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 modeled. The stochastic behavior of ocean velocities is a function of time and location, which is modeled 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 computationally feasible considering the required time for changing operating depth. On the other hand, the cumulative energy production of the RL algorithm is higher than the MPC method, which verifies that the 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 an 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

Research reported in this publication was supported in part by the National Science Foundation under grants ECCS-1809164 and OAC-2017597, and an Early-Career Research Fellowship from the National Academies Gulf Research Program.

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).

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

Fig. 1. Sample daily shear profile recorded at 4-hour intervals.

TABLE I. Comparing different algorithms 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

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 $\mathbf{V}$ matrix as follows:

$$\mathbf{V}_{n \times m} = \begin{bmatrix} v(z_1, t_1) & \dots & v(z_1, t_m) \\ \vdots & \ddots & \vdots \\ v(z_n, t_1) & \dots & v(z_n, t_m) \end{bmatrix} \quad (1)$$

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 $\mathbf{V}$ 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:

$$\mathbf{X} = \begin{bmatrix} x_1 & \dots & x_m \\ \vdots & \ddots & \vdots \\ x_n & \dots & x_{n \times m} \end{bmatrix} = \begin{bmatrix} (z_1, t_1) & \dots & (z_1, t_m) \\ \vdots & \ddots & \vdots \\ (z_n, t_1) & \dots & (z_n, t_m) \end{bmatrix} \quad (2)$$

$$V = \begin{bmatrix} v_1 & \dots & v_m \\ \vdots & \ddots & \vdots \\ v_n & \dots & v_{n \times m} \end{bmatrix} \quad (3)$$

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 $\mathcal{Z}$ denote the domain of allowable ocean depth choices and $\mathcal{T}$ represent the temporal space. Our goal is to model ocean current velocity at depth $z \in \mathcal{Z}$ and time $t \in \mathcal{T}$, by constructing a function $f : \mathcal{Z} \times \mathcal{T} \rightarrow \mathcal{R}$ whose output is a prediction y . 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). 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:

$$f(z, t) \sim \mathcal{GP}(m(z, t), k((z, t), (z', t')))) \quad (4)$$

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 n velocities V and prediction (v_*) is defined as [30]:

$$\begin{bmatrix} V \\ v_* \end{bmatrix} = \left(\begin{bmatrix} f \\ f_* \end{bmatrix} + \begin{bmatrix} \epsilon \\ \epsilon_* \end{bmatrix} \right) \sim \mathcal{N}\left(0, \begin{bmatrix} K_v & k_* \\ k_*^T & k_{**} + \sigma^2 \end{bmatrix}\right) \quad (5)$$

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

$$k_* = [k((z_1, t_1), (z_*, t_*)), \dots, k((z_n, t_n), (z_*, t_*)))] \quad (6)$$

$$k_{**} = k((z_*, t_*), (z_*, t_*)) \quad (7)$$

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

$$m(z_*, t_*) = k_*^T K_v^{-1} V \quad (8)$$

$$\sigma^2(z_*, t_*) = k_{**} - k_*^T K_v^{-1} k_* + \sigma^2 \quad (9)$$

$$f(z, t) \sim \mathcal{GP}(m(z_*, t_*), \sigma^2(z_*, t_*)) \quad (10)$$

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 OCT with a single 20 m diameter variable pitch rotor. 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 for the creation the linear models used in the presented formulations. When buoyancy controlled OCTs are

maneuvering, the total output power of the system consist of two primary parts:

- 1) Generated power of the OCT system from hydrokinetic energy extraction, denoted as P_{OCT} ;
- 2) Power consumed by ballast pumps to control the operating depth, denoted as $P_{ballast}$.

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

$$P_{net} = P_{OCT} - P_{ballast} \quad (11)$$

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

$$P_{OCT} = \frac{1}{2} \rho A C_p v^3 = \frac{1}{2} \times 1030 \times 100\pi \times 0.415 \times v^3 \quad (12)$$

where ρ is the water density (1030 kg/m^3), A is the swept area of the OCT rotor (100π), C_p is the average power coefficient of 41.5 % from [31], and v is the magnitude of the ocean velocity.

Second term, $P_{ballast}$: To calculate $P_{ballast}$, the ballast model and its average consumed power 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., $P_{abs} \cong 0 \text{ kPa}$). 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/A) through the orifice divided by pump efficiency:

$$P_B^{fill} = 0 \quad (13)$$

$$P_B^{empty} = \frac{F \cdot V}{\eta_{pump}} = \frac{(\Delta P \cdot A)(\frac{Q_B}{A})}{\eta_{pump}} = \frac{\Delta P \cdot Q_B}{\eta_{pump}} \quad (14)$$

$$\Delta P = P_{atm} + P_{HS} \quad (15)$$

where η_{pump} denotes the pump efficiency (0.75), P_{atm} is atmospheric pressure (101 kPa), P_{HS} is hydrostatic pressure ($P_{HS} = \rho \cdot g \cdot z$), and $g = 9.81 \text{ m/s}^2$ is gravity. Then the P_B^{empty} is rewritten as:

$$\begin{aligned} P_B^{empty} &= \frac{(P_{atm} + P_{HS})Q_B}{\eta_{pump}} \\ &= \frac{(101 + 10.1z)Q_B}{0.75} \end{aligned} \quad (16)$$

To put things in perspective, a volumetric flow rate Q_B of $0.023 \text{ m}^3/\text{s}$ is achieved at a depth of 50m assuming $P_B^{empty} = 18.8 \text{ kW}$, which is the ballast pump power associated with WWII submarines [36]. Since each of the two ballast tanks have volumes of 31.251 m^3 [32], these tanks can be completely emptied of water in 22.2 minutes 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 which is independent of the fill rate.

To change OCT depth, ballast tank fill levels are changed every time step, Δt , 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 the change in depth and change in fill fraction:

$$\Delta F_F = \Delta z \frac{dF_F}{dz} \quad (17)$$

In our paper, a simple quasi-static relationship is estimated between these states for a steady and homogeneous flow field using the nonlinear simulation in [32]. The fill fraction is changed in each tank by $\Delta F_F = 0.55 - 0.5 = 0.05$, and the quasi-static change is recorded in $\Delta z = 31.00 - 50.36 = -19.36$ m. 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:

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

For the calculations in (17) Δz is calculated each time step Δt based on the specified depth from the previous time step and the change in depth based on the change in current speed once the previous time step:

$$\Delta z = z^{\Delta t} - z^{\Delta t-1} - \Delta v \frac{dz}{dv} \quad (19)$$

A simple quasi-static relationship is estimated between depth and velocity $\frac{dz}{dv}$ for a steady and homogeneous flow field. The ocean velocity is changed by $\Delta v = 1.7 - 1.6 = 0.1$ m/s, and the quasi-static change is recorded as $\Delta z = 75.25 - 50.36 = 24.89$ m. Therefore, for an ocean velocity of 1.6 m/s and fill fractions of 0.5 , the quasi-static linear relationship between the change in depth and water velocity magnitude is:

$$\frac{dz}{dv} \cong \frac{\Delta z}{\Delta v} = 248.9 [m/(m/s)] \quad (20)$$

The average power consumed over a time step can then be calculated using (17), (18), the total energy used to fill the tanks with water, and Δt :

$$P_{ballast}^{\Delta t} = \begin{cases} 0, & \Delta z > 0 \\ \frac{(14.02[kWh])(-0.0026\Delta z)}{\Delta t}, & \Delta z < 0 \end{cases} \quad (21)$$

so, based on (11), (12), and (21), the net power of the OCT system is calculated as:

$$P_{net} = \frac{1}{2} \times 1030 \times 100\pi \times 0.415v^3 - P_{ballast}^{\Delta t} \quad (22)$$

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

Fig. 2. Schematic diagram of the studied problem. The OCT will be controlled spatiotemporally, and the harvested net power will be transmitted to the onshore substation for grid integration.

Algorithm 1 Proposed solution procedure for spatiotemporal optimization

- 1: Initialize n data sample $v(z, t)$ and optimization method parameters (MPC and RL methods)
- 2: **for** each sampling time Δt **do**
- 3: Predict $v(z, t)$ using GP model;
- 4: $f(z, t) \sim \mathcal{GP}(m(z_*, t_*), \sigma^2(z_*, t_*))$
- 5: **for** Prediction Horizon T **do**
- 6: Apply optimization algorithm (MPC or RL)
- 7: Obtain optimal z
- 8: **end for**
- 9: **end for**

fill fraction of the aft tank), and the electromechanical torque T_{em} . Moreover, electromechanical torque T_{em} , current I and rotor angular speed ω 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], maximizing the output power of an autonomous underwater vehicle [37], and minimizing the operational cost of wind turbine considering the fault-tolerant method [38]. 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

Fig. 3. Proposed spatiotemporal optimization structure to maximize OCT system power.

objective function is presented as follows:

$$\min_{z(p)} J(z(p)) = \sum_{i=p}^{p+T-1} [\beta_1 J_{\text{exploitation}}(z(i|p)) + \beta_2 J_{\text{exploration}}(z(i|p))] \quad (23)$$

subject to

$$40 \leq z(i|p) \leq 200 \quad (24)$$

$$f(z, t) \sim \mathcal{GP}(m(z_*, t_*), \sigma^2(z_*, t_*)) \quad (25)$$

$$\frac{z(i|p) - z(i-1|p)}{\Delta t} \leq r \quad (26)$$

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 $\sigma^2(z_*, t_*)$, r denotes the rate limitation on the speed, and β_1 and β_2 are the multiples of two terms. The exploitation term of the objective function $J_{\text{exploitation}}$ is defined as follows:

$$J_{\text{exploitation}}(z(i|p)) = P_{\text{rated}} - E(P(z(i|p))) \quad (27)$$

where $E(P(z(i|p)))$ shows the expected power according to the predicted ocean velocities, and P_{rated} shows the rated power calculated by (22) 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_{rated} and $E(P(z(i|p)))$. The exploration term of the objective function is defined as follows:

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

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 2 Proposed model predictive control algorithm to solve the spatiotemporal optimization

- 1: Initialize n data sample $v(z, t)$, $i = 0$, Δt , and T
- 2: **for** each sampling time Δt **do**
- 3: Predict $v(z, t)$ using 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**

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.

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*, $s_i \in \mathcal{S}$, *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 in order 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 $(0, 1]$. 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 (22), 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:

$$\mathcal{S} = \{s_i | ((z, t), P)\} \quad (29)$$

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:

$$\mathcal{A} = \{a_i | +\Delta z, -\Delta z, 0\} \quad (30)$$

Algorithm 3 Proposed reinforcement learning algorithm to solve the spatiotemporal optimization

```

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

```

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 considered as the desirable goal. Hence, any action (water depth change) should be done in order to increase the net power at each time step. The reward function is defined as follows:

$$R = \begin{cases} \Delta P, & \Delta P > \delta \text{ or } \Delta P < -\delta \\ 0, & \Delta P = \delta \end{cases} \quad (31)$$

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. As the final goal is to maximize the net power, the reward is larger for higher power increase.

We use Q-learning algorithm to solve this RL problem. Q-value $Q(s, a)$ denotes the expected return of taking action a in state s , and the Q-values table is updated as follows:

$$Q_{t+1}(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[R_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (32)$$

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.

In order to balance the exploration vs the exploitation, ϵ -greedy policy is used for action selection. Two different options are available for selecting the next action at each time step, including (i) choosing the actions with the highest

TABLE II. Simulation parameters.

Parameter	Value	Parameter	Value
Time	200 hour	δ	0.01
γ	0.8	β_1	1
β_2	100	T	??
Δt	1 hour and 2 hour	v_{rated}	1.6 m/s

estimated reward; or (ii) choosing a random action. In ϵ -greedy policy, the random actions are chosen as follows:

$$Action = \begin{cases} A \leftarrow \underset{a}{\operatorname{argmax}} Q, & \text{with probability } 1 - \epsilon \\ A \leftarrow a(\text{random}), & \text{with probability } \epsilon \end{cases} \quad (33)$$

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

The Q-learning algorithm using ϵ greedy approach is presented in Algorithm 3, which is used for maximizing the net power of OCT system according to the state space (29), the action space (30), and the reward function (31). As it is presented in this algorithm, the action (depth change) is selected based on ϵ -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 (23), 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,

Fig. 4. Average the Q-values over 100 trial episodes shown for three first sampling time, second sampling time, and ?? sampling time. Solid lines show average Q-values, and shaded region determines a 95% confidence interval.

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 ΔZ) 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., first sampling time, second sampling time, and ?? sampling time). 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.

In Fig. 6, the harvested power of the OCT system is presented, which shows that the output power of OCT system is increasing where the spatiotemporal optimization is applied. MPC based optimization and RL based method follow the similar trend in power increase. However, RL based spatiotemporal optimization surpasses MPC method at some time samples.

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

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

Fig. 6. Comparing optimal power under MPC based optimization, RL based optimization and case without spatiotemporal optimization.

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 $J_{exploration}$ in 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, 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

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.

MPC Algorithm		RL Algorithm	
MWh	%	MWh	%
Accurate GP Model			
70.625	-	70.618	-
Inaccurate GP Model			
69.896	1.03	70.409	0.29
66.199	6.2	67.888	3.8

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

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 for MPC based optimization, RL based optimization, and case without applying spatiotemporal optimization are equal to 70.618 MWh, 70.625 MWh and 59.612 MWh, respectively.

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

- [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] G. Lavidas, "Energy and socio-economic benefits from the development of wave energy in greece," *Renewable Energy*, vol. 132, pp. 1290–1300, march 2019.
- [3] K. Haas, "Assessment of energy production potential from ocean currents along the united states coastline," Georgia Tech Research Corporation, Tech. Rep., 2013.
- [4] 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.
- [5] Y. Geong, Z. Yang, X. Shan, Y. Sun, T. Xie, and Y. Zi, "Capturing flow energy from ocean and wind," *Energies*, vol. 12, no. 11, june 2019.
- [6] D. Magagna and A. Uihlein, "Ocean energy development in europe: Current status and future perspectives," *International Journal of Marine Energy*, vol. 11, pp. 84–104, September 2015.
- [7] 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.
- [8] M. Kehe, 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, March 2018.
- [9] X. Yin and X. Zhao, "Sensorless maximum power extraction control of a hydrostatic tidal turbine based on adaptive extreme learning machine," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 1, pp. 426–434, January 2020.
- [10] Y. Saidi, A. Mezouar, Y. Miloud, K. Djamel Eddine Kerrouche, B. Brahm, and M. A. Benmahdjoub, "Advanced non-linear backstepping control design for variable speed wind turbine power maximization based on tip-speed-ratio approach during partial load operation," *International Journal of Dynamics and Control*, pp. 1–14, July 2019.
- [11] S. Li and J. Li, "Output predictor-based active disturbance rejection control for a wind energy conversion system with pmsg," *IEEE Access*, vol. 5, pp. 5205–5214, March 2017.
- [12] X. Yin, Z. Jiang, and L. Pan, "Recurrent neural network based adaptive integral sliding mode power maximization control for wind power systems," *Renewable Energy*, vol. 145, pp. 1149–1157, January 2020.
- [13] D. R. Song, Q. A. Li, Z. Cai, L. Li, J. Yang, and H. J. Y. Su, Mei, "Model predictive control using multi-step prediction model for electrical yaw system of horizontal-axis wind turbines," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 4, pp. 2084–2093, October 2019.
- [14] T. Jain and J. Julien Yame, "Fault-tolerant process-economic model predictive control for wind turbines," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 4, pp. 1696–1704, October 2019.
- [15] 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, December 2017.
- [16] A. Bafandeh, S. Bin-Karim, A. Baheri, and C. Vermillion, "A comparative assessment of hierarchical control structures for spatiotemporally-varying systems, with application to airborne wind energy," *Control Engineering Practice*, vol. 74, pp. 71–83, May 2018.

- [17] C. Wei, Z. Zhang, W. Qiao, and L. Qu, "Reinforcement-learning-based intelligent maximum power point tracking control for wind energy conversion systems," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6360–6370, October 2015.
- [18] C. Wei, Z. Zhang, W. Qiao, and Qu, "An adaptive network-based reinforcement learning method for mppt control of pmsg wind energy conversion systems," *IEEE Transactions on Power Electronics*, vol. 31, no. 11, pp. 7837–7848, November 2016.
- [19] E. Anderlini, D. I. M. Forehand, E. Bannon, and M. Abusara, "Control of a realistic wave energy converter model using least-squares policy iteration," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1618–1628, October 2017.
- [20] E. Anderlini, D. I. M. Forehand, P. Stansell, Q. Xiao, and M. Abusara, "Control of a point absorber using reinforcement learning," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 4, pp. 1681–1690, 2016.
- [21] T. D. Ngo, C. Sultan, J. H. VanZwieten, and N. I. Xiros, "Variance constrained cyclic blade control of moored ocean current turbines," *2016 American Control Conference (ACC)*, July 2016.
- [22] T. Ngo, C. Sultan, J. H. VanZwieten, and N. I. Xiros, "Model predictive control for moored ocean current turbines," *2017 American Control Conference (ACC)*, May 2017.
- [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] 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," *2018 American Control Conference (ACC)*, June 2018.
- [25] 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, November 2018.
- [26] M. C. P. Machado, J. H. VanZwieten, A. P. Callou, I. Pinos, *et al.*, "A measurement based analyses of the hydrokinetic energy in the gulf stream," in *The Twenty-fifth International Ocean and Polar Engineering Conference*. International Society of Offshore and Polar Engineers, 2015.
- [27] B. Pearson and B. Fox-Kemper, "Log-normal turbulence dissipation in global ocean models," *Physical review letters*, vol. 120, no. 9, March 2018.
- [28] 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.
- [29] 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.
- [30] 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, October 2016.
- [31] 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.
- [32] 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.
- [33] 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.
- [34] 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.
- [35] D. Coiro, G. Troise, F. Scherillo, A. De Marco, G. Calise, and N. Biz-zarrini, "Development, deployment and experimental test on the novel tethered system gem for tidal current energy exploitation," *Renewable Energy*, vol. 114, pp. 323–336, 2017.
- [36] The fleet type submarine online submarine trim and drain systems (chapter 2). [Online]. Available: <https://maritime.org/doc/fleetsub/trim/chap2.htm>
- [37] 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.
- [38] T. Jain and J. J. Yame, "Fault-tolerant economic model predictive control for wind turbines," *IEEE transactions on sustainable energy*, vol. 10, no. 4, pp. 1696–1704, October 2019.