision variables and the corresponding slew rate. Note that the decision variable is defined as the operating depth of the OCT system z in our spatiotemporal optimization problem.

To be brief, at the lower-level, mechanical and electrical controllers are considered. The control inputs of the mechanical controller (i.e., flight controller) are the two ballast tank fill fractions (B_{f} is the fill fraction of the forward tank and B_{a} is the fill fraction of the aft tank) and the electromechanical torque T_{em}. The control inputs of the electrical controller (i.e., generator controller) are the electromechanical torque T_{em}, generator current I, and rotor angular speed ω. We develop controllers for the lower-level with a full turbine dynamics for a similar OCT system [39], [43]–[45]. To better understand the dynamic and linear model of the OCT system, we briefly review the OCT model here, and the readers are referred to the authors' previous works [42], [45] for details.

The investigated OCT system is represented by 14 states $x = [u\ v\ w\ p_{\text{b}}\ p_{\text{r}}\ q\ r\ x\ y\ z\ \phi_{\text{b}}\ \phi_{\text{r}}\ \theta\ \psi]$, consisting of the linear velocity of the OCT body $[u\ v\ w]$, the angular velocity of the OCT body $[p_b\ q\ r]$, the position of the OCT body $[x\ y\ z]$, the Euler angles of the OCT body $[\phi_{\text{b}}\ \theta\ \psi]$, and the angular velocity and rotation angle of the rotor $[p_{\text{r}}\ \phi_{\text{r}}]$. This system is described with seven degree-of-freedom (DOF) equations of motion as discussed in [46], including six-DOF motion of the OCT body and one-DOF rotation of the rotor about the x-axis of the body frame. The flight controller of our previous work is designed subject to the linear model of the OCT system with the main objective of following the optimal path with a minimized error. To formulate the linear model of the OCT system, the system response is linearized about the equilibrium point specified for a specific flow speed and desired operating depth with the corresponding controller inputs. During this process, the rotor rotation angle state ϕ_r is eliminated by averaging the effects of ϕ_r over one rotor rotation. As mentioned earlier, our main focus is on the upper-level spatiotemporal optimization assuming that the flight controller is able to follow the optimal depth. A feasible vertical path for the OCT system is ensured by introducing the primary constraints of allowable operating depth and limit on the slew rate to the optimization problem.

III. Environment and OCT Modeling

The uncertainties in the ocean current velocity field are addressed, and the future ocean velocity is predicted that is further used in the spatiotemporal optimization. The net generated power of an OCT system, which is a function of the ocean velocity, is then modeled in detail.

A. Statistical Ocean Current Shear Profile Characterization

The ocean velocity field varies with time in a 3D space. However, moored OCTs can primarily vary their vertical lo-

cation (i.e., depth), and current shear is most prominent in the vertical direction. Therefore, the ocean velocity's dependence on both time and water depth is of primary importance; in other words, the ocean velocity should be determined at specific times t and operating depths z. The observed (recorded) ocean velocity data are represented as:

$$\mathbf{V_{n\times m}} = \begin{bmatrix} v(z_1,t_1) & \dots & v(z_1,t_m) \\ \vdots & \ddots & \\ v(z_n,t_1) & \dots & v(z_n,t_m) \end{bmatrix} \tag{2}$$

where n and m are the numbers of the discrete depths and the time samples, and $v(z_i,t_j)$ is the recorded ocean velocity at z_i and t_j. Matrix $\mathbf{V}$ includes $n\times m$ recorded ocean velocities.

Let $\mathcal{Z}$ denote the domain of allowable ocean depth choices, and $\mathcal{T}$ represent the temporal space. The goal here is to model ocean current velocity at depth $z\in\mathcal{Z}$ and time $t\in\mathcal{T}$, by constructing a function $f:\mathcal{Z}\times\mathcal{T}\rightarrow\mathcal{R}$ whose output is a predicted velocity. The ocean velocity and especially the ocean turbulence are usually modeled with probability distributions, including log-normal distribution [47] and Burr distribution [48]. In this paper, different methods for modeling the velocity are developed and compared, as shown in Table I. Linear regression, regression trees, support vector machines, and the GP are compared and quantified using root mean square error (RMSE), mean square error (MSE), and mean absolute error (MAE), by assuming 70% of the recorded velocities as a training set, 15% for validation, and 15% for testing. The GP modeling shows the best performance and is briefly introduced in the following.

Gaussian Process Model: The GP is a probabilistic approach used to define a prior probability distributions over latent functions directly, which is extensively applied in wind speed forecasting [23], [49], [50] and ocean current flow velocity prediction [51], [52]. To predict the ocean current velocity, the GP model can be addressed through either a pure learning approach by finding the parameters for the Gaussian predictive distribution [51] or combining the laws of physics (e.g., nonlinear dynamics described by the partial differential equations) with the learning approach and propagating the uncertainty through time with the time-stepping method [32]. We proceed with the first approach when constructing the GP model through learning from the observed velocity, while the latter is beyond the scope of our study and calls for a dynamic nonlinear model to describe the ocean current velocity. Another potential approach for modeling the ocean current flow velocity is to develop a physics-based deep neural network to learn the spatiotemporal features of the data [53], but this is not the focus of the current study.

To approach the real-world application and given that an ADCP will likely be installed just upstream from a turbine, an online learning approach that assumes water measurements velocities are available throughout the water column is used to predict the ocean current velocity. The aforementioned online approach deals with a finite size of input data to avoid the scalability issue as suggested in [51]. It is worth noting that the scalability of the GP model is well studied in the literature, proposing several precautions to avoid this issue, e.g., a sparse representation to recursively update the GP [54] or using a stochastic variational inference approach [55]. The computational complexity is intensified for our application due to its nature for a long-term and persistent operation when the size of input data increases. Hence, the prediction is generated according to a fixed-size rolling window of the observed ocean current velocity, which is accordingly rolled to complete the prediction over the prediction horizon, thereby removing the old data by the arrival of a new observation in the online application to maintain a fixed-size window. Since the accuracy degrades for predicting distant data when proceeding with a fixed GP model, the GP model is periodically regenerated.

TABLE I: Comparing different methods for modeling ocean velocities. Results are quantified by RMSE, MSE, and MAE.

Algorithms	RMSE	MSE	MAE
Linear regression	0.0369	0.0014	0.0297
Regression trees	0.0614	0.0038	0.0302
Support vector machines	0.0232	0.0005	0.0205
Gaussian process model	0.0031	9.6×10^{-6}	0.0022

Define $\mathbf{x}$ as the input set of the recorded ocean velocity associated with the target ocean velocity of v. The GP model with mean $m(\mathbf{x})$ (i.e., encodes the central tendency) and covariance $k(\mathbf{x},\mathbf{x}^{'})$ (i.e., denotes the shape and structure between any two input sets of $\mathbf{x}$ and $\mathbf{x}^{'}$) is defined as:

$$f(\mathbf{x})\sim\mathcal{GP}(m(\mathbf{x}),k(\mathbf{x},\mathbf{x}^{'})) \tag{3}$$

The ocean velocity is predicted as $v=f(\mathbf{x})+\epsilon$, while ϵ denotes a Gaussian distribution $\mathcal{N}(0,\sigma^2)$. The joint distribution over the input set $\mathbf{x}$ and a prediction of the target ocean velocity v_* is defined as [56]:

$$\begin{bmatrix}\mathbf{v}\\ v_*\end{bmatrix} = (\begin{bmatrix}f\\ f_*\end{bmatrix}+\begin{bmatrix}\epsilon\\ \epsilon_*\end{bmatrix})\sim\mathcal{N}(0,\begin{bmatrix}K_v & \mathbf{k}_*\\ \mathbf{k}_*^{\mathrm{T}} & k_{**}+\sigma^2\end{bmatrix}) \tag{4}$$

where $f_*\triangleq f(\mathbf{x}_*)$ shows the latent function based on a new input vector $\mathbf{x}_*$ with corresponding noise ϵ_*. $\mathbf{v}$ denotes an output set of the ocean velocity. Define $K_v=k(\mathbf{x},\mathbf{x})$; $\mathbf{k}_*$ and k_{**} are calculated as:

$$\mathbf{k}_*=[k(\mathbf{x}_*,\mathbf{x}_1),k(\mathbf{x}_*,\mathbf{x}_2),...,k(\mathbf{x}_*,\mathbf{x}_m)] \tag{5}$$

$$k_{**}=k(\mathbf{x}_*,\mathbf{x}_*) \tag{6}$$

To represent the GP modeling, the mean $m(\mathbf{x}_*)$ and covariance $\sigma^2(\mathbf{x}_*)$ are defined as follows:

$$m_*\triangleq m(\mathbf{x}_*)=\mathbf{k}_*^{\mathrm{T}}K_v^{-1}\mathbf{v} \tag{7}$$

$$\sigma_*^2\triangleq\sigma^2(\mathbf{x}_*)=k_{**}-\mathbf{k}_*^{\mathrm{T}}K_v^{-1}\mathbf{k}_*+\sigma^2 \tag{8}$$

Finally, the GP model for predicting new ocean velocities is determined using mean and covariance in (7) and (8):

$$f_*\sim\mathcal{GP}(m_*,\sigma_*^2) \tag{9}$$

B. Mathematical Model of OCT Output Power

The output power of an OCT with variable blade pitch is investigated numerically in [57], and we use these rotor performance characteristics in this study. These model parameters are for a 700 kW OCT with a single 20 m diameter

variable pitch rotor. This OCT model is extended to include two variable ballast tanks, and a 607 m long mooring cable that attaches to the seafloor at a depth of 325 m [42]. These model parameters roughly follow the prototype systems from IHI Corp. [58], [59] and the University of Naples [60], but with a single rotor and variable ballast tanks sized to operate at a depth of 50 m when half filled with ballast water at the Gulf Stream's mean flow speed off Southeast Florida of 1.6 $\mathrm{m/s}$. The nonlinear OCT modeling techniques presented in [42], [57] are utilized when simulating this system for the creation of the linear models used in the presented formulations. The output power of the system consists of three primary parts:

1) Generated power from hydrokinetic energy extraction, denoted as P_{OCT};
2) Power consumed by ballast pumps to Hold Depth, denoted as $P_{\text{ballast}}^{\text{HD}}$; and
3) Power consumed by ballast pumps to Change Depth, denoted as $P_{\text{ballast}}^{\text{CD}}$.

The total harvested power from the OCT system is calculated by (10):

$$P_{\text{net}} = P_{\text{OCT}} - P_{\text{ballast}}^{\text{HD}} - P_{\text{ballast}}^{\text{CD}} \tag{10}$$

where P_{net} is the total harvested power from the OCT system, P_{OCT}, $P_{\text{ballast}}^{\text{HD}}$, and $P_{\text{ballast}}^{\text{CD}}$ denote three parts of the harvested power.

First term, P_{OCT}: The generated power of OCT system is related to ocean velocity according to [61]:

$$P_{\text{OCT}} = \frac{1}{2}\rho A C_{\text{p}} v^3 \tag{11}$$

where ρ is the water density, A is the swept area of the OCT rotor, C_{p} is the average power coefficient, and v is the magnitude of the ocean velocity.

Second term, $P_{\text{ballast}}^{\text{HD}}$: To calculate $P_{\text{ballast}}^{\text{HD}}$, the average consumed power for maintaining a near constant depth in a time-varying current is determined. This model assumes that ballast tank water fill levels, which are defined as a fraction of ballast capacity F_{F}, are adjusted every Δt_1 to counteract changes in the flow velocity Δv and maintain the desired operating depth z. It should be noted that changes in flow velocity impact the mooring cable force (i.e., downward force in the OCT), resulting in an OCT elevation change unless counteracted by an equal and opposite change in the buoyancy force. Hence, the ballast levels must be changed to hold a constant depth when the flow velocities change.

For these adjustments, the model assumes that a pump drives water through an opening such that the pressure in the tank is at vacuum pressure (i.e., $P_{\text{abs}} \cong 0\,\text{kPa}$). Using this approach, very little power is used when the tanks are being filled with water since this can be driven by the natural pressure difference between ambient pressure and vacuum pressure. The power required to pump sea water out of the tank can be calculated from the product of the force F (i.e., the product of pressure and area, $\Delta P A$) and velocity V (i.e., the quotient of volumetric flow rate over the area, Q_{B}/A) through the orifice divided by pump efficiency. Accordingly, the power to fill the tank $P_{\text{B}}^{\text{fill}}$ and the power to empty the tank $P_{\text{B}}^{\text{empty}}$ are defined as follows:

$$P_{\text{B}}^{\text{fill}} = 0 \tag{12}$$

$$P_{\text{B}}^{\text{empty}} = \frac{FV}{\eta_{\text{pump}}} = \frac{\Delta P Q_{\text{B}}}{\eta_{\text{pump}}} \tag{13}$$

where η_{pump} denotes the pump efficiency. Given that $\Delta P = P_{\text{atm}} + P_{\text{HS}}$, the power to empty the tank defined in (13) can be rewritten as:

$$P_{\text{B}}^{\text{empty}} = \frac{(P_{\text{atm}} + P_{\text{HS}})Q_{\text{B}}}{\eta_{\text{pump}}} \tag{14}$$

where P_{atm} is atmospheric pressure, $P_{\text{HS}} = \rho.g.z$ is hydrostatic pressure, and g is gravity.

Assuming a constant value for Q_{B}, two ballast tanks with volume of ν_{B} can be completely emptied of water in $\Delta t_{\text{e}} = \frac{Q_{\text{B}}}{\nu_{\text{B}}}$. Hence, the energy required to empty either ballast tank $E_{\text{tank}}^{\text{empty}}$ and fill either ballast tank $E_{\text{tank}}^{\text{fill}}$ at an operating depth of z_{opt} are calculated by:

$$E_{\text{tank}}^{\text{fill}} = 0 \tag{15}$$

$$E_{\text{tank}}^{\text{empty}} = \frac{P_{\text{B}}^{\text{empty}}}{\Delta t_{\text{e}}} \tag{16}$$

To maintain OCT depth, ballast tank fill levels are changed every Δt_1 by a fraction of their total fill levels ΔF_{F} to counteract the changes in flow velocity Δv. These changes in ballast fill levels occur more frequently than the changes in desired depth calculated by the spatiotemporal optimization algorithm, with the associated fill level changes calculated using linear estimates of the relationships between fill level changes and equilibrium depth changes $\frac{dF_{\text{F}}}{dz}$ as well as between flow speed changes and equilibrium depth changes $\frac{dv}{dz}$. Assuming linearity, the following relationship exists between flow speed changes and ballast level changes necessary to maintain a constant depth:

$$\frac{\Delta F_{\text{F}}}{\Delta v} = \frac{dF_{\text{F}}}{dz}\frac{dz}{dv} \Rightarrow \Delta F_{\text{F}} = \frac{dF_{\text{F}}}{dz}\frac{dz}{dv}\Delta v = \zeta \Delta v \tag{17}$$

where $\zeta = \frac{dF_{\text{F}}}{dz}\frac{dz}{dv}$ denotes a constant coefficient to relate the flow speed changes and ballast level changes. It should be noted that in this paper, a quasi-static relationship is assumed between the states and a steady and homogeneous flow field when running the nonlinear simulation [42].

The average power consumed over each Δt_1 to maintain the operating depth can be calculated using (17):

$$P_{\text{ballast}}^{\text{HD}} = \begin{cases} 0, & \text{if } \Delta v < 0 \\ \frac{(E_{\text{B}}^{\text{empty}})(\zeta \Delta v)}{\Delta t_1}, & \text{if } \Delta v > 0 \end{cases} \tag{18}$$

Third term, $P_{\text{ballast}}^{\text{CD}}$: To calculate $P_{\text{ballast}}^{\text{CD}}$, the ballast model and its average consumed power for changing depth can be determined using many of the assumptions and models introduced for the second term of this formulation. To change OCT depth, ballast tank fill levels are changed every time step Δt_2 by a fraction of their total fill levels ΔF_{F}. This change in fill level is based on the desired change in depth Δz using the linear relationship between Δz and ΔF_{F}:

$$\Delta F_{\text{F}} = \frac{dF_{\text{F}}}{dz}\Delta z = \kappa \Delta z \tag{19}$$

Algorithm 1 Upper-level Spatiotemporal Optimization

```
Initialize n data sample v(z,t) and optimization method parameters
for Each t do
    Construct the GP model (f_*) based on a fixed-size window of the observed data;
    for Prediction Horizon T do
        Predict v_* over T using the GP model f_* ~ GP(m_*, σ_*^2);
    end for
    Apply optimization algorithm (e.g., MPC or RL);
    Output optimal depth z^*;
end for
```

where $\kappa = \frac{dF_{\mathrm{F}}}{dz}$ denotes a constant coefficient to relate the fill level changes and depth changes.

The average consumed power to satisfy Δz depth change is calculated with respect to the energy required to fill the ballast tank (15), the energy required to empty the ballast tank (16), and the relation between fill fraction change and depth change (19) as follows:

$$P_{\text{ballast}}^{\text{CD}} = \begin{cases} 0, & \text{if } \Delta z > 0 \\ \frac{(E_{\text{B}}^{\text{empty}})(\kappa \Delta z)}{\Delta t_2}, & \text{if } \Delta z < 0 \end{cases} \tag{20}$$

Therefore, the total harvested power, which is defined in (10) is calculated based on the expressions found for each term of output power (11), (18), and (20).

IV. Proposed Methodology

In this paper, we focus on the upper-level spatiotemporal optimization shown in Fig. 2. The pseudocode for this process is presented in Algorithm 1. Two approaches, model predictive control (MPC) as a model-based approach and reinforcement learning (RL) as a learning-based approach, are considered in this paper. Also, the obtained path planning results are justified with a baseline A* approach.

A. Baseline Approach: A*-based Vertical Path Planning

The A* algorithm is utilized as a search engine to find the optimum vertical path that maximizes harnessed power, where the vertical positions are shown as a discretized vector $\mathcal{Z} = [z_1, ..., z_n]$. Given a start depth $s \in \mathcal{Z}$ and the predicted ocean current velocities v_* over the $\mathcal{Z}$, the cost of transition (vertical movement) from the start depth to any depths is defined by $P_{\text{r}} - E(P_{\text{net}}(z(i), v_*(z(i), i)))$, with i denoting the time. The A* approach seeks feasible vertical positions (illustrated as a so-called open list in the A* algorithm), considering that the optimal vertical position at each time step justifies the minimum cost.

B. MPC-based Vertical Path Planning

MPC is considered in this paper because of its capability to handle constrained problems. MPC is a powerful method for optimizing some objective functions by using a model of the system to be controlled to predict future states and actions. In this paper, the objective is to maximize the output power of the OCT system. To fully embrace the uncertainties in the ocean velocity, the objective function is defined with two terms interpreting the exploration and exploitation.

The exploitation term in the objective function is determined to find the optimal water depth where net power is maximized. In fact, different water depths $\mathbf{z}$ are explored to find the optimal one over the prediction horizon. At the same time, the uncertainties in ocean velocity prediction should also be considered in the objective function, so the exploration term is included in the objective function to penalize the predicted velocity with higher variance. The objective function is formulated as a nonconvex optimization problem, which interprets the real power maximization objective of the OCT system with respect to the spatiotemporal ocean current velocity model, still facing nonlinearity and computational complexity. The constrained optimization problem which has to be solved by the MPC design is:

$$\mathbf{z}^*(p) = \underset{\mathbf{z}(p)}{\arg\min} \sum_{i=p}^{p+T-1} \left[J_{\text{exploit}}(z(i|p)) + \beta J_{\text{explore}}(z(i|p)) \right] \tag{21a}$$

subject to

$$v_*(z(i|p), i|p) \sim f_* \sim \mathcal{GP}(m_*, \sigma_*^2) \tag{21b}$$

$$z^{\min} \leq z(i|p) \leq z^{\max} \tag{21c}$$

$$\frac{z(i|p) - z(i-1|p)}{\Delta t_2} \leq r \tag{21d}$$

where p denotes the p-th sampling time, T represents the prediction horizon. Define $\mathbf{z}(p) \triangleq [z(p|p), ..., z(p+T-1|p)]$, $\mathbf{z}^*(p) \triangleq [z^*(p|p), ..., z^*(p+T-1|p)]$. $v_*(z(i|p), i|p)$ denotes the predicted velocity. Constraint (21b) shows the GP model of ocean velocity to generate the ocean velocity prediction over the prediction horizon with mean m_* and covariance σ_*^2 (as explained in Section III-A), r denotes the rate limitation on the speed, β is the gain of the exploration term, and $z^{\min}$ and $z^{\max}$ are the minimum and maximum allowable depth. Note that the current optimal depth is chosen as the first element of the optimal decision variable sequence $\mathbf{z}^*(p)$.

The exploitation term J_{exploit} is defined as:

$$J_{\text{exploit}}(z(i|p)) = P_{\text{r}} - E(P_{\text{net}}(z(i|p), v_*(z(i|p), i|p))) \tag{22}$$

where $E(P_{\text{net}}(z(i|p), v_*(z(i|p), i|p)))$ is the expected power calculated by (10) based on the predicted ocean velocities estimated through GP modeling as shown in (21b), and P_{r} is the rated power imposed by the OCT model and its generator rated power. The main objective for this term is to reach the rated power at each time step, which is defined as minimizing the difference between P_{r} and the expected power.

The exploration term J_{explore} is defined as the standard deviation of the predicted ocean velocity, conditioned upon previously recorded ocean velocities, representing the uncertainties in the ocean velocity:

$$J_{\text{explore}}(z(i|p)) = \sum_{\mathbf{z}} \sigma^2(\mathbf{v}_*(\mathbf{z}, i|p) | \mathbf{V}, \mathbf{V}') \tag{23}$$

Algorithm 2 Proposed MPC-based Design

```
Initialize n data sample v(z,t), i = 0, Δt1, Δt2, and T
for Each t do
    Construct the GP model (f*) based on a fixed-size window of the observed data;
    for Prediction Horizon T do
        Predict v* over T using the GP model f* ~ GP(m*, σ*²);
    end for
    Solve (21a) and obtain optimal depth trajectory z*(p);
    Output optimal depth z*;
end for
```

where $\sigma^2(\mathbf{v}_*)$ shows the variance of predicted ocean velocities for all operating water depths $\mathbf{z}$, $\mathbf{V}$ is the recorded velocity at the operating depth and time defined in (2), and $\mathbf{V}^{'}$ denotes all new velocity measurements over the future horizon between p taken up to step i.

Inspired by [23], the exploration term is included in the optimization problem to ensure a continued exploration in the feasible spatial domain. The exploration term evaluates the uncertainty in the ocean current velocity, and the main intuition behind the exploration term is to penalize the increased uncertainty in the ocean current velocity. Note that the major challenge for (23) is to find an appropriate weight β chosen through experimental testing. Fine-tuning through trial and error is used to determine the exploration weight to balance the under exploration (i.e., small β) along with the over exploration (i.e., large β).

The algorithm of OCT path optimization using MPC is presented in Algorithm 2. Firstly, all recorded ocean velocities and the prediction horizon T should be initialized. Then, v_* is predicted for each sampling time using (21b). The objective function is calculated over the prediction horizon while optimal depth is determined. Note that a sliding window is applied here, and only the first element from the optimal depth trajectory will be used in each sampling time. To solve the MPC-based optimization problem, we use dynamic programming (DP) by forward recursion, and to maintain the computational tractability, the optimization problem is defined over the depth-time grid (Fig. 3). It should be noted that by using the DP, convergence to the global optimum is guaranteed subject to the ocean environment grid resolution.

C. RL-based Vertical Path Planning

The RL is adopted here for its capability of learning a policy from historical data (i.e., data-driven), which could be robust to the environment model errors. The RL method takes the perspective of an agent (i.e., an OCT system in this study) that optimizes its behavior by interacting with the environment and learning from the feedback received. In the RL approach, the set of actions is done by the agent, and it receives the reward from the environment. Therefore, the learning procedure is completed for the agent from observing its resulted reward.

It is critical to define the *set of states,* $s_i \in \mathcal{S}$, *actions,* $a_i \in \mathcal{A}$ and *rewards,* $r_i \in \mathcal{R}$. The long-term performance is optimized by learning a *policy* $\pi_\theta(a_i|s_i)$ for picking actions in state transition to maximize the total accumulated reward $R_T^\pi = \sum_{\tau=0}^{T} \gamma^\tau r_{i+\tau}$, where γ is the discount factor in $(0, 1]$. As a reminder, the action of depth change at each time step should be determined to maximize the OCT net power.

1) State space: The net power of the OCT system is calculated in (10) as a function of the water depth z. We realize that the optimization gets more complicated when the prediction horizon T is considered. Transitions between different water depths $\mathbf{z}$ over the prediction horizon result in different net power. Therefore, the set of states is defined as:

$$S = \{s_i|((z,t),P)\} \tag{24}$$

in which, the number of states is equal to n^T if the number of operating depths is n, and the prediction horizon is T. Hence, increasing the prediction horizon can increase the complexity of the problem exponentially, so it is important to choose a reasonable prediction horizon to avoid the curse of dimensionality.

2) Action space: The set of action defines the possible operation at each state. As the OCT is described with a single state (i.e., $((z,t),P)$), the depth change modifies the state of the system. Therefore, the action is defined as the water depth change, resulting in the power change. The action space is defined as follows:

$$A = \{a_i| + \Delta z, -\Delta z, 0\} \tag{25}$$

where $+\Delta z$ shows the depth increase, $-\Delta z$ is the depth decrease, and 0 determines no change in the water depth.

3) Reward function: The reward function determines the reward received after each action. Further, the reward function should be defined to find the desirable actions. In our problem, increasing the net OCT power is considered as the desirable goal. Hence, any action (water depth change) should be done to increase the net power at each time step. The reward function is defined as:

$$R = \begin{cases} \Delta P, & \Delta P > \delta \text{ or } \Delta P < -\delta \\ 0, & \text{otherwise} \end{cases} \tag{26}$$

where ΔP is the power change, and δ is a small positive number rather than 0. Hence, if ΔP is positive, the power is increasing, and the action is rewarded with positive numbers.

Q-learning: Q-learning algorithm [62] is used to solve our RL problem. Q-value $Q(s,a)$ denotes the expected return of taking action a in state s, and the Q-value table is updated as follows:

$$\begin{aligned} Q_{t+1}(s_t,a_t) \leftarrow Q(s_t,a_t) + \\ \alpha[R_{t+1} + \gamma \max_a Q(s_{t+1},a) - Q(s_t,a_t)] \end{aligned} \tag{27}$$

where s_t is the state visited at t, R_{t+1} denotes the observed reward at $t+1$, and $\alpha \in (0,1]$ is the learning rate.

For action selection from the Q-value table, ϵ-greedy policy is used. Two different options are available for selecting the next action at each time step, including choosing the actions with the highest estimated reward or choosing a random action:

$$a = \begin{cases} \arg\max_a Q, & \text{with probability 1-}\epsilon \\ \text{random value}, & \text{with probability } \epsilon \end{cases} \tag{28}$$

Algorithm 3 Proposed RL-based Design Offline Training

Initialize n training data sample $v(z,t)$, Δt_1, Δt_2, and $Q(s,a)$ for all $s \in S$ and $a \in A$
for N sampling time **do**
  Input recorded velocity $v(z,t)$;
  for episode $= 1, 2, \ldots, N_{\text{episode}}$ **do**
    Initialize s $((z,t), P)$;
    for step$= 1, 2, \ldots, N_{\text{step}}$ **do**
      Select action a according to ϵ-greedy (28);
      Take action a and obtain R_{t+1} by (26);
      Update the Q-value according to (27);
    end for
  end for
end for
Output offline trained optimal Q-value table Q^*;

Algorithm 4 Proposed RL-based Design Online Testing with Offline Learning

Initialize Δt_1, Δt_2, and current time t
Load optimal Q-value table Q^* obtained through an offline training;
for Each t **do**
  Select action a according to Q^*;
  Output optimal depth z^*;
end for

where ϵ is usually set as a small value, such as 0.05.

In our spatiotemporal optimization problem, the ocean environment is modeled as a grid capturing the spatial, temporal, and ocean current velocity features. The main objective in vertical path planning is to find the optimal vertical path (i.e., the depth) with the maximized net power over the prediction horizon T; hence, we should visit all feasible vertical paths to find the globally optimal path. This grid and the feasible vertical paths are visualized in Fig. 3. As shown in this figure, the ocean environment has n^T feasible vertical paths assuming n discrete depths. The information of these vertical paths is transferred to the Q-value table through a Q-vector of size n^T.

The Q-learning offline training is presented in Algorithm 3. An action a (depth change) is selected based on the ϵ-greedy policy at each time step. After taking action, a new state (a new water depth) will be observed, and the reward will be calculated by (26). Finally, the Q-function (27) is updated based on the new state and the reward. Note that after the offline training on the training dataset, the obtained optimal Q-value table can be applied to find the optimal vertical path as presented in Algorithm 4. On the other hand, the attained optimal Q-value table Q^* with the offline training can be deployed online while it keeps updating according to the observed ocean velocity through an incremental learning strategy (see Algorithm 5). The concept of incremental learning is also applied for the autonomous navigation of mobile robots [63].

To enable the incremental learning, the optimal Q-value table is initially determined through the offline training, which is then slightly updated through the online training. Since the shear profile of the ocean current does not show an aggressive

Algorithm 5 Proposed RL-based Design Online Testing with Incremental Learning

Initialize Δt_1, Δt_2, and current time t $(t \geq 1)$
Load optimal Q-value table Q^* obtained through an offline training;
for Each t **do**
  if $t = 1$ **then**
    Select action a according to Q^*;
  else
    Construct the GP model (f_*) based on a fixed-size window of the observed data;
    for Prediction Horizon T **do**
      Predict v_* over T using the GP model $f_* \sim \mathcal{GP}(m_*, \sigma_*^2)$;
    end for
    Update optimal Q-value table Q^* according to the predicted velocity over T;
    Select action a according to the updated Q^*;
  end if
  Output optimal depth z^*;
end for

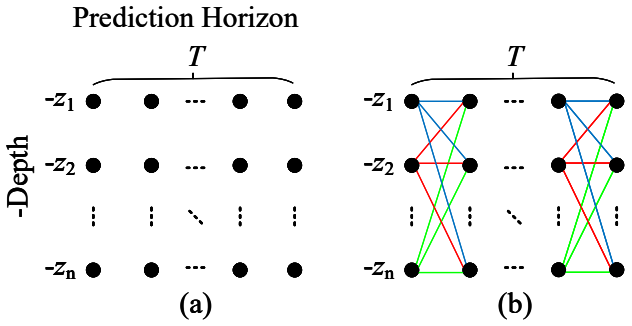


Fig. 3: Schematic of the ocean environment. (a) Ocean environment grid; (b) Feasible vertical paths.

change outside of conditions seen during hurricanes [42], which is beyond our study, the obtained Q-value table after training on a large dataset is moderately affected by a set of the observed ocean velocity, thereby avoiding any instability and aggressive control policy. The Q-value table obtained by the offline training is sufficiently precise to find the optimal action; however, updating the Q-value table with incremental learning may result in an even better result, i.e., finding the optimal path with the highest reward or net power. The safe and acceptable performance of the online training is justified with the experimental results obtained in Section V, resulting in a slightly higher power than the case developed with the offline trained Q-value table. The major parameters of the Q-learning, including γ and α, are tuned during offline training through trial and error.

Note that the whole Q-learning algorithm is completely coded in Matlab, which includes three major sections (i) Environment: implementing the ocean environment defined by a grid of $n \times T$; (ii) Training: constructing the Q-value

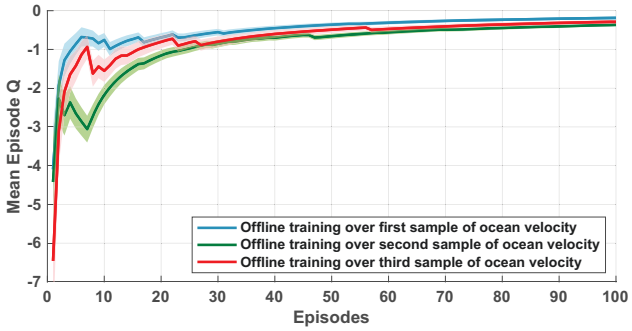


Fig. 4: Average Q-values over 100 trial episodes are shown for offline training over the first sample of ocean velocity, second sample of ocean velocity, and third sample of ocean velocity. Solid lines show average Q-values, and the shaded region determines a 95% confidence interval.

table and performing the training by updating the Q-value (Algorithm 3); and (iii) Testing: evaluating the constructed Q-value table through either direct testing (Algorithm 4) or performing simultaneous testing and updating through incremental learning (Algorithm 5).

It is noteworthy to mention that the Q-value table approach suffers from the "curse of dimensionality" when scaling to the problems with high-dimensional action and state spaces. To tame this deficiency, a deep Q-learning approach as a combination of the Q-learning and deep neural network can be applied to approximate the Q-value function, which is appropriate for a larger environment (especially a grid with a longer prediction horizon), as well as, enlarging the action decision variables (i.e., increasing the number of decision variables). Although the Q-value table is sufficient for the size of the ocean environment grid presented in the current study, we also apply a particular type of deep Q-learning called deep Q-network (DQN) to solve the problem at hand. A detailed discussion on the application of the DQN to the larger ocean environment is presented in other authors' work [64], [65].

Deep Q-network (DQN): In the DQN, the deep neural network is applied to find the optimal Q-value Q^*, which is previously determined through the Q-value table. Two neural networks with the same structure but different weights ϑ named Q-network $Q(.;\vartheta)$ and target network $Q_t(.;\vartheta^-)$ are employed to build the DQN. The main idea behind training the DQN is to find the optimal weights for these two networks through minimizing a loss function, where the weights are updated through gradient descent. The loss function is defined as:

$$L(\theta) = [Q_t(s_t, a_t; \vartheta^-) - Q(s_t, a_t; \vartheta)]^2 \tag{29}$$

where $Q_t(s_t, a_t; \vartheta^-)$ is defined as:

$$Q_t(s_t, a_t; \vartheta^-) \triangleq R_{t+1} + \gamma \max_a Q(s_{t+1}, a; \vartheta^-) \tag{30}$$

V. Results and Discussions

To evaluate the performance of the proposed methods for OCT vertical path planning, the following four approaches are compared:

- Case without Spatiotemporal Optimization: The OCT is located at a fixed depth, which means no change in operating depth, i.e., the output power of the system is determined only by the ocean velocity at the operating depth.
- Path Planning with A* Algorithm: One commonly used algorithm for grid-based path planning is the A* algorithm, which can efficiently seek the ocean environment grid and find the optimal path [10]. The A* algorithm applies a greedy strategy, where the heuristic function is defined by the difference between P_{r} and the expected power. The heuristic function is applied to label each path between two depths as shown in Fig. 3 and generate a weighted grid (or graph), which is further employed to find the optimal path.
- MPC-based Spatiotemporal Optimization: The optimal operating depth z^* is determined using objective function $J(z(p))$ in (21a), which aims to minimize the difference between the maximum power and the harvested power at each time step. Further, predicting the ocean velocity results in uncertainties, which is addressed through the exploration term J_{explore}.
- RL-based Spatiotemporal Optimization: Learning from different experiments should be considered to update the Q-value table. Specifically, the Q-function is calculated for each state and action pair, and the Q-value table is then updated. At each state (z, t), the optimal action (i.e., depth change Δz) should be determined to maximize the total power. The algorithm is trained offline by multiple trials, and the cumulative Q-values over trial episodes are shown in Fig. 4, which verifies the convergence of Q over 100 episodes for different sampling time. After training the RL algorithm, it is applied for online testing with both offline learning and incremental learning. Moreover, the DRL algorithm based on the DQN is applied to approximate the optimal Q-value (Q^*) through training and finding the optimal network, which is able to detect the optimal action (i.e., Δz).

A. Simulation Setup

We use real ocean velocity data and the example buoyancy-controlled OCT that is discussed in Section III. Recorded ocean velocities $\mathbf{V}$ are modeled by the GP model in (9). At each time step, the ocean velocity v_* is predicted over the prediction horizon. The predicted velocity is then considered as the observed velocity, and the next velocity is predicted by a sliding prediction window. Based on predicted ocean velocities, the optimal ocean depth z^* is determined. Key parameters used in the simulation are presented in Table II, and for the DQN, a network with two hidden layers is selected with a batch size of 64. All the experiments were conducted on a PC equipped with a 2.6 GHz CPU and 16 GB of RAM.

OCT system specification: The theoretical model of the investigated OCT system is presented in Section III-B. Assuming an OCT operating depth of $z_{\text{opt}} = 50\,\text{m}$, the relationship between pump power and volumetric flow rate is determined

TABLE II: Key parameters used in the optimization.

Parameter	Description	Value	Unit
P_{r}	Rated power	700	kW
ρ	Water density	1030	$\mathrm{kg/m^3}$
A	Swept area of the OCT rotor	100π	$\mathrm{m^2}$
C_{p}	Average power coefficient	41.5	%
η_{pump}	Pump efficiency	0.75	-
P_{atm}	Atmospheric pressure	101	kPa
g	Gravity	9.81	$\mathrm{m/s^2}$
Q_{B}	Volumetric flow rate	0.023	$\mathrm{m^3/s}$
ν_{B}	Volume of each ballast tank	31.251	$\mathrm{m^3}$
ζ	Coefficient in (17)	0.65	s/m
Δt_1	Sampling time to hold depth in (18)	0.25	hours
κ	Coefficient in (19)	-0.0026	1/m
Δt_2	Sampling time to change depth in (20)	1	hour
β	Gain of exploration term in (21a)	100	-
T	Prediction horizon	2	hours
n	Number of discrete depths	27	-
$z^{\min}$	Minimum allowable depth	40	m
$z^{\max}$	Maximum allowable depth	200	m
γ	Discount factor	0.8	-
δ	Constant value in (26)	1	kW

using (14):

$$P_{\mathrm{B}_{\max}}^{\text{empty}} = \frac{606 Q_{\mathrm{B}_{\max}}}{0.75} = 808 Q_{\mathrm{B}_{\max}} \tag{31}$$

It should be noted that the volumetric flow rate and ballast pump power are calculated associated with WWII submarines [66]. Accordingly, the power and energy to empty two ballast tanks are calculated as $P_{\mathrm{B}}^{\text{empty}} = 18.8$ and $E_{\mathrm{B}}^{\text{empty}} = 14.02$, respectively, with respect to (31) and values presented in Table II. The linear quasi-static relationship is interpreted for a flow speed of 1.6 m/s. In our application, the depth is limited within 40 m to 200 m since the maximum ocean currents nearly always occur in the top 200 m depths [67].

Ocean current shear profile: In this paper, we use the data recorded by a 75 kHz acoustic Doppler current profiler (ADCP) at a latitude of $26.09°\mathrm{N}$ and longitude of $-79.80°\mathrm{E}$. ADCPs measuring water velocity data over $300+$ m can make approximately one measurement per second, have a vertical spatial resolution of around 6 m, and typically save time-averaged water velocity measurements every 3 to 15 minutes to reduce measurement noise. The data include measures of northward current velocity, eastward current velocity, and data quality related parameters. To save battery, data used in this study were recorded every 15 minutes with 25 measurements averaged to create each recording, yielding an estimated measurement error standard deviation of 0.02 m/s throughout the water column (based on the ADCP's manufacturer specification found in [68]). Recorded data were filtered to remove bad data through the method suggested in [67], with "bad" data primarily measured above a depth of 50 m. "Bad" data were identified using ADCP correlation and percent good, as well as the echo intensity spikes characteristic of the initial return of side lobe acoustics reflecting from the sea surface. The ocean shear profile calculated from these data over a sample one-week period is shown in Fig. 5. To expand the spatiotemporal path planning algorithms from this paper to real turbine applications, an ADCP will likely be installed just upstream from an ocean current turbine, or array of turbines, so that near real-time current velocity data are accessible.

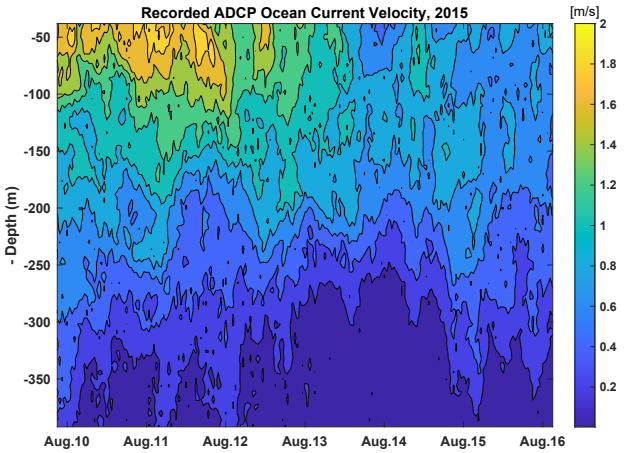


Fig. 5: Sample ocean current shear profile recorded by an ADCP in the Gulf Stream over one week.

B. OCT Operating Depth Assessment for Linear Model and Nonlinear Model

To derive the OCT output power model, certain approximations are made to represent the nonlinear model presented in [42] with linear quasi-static relationships. The major approximations are included in (17) and (19) to determine the relationship between the fill fraction change and flow speed change as well as depth change. Hence, the fill fraction is affected by two parameters of depth and velocity, and the whole fill fraction equation can be rewritten as follows:

$$\Delta F_{\mathrm{F}} = \zeta \Delta v + \kappa \Delta z \tag{32}$$

where $\Delta \triangleq (.) - (.)_{\mathrm{eq}}$ given that $F_{\mathrm{eq}} = 0.5$, $z_{\mathrm{eq}} = 50$ m, and $v_{\mathrm{eq}} = 1.6$ m/s.

To justify the accuracy of the presented linear model in (32), the OCT operating depth obtained through the nonlinear model [42] is compared with those calculated by (32). The operating depths associated with the specified flow speeds and fill fractions (equal values for both tanks) are shown in Fig. 6. Note that the operating depths are illustrated under the sea surface, where the OCT's ballast tank hits the surface if the OCT system is located at a depth of 15.34 m. As this figure shows, the operating depth calculated by the linear model (32) matches those obtained through the nonlinear model, where a negligible difference is anticipated due to approximations made to attain the linear model from a nonlinear model.

C. Comparative Results

The obtained results over a sample simulation time of 200 hours are presented in Fig. 7 to Fig. 9. The optimal path obtained over 200 hours with A* algorithm, MPC algorithm, online testing of RL algorithm with offline learning, online testing of RL algorithm with incremental learning, and DRL algorithm are illustrated in Fig. 7. The control action is defined as the depth change in this study, and the optimal depth and the corresponding velocity are shown under the A* method, MPC-based method, RL-based methods, DRL method, and the case without spatiotemporal optimization. As presented in Fig. 8(a), the optimal paths show similarities, especially for

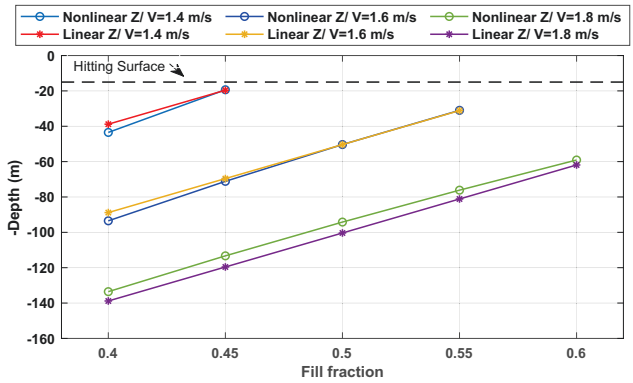


Fig. 6: Comparison of equilibrium depths associated with the specified flow speeds and fill fractions (equal values for both tanks), found using both the non-linear and linear models. It is noted that fill fractions of 0.5 for each tank and a flow speed of 1.6 m/s are utilized when creating the linear model.

the RL algorithm and MPC algorithm. However, there exist noticeable differences in selecting the next control action, such as at $t = 130$. Although different z are selected as the next optimal depth for each algorithm, the velocities at these depths are nearly the same (Fig. 8(b)), verifying that these algorithms can make optimal decisions.

The harvested power of the OCT system is presented in Fig. 9, which shows that the output power of OCT increases when the spatiotemporal optimization is applied. Different methods for path planning follow a similar trend in power increase; however, RL-based spatiotemporal optimization with incremental learning surpasses the remaining methods at some time samples. In addition, the average power of the OCT system in terms of P_{OCT}, $P_{\text{ballast}}^{\text{HD}}$, $P_{\text{ballast}}^{\text{CD}}$, and P_{net} for three different operating depths are presented in Table III.

D. Robustness Analysis

For an evaluation of the spatiotemporal optimization robustness, the proposed methods are implemented by the perturbed ocean velocity, in which the ocean velocity is not correctly modeled due to the scale error in the velocity sensors, the data loss from measurement, and so on [2]. Fig. 10 shows the perturbation of cumulative energy from baseline obtained by the A* algorithm, the MPC algorithm, the RL algorithm with offline learning, the RL with incremental learning, and the DRL algorithm in response to 5% noise disturbances of the same intensities for 100 test cases. We can observe that the RL algorithms are sufficiently robust, where a similar distribution of the results (i.e., cumulative energy) is obtained for the perturbed ocean velocity model. Moreover, the RL algorithm with incremental learning outperforms the MPC algorithm, and the perturbed results (i.e., that in a small interval [39.348 $\sim$ 39.978] MWh) is very close to the obtained cumulative energy for baseline case without noise (40.366 MWh).

To further show the robustness of our proposed methods, the cumulative energy of the A* algorithm, the MPC algorithm, the RL algorithm with offline learning, the RL with incremental learning, and the DRL algorithm in response to

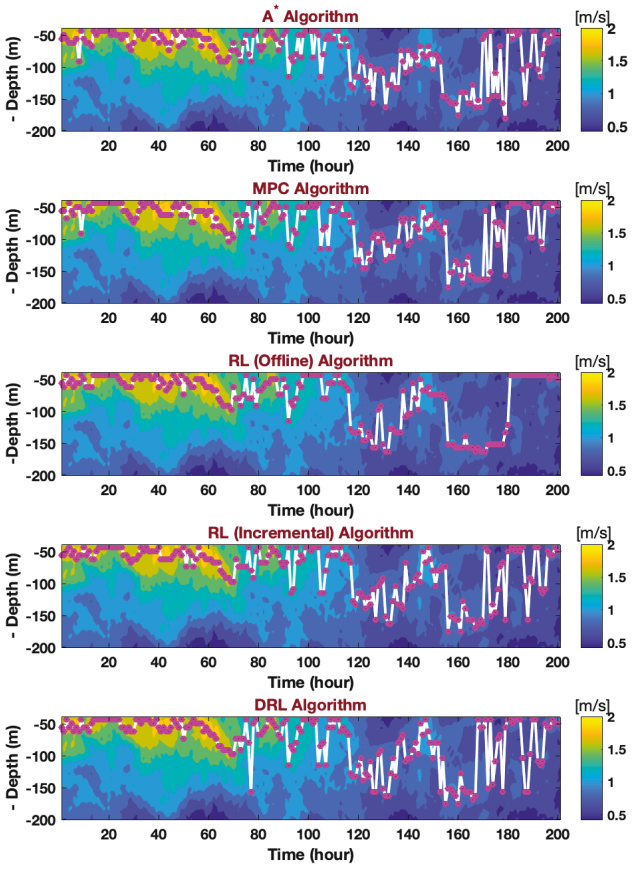
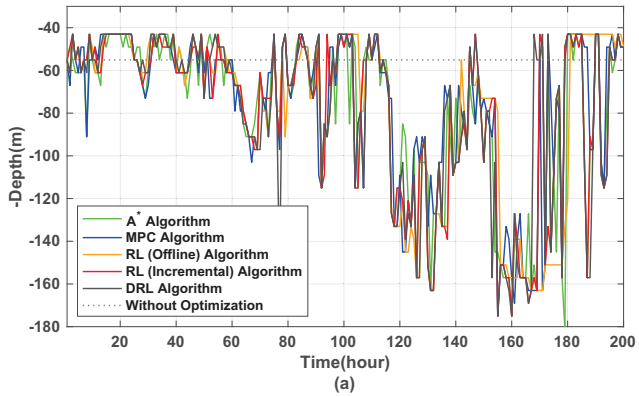


Fig. 7: Comparing optimal path obtained over 200 hours under A* algorithm, MPC-based optimization, RL-based optimization with offline learning, RL-based optimization with incremental learning, and DRL algorithm.

noise disturbances, ranging from 5% to 20%, for 100 test cases are presented in Table IV. Note that the results are reported as an average of the cumulative energy obtained by 100 tests. As can be seen from the table, the superiority of the RL-based method with incremental learning over other methods is verified under these four cases, justifying that the A* algorithm, MPC algorithm, and DRL algorithm are more sensitive to the increased ocean velocity modeling error.

E. Discussions

The RL algorithm as a learning-based method usually has offline training and online deployment phases, which is introduced as the RL algorithm with offline learning in our study. In another proposed approach (named the RL algorithm with incremental learning), once the offline training is finished, it will be deployed online while it keeps updating the policy (e.g., the Q-value table). The RL algorithm with incremental learning is slightly better than the one with offline learning since it can adapt better to minor changes in the environment. The DRL algorithm estimates the Q-value function through the neural network trained offline using the recorded ocean velocity; hence, a pretty good DRL can generate an estimate from the optimal path instead of a precise optimal path. However, the DRL algorithm can address the vertical path planning



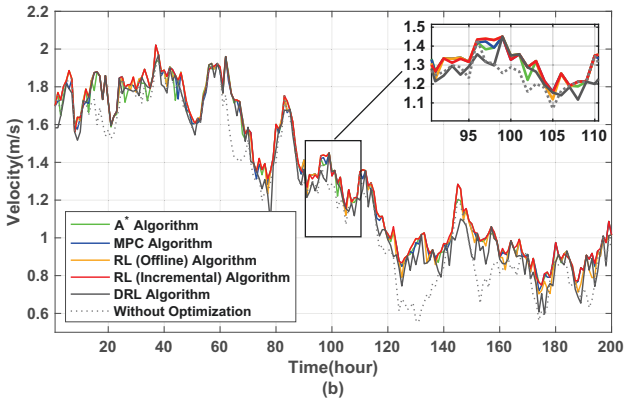


Fig. 8: Comparing optimal depth and velocity over 200 hours obtained by case without spatiotemporal optimization, A* algorithm, MPC algorithm, RL-based algorithm with offline learning, RL-based algorithm with incremental learning, and DRL algorithm. (a) Optimal depth; (b) Optimal velocity.

with a long prediction horizon, which can not be solved by a Q-value table (RL algorithm) due to high computational complexity. Different learning-based approaches are feasible for vertical path planning to maximize the output power of the OCT system, while the RL algorithm with incremental learning shows a better performance in the problem at hand. Another important feature of the RL method is its robustness to errors and uncertainties. In the MPC algorithm, the exploration term J_{explore} in the objective function $J(z)$ is defined to address this issue. However, the main challenge is to determine the corresponding weight β of this term J_{explore}, which is chosen experimentally. RL algorithm is directly trained by taking different actions in each state, which means it will choose the best action at each state based on its experience. Through learning from historical data and online incremental learning, the robustness of the algorithm will increase.

The cumulative produced energy of the OCT system using the A* algorithm, the MPC algorithm, the RL algorithm with offline learning, the RL with incremental learning, and the DRL algorithm are compared, as shown in Fig. 11. The cumulative energy shows a high increase compared to the case without optimization, highlighting the importance of applying path planning algorithms. We also observe that the cumulative

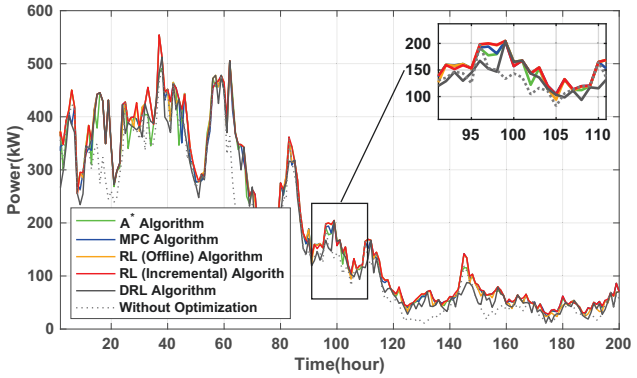


Fig. 9: Comparing optimal power under case without spatiotemporal optimization, A* algorithm, MPC algorithm, RL-based algorithm with offline learning, RL-based algorithm with incremental learning, and DRL algorithm.

TABLE III: Comparing the detailed average power terms of the OCT for three different operating depths.

z (Time=0)	P_{OCT}	$P^{\text{HD}}_{\text{ballast}}$	$P^{\text{CD}}_{\text{ballast}}$	P_{net}
m	kW	kW	kW	kW
Without Optimization				
45	167.5736	11.2947	-	156.2789
50	171.1748	11.3457	-	159.8291
55	170.6034	10.4052	-	160.1982
A* Algorithm				
45	196.6085	9.8111	0.3554	186.4420
50	196.7481	9.3609	0.3160	187.0712
55	196.64	9.6871	0.3368	186.6161
MPC Algorithm				
45	196.8896	9.3700	0.3488	187.1708
50	196.8848	9.4411	0.3018	187.1419
55	197.1997	9.8767	0.2679	187.0551
RL Algorithm with Offline Learning				
45	199.2042	10.0061	0.1968	189.0013
50	199.1987	10.1446	0.1761	188.8781
55	199.5513	10.2631	0.1684	189.1198
RL Algorithm with Incremental Learning				
45	201.5854	9.1403	0.2996	192.1455
50	201.9442	8.6938	0.3051	192.9453
55	201.9723	9.1531	0.2942	192.525
DRL Algorithm				
45	184.6497	10.9068	0.1640	173.5789
50	184.8559	10.7460	0.1651	173.9448
55	184.8559	10.7952	0.1662	173.8945

energy productions of the OCT system are very close for different compared optimization approaches. RL algorithm with incremental learning outperforms the remaining methods, and the final energy production for the A* algorithm, the MPC algorithm, the RL algorithm with offline learning, the RL with incremental learning, the DRL algorithm, and case without applying spatiotemporal optimization are 39.262 MWh, 39.653 MWh, 39.877 MWh, 40.366 MWh, 36.965 MWh, and 34.121 MWh, respectively.

To present a comprehensive and fair perspective of our proposed spatiotemporal optimization, its limitations are summarized. As discussed in Section II, we focus on the upper-level vertical path planning, assuming that the lower-level control (flight control) is able to track the commanded optimal path. This assumption arises from our experiments and

TABLE IV: Comparing the robustness in percent decrease of cumulative energy using accurate ocean velocity model vs. perturbed ocean velocity models.

Noise	A* Algorithm		MPC Algorithm		RL (Offline) Algorithm		RL (Incremental) Algorithm		DRL Algorithm	
%	MWh	%	MWh	%	MWh	%	MWh	%	MWh	%
Baseline Ocean Velocity Model										
-	39.262	-	39.653	-	39.877	-	40.366	-	36.965	-
Ocean Velocity Model Perturbed with Noise										
5	38.230	2.63	38.364	3.25	39.641	0.59	39.658	1.75	36.507	1.24
10	37.273	5.07	37.594	5.19	38.737	2.86	38.720	4.08	34.351	7.07
15	36.390	7.31	36.998	6.70	38.066	4.54	38.158	5.47	32.938	10.89
20	35.736	8.98	36.430	8.13	37.052	7.08	37.195	7.85	32.655	11.66

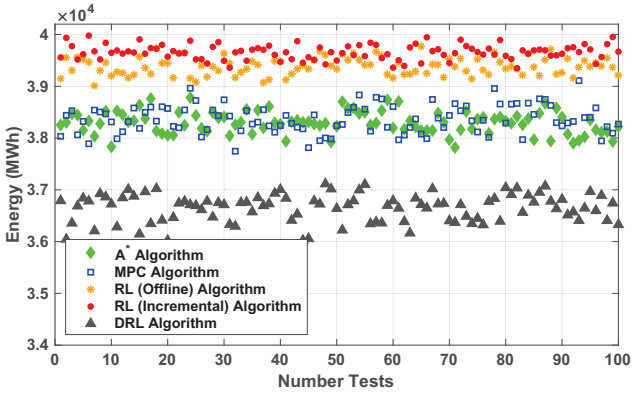


Fig. 10: Robustness comparison of the A* algorithm, MPC algorithm, RL-based algorithm with offline learning, RL-based algorithm with incremental learning, and DRL algorithm through cumulative energy under noise.

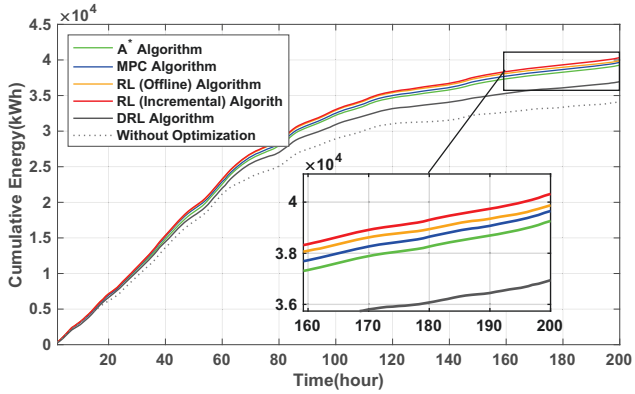


Fig. 11: Comparing cumulative energy under case without spatiotemporal optimization, A* algorithm, MPC algorithm, RL-based algorithm with offline learning, RL-based algorithm with incremental learning, and DRL algorithm.

knowledge on the dynamics of the OCT system and the previously designed flight controllers [44], [45], [69], ignoring the detailed model of the OCT system. Two major operational constraints of the OCT system directly affecting the vertical path planning, including the range of operating depth and the constraint on the slew rate (i.e., highly related with the mooring system design), as well as the real recorded data from the Gulf Stream, are included in our spatiotemporal model to ensure a feasible vertical path according to the OCT model and real ocean environment. We are further developing an integrated path planning and tracking control framework for the investigated OCT system [64]. Preliminary results verify the ability of the OCT system to follow the commanded vertical path by the proposed spatiotemporal optimization, where the detailed model of the OCT system and constraints on its actuators (B_{f}, B_{a}, and τ_{em}) are included in the flight controller design.

VI. Conclusions

In this study, a novel spatiotemporal optimization approach was presented for OCT vertical path planning to maximize the net power. The GP model was first developed to model the ocean velocity based on real data. Then, the OCT power models were formulated, including the directly generated power from hydrokinetic energy, the consumed power for stabilizing, and the consumed power for changing the operating depth. Two types of methodologies, including MPC- and RL-based algorithms, were developed to solve the proposed spatiotemporal problem. Compared with the baseline approaches, the obtained results verified that both methods are efficient in finding the optimal path to maximize the output power. Moreover, the deep RL-based method with online incremental learning showed better performance in terms of cumulative energy and robustness.

Future work is needed to fully investigate the interaction between the lower-level controllers (i.e., flight controller or generator controller) with the upper-level controller (i.e., spatiotemporal optimization). It is also critical to extend the proposed spatiotemporal optimization to an OCT array to maximize the total power of the array considering the wake effects among the OCTs and integrate the harnessed power into power grids using energy storage.

References

[1] M. I. Yuce and A. Muratoglu, "Hydrokinetic energy conversion systems: A technology status review," *Renewable and Sustainable Energy Reviews*, vol. 43, pp. 72–82, 2015.

[2] Y. Tang, J. VanZwieten, B. Dunlap, D. Wilson, C. Sultan, and N. Xiros, "In-stream hydrokinetic turbine fault detection and fault tolerant control-a benchmark model," in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 4442–4447.

[3] A. C. O'Sullivan and G. Lightbody, "Co-design of a wave energy converter using constrained predictive control," *Renewable energy*, vol. 102, pp. 142–156, 2017.

[4] A. Hasankhani, Y. Tang, A. Snyder, J. VanZwieten, and W. Qiao, "Control co-design for buoyancy-controlled mhk turbine: A nested optimization of geometry and spatial-temporal path planning," *arXiv preprint arXiv:2110.06470*, 2021.

[5] T. Wilberforce, Z. El Hassan, D. A., J. Thompson, B. Soudan, and A. Olabi, "Overview of ocean power technology," *Energy*, vol. 175, pp. 165–181, may 2019.

[6] S. Ponda, R. Kolacinski, and E. Frazzoli, "Trajectory optimization for target localization using small unmanned aerial vehicles," in *AIAA guidance, navigation, and control conference*, 2009, p. 6015.

[7] A. Tsourdos, B. White, and M. Shanmugavel, *Cooperative path planning of unmanned aerial vehicles*. John Wiley & Sons, 2010, vol. 32.

[8] H. Yu, K. Meier, M. Argyle, and R. W. Beard, "Cooperative path planning for target tracking in urban environments using unmanned air and ground vehicles," *IEEE/ASME Transactions on Mechatronics*, vol. 20, no. 2, pp. 541–552, 2015.

[9] P. Yang, K. Tang, J. A. Lozano, and X. Cao, "Path planning for single unmanned aerial vehicle by separately evolving waypoints," *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1130–1146, 2015.

[10] B. Garau, A. Alvarez, and G. Oliver, "Path planning of autonomous underwater vehicles in current fields with complex spatial variability: an A* approach," in *Proceedings of the 2005 IEEE international conference on robotics and automation*. IEEE, 2005, pp. 194–198.

[11] K. P. Carroll, S. R. McClaran, E. L. Nelson, D. M. Barnett, D. K. Friesen, and G. N. William, "Auv path planning: an a* approach to path planning with consideration of variable vehicle speeds and multiple, overlapping, time-dependent exclusion zones," in *Proceedings of the 1992 symposium on autonomous underwater vehicle technology*. IEEE, 1992, pp. 79–84.

[12] A. Bafandeh and C. Vermillion, "Real-time altitude optimization of airborne wind energy systems using lyapunov-based switched extremum seeking control," in *2016 American Control Conference (ACC)*. IEEE, 2016, pp. 4990–4995.

[13] A. Bafande and C. Vermillion, "Altitude optimization of airborne wind energy systems via switched extremum seeking—design, analysis, and economic assessment," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 6, pp. 2022–2033, 2016.

[14] A. Bafandeh and C. Vermillion, "Optimal altitude control of an integrated airborne wind energy system with globalized lyapunov-based switched extremum seeking," in *2018 European Control Conference (ECC)*. IEEE, 2018, pp. 2089–2094.

[15] F. Mandić, N. Mišković, and Z. Vukić, "Range–only navigation–maximizing system observability by using extremum seeking," *IFAC-PapersOnLine*, vol. 48, no. 16, pp. 101–106, 2015.

[16] J. Cochran, E. Kanso, S. D. Kelly, H. Xiong, and M. Krstic, "Source seeking for two nonholonomic models of fish locomotion," *IEEE Transactions on Robotics*, vol. 25, no. 5, pp. 1166–1176, 2009.

[17] A. Hernandez-Sanchez, I. Chairez, and A. Poznyak, "Extended integral sliding mode robust sub-gradient extremum seeking control for tracking trajectory of autonomous underwater vehicle," in *2020 7th International Conference on Control, Decision and Information Technologies (CoDIT)*, vol. 1. IEEE, 2020, pp. 433–438.

[18] T. B. von See, T. Meurer, and J. Greinert, "Marine boundary layer tracking using an auv with ukf based extremum seeking," *IFAC-PapersOnLine*, vol. 54, no. 16, pp. 320–326, 2021.

[19] C. Liu, S. Lee, S. Varnhagen, and H. E. Tseng, "Path planning for autonomous vehicles using model predictive control," in *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2017, pp. 174–179.

[20] T. Shim, G. Adireddy, and H. Yuan, "Autonomous vehicle collision avoidance system using path planning and model-predictive-control-based active front steering and wheel torque control," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of automobile engineering*, vol. 226, no. 6, pp. 767–778, 2012.

[21] J. Ji, A. Khajepour, W. W. Melek, and Y. Huang, "Path planning and tracking for vehicle collision avoidance based on model predictive control with multiconstraints," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 952–964, 2016.

[22] M. Graf Plessen, D. Bernardini, H. Esen, and A. Bemporad, "Spatial-based predictive control and geometric corridor planning for adaptive cruise control coupled with obstacle avoidance," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 1, pp. 38–50, 2018.

[23] S. Bin-Karim, A. Bafandeh, A. Baheri, and C. Vermillion, "Spatiotemporal optimization through gaussian process-based model predictive control: A case study in airborne wind energy," *IEEE Transactions on Control Systems Technology*, vol. 27, no. 2, pp. 798–805, 2017.

[24] M. Kehs, C. Vermillion, and H. Fathy, "Online energy maximization of an airborne wind energy turbine in simulated periodic flight," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 2, pp. 393–403, 2017.

[25] S. Aradi, "Survey of deep reinforcement learning for motion planning of autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–20, 2020.

[26] B. Wang, Z. Liu, Q. Li, and A. Prorok, "Mobile robot path planning in dynamic environments through globally guided reinforcement learning," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6932–6939, 2020.

[27] J. Xie, Z. Shao, Y. Li, Y. Guan, and J. Tan, "Deep reinforcement learning with optimized reward functions for robotic trajectory planning," *IEEE Access*, vol. 7, pp. 105 669–105 679, 2019.

[28] Q. Zhang, K. Wu, and Y. Shi, "Route planning and power management for phevs with reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 4751–4762, 2020.

[29] P. F. Lermusiaux, C.-S. Chiu, G. G. Gawarkiewicz, P. Abbot, A. R. Robinson, R. N. Miller, P. J. Haley, W. G. Leslie, S. J. Majumdar, A. Pang, *et al.*, "Quantifying uncertainties in ocean predictions," HARVARD UNIV CAMBRIDGE MA, Tech. Rep., 2006.

[30] M. Deisenroth and C. E. Rasmussen, "Pilco: A model-based and data-efficient approach to policy search," in *Proceedings of the 28th International Conference on machine learning (ICML-11)*. Citeseer, 2011, pp. 465–472.

[31] Y. Pan, G. I. Boutselis, and E. A. Theodorou, "Efficient reinforcement learning via probabilistic trajectory optimization," *IEEE transactions on neural networks and learning systems*, vol. 29, no. 11, pp. 5459–5474, 2018.

[32] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Numerical gaussian processes for time-dependent and nonlinear partial differential equations," *SIAM Journal on Scientific Computing*, vol. 40, no. 1, pp. A172–A198, 2018.

[33] S. Bin-Karim, M. Muglia, A. Mazzoleni, and C. Vermillion, "Control of a relocatable energy-harvesting autonomous underwater vehicle in a spatiotemporally-varying gulf stream resource," in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 2575–2580.

[34] T. Wang, R. M. Lima, L. Giraldi, and O. M. Knio, "Trajectory planning for autonomous underwater vehicles in the presence of obstacles and a nonlinear flow field using mixed integer nonlinear programming," *Computers & Operations Research*, vol. 101, pp. 55–75, 2019.

[35] D. Kularatne, H. Hajieghrary, and M. A. Hsieh, "Optimal path planning in time-varying flows with forecasting uncertainties," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 4857–4864.

[36] A. Ramos, V. García-Garrido, A. Mancho, S. Wiggins, J. Coca, S. Glenn, O. Schofield, J. Kohut, D. Aragon, J. Kerfoot, *et al.*, "Lagrangian coherent structure assisted path planning for transoceanic autonomous underwater vehicle missions," *Scientific reports*, vol. 8, no. 1, pp. 1–9, 2018.

[37] A. Baheri, P. Ramaprabhu, and C. Vermillion, "Iterative 3d layout optimization and parametric trade study for a reconfigurable ocean current turbine array using bayesian optimization," *Renewable energy*, vol. 127, pp. 1052–1063, 2018.

[38] S. Bin-Karim, M. Muglia, and C. Vermillion, "Centralized position optimization of multiple agents in spatiotemporally-varying environment: a case study with relocatable energy-harvesting autonomous underwater vehicles in the gulf stream," in *2019 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE, 2019, pp. 264–269.

[39] Y. Tang, Y. Zhang, A. Hasankhani, and J. VanZwieten, "Adaptive super-twisting sliding mode control for ocean current turbine-driven permanent magnet synchronous generator," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 211–217.

[40] Z. Zhou, F. Scuiller, J. F. Charpentier, M. E. H. Benbouzid, and T. Tang, "Power smoothing control in a grid-connected marine current turbine system for compensating swell effect," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 3, pp. 816–826, 2013.

[41] Z. Zhou, S. B. Elghali, M. Benbouzid, Y. Amirat, E. Elbouchikhi, and G. Feld, "Tidal stream turbine control: An active disturbance rejection control approach," *Ocean Engineering*, vol. 202, p. 107190, 2020.

[42] A. Hasankhani, J. VanZwieten, Y. Tang, B. Dunlap, A. De Luera, C. Sultan, and N. Xiros, "Modeling and numerical simulation of a buoyancy controlled ocean current turbine," *International Marine Energy Journal*, vol. 4, no. 2, pp. 47–58, 2021.

[43] T. D. Ngo, C. Sultan, J. H. VanZwieten, and N. I. Xiros, "Variance constrained cyclic blade control of moored ocean current turbines," in *2016 American Control Conference (ACC)*. IEEE, 2016, pp. 6622–6627.

[44] T. D. Ngo., C. Sultan, J. H. VanZwieten, and N. I. Xiros, "Model predictive control for moored ocean current turbines," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 875–880.

[45] T. D. Ngo, C. Sultan, J. H. VanZwieten, and N. I. Xiros, "Constrained control of moored ocean current turbines with cyclic blade pitch variations," *IEEE Journal of Oceanic Engineering*, 2020.

[46] J. H. VanZwieten, N. Vanrietvelde, and B. L. Hacker, "Numerical simulation of an experimental ocean current turbine," *IEEE Journal of Oceanic Engineering*, vol. 38, no. 1, pp. 131–143, 2012.
[47] B. Pearson and B. Fox-Kemper, "Log-normal turbulence dissipation in global ocean models," *Physical review letters*, vol. 120, no. 9, 2018.
[48] I. Lozovatsky, K. Shearman, A. Pirro, and H. J. S. Fernando, "Probability distribution of turbulent kinetic energy dissipation rate in stratified turbulence: Microstructure measurements in the southern california bight," *JGR Oceans*, vol. 124, no. 7, pp. 4591–4604, July 2019.
[49] J. Yan, K. Li, B. Er-Wei, J. Deng, and A. M. Foley, "Hybrid probabilistic wind power forecasting using temporally local gaussian process," *IEEE transactions on sustainable energy*, vol. 7, no. 1, pp. 87–95, Jan. 2016.
[50] H. Cai, X. Jia, J. Feng, W. Li, Y.-M. Hsu, and J. Lee, "Gaussian process regression for numerical wind speed prediction enhancement," *Renewable Energy*, vol. 146, pp. 2112–2123, 2020.
[51] K.-C. Ma, L. Liu, and G. S. Sukhatme, "Informative planning and online learning with sparse gaussian processes," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 4292–4298.
[52] 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," *Journal of Field Robotics*, vol. 33, no. 1, pp. 47–66, 2016.
[53] Y. Huang, Y. Tang, H. Zhuang, J. VanZwieten, and L. M. Cherubin, "Physics-informed tensor-train convlstm for volumetric velocity forecasting of loop current," *Frontiers in Artificial Intelligence*, p. 197, 2021.
[54] L. Csató and M. Opper, "Sparse on-line gaussian processes," *Neural computation*, vol. 14, no. 3, pp. 641–668, 2002.
[55] J. Hensman, N. Fusi, and N. D. Lawrence, "Gaussian processes for big data," *arXiv preprint arXiv:1309.6835*, 2013.
[56] C. Zhang, H. Wei, X. Zhao, T. Liu, and K. Zhang, "A gaussian process regression based hybrid approach for short-term wind speed prediction," *Energy Conversion and Management*, vol. 126, pp. 1084–1092, 2016.
[57] J. VanZwieten, P. Pyakurel, T. Ngo, C. Sultan, and N. I. Xiros, "An assessment of using variable blade pitch for moored ocean current turbine flight control," *International Journal of marine energy*, vol. 13, pp. 16–26, 2016.
[58] K. Yahagi and K. Takagi, "Moment loads acting on a blade of an ocean current turbine in shear flow," *Ocean Engineering*, vol. 172, pp. 446–455, 2019.
[59] T. Ueno, S. Nagaya, M. Shimizu, H. Saito, S. Murata, and N. Handa, "Development and demonstration test for floating type ocean current turbine system conducted in kuroshio current," in *2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO)*. IEEE, 2018, pp. 1–6.
[60] D. Coiro, G. Troise, F. Scherillo, A. De Marco, G. Calise, and N. Bizzarrini, "Development, deployment and experimental test on the novel tethered system gem for tidal current energy exploitation," *Renewable Energy*, vol. 114, pp. 323–336, 2017.
[61] S. Couch and I. Bryden, "Tidal current energy extraction: Hydrodynamic resource characteristics," *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, vol. 220, no. 4, pp. 185–194, 2006.
[62] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
[63] M. Luong and C. Pham, "Incremental learning for autonomous navigation of mobile robots based on deep reinforcement learning," *Journal of Intelligent & Robotic Systems*, vol. 101, no. 1, pp. 1–11, 2021.
[64] A. Hasankhani, E. B. Ondes, Y. Tang, C. Sultan, and J. VanZwieten, "Integrated path planning and tracking control of marine current turbine in uncertain ocean environments," *2022 Annual American Control Conference (ACC)*, Accepted.
[65] A. Hasankhani, Y. Tang, J. VanZwieten, and C. Sultan, "Comparison of deep reinforcement learning and model predictive control for real-time depth optimization of a lifting surface controlled ocean current turbine," in *2021 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE, 2021, pp. 301–308.
[66] The fleet type submarine online submarine trim and drain systems (chapter 2). [Online]. Available: https://maritime.org/doc/fleetsub/trim/chap2.htm
[67] M. C. P. M. Machado, J. H. VanZwieten, and I. Pinos., "A measurement based analyses of the hydrokinetic energy in the gulf stream," *Journal of Ocean and Wind Energy*, vol. 3, no. 1, pp. 25–30, 2016.
[68] T. Marine, "Workhorse long ranger 75 khz adcp," Brochure, 2020. [Online]. Available: http://www.teledynemarine.com/Lists/Downloads/long-ranger-datasheet-lr.pdf
[69] E. B. Ondes, C. Sultan, A. Hasankhani, J. VanZwieten, and N. I. Xiros, "Nu-gap metric based multi-model predictive control of an ocean current turbine system with blade pitch failures," *IEEE Journal of Oceanic Engineering*, Submitted.

Arezoo Hasankhani received the B.S. and M.S. degrees from Amirkabir University in electrical engineering in 2012 and 2014, respectively. Since 2019, she has been a Ph.D. student in the Department of Electrical Engineering and Computer Science, and a research assistant in the Institute for Sensing and Embedded Network Systems Engineering, Florida Atlantic University, Boca Raton, FL, USA. Her research interests include optimization, control, renewable energy system, and reinforcement learning.

Yufei Tang (M'16) received the Ph.D. degree in Electrical Engineering from the University of Rhode Island, Kingston, RI, USA, in 2016. He is currently an Assistant Professor with the Department of Electrical Engineering and Computer Science, and a Faculty Fellow with the Institute for Sensing and Embedded Network Systems Engineering, Florida Atlantic University, Boca Raton, FL, USA. His research interests include machine learning, data mining, dynamical systems, and renewable energy.

Dr. Tang was a recipient of the IEEE International Conference on Communications Best Paper Award in 2014, the National Academies Gulf Research Program Early-Career Research Fellowship in 2019, and the National Science Foundation CAREER Award in 2022.

James VanZwieten received the B.S., M.S., and Ph.D. degrees in Ocean Engineering from Florida Atlantic University, Boca Raton, in 2001, 2003, and 2007, respectively. Currently, he is an Associate Research Professor at the Southeast National Marine Renewable Energy Center, Florida Atlantic University. His research interests include modeling and control of marine vehicles, in-stream hydrokinetic energy production, ocean thermal energy conversion, and sea water air conditioning. Dr. VanZwieten is a member of the American Society of Civil Engineers (ASCE) Marine Renewable Energy Committee and Chair of the Marine Hydrokinetic Subcommittee.

Cornel Sultan received the engineering diploma in aerospace engineering from Bucharest Polytechnic University, Bucharest, Romania, and the M.S. degree in mathematics and the Ph.D. degree in aerospace engineering from Purdue University, West Lafayette, IN, USA, in 1999. He was affiliated with a start-up company, Tensegra, Inc. (1999–2001), Harvard Medical School (2001–2003), Scientific Systems Company, Inc., (2001–2004), and United Technologies Research Center (2004–2007). He is currently a Professor with the Aerospace and Ocean Engineering Department, Virginia Tech, Blacksburg, VA, USA. His research interests include dynamics and control of tensegrity structures, membranes, and aerospace and marine systems.