Spatiotemporal Optimization for Path Planning, with Application to 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 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 consumption (for controlling the depth) and generation (from hydrokinetic) are modeled. The stochastic behavior of ocean velocities is a function of spatial and temporal parameters, as modeled through a Gaussian process approach. The net power of the OCT system is composed of three parts, generated power, power for maintaining the system at an operating depth, and power consumed for navigating the turbine to the optimal water depth. Two different algorithms, including model predictive control (MPC) as a model-based method and reinforcement learning (RL) as a learning-based method, are developed to solve the formulated spatiotemporal optimization problem with constraints. Comparative studies show that the MPC and the RL based methods are computationally feasible considering the required time for changing operating depth. Analysis of the robustness is further carried out under the inaccurate ocean velocity predictions. Results verify the efficiency of both 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 an over 18% increase compared to the baseline.**

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

I. Introduction

MARINE hydrokinetic (MHK) energy has been 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] or optimizing the operational strategy [4]. This study focuses on maximizing the net output power of an ocean current turbine (OCT) through optimal path planning.

Path planning has been studied for many other applications, such as unmanned vehicle trajectory optimization [5], [6], however, the challenge for OCT path planning is unique - the turbine will be tethered to the seafloor through a mooring system and be 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 path planning, model predictive control (MPC) and reinforcement learning (RL) have gained increasing attention in recent years due to their success in both academia and industry. To address the path planning for autonomous vehicles in a dynamic environment, MPC-based approaches have been applied to determine the optimal path and operation mode of vehicles [7], [8]. A spatial-based path planning method for autonomous vehicles has been introduced in [9], verifying autonomous driving in an obstacle-free environment and in the presence of obstacles. In similar research to our problem, an airborne wind energy system has employed MPC-based techniques to set the optimal altitude [10], [11], in which the future wind speeds were predicted and the output power was optimized by changing the location of the system to access the optimal velocities. On the other hand, the path planning of autonomous vehicles has been justified through RL approaches considering safety, security, and communication issues [12]. An RL-based path planning of mobile robots with obstacles has been proposed to avoid collision and determine the optimal path through identifying the environmental spatiotemporal data [13], [14]. To realize the minimum power consumption for plug-in hybrid electric vehicles, path planning and energy management have been addressed through a RL-based approach [15]. However, due to the error induced by the prediction modeling, spatiotemporal uncertainties, and etc. [16], the feasibility and robustness of these methods applied to OCT path planning have not been investigated.

It is critical to study the optimal path planning for agents in an oceanic environment through predictive methodology, due to the highly stochastic oceanic environment and the

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

A. Hasankhani and Y. Tang are with the Department of Computer & 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).

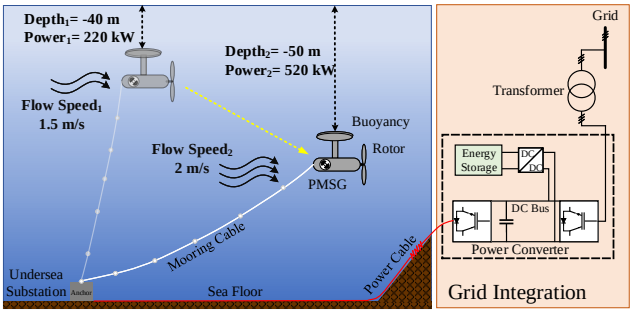

Fig. 1: Schematic diagram of the studied problem. The OCT is controlled spatiotemporally, where the optimal path is planned through changing the depth, e.g., move from $depth_1 = -40\,m$ to $depth_2 = -50\,m$ over one sampling time, resulting in a 300 kW power increase.

absence of human intervention (e.g., remote areas and deep-sea). An estimator has been designed to predict the spatially dependent ocean velocity, and trajectory planning then further been developed to determine the location of an AUV that used as an energy harvester [17]. Trajectory planning of AUVs has been addressed to avoid collisions in the presence of dynamic obstacles [18] and to reduce expected cost [19]. Transoceanic gliders treated as AUVs have been studied, aiming to a decreased battery consumption on long-duration missions [20]. Bayesian optimization has been used to specify the configuration and location of an OCT array to maximize harvested energy [21], however, the research has been focused on the economic side and a more detailed and realistic model is desired.

This paper will advance the knowledge of optimal path planning of an OCT described by a highly nonlinear dynamic model [22] in an uncertain oceanic environment. An analytical and closed-form expression will be 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 will be 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 path planning, which distinguishes us from the studies of maximum power point tracking control in the literature [23], [24]. Furthermore, since the OCT will vary its vertical location at each time step, our path optimization problem is different from the tidal turbine control and optimization [25].

The contribution of this study is two-fold:

- This paper proposes, for the first time, a precise 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.
- This paper further formulates a novel spatiotemporal optimization problem for path planning to maximize the total harvested power of an OCT system. Under this problem formulation, a RL-based method is designed to explore the optimal control actions and results are compared with an MPC-based strategy.

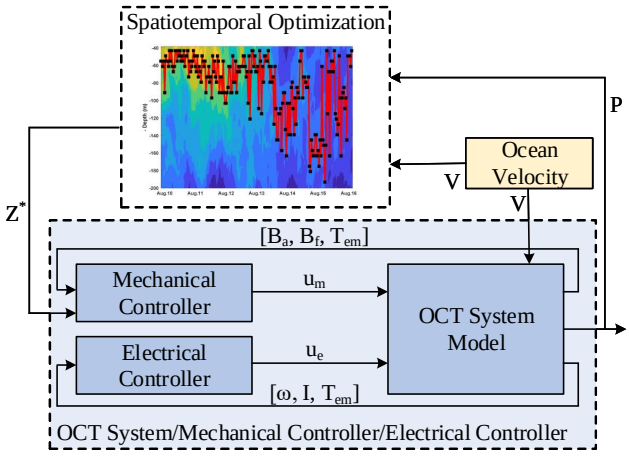

Fig. 2: Proposed spatiotemporal optimization for OCT path planning. The system will be controlled hierarchically, where the upper-level is designed to enable spatiotemporal optimization, and the lower-level is a dynamic tracking control. This paper will focus on the spatiotemporal optimization.

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 usage 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 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.

To address the complexity of this problem, a novel hierarchical spatiotemporal optimization and control framework is proposed, shown in Fig. 2. The ocean velocity is first modeled with a Gaussian process (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 will track the prescribed set-points from the upper-level and adjust the turbine dynamics.

We will 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. To be brief, at the lower-level, mechanical and electrical controllers are considered. The control inputs of the mechanical controller (i.e., buoyancy controller) are the two buoyancy 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 have developed controllers for the lower-level with a full turbine dynamics [23], [26]–[28], and the readers are referred to our previous work.

III. Environment and OCT Modeling

The high uncertainties in the ocean current velocity field are addressed, and the future ocean velocity is predicted that will be 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 location (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. In this paper, we use the data recorded by a 75 kHz acoustic Doppler current profiler (ADCP) at a latitude of $26.09°N$ and longitude of $-79.80°E$. The ocean shear profile of these data over a sample one-week period is shown in Fig. 3. These observed ocean velocity data are represented as:

$$\mathbf{X} = \begin{bmatrix} x_1 & \dots & x_m \\ \vdots & \ddots & \\ x_n & \dots & x_{n\times m} \end{bmatrix} = \begin{bmatrix} (z_1,t_1) & \dots & (z_1,t_m) \\ \vdots & \ddots & \\ (z_n,t_1) & \dots & (z_n,t_m) \end{bmatrix} \quad (1)$$

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

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 $\mathbf{X}$ contains the set of ocean depth and time, and $\mathbf{V}$ 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 $f : \mathcal{Z} \times \mathcal{T} \to \mathcal{R}$ whose output is a prediction y. The ocean velocity and especially the ocean turbulence are usually modeled with probability distributions, including log-normal distribution [29] and Burr distribution [30]. 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 Gaussian process (GP) have been 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 will be briefly introduced in the following.

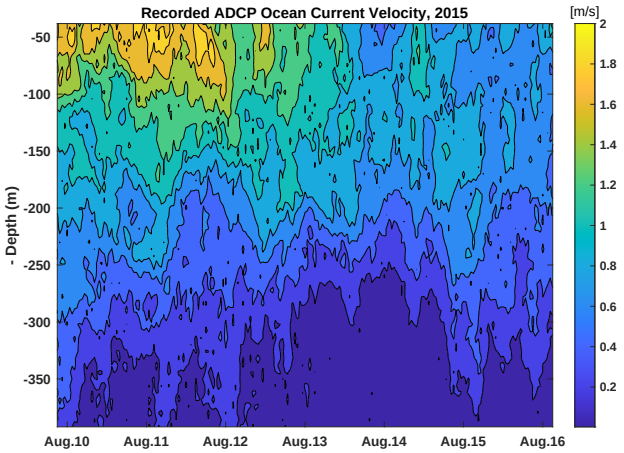

Fig. 3: Sample ocean current shear profile recorded by an ADCP in the Gulf Stream over one week.

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

Gaussian Process Model: GP is a probabilistic approach used to define a prior probability distributions over latent functions directly, which has been extensively applied in wind speed forecasting [31], [32] and ocean current flow velocity prediction [33], [34]. The prediction y is defined by a log-normal distribution so that $y = log[v(z,t)/v_c(z_1,t_1)]$, where $v_c(z_1,t_1)$ is a reference ocean current velocity. This transformation results in a random variable that can be reasonably modeled as the Gaussian process. GP model with mean $m(z,t)$ (i.e., encodes the central tendency) and covariance $k((z,t),(z^{'},t^{'}))$ (i.e., denotes the shape and structure) is defined as:

$$f(z,t) \sim \mathcal{GP}(m(z,t), k((z,t),(z^{'},t^{'}))) \quad (3)$$

The ocean velocity is predicted as $V = f(z,t) + \epsilon$, while ϵ denotes a Gaussian distribution $\mathcal{N}(0,\sigma^2)$. The joint distribution over recorded velocities $\mathbf{V}$ and prediction v_* is defined as [35]:

$$\begin{bmatrix} V \\ v_* \end{bmatrix} = (\begin{bmatrix} f \\ f_* \end{bmatrix} + \begin{bmatrix} \epsilon \\ \epsilon_* \end{bmatrix}) \sim \mathcal{N}(0, \begin{bmatrix} K_v & k_* \\ k_*^T & k_{**} + \sigma^2 \end{bmatrix}) \quad (4)$$

where f_* shows the latent function based on new input (z_*,t_*) with corresponding noise ϵ_*. k_* and k_{**} are calculated as:

$$k_* = [k((z_1,t_1),(z_*,t_*)), ..., k((z_n,t_m),(z_*,t_*)))] \quad (5)$$

$$k_{**} = k((z_*,t_*),(z_*,t_*)) \quad (6)$$

Finally, the GP model for predicting new ocean velocities is determined using the following mean $m(z_*,t_*)$ and covariance $\sigma^2(z_*,t_*)$:

$$m(z_*,t_*) = k_*^T K_v^{-1} V \quad (7)$$

$$\sigma^2(z_*,t_*) = k_{**} - k_*^T K_v^{-1} k_* + \sigma^2 \quad (8)$$

$$f(z,t) \sim \mathcal{GP}(m(z_*,t_*), \sigma^2(z_*,t_*)) \quad (9)$$

B. Mathematical Model of OCT Output Power

The output power of an OCT with variable blade pitch has been investigated numerically in [36], and we will 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 has been extended to include two variable buoyancy tanks, and a 607 m long mooring cable that attaches to the seafloor at a depth of 325 m [22]. These model parameters roughly follow the prototype systems from IHI Corp. [37], [38] and the University of Naples [39], 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 m/s. This allows for dynamic spatial control and response. The nonlinear OCT modeling techniques presented in [22], [36] are utilized when simulating this system for the creation of the linear models used in the presented formulations. When buoyancy controlled OCTs are operating, 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 $P^{HD}_{ballast}$; and
3) Power consumed by ballast pumps to Change Depth, denoted as $P^{CD}_{ballast}$.

The total harvested power from the OCT system can be calculated as:

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

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

$$P_{OCT} = \frac{1}{2}\rho A C_p v^3 = \frac{1}{2} \times 1030 \times 100\pi \times 0.415 \times v^3 \tag{11}$$

where $\rho = 1030\,kg/m^3$ is the water density, $A = 100\pi$ is the swept area of the OCT rotor, $C_p = 41.5\%$ is the average power coefficient from [36], and v is the magnitude of the ocean velocity.

Second term, $P^{HD}_{ballast}$: To calculate $P^{HD}_{ballast}$, 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., $P_{abs} \cong 0\,kPa$). Using this approach, very little power is used (i.e., $P_{fill} \cong 0$) 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 (i.e., the product of pressure and area, ΔPA) and velocity (i.e., the quotient of volumetric flow rate over the area, Q_B/A) through the orifice divided by pump efficiency:

$$P^{fill}_B = 0 \tag{12}$$

$$P^{empty}_B = \frac{F \cdot V}{\eta_{pump}} = \frac{(\Delta P \cdot A)(\frac{(Q_B)}{A})}{\eta_{pump}} = \frac{\Delta P \cdot Q_B}{\eta_{pump}} \tag{13}$$

$$\Delta P = P_{atm} + P_{HS} \tag{14}$$

where $\eta_{pump} = 0.75$ denotes the pump efficiency, $P_{atm} = 101\,kPa$ is atmospheric pressure, $P_{HS} = \rho.g.z$ is hydrostatic pressure in kPa, and $g = 9.81\,m/s^2$ is gravity. P^{empty}_B in kW can be rewritten as:

$$P^{empty}_B = \frac{(P_{atm} + P_{HS})Q_B}{\eta_{pump}} = \frac{(101 + 10.1z)Q_B}{0.75} \tag{15}$$

Assuming an OCT operating depth of $z = 50\,m$, the relationship between pump power and volumetric flow rate is determined using Eq. (15):

$$P^{empty}_{B_{max}} = \frac{606.Q_{B_{max}}}{0.75} = 808 Q_{B_{max}}[\,kW] \tag{16}$$

To add perspective, a volumetric flow rate Q_B of $0.023\,m^3/s$ is achieved at a depth of $50m$ assuming $P^{empty}_B = 18.8\,kW$, which is the ballast pump power associated with WWII submarines [40]. Since each of the two ballast tanks has a volume of 31.251 m^3 [22], these tanks can be completely emptied of water in 22.2 minutes ($\Delta F_F = -1$) with a power consumption of $37.6\,kW$, 5.4% of the simulated OCT's rated power, and a total energy usage of $14.02\,kWh$ that is independent of the fill rate.

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, $\frac{dF_F}{dz}$, 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:

$$\frac{\Delta F_F}{\Delta v} = \frac{dF_F}{dz}\frac{dz}{dv} => \Delta F_F = \Delta v \frac{dF_F}{dz}\frac{dz}{dv} \tag{17}$$

In this paper, a quasi-static relationship is assumed between these states and a steady and homogeneous flow field when running the nonlinear simulation [22]. The fill fraction in each tank is changed by $\Delta F_F = 0.55 - 0.5 = 0.05$, and the resulting quasi-static depth change is $\Delta z = 31.00 - 50.36 = -19.36\,m$. Therefore, for a flow speed of $1.6\,m/s$, the linear quasi-static relationship between Δz and ΔF_F is:

$$\frac{dF_F}{dz} \cong \frac{\Delta F_F}{\Delta z} = -0.0026[\,1/m] \tag{18}$$

Similarly, a quasi-static relationship is estimated between Δz and velocity ($\frac{dz}{dv}$) for a steady and homogeneous flow field.

The water velocity is changed by $\Delta v = 1.7 - 1.6 = 0.1\, m/s$, and the resulting quasi-static depth change is $\Delta z = 75.25 - 50.36 = 24.89$. Therefore, for an water velocity of $1.6\, m/s$ and fill fractions of 0.5, the linear quasi-static relationship between Δz and Δv is:

$$\frac{dz}{dv} \cong \frac{\Delta z}{\Delta v} = 248.9[\, m/(m/s)] \tag{19}$$

Therefore, using Eqs. (17), (18), and (19), the linear relationship between the fill fraction change necessary to maintain depth and the corresponding Δv is:

$$\Delta F_F = \Delta v(-0.0026)(248.9) = -0.65\Delta v \tag{20}$$

The average power consumed in kW over each Δt_1 in hours for a Δv in $\frac{m}{s}$ can be calculated using Eq. (20), the total energy required to fill the tanks with water in kWh, and Δt_1 is:

$$P_{ballast}^{HD} = \begin{cases} 0, & \Delta v < 0 \\ \frac{(14.02)(0.65\Delta v)}{\Delta t_1}, & \Delta v > 0 \end{cases} \tag{21}$$

Third term, $P_{ballast}^{CD}$: To calculate $P_{ballast}^{CD}$, 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 from Eq. (18):

$$\Delta F_F = \Delta z \frac{dF_F}{dz} = -0.0026\Delta z \tag{22}$$

Since each ballast tank has a volume of 31.251 m^3, we can relate the volumetric empty rate to both ballast tanks fractional fill rate $\dot{F}_F$ (i.e., $31.251\dot{v}_{B_{max}} = \dot{F}_F$) and relate $P_{B_{max}}^{empty}$ to $\dot{F}_F$ as $P_{B_{max}}^{empty} = 808 \times 31.251 \times \dot{F}_F = 25.25\dot{F}_F[\, MW]$. This can also be viewed in terms of the energy required to completely empty/fill each ballast tank. To empty either ballast tank E_{tank}^{empty} and fill either ballast tank E_{tank}^{fill} at a depth of $50\, m$ will require:

$$E_{tank}^{fill} = 0 \tag{23}$$

$$E_{tank}^{empty} = 25.25[\, MW.s] = 7.01[\, kWh] \tag{24}$$

The average power in kW consumed over Δt_2 in hours for Δz in meters can be calculated using Eq. (22), the total energy used to fill the tanks with water, and Δt_2 in hours:

$$P_{ballast}^{CD} = \begin{cases} 0, & \Delta z > 0 \\ \frac{(14.02)(-0.0026\Delta z)}{\Delta t_2}, & \Delta z < 0 \end{cases} \tag{25}$$

Therefore, based on Eqs. (10), (11), (21), and (25), the average net power of the OCT system is calculated as:

$$P_{net} = \frac{1}{2} \times 1030 \times 100\pi \times 0.415v^3 - P_{ballast}^{HD} - P_{ballast}^{CD} \tag{26}$$

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.

Algorithm 1 Upper-level Spatiotemporal Optimization

```
Initialize n data sample v(z,t) and optimization method parameters
for each sampling time Δt_2 do
    Predict v(z,t) using GP model;
    f(z,t) ~ GP(m(z_*,t_*), σ^2(z_*,t_*))
    for Prediction Horizon T do
        Apply optimization algorithm (e.g., MPC or RL)
    end for
    Output optimal depth z^*
end for
```

A. MPC-based Optimal Path Planning

MPC is considered in this paper because of its capability to handle highly 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, objective function to be optimized is defined with two terms, including 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 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 constrained optimization problem which has to be solved by the MPC design is::

$$J(z(p)) = \min_{z(p)} \sum_{i=p}^{p+T-1} [\beta_1 J_{exploitation}(z(i|p)) + \beta_2 J_{exploration}(z(i|p))] \tag{27}$$

subject to

$$z(i+1) = z_{sp}^{'}(i) \tag{28}$$

$$z^{min} \leq z(i|p) \leq z^{max} \tag{29}$$

$$f(z,t) \sim \mathcal{GP}(m(z_*,t_*), \sigma^2(z_*,t_*)) \tag{30}$$

$$\frac{z(i|p) - z(i-1|p)}{\Delta t_2} \leq r \tag{31}$$

where $J(z(p))$ is the overall objective, p denotes the p-th sampling time, T represents the prediction horizon, $z(p)$ denotes the depth trajectory at p-th sampling time, $z_{sp}(i) = [z_{sp}(p|p), ..., z_{sp}(i+T-1|p)]$ denotes a vector of depth set-points for the OCT over the prediction horizon T, and $z_{sp}^{'}(i)$ corresponds to the first element of the optimal control sequence $z_{sp}(p|p)$. $f(z,t)$ shows the GP model of ocean velocity with mean $m(z_*,t_*)$ and covariance $\sigma^2(z_*,t_*)$, r denotes the rate limitation on the speed, β_1 and β_2 are the gains of the two terms, and z^{min} and z^{max} are the minimum and maximum allowable depth. It should be noted that in our application the depth is limited within 40 m to 200 m since

Algorithm 2 Proposed MPC-based Design

```
Initialize n data sample v(z,t), i = 0, Δt1, Δt2, and T
for each sampling time Δt2 do
    Predict v(z,t) using GP method;
    f(z,t) ~ GP(m(z*,t*), σ²(z*,t*))
    for Prediction Horizon T do
        Solve J(z(p)) in (27) and obtain optimal depth
        trajectory z(p);
    end for
    Output optimal depth z*;
end for
```

the maximum ocean currents nearly always occur in the top 200 m depths [41].

The exploitation term $J_{exploitation}$ is defined as:

$$J_{exploitation}(z(i|p)) = P_{rated} - E(P(z(i|p))) \tag{32}$$

where $E(P(z(i|p))$ is the expected power based on the predicted ocean velocities, and P_{rated} is the rated power calculated by Eq. (26). 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_{rated} and $E(P(z(i|p)))$.

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

$$J_{exploration}(z(i|p)) = \sum_{z(i)} \sigma^2(v(z(i|p))|V, V^*) \tag{33}$$

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

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 Eq. (9). The objective function is calculated over the prediction horizon, while optimal depth is determined. Note that a sliding window is applied here, meaning that only the first element in the predicted optimal depth z^* will be used in each sampling time.

B. RL-based Optimal Path Planning

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 Eq. (26) 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 z over the prediction horizon results in different net power. Therefore, the set of states is defined as:

$$S = \{s_i|((z, t), P)\} \tag{34}$$

in which, the number of states is equal to n^T if the number of operating depth z 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:

$$A = \{a_i| + \Delta z, -\Delta z, 0\} \tag{35}$$

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:

$$R = \begin{cases} \Delta P, & \Delta P > \delta \text{ or } \Delta P < -\delta \\ 0, & otherwise \end{cases} \tag{36}$$

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 algorithm [42] 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:

$$\begin{aligned} Q_{t+1}(s_t, a_t) \leftarrow Q(s_t, a_t) + \\ \alpha[R_{t+1} + \gamma max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \end{aligned} \tag{37}$$

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:

$$Action = \begin{cases} A \leftarrow \underset{a}{argmax}\, Q, & \text{with probability } 1 - \epsilon \\ A \leftarrow a(random), & \text{with probability } \epsilon \end{cases} \tag{38}$$

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

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 $episode = 1, 2, \ldots, N_{episode}$ **do**
    Initialize s $((z,t), P)$;
    for $step = 1, 2, \ldots, N_{step}$ **do**
      Select action a according to ϵ-greedy Eq. (38);
      Take action a and obtain R_{t+1} by Eq. (36);
      Update the Q-value according to Eq. (37);
    end for
  end for
end for
Output offline trained optimal Q-value table Q^*;

Algorithm 4 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^*;
if $t = 1$ **then**
  Select action a according to Q^*;
else
  for Each t **do**
    Observe previous recorded velocity $v(z, t-1)$;
    for Prediction Horizon T **do**
      Predict $v(z,t)$ using GP method;
      $f(z,t) \sim \mathcal{GP}(m(z_*, t_*), \sigma^2(z_*, t_*))$
    end for
    Update optimal Q-value table Q^* according to the observed velocity and predicted velocity over T $(v(z, t:t+T))$;
    Select action a according to the updated Q^*;
  end for
end if
Output optimal depth z^*;

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 the action, a new state (a new water depth) will be observed, and the reward will be calculated by Eq. (36). Finally, the Q-function Eq. (37) is updated based on the new state and the reward. Note that after trained on all the training dataset, the optimal Q-value table Q^* is obtained and will be deployed online while it keeps updating according to the observed ocean velocity through an incremental learning strategy.

V. Results and Discussions

A. Simulation Setup

We will use real ocean velocity data and the example buoyancy-controlled OCT that was discussed in Section III. Recorded ocean velocities V are modeled by GP model in Eq. (9). At each time step, the next ocean velocity v_{n+1} is predicted over the prediction window. 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.

TABLE II: Key parameters used in the optimization.

Parameter	Value	Parameter	Value
Simulation Time	200 hour	δ	1 kW
γ	0.8	β_1	1
β_2	100	Δt_1	0.25 hour
Δt_2	1 hour	T	$2\Delta t_2 = 2$ hours

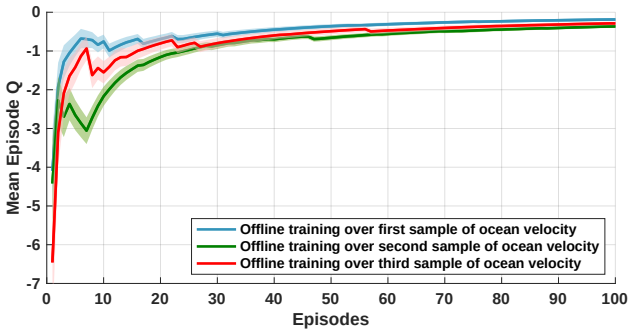

Fig. 4: Average Q-values over 100 trial episodes are shown for offline training over the first sample of ocean velocity, second sample of ocean velocity, and third sample of ocean velocity. Solid lines show average Q-values, and the shaded region determines a 95% confidence interval.

To evaluate the performance of the proposed methods for optimal path planning, the following three approaches are compared:

- **Case without Spatiotemporal Optimization:** The OCT is located at a fixed depth, 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.
- **MPC-based Spatiotemporal Optimization:** The optimal operating depth z^* is determined using objective function $J(z(p))$ in Eq. (27), 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 $J_{exploration}$.
- **RL-based Spatiotemporal Optimization:** Learning from different experiments should be considered to update the Q-values table. Specifically, the Q-function is calculated for each state and action pair and the Q-values 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 is shown in Fig. 4, which verifies the convergence of Q over 100 episodes for different sampling time.

Other key parameters used in the simulation are presented in Table II. All the experiments were conducted on a machine equipped with a 2.6 GHz CPU and 16 GB of RAM.

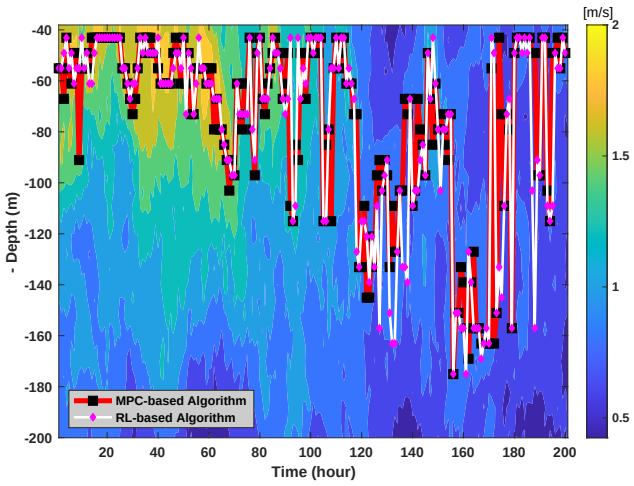

Fig. 5: Comparing optimal path obtained over 200 hours under MPC-based optimization and RL-based optimization.

B. Comparative Results

The obtained results over a sample simulation time of 200 hours are presented in Fig. 5 to Fig. 7. The optimal path obtained over 200 hours with MPC-based optimization and online testing of RL-based optimization are illustrated in Fig. 5. The control action is defined as the depth change in this study, and the optimal depth and the corresponding velocity are shown under the MPC-based method, RL-based method, and the case without spatiotemporal optimization. As presented in Fig. 6(a), the optimal depth is similar for both MPC-based and RL-based algorithms. However, there exist some differences in selecting the next control action, such as at $t = 130$. Although two different z are selected as the next optimal depth for MPC and RL algorithms, the velocities at the two depths are nearly the same (Fig. 6(b)), verifying that both algorithms can make optimal selections.

The harvested power of the OCT system is presented in Fig. 7, which shows that the output power of OCT increases when the spatiotemporal optimization is applied. MPC-based optimization and RL-based method follow a similar trend in power increase. However, RL-based spatiotemporal optimization surpasses the MPC method at some time samples. In addition, the average power of the OCT system in terms of P_{OCT}, $P^{HD}_{ballast}$, $P^{CD}_{ballast}$, and P_{net} for three different operating depths are presented in Table III.

C. 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. 8 shows the perturbation of cumulative energy from baseline obtained by the MPC algorithm and the RL algorithm in response to 5% noise disturbances of the same intensities for 100 test cases. We can observe that both algorithms are sufficiently robust, where similar distribution of the results (i.e., cumulative energy) are

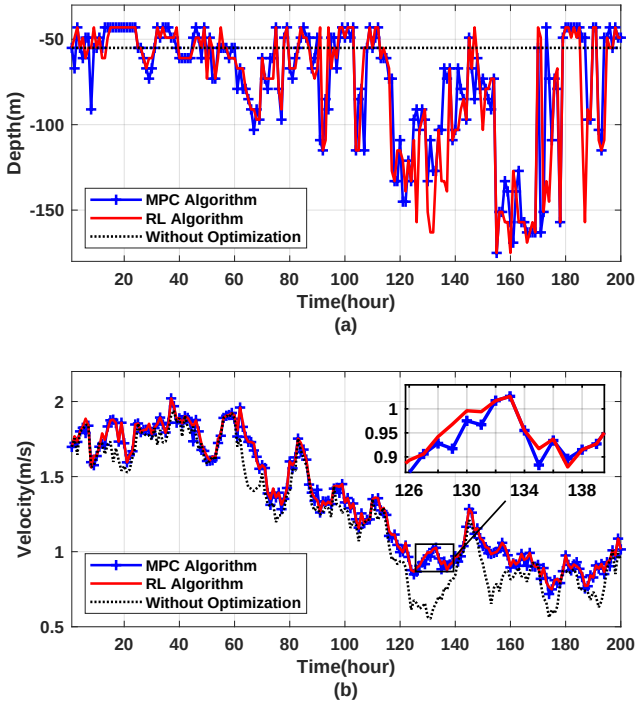

Fig. 6: Comparing optimal depth and velocity over 200 hours obtained by MPC algorithm; online testing of RL algorithm, and case without spatiotemporal optimization. (a) Optimal depth; (b) Optimal velocity.

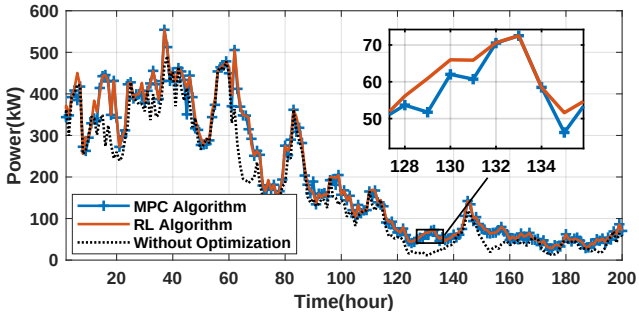

Fig. 7: Comparing optimal power under MPC-based optimization, online testing of RL-based optimization and case without spatiotemporal optimization.

obtained for perturbed ocean velocity model. Moreover, the RL algorithm 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 both MPC and RL algorithms in response to 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. Four cases of perturbed ocean velocity are generated with 5%, 10%, 15%, and 20% noise. As can be seen from the table, the superiority of the RL-based method over the MPC-based method is verified under these four cases, justifying that

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

$z\ (Time=0)$	P_{OCT}	$P^{HD}_{ballast}$	$P^{CD}_{ballast}$	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
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				
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

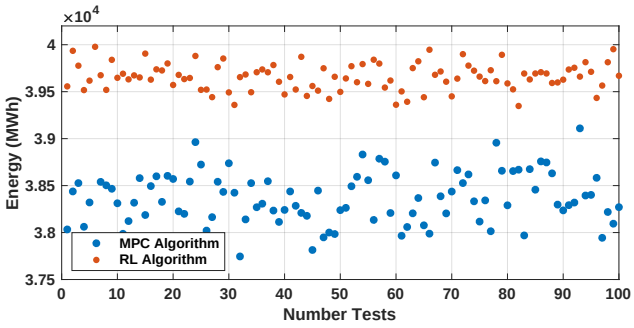

Fig. 8: Robustness comparison of the MPC algorithm and RL algorithm through cumulative energy under noise.

the MPC algorithm is more sensitive to the increased ocean velocity modeling error.

D. Discussions

RL algorithm as a learning-based method usually has offline training and online deployment phases. Once the offline training is finished, it will be deployed online while it keeps updating the policy (e.g., the Q-value table). Both RL and MPC algorithms are feasible for path planning to maximize the output power of the OCT system. Another important feature of the presented method is its robustness to errors and uncertainties. In the MPC algorithm, the exploration term $J_{exploration}$ in the objective function $J(z)$ is defined to address this issue. However, the main challenge is to determine the corresponding weight β_2 of this term $J_{exploration}$, 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.

Finally, the cumulative produced energy of the OCT system using MPC and RL algorithms are compared, as shown in Fig. 9. The cumulative energy shows a high increase compared to the case without optimization, highlighting the importance of applying spatiotemporal optimization. We also observe that the cumulative energy productions of the OCT system are very close for the two compared optimization approaches. RL algorithm outperforms the MPC method, and

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

Noise	MPC Algorithm		RL Algorithm	
%	MWh	%	MWh	%
Baseline Ocean Velocity Model				
-	39.653	-	40.366	-
Ocean Velocity Model Perturbed with Noise				
5	38.364	3.25	39.658	1.75
10	37.594	5.19	38.720	4.08
15	36.998	6.70	38.158	5.47
20	36.430	8.13	37.195	7.85

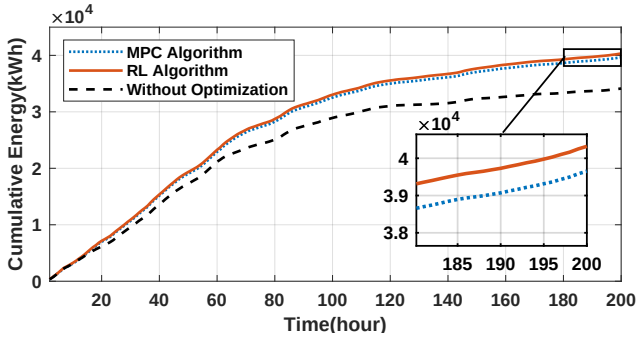

Fig. 9: Comparing cumulative energy under MPC-based optimization, online testing of RL-based optimization, and case without spatiotemporal optimization.

the final energy production for MPC-based optimization, RL-based optimization, and case without applying spatiotemporal optimization are 39.653 MWh, 40.366 MWh, and 34.121 MWh, respectively.

VI. Conclusions

In this study, a novel spatiotemporal optimization approach was presented for OCT 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, the consumed power for stabilizing, and the consumed power for changing the operating depth. Two approaches, including MPC- and RL-based algorithms, were developed to solve the proposed spatiotemporal problem. The obtained results verified that both methods are efficient in finding the optimal path to maximize the output power. Moreover, the RL-based method showed better performance than the MPC-based method 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] 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.

[5] W. Caharija, K. Y. Pettersen, M. Bibuli, P. Calado, E. Zereik, J. Braga, J. T. Gravdahl, A. J. Sørensen, M. Milovanović, and G. Bruzzone, "Integral line-of-sight guidance and control of underactuated marine vehicles: Theory, simulations, and experiments," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 5, pp. 1623–1642, 2016.

[6] C. Shen, Y. Shi, and B. Buckham, "Path-following control of an auv: A multiobjective model predictive control approach," *IEEE Transactions on Control Systems Technology*, vol. 27, no. 3, pp. 1334–1342, 2018.

[7] 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.

[8] D. L. Chen and G. P. Liu, "Coordinated path-following control for multiple autonomous vehicles with communication time delays," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 5, pp. 2005–2012, 2020.

[9] 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.

[10] 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.

[11] 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.

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

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] A. Hasankhani, J. VanZwieten, Y. Tang, B. Dunalp, A. D. Luera, C. Sultan, and N. Xiros, "Modeling and numerical simulation of a buoyancy controlled ocean current turbine," *International Journal of Marine Energy*, Submitted.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] B. Pearson and B. Fox-Kemper, "Log-normal turbulence dissipation in global ocean models," *Physical review letters*, vol. 120, no. 9, 2018.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] The fleet type submarine online submarine trim and drain systems (chapter 2). [Online]. Available: https://maritime.org/doc/fleetsub/trim/chap2.htm

[41] 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.

[42] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.