Ocean Current Turbine Power Maximization: A Spatiotemporal Optimization Approach

Arezoo Hasankhani, Student Member, IEEE, Yufei Tang, Member, IEEE, and James VanZwieten

Abstract—This paper presents a novel spatiotemporal optimization approach for maximizing the output power of an ocean current turbine (OCT) under uncertain ocean velocities. In order to determine output power, ocean velocities and the power consumed and generated by an OCT system are modelled. The stochastic behavior of ocean velocities is a function of time and location, which is modelled as a Gaussian process. The power of the OCT system is composed of three parts, including generated power, power for maintaining the system at an operating depth, and power consumed for changing the water depth to reach the maximum power. Two different algorithms, including model predictive control (MPC) as a model-based method and reinforcement learning (RL) as a learning-based method are proposed to design the optimization structure, and comparative studies are presented. On one hand, the MPC based controller is faster in finding the optimal water depth, while the RL is also computational feasible considering the required time for changing operating depth. On the other hand, the cumulative energy production of the RL algorithm is higher than MPC method, which verifies that learning-based RL algorithm can provide a better solution to address the uncertainties in renewable energy systems. Results verify the efficiency of both presented methods in maximizing the total power of an OCT system, where the total harnessed energy after 200 hours shows over 18% increase compared to the baseline.

Index Terms—Ocean current turbine, power maximization, 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] and promises great socioeconomic impacts [2]. 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), which are mostly located in the coastal areas with high population densities [3], [4]. The main challenges in developing MHK based energy are high investment and maintenance costs (e.g., hostile operating environment) and difficulty of integration their produced electricity to the electrical grid (e.g., remote areas and deep sea) [5], [6]. There are many approaches to address the high cost challenges, such as lower the cost of turbine design or optimized operational strategy [7], and this study focuses on maximizing the output power of an ocean current turbine (OCT).

Power maximization algorithms studied for similar renewable energy systems, such as wind turbines, mainly focused on generator control or blade pitch/yaw control. Power maximization of an airborne wind turbine is studied using a hierarchical controller [8], in which the upper level is responsible for setting the trajectory and the lower level controls the generator. Extreme learning machine is used to control and maximize the output power of a hydro-static tidal turbine, which shows better performance under uncertain and turbulent environments compared to conventional controllers [9]. Nonlinear backstepping control is applied for maximum energy harvesting from a variable speed wind turbine with claimed advantages of fast performance and reduced mechanical stress [10]. An adaptive controller is proposed to maximize the output power of wind turbines considering the uncertainties in wind and turbine dynamics simultaneously [11], where the controller includes Smith predictor to reduce the stochastic behavior. In a recent study, a recurrent neural network is combined with integral sliding mode controller to minimize the uncertainties in wind turbine dynamics and maximize the output power [12].

Among many of the turbine controller designs, model predictive control (MPC) and reinforcement learning (RL) have gained increasing attentions in recent years. MPC controller is applied for setting the yaw system in wind turbines in order to set wind turbine in the forecasted wind direction [13]. A fault-tolerant MPC controller is designed for wind turbines to maximize output power and minimize fatigue at the same time [14]. An airborne wind energy system employs MPC-based techniques to set the optimal altitude [15], in which future wind speeds are predicted and output power is optimized by changing the location of the system to optimal velocities. In another close study [16], performance of the MPC method and a Lyapunov-based controller are compared for optimizing the output power of an airborne wind energy system. On the other hand, RL based design is applied for wind turbine power optimization [17], [18], where the Q-learning algorithm is used to find the optimal generator speed. In another application, RL is used to maximize the harvested energy from wave energy converters due to high nonlinearity of these systems [19], [20].

OCT control has been investigated from a different point of view in previous studies. The blade pitch control of OCT system is addressed in [21] in order to stabilize the mechanical operation of the system. A MPC-based controller is designed using blade pitch and torque as the input to stabilize the operation of an OCT system [22]. Further, a permanent magnet
synchronous generator is controlled through an adaptive super-twisting sliding mode algorithm to assure stable operation of the OCT under stochastic ocean velocity [23]. An estimator is designed to predict the spatially dependent velocity of ocean, and a MPC-based controller is used to optimize the location of an underwater vehicle used as an energy harvesting system [24]. A Bayesian optimization is used to specify the configuration and location to harvest maximum energy from an OCT array [25]. However, the relation between OCT systems and the high uncertainties in the ocean velocity are not precisely described.

Based on aforementioned discussions, critical scientific gaps are identified in ocean current energy harvesting: the lack of a model to precisely describe the power usage in OCT systems, and the need of an optimization framework with effective solution strategy. The contribution of this study is two-fold:

- This paper proposes, for the first time, a precise model to represent the harvested power of an OCT system, which includes directly generated power from hydrokinetic, consumed power for stabilizing the system at specified water depths, and consumed power to reach the new optimal operating depths.
- This paper further formulates a novel spatiotemporal optimization problem for maximizing the total harvested power of an OCT system. Specifically, a RL-based method is designed to explore the optimal control actions and the results are compared with a MPC-based strategy. This paper compares a learning-based method (i.e., RL) with a model-based method (i.e., MPC) for renewable energy optimization problem. Both methods can address the dynamic response and uncertainties in renewable energies, and can provide the power optimization over different prediction horizons while considering the prediction error.

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

II. ENVIRONMENT AND OUTPUT POWER MODELING

In this section, the ocean velocity is first modelled. The high uncertainties in the ocean current velocity field should be addressed by recognizing the pattern of velocity spatially and temporally. The pattern recognition of ocean velocity is useful for predicting the future ocean velocities. The total generated power of an OCT system, which is a function of the ocean velocity, is then modelled in detail.

A. Statistical Ocean Current Shear Profile Characterization

In order to increase the robustness of the designed spatiotemporal optimization method, the ocean velocity should be modelled. The ocean velocity field varies both with time and 3-D space. However, moored OCTs can primarily vary their vertical location 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 in Southeast Florida by a bottom mounted acoustic Doppler current profile (ADCP) between February 2009 and March 2010 [26]. The ocean shear profile of these data with 4-hour intervals for one sample day is shown in Fig. 2. The ocean velocity data are stored in \( V \) matrix as follows:

\[
V_{n \times m} = \begin{bmatrix}
v(z_1, t_1) & \ldots & v(z_1, t_m) \\
v(z_2, t_1) & \ldots & v(z_2, t_m) \\
\vdots & \ddots & \vdots \\
v(z_n, t_1) & \ldots & v(z_n, t_m) 
\end{bmatrix}
\]  

(1)

where \( n \) and \( m \) show the number of discrete depths and the number of time samples, and \( v(z_1, t_1) \) is the measured ocean velocity at \( z_1 \) and \( t_1 \). Matrix \( V \) includes \( n \times m \) recorded ocean velocities (i.e., \( m \) velocity vectors each as a function of depth).

The ocean velocity and especially the ocean turbulence are modeled with different probability distributions, including log-normal distribution [27] and Burr distribution [28]. In this paper, different methods for modelling the ocean velocity are developed. The Gaussian process modeling shows the best performance, which is compared with other methods in Table I. The Gaussian process modeling is presented in the following.

**Gaussian Process Model:** The Gaussian process (GP) is a non-parametric probabilistic approach used to define a prior probability distributions over latent functions directly, which has been extensively applied in wind forecasting [15], [29]. As we will show, GP can also be utilized to predict the ocean velocity. The observed data are first represented as follows:

\[
X = \begin{bmatrix}
x_1 & \ldots & x_m \\
\vdots & \ddots & \vdots \\
x_n & \ldots & x_{n \times m}
\end{bmatrix} = \begin{bmatrix}
(z_1, t_1) & \ldots & (z_1, t_m) \\
\vdots & \ddots & \vdots \\
(z_n, t_1) & \ldots & (z_n, t_m)
\end{bmatrix}
\]  

(2)

![Fig. 1. Sample daily shear profile recorded at 4-hour intervals.](image)

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>RMSE</th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear regression</td>
<td>0.0369</td>
<td>0.0144</td>
<td>0.0297</td>
</tr>
<tr>
<td>Regression trees</td>
<td>0.0614</td>
<td>0.0038</td>
<td>0.0302</td>
</tr>
<tr>
<td>Support vector machines</td>
<td>0.0232</td>
<td>0.0005</td>
<td>0.0205</td>
</tr>
<tr>
<td>Gaussian process model</td>
<td>0.0031</td>
<td>9.6 × 10⁻⁶</td>
<td>0.0022</td>
</tr>
</tbody>
</table>
where $X$ denotes the set of ocean depth and time, and $V$ shows recent $n \times m$ current velocity observations at the corresponding time and depth. Let $Z$ denote the domain of allowable ocean depth choices and $T$ represent the temporal space. Our goal is to model ocean current velocity at depth $z \in Z$ and time $t \in T$, by constructing a function $f : Z \times T \to R$ whose output is a prediction $y$. The prediction $y$ is defined by a log-normal distribution so that $y = \log[v(z,t)/v_0(z_1,t_1)]$, where $v_0(z_1,t_1)$ is a reference ocean current velocity. This transformation results in a random variable that can be reasonably modeled as the Gaussian process (GP). 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:

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

(4)

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

$$[V, v_\ast] = ([f, f_\ast] + [\epsilon, \epsilon_\ast]) \sim \mathcal{N}(0, \begin{bmatrix} K_{\epsilon} & k_{\epsilon,\ast} \\ k_{\ast,\epsilon} & K_{\ast,\ast} + \sigma^2 \end{bmatrix})$$

(5)

where $f_\ast$ shows the latent function based on new input $(z_\ast, t_\ast)$ with corresponding noise $\epsilon_\ast$, $k_\ast$ and $k_{\ast,\ast}$ are calculated as:

$$k_\ast = k((z_1,t_1),(z_\ast, t_\ast)), ..., k((z_n,t_n),(z_\ast, t_\ast)))$$

(6)

$$k_{\ast,\ast} = k((z_\ast, t_\ast),(z_\ast, t_\ast))$$

(7)

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

$$m(z_\ast, t_\ast) = k_{\ast}^T K_\nu^{-1} V$$

(8)

$$\sigma^2(z_\ast, t_\ast) = k_{\ast,\ast} - k_{\ast}^T K_\nu^{-1} k_{\ast} + \sigma^2$$

(9)

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

(10)

B. Mathematical Model of Output Power of OCT

The output power of an OCT system with variable blade pitch is investigated numerically in [31], and we will use these rotor performance characteristics in this study. These model parameters are for a 700 kW single 20 m diameter variable pitch rotor turbine. This OCT has been extended to include two variable buoyancy tanks and a 607 m long mooring cable that attaches to the sea floor at a depth of 325 m [32]. These model parameters roughly follow the prototype systems from IHI Corp. [33], [34] and the University of Naples [35], but with the variable ballast tanks sized to operate at a depth of 50 m when half filled with ballast water and operating at the Gulf Stream’s mean flow speed off Southeast Florida of 1.6 m/s. This allows for the dynamic spatial control and response. The nonlinear OCT modeling techniques presented in [31], [32] are utilized for dynamically simulating this system which are used to create the linear models used in the present formulations.

When buoyancy controlled OCTs are maneuvering, the total output power of the system consist of three primary parts:

1) Generated power of the OCT system from hydrokinetic energy extraction, denoted as $P_{OCT}$;
2) Power for maintaining the OCT system at a targeted operating depth in a time varying current, which is related to pump power required to change ballast fill values, denoted as $P_{ballast}$;
3) Power consumed ballast pumps in order to change the water depth to reach the maximum output power, denoted as $P_{\Delta z}$.

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

$$P_{net} = P_{OCT} - P_{ballast} - P_{\Delta z}$$

(11)

First term $P_{OCT}$: The generated power of OCT system is typical and related to ocean velocity shown in the following:

$$P_{OCT} = \frac{1}{2} \rho Av^3 C_p = \frac{1}{2} \times 1030 \times 100 \times 0.415 \times v^3$$

(12)

where $\rho$ is the water density (1030 kg/m³), $A$ is the effective cross-sectional area related to OCT (100π), $v$ determines the ocean velocity, and $C_p$ is the average power coefficient of 41.5 % from [31].

Second term $P_{ballast}$: In order to calculate $P_{ballast}$, the ballast model and its consumed power and energy should be determined. This 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 force (i.e., product of pressure and area, $\Delta P A$) and velocity (i.e., quotient of volumetric flow rate over area, $Q_{B_{max}}/A$) through the orifice divided by pump efficiency:

$$P_{fill} = \frac{F \cdot V}{\eta_{pump}} = \frac{(\Delta P A)(Q_{B_{max}})}{\eta_{pump}} = \Delta P \cdot Q_{B_{max}} \eta_{pump}$$

(14)

$$\Delta P = P_{\text{atm}} + P_{\text{HS}}$$

(15)

where $\eta_{pump}$ denotes the pump efficiency assumed 0.75 in our study. $P_{\text{atm}}$ is atmospheric pressure (101 kPa), $P_{\text{HS}}$ is hydrostatic pressure ($P_{\text{HS}} = \rho g z$), and $g = 9.81$ m/s² is gravity. Then the $P_{empty}^{B_{max}}$ is rewritten as:

$$P_{empty}^{B_{max}} = \frac{(P_{\text{atm}} + P_{\text{HS}}) \cdot Q_{B_{max}}}{\eta_{pump}}$$

(16)

Assuming an OCT operating depth of $z \cong 50$ m, the relationship between pump power and volumetric flow rate is determined using (16):

$$P_{empty}^{B_{max}} = \frac{606 Q_{B_{max}}}{0.75} = 808 Q_{B_{max}} [kW]$$

(17)

The volumetric flow rate $Q_{B_{max}}$ is estimated 0.023 m³/s assuming $P_{empty}^{B_{max}} = 18.8$ kW [36]. Since each of the two
ballast tanks have volumes of 31.251 m³; these tanks can be completely emptied of water in 22.2?? minutes with a power consumption of 37.6 kW, and a total energy usage of 14.02 kWh.

\[ P_{ballast} = ??? \tag{18} \]

**Third term \( P_{\Delta z} \):** In order to model the third part of generated power \( P_{\Delta z} \), the linear relation between fill fraction \( F_F \) and depth change \( \Delta z \) should be determined. In our paper, a simple quasi-static relationship is estimated between these states for a steady and homogeneous flow field. The fill fraction is changed in each tank by \( \Delta F_F \), and the quasi-static change is recorded in \( \Delta z \). Therefore, for a flow speed of 1.6 m/s, we can calculate the quasi-static linear relationship between the change in depth and the change in fill fraction according to the following:

\[ \frac{d\Delta z}{dF_F} \cong \frac{d\Delta z^+ - d\Delta z^-}{dF^+_F - dF^-_F} \tag{19} \]

The linear relation between \( \Delta F_F \) and \( \Delta z \) is calculated at \(-179.78 m\) using (19). Since each ballast tank has a volume of 31.251 m³, we can relate the volumetric empty rate to the both ballast tanks fractional fill rate \( \dot{F}_F \) (i.e., 31.251 \( \dot{v}_{\text{max}} = \dot{F}_F \)) and we can relate \( P_{\text{empty}}^F \) to \( F_F \) as \( P_{\text{empty}}^F = 803 \times 31.251 \times \dot{F}_F = 25.25\dot{F}_F [MW] \). This can also be viewed in terms of the energy required to completely empty/fill each ballast tank.

To empty either ballast tank \( E_{\text{tank}}^\text{empty} \) and fill either ballast tank \( E_{\text{tank}}^\text{fill} \) at a depth of 50 m will require:

\[ E_{\text{tank}}^\text{fill} = 0 \quad \tag{20} \]

\[ E_{\text{tank}}^\text{empty} = 25.25[MW.s] = 7.01[kW.h] \quad \tag{21} \]

Finally, the power \( P_{\Delta z} \) required to move the OCT system \( \Delta z \) upwards can be estimated as:

\[ P_{\Delta z} = \frac{2 \times E_{\text{tank}}^\text{empty} \Delta z}{179.9} = \frac{50.50 \Delta z}{179.78} = 280.9\Delta z[kW.s] \quad \tag{22} \]

so, based on (11), (12), (18), and (22), the net power of the OCT system is calculated as:

\[ P_{\text{net}} = \frac{1}{2} \rho Av^3C_p - (101+10.1z) \times 0.03067 - 280.9\Delta z \tag{23} \]

**III. PROPOSED METHODOLOGY**

**A. Problem Formulation and Overall Control Framework**

Schematic diagram of the studied problem for maximizing generated power of an OCT system through a spatiotemporal approach is shown in Fig. 2. According to the predicted ocean velocity, the OCT system location should be determined at each time step to maximize the output power.

The proposed control structure for spatiotemporal optimization is shown in Fig. 3, and the solution procedure is presented in Algorithm 1. Two approaches including RL algorithm and MPC method are considered in our paper. The mechanical and electrical controllers are considered in relation with OCT system using the control commands (i.e., \( u_m \) and \( u_e \)) and control inputs in the low-level. The control inputs to the mechanical controller (i.e., flight controller using the variable buoyancy control) include 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 \( T_{em} \). Moreover, electromechanical torque \( T_{em} \), current \( I \) and rotor angular speed \( \omega \) are the control inputs to the electrical controller. The ocean velocity is modeled with \( GP \) method, and the forecasted velocity is used to calculate the output power of OCT system. The output power \( P \) is received by the spatiotemporal optimization at each time step in the high-level, while the optimal water depth \( z^* \) is determined according to our proposed optimization method.

**B. MPC-based Optimization Design**

MPC is a powerful method for optimizing the objective function by predicting the future actions. The MPC-based algorithm has been used for maximizing the output power of an airborne system [15] and minimizing the operational cost of wind turbine considering the fault-tolerant method [37]. In this paper, MPC method is used in order to maximize the output power of the OCT system. Firstly, the objective function should be defined, which includes both exploration and exploitation.

The exploitation term in the objective function is determined in order to find the optimal water depth where the net power is maximized. In fact, different water depth \( z \) are explored in order to find the optimal one over the prediction horizon. At the same time, the uncertainties in ocean velocity prediction should be considered in the objective function, so the exploration term is determined in the objective function to penalize the predicted velocity with higher variance. The objective function is presented as follows:
Fig. 3. Proposed spatiotemporal optimization structure to maximize OCT system power.

\[ \min_{z(p)} J(z(p)) = \sum_{i=p}^{p+T-1} [\beta_1 J_{\text{exploitation}}(z(i)|p) + \beta_2 J_{\text{exploration}}(z(i)|p)] \tag{24} \]

subject to

\[ 40 \leq z(i)|p \leq 200 \tag{25} \]

\[ f(z,t) \sim \mathcal{GP}(m(z_*, t_*), \sigma^2(z_*, t_*)) \tag{26} \]

where \( J(z(p)) \) is the overall objective, \( p \) denotes \( p \)-th sampling time, \( T \) represents the prediction horizon, \( z(p) \) denotes the depth trajectory at \( p \)-th sampling time, \( f(z,t) \) shows the GP model of ocean velocity with mean \( m(z_*, t_*) \) and covariance \( \sigma^2(z_*, t_*) \), and \( \beta_1 \) and \( \beta_2 \) are the multiples of two terms. The exploitation term of the objective function \( J_{\text{exploitation}} \) is defined as follows:

\[ J_{\text{exploitation}}(z(i)|p) = P_{\text{rated}} - E(P(z(i)|p)) \tag{27} \]

where \( E(P(z(i)|p)) \) shows the expected power according to the predicted ocean velocities, and \( P_{\text{rated}} \) shows the rated power calculated by (23) at rated velocity (i.e., 2 m/s). Therefore, the main objective is to reach to the rated power at each time step, which is defined as minimizing the difference between \( P_{\text{rated}} \) and \( E(P(z(i)|p)) \). The exploration term of the objective function is defined as follows:

\[ J_{\text{exploration}}(z(i)|p) = \sum_{z(i)} \sigma^2(v(z(i)|p))|V, V^* \tag{28} \]

where \( \sigma^2(v) \) shows the variance of predicted ocean velocities for all operating water depth \( z(i) \), \( V \) is the recorded velocity at operating depth and time defined in (1), and \( V^* \) denotes all new velocity measurements over the future horizon between \( p \) taken up to step \( i \).

Algorithm of OCT system power maximization using MPC is presented in Algorithm 2. Firstly, all recorded ocean velocities and prediction horizon \( T \) should be initialized. Then, \( v \) is predicted for each time sample using (10). The objective function is calculated over the prediction horizon, while optimal depth is determined.

Algorithm 2 Proposed model predictive control algorithm to solve the spatiotemporal optimization

1: Initialize \( n \) data sample \( v(z, t), i = 0 \) and \( T \)
2: for each sampling time do
3: Predict \( v(z, t) \) using \( \mathcal{GP} \) method;
4: \( f(z, t) \sim \mathcal{GP}(m(z_*, t_*), \sigma^2(z_*, t_*)) \)
5: for Prediction Horizon \( T \) do
6: Solve \( J(z(p)) \) in (24) and obtain optimal depth trajectory \( z(p) \);
7: end for
8: Output optimal depth;
9: end for

C. RL-based Optimization Design

The RL method takes the perspective of an agent (i.e., an OCT system in this study) that optimizes its behavior by interacting with its environment and learning from the feedback received. In RL approach, the set of action is done by the agent, and it receives the reward from environment. Therefore, the learning procedure is completed for agent from observing its resulted reward. In our problem, it is critical to define the set of states, \( S \), actions, \( A \), and rewards, \( R \). The long-term performance is optimized by learning a policy \( \pi_{\theta}(a_i|s_i) \) for picking actions in state transition in order to maximize the total accumulated reward\

\[ R_{T} = \sum_{t=0}^{T} \gamma^t r_{i+t}, \]\

where \( \gamma \) is the discount factor in (0, 1]. In another word, the action of depth change at each time step should be determined to maximize the net power by the OCT.

1) State space: The net power of the OCT system is calculated in (23), while it is dependent on the water depth \( z \). As a result, the specific power is calculated for each \( z \). Further, the problem gets more complicated when the prediction horizon \( T \) is considered. Transition between different water depth \( z \) over the prediction horizon results in different net power. Therefore, the set of states is defined as follows:

\[ S = \{s_i|((z, t), P)\} \tag{29} \]

in which, if the number of operating depth \( z \) is equal to \( n \), and the prediction horizon is considered \( T \) time step, the number of states is equal to \( n^T \). Hence, increasing the prediction horizon can increase the complexity of the problem exponentially, so it is important to choose the prediction horizon precisely in order to avoid the curse of dimensionality.

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

\[ A = \{a_i|+\Delta z, -\Delta z, 0\} \tag{30} \]

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: Reward function determines the reward received after each action. Further, the reward function should be defined to find the desirable actions. In our spatiotemporal power maximization problem, increasing power
is presented in this algorithm, the action (depth change) is selected based on $\epsilon$-greedy policy at each time step. After taking action $a$ (depth change), the new state (i.e., new water depth) should be observed, and the reward should be calculated by (31). Finally, the Q-function (32) is updated based on the new state and the reward.

### IV. RESULTS AND DISCUSSIONS

#### A. Simulation Setup

In order to show the efficiency of the proposed method, it is tested with real ocean velocity data and the example OCT system discussed in Section II. The recorded ocean velocities $V$ are modeled by GP model in (10). After modeling ocean velocities, the future velocities are predicted, and the optimal ocean depth $z$ is determined to maximize the total power. 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 sliding prediction window. The parameters of simulation are presented in Table II.

The applications of the spatiotemporal optimization for maximizing output power of the OCT system are verified through three comparative approaches as follows:

**Case without Spatiotemporal Optimization:** In this case, the OCT system is located in a constant depth, so there is no change in operating depth, i.e., the output power of the system is determined only through variations in the ocean velocity at the operating depth.

**Spatiotemporal Optimization by MPC based Algorithm:** In this algorithm, the optimal water depth $z$ is determined using objective function $J(z(p))$ in (24), which aims to minimize the difference between the maximum power (i.e., rated power) and the harvested power at each time step. Further, predicting the ocean velocity results in uncertainties, which is addressed through the exploration term $J_{exploration}$ of the objective function.

**Spatiotemporal Optimization by RL based Algorithm:** RL algorithm as a learning-based method should learn from different experiments, while the online learning procedure should be considered for the OCT system to choose the optimal water depth $z$. As it is explained in Section III, the OCT system can choose between three different actions at each time step, including depth increase, depth decrease, and no change in operating depth. Hence, the Q-function is calculated for each state and possible actions by (32) and Q-values table is updated. At each state $(z, t)$, the optimal action (i.e., depth change $\Delta Z$) should be determined to maximize the total power. OCT is trained continuously while the system is online, improving performance all the time. The cumulative Q-values over trial episodes is shown in Fig. 4, which verifies the convergence of $Q$ over 100 episodes for three different sampling time (i.e., $Sampling Time = 1$, $Sampling Time = 2$ hour)

#### Table II. Simulation parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>200 hour</td>
<td>$\delta$</td>
<td>0.01</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.8</td>
<td>$\beta_1$</td>
<td>1</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>100</td>
<td>$T$</td>
<td>2 hour</td>
</tr>
</tbody>
</table>

#### Algorithm 3 Proposed reinforcement learning algorithm to solve the spatiotemporal optimization

1. Initialize $n$ data sample $v(z, t)$
2. for each sampling time do
3. Predict $v(z, t)$ using GP method;
4. $f(z, t) \sim GP(m(z, t_s), \sigma^2(z, t_s))$
5. Initialize $Q(s, a)$ for all $s \in S$ and $a \in A$
6. for episode = $1, 2, \ldots, N_{episode}$ do
7. Initialize $s ((z, t), P)$
8. for step = $1, 2, \ldots, N_{step}$ do
9. Select action $a$ according to $\epsilon$-greedy policy;
10. $\text{Action} = \begin{cases} A \leftarrow \text{argmax}_a Q, & 1 - \epsilon \\ A \leftarrow a(\text{random}), & \epsilon \end{cases}$
11. Take action $a$ and observe resulted state $s_{t+1}$ and obtained $R_{t+1}$ by (31);
12. $Q_{t+1}(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[R_{t+1} + \gamma Q_{t+1}(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$
13. $s \leftarrow s_{t+1}$
14. end for
15. end for
16. Output optimal depth;
17. end for

- Table II: Simulation parameters.
- Algorithm 3: Proposed reinforcement learning algorithm to solve the spatiotemporal optimization.
Fig. 4. Average the Q-values over 100 trial episodes shown for three Sampling Time = 1, Sampling Time = 50 and Sampling Time = 100. Solid lines show average Q-values, and shaded region determines a 95% confidence interval.

Fig. 5. Comparing optimal depth and velocity for Time = 200 hours. (a) Optimal depth; (b) Optimal velocity.

Time = 50, and Sampling Time = 100). Further, reducing the 95% confidence interval (i.e., shaded region) over training episodes verifies the efficiency of our presented method.

B. Comparative Results

The obtained results under three approaches are shown in Fig. 5. The control action is defined as the depth change in our study, and the optimal depth and the corresponding velocity are shown under MPC-based method, RL-based algorithm, and the case without spatiotemporal optimization. As it is shown in Fig. 5 (a), the optimal depth are almost similar in both MPC-based method and RL-based algorithm. However, some differences in selecting the next control action occur, such as \( t = 120 \). Although two different \( z \) are selected as the next optimal depth for MPC and RL algorithms, the velocities are almost the same (Fig. 5 (b)), which verifies the correct selection of two algorithms. In another word, the next operating depth in the spatiotemporal optimization is selected to maximize the total power, so it is important to consider the ocean velocity at different operating depth.

C. Discussions

RL algorithm as the learning-based method needs more time for finding the optimal water depth \( z \) due to the learning procedure, but it does not cause any challenges for our optimization problem. Hence, both RL and MPC algorithms are temporally feasible in maximizing the output power of the OCT system.

Another important feature of the presented method is its robustness to cope with the errors and uncertainties. In MPC algorithm, the exploration term \( J_{\text{exploration}} \) in objective function \( J(z) \) is defined to address this issue. However, the main challenge is to determine the corresponding weight \( \beta_2 \) of this term \( J_{\text{exploration}} \), which is chosen experimentally. RL algorithm is directly trained by taking different actions in each state, so its robustness is increasing, and the OCT system learns to choose the best action at each state based on its experience. To show the robustness of our proposed spatiotemporal method, the cumulative energy of both MPC and RL algorithms are presented in Table III where the ocean velocity is not correctly predicted due to the scale error in the velocity sensors, the data loss from measurement, and so on [4]. Therefore, the superiority of RL based method compared to MPC method is verified under two inaccurate ocean velocity predictions.

Finally, the cumulative produced energy of the OCT system using MPC and RL algorithms are compared, shown in Fig. 7. The cumulative energy shows high increase compared to the case without optimization, which highlights the importance of applying spatiotemporal optimization. We also observe that the cumulative energy productions of OCT system are very close for two optimization approaches. However, RL algorithm outperforms MPC method, and the final energy production
TABLE III. Comparing the robustness of RL based and MPC based spatiotemporal optimizations according to the percent decrease in cumulative energy using accurate GP model vs inaccurate GP model for ocean velocity prediction.

<table>
<thead>
<tr>
<th>Time(hour)</th>
<th>MPC Algorithm</th>
<th>RL Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>184</td>
<td>70.625</td>
<td>70.618</td>
</tr>
<tr>
<td>185</td>
<td>70.409</td>
<td>69.896</td>
</tr>
<tr>
<td>186</td>
<td>67.888</td>
<td>66.199</td>
</tr>
</tbody>
</table>

Fig. 7. Comparing cumulative energy under MPC based optimization, RL based optimization and case without spatiotemporal optimization.

V. CONCLUSIONS

In this study, a novel spatiotemporal optimization approach was presented for maximizing the output power of an OCT system. The Gaussian process modeling was first developed for modeling the real ocean velocity data and predicting the future ocean velocities. Then, the harvested power of an OCT system was formulated considering directly generated power from hydrokinetic and consumed power for stabilizing and changing the system at different operating depth. Two approaches, including MPC and RL based algorithms were developed to address the proposed spatiotemporal power maximization problem. The obtained results verified that both methods are efficient in maximizing the output power compared to the case without using spatiotemporal optimization. Moreover, RL-based method shows better performance than MPC-based method in terms of cumulative energy and robustness.

Future work is needed to test the robustness of our proposed method over different uncertain conditions. In addition, the mechanical controller (i.e., flight controller) and electrical controller will be designed in the next work, while their interactions with the high-level (i.e., spatiotemporal optimization) is going to be investigated through a hierarchical structure. Further, it is also critical to extend this proposed spatiotemporal optimization to an OCT array in order to maximize the total power of the array considering the wake effects among the OCTs and integrate the harnessed power into smart grids using energy storage.

REFERENCES


