Ocean Current Turbine Power Maximization: A Spatiotemporal Optimization Approach

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

02-08-2020 / 24-08-2020

CITATION

Tang, Yufei (2020): Ocean Current Turbine Power Maximization: A Spatiotemporal Optimization Approach. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.12751181.v3>

DOI

[10.36227/techrxiv.12751181.v3](https://doi.org/10.36227/techrxiv.12751181.v3)

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. To determine output power, ocean velocities and OCT power consumption and generation 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. Comparative studies show that the MPC and the RL based controllers are both computationally feasible considering the required time for changing operating depth. Moreover, the cumulative energy production of the RL algorithm is higher than the MPC method, which verifies that the learning-based 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). There are many approaches to address the high cost challenges, such as lowering the cost of turbine design or optimizing the operational strategy [5], 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 has been studied using a hierarchical controller [6], in which the upper level is responsible for setting the trajectory and the lower level controls the generator. An extreme learning machine has been used to maximize the output power of a hydro-static tidal turbine, which shows good performance under uncertain and turbulent environments [7]. Nonlinear backstepping control has been applied for maximum energy harvesting from a variable speed wind turbine with claimed advantages of fast performance and reduced mechanical stress [8]. An adaptive controller has been proposed to maximize the power output of wind turbines considering uncertainties in wind and turbine dynamics simultaneously [9], where the controller includes Smith predictor to reduce the stochastic behavior.

Among many of the turbine controller designs, model predictive control (MPC) and reinforcement learning (RL) have gained increasing attentions in recent years. MPC has been applied for setting the yaw angle of wind turbines into the forecasted wind direction [10]. A fault-tolerant MPC controller has been designed for wind turbines to maximize output power and minimize fatigue simultaneously [11]. An airborne wind energy system has employed MPC-based techniques to set the optimal altitude [12], in which future wind speeds are predicted and output power is optimized by changing the location of the system to access the optimal velocities. In another close study [13], performance of an MPC method and a Lyapunov-based controller have been compared for optimizing the output power of an airborne wind energy system. On the other hand, RL based design has been applied for wind turbine power optimization [14], [15], where the Q-learning algorithm has been used to find the optimal generator speed. In another application, RL has been used to maximize the harvested energy from wave energy converters due to high nonlinearity of these systems [16], [17].

OCT control has been investigated from a different point of view in previous studies. The blade pitch control of an OCT has been addressed in [18] in order to stabilize its mechanical operation. A MPC-based controller has been designed using blade pitch and torque as the input to stabilize the operation of an OCT system [19]. Further, a permanent magnet synchronous generator has been controlled through an adaptive super-twisting sliding mode algorithm to assure stable operation of the OCT under stochastic ocean velocity [20]. An estimator has been designed to predict the spatially dependent velocity of ocean, and a MPC-based controller has been used to optimize the location of an underwater vehicle used as an energy harvesting system [21]. A Bayesian optimization has been used to specify the configuration and location to harvest maximum energy from an OCT array [22]; however,

Research reported in this publication was supported in part by the National Science Foundation under grants ECCS-1809164 and OAC-2017597.

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).

the model has been focused on the economic side; a more realistic dynamic model is desired.

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 ocean currents, 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. Moreover, a RL-based method is designed to explore the optimal control actions and 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 a 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 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 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 in Southeast Florida by a bottom mounted acoustic Doppler current profile (ADCP) between February 2009 and March 2010 [23]. The ocean shear profile of these data with 4-hour intervals for one sample day is shown in Fig. 1. These observed ocean velocity data are

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

represented as:

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

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

where n and m show the number of discrete depths and the number of time samples, and $v(z_i, t_j)$ is the measured ocean velocity at z_i and t_j . Matrix $\mathbf{X}$ denotes the set of ocean depth and time, and $\mathbf{V}$ includes $n \times m$ recorded ocean velocities (i.e., m velocity vectors each as a function of depth).

Let $\mathcal{Z}$ denote the domain of allowable ocean depth choices and $\mathcal{T}$ represent the temporal space. The goal here is to model ocean current velocity at depth $z \in \mathcal{Z}$ and time $t \in \mathcal{T}$, by constructing a function $f : \mathcal{Z} \times \mathcal{T} \rightarrow \mathcal{R}$ whose output is a prediction y . The ocean velocity and especially the ocean turbulence are usually modeled with probability distributions, including log-normal distribution and Burr distribution. In this paper, different methods for modelling the velocity are developed and compared, shown in Table I. Linear regression, regression trees, support vector machines, and Gaussian process (GP) modeling have been compared with root mean square error (RMSE), mean square error (MSE), and mean absolute error (MAE). The GP modeling shows the best performance, and will be briefly introduced in the following.

Gaussian Process Model: 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 [12], [24]. As we will show, GP can also be utilized to predict the ocean velocity. 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

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

over recorded velocities $\mathbf{V}$ and prediction v_* is defined as [25]:

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

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

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

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

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 is investigated numerically in [26], 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 sea floor at a depth of 325 m [27]. These model parameters roughly follow the prototype systems from IHI Corp. [28], [29] and the University of Naples [30], but with a single rotor and 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 [26], [27] are utilized for simulating this system for the creation the linear models used in the presented formulations. When buoyancy controlled OCTs are operating, the output power of the system consist 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 $P_{ballast}^{HD}$.
- 3) Power consumed by ballast pumps to Change Depth, denoted as $P_{ballast}^{CD}$.

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

$$P_{net} = P_{OCT} - P_{ballast}^{HD} - P_{ballast}^{CD} \quad (10)$$

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

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

where $\rho = 1030 \text{ kg/m}^3$ 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 [26], and v is the magnitude of the ocean velocity.

Second term, $P_{ballast}^{HD}$: To calculate $P_{ballast}^{HD}$, the average consumed power for maintaining a near constant depth in a time-varying current is determined. This model assumes that ballast tank water fill levels 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 . 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., $P_{abs} \cong 0 \text{ kPa}$). 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, $\Delta P \cdot A$) and velocity (i.e., quotient of volumetric flow rate over area, Q_B/A) through the orifice divided by pump efficiency:

$$P_B^{fill} = 0 \quad (12)$$

$$P_B^{empty} = \frac{F \cdot V}{\eta_{pump}} = \frac{(\Delta P \cdot A) \left(\frac{Q_B}{A} \right)}{\eta_{pump}} = \frac{\Delta P \cdot Q_B}{\eta_{pump}} \quad (13)$$

$$\Delta P = P_{atm} + P_{HS} \quad (14)$$

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

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

To add 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 [31]. Since each of the two ballast tanks has a volume of 31.251 m^3 [27], 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:

$$\frac{\Delta F_F}{\Delta v} = \frac{dF_F}{dz} \frac{dz}{dv} \Rightarrow \Delta F_F = \Delta v \frac{dF_F}{dz} \frac{dz}{dv} \quad (16)$$

In this paper, a quasi-static relationship is assumed between these states and a steady and homogeneous flow field when

running the nonlinear simulation [27]. 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:

$$\frac{dF_F}{dz} \cong \frac{\Delta F_F}{\Delta z} = -0.0026 [1/m] \quad (17)$$

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:

$$\frac{dz}{dv} \cong \frac{\Delta z}{\Delta v} = 248.9 [m/(m/s)] \quad (18)$$

Therefore, using (16), (17), and (18) the linear relationship between the fill fraction change necessary to maintain depth and the corresponding Δv is:

$$\Delta F_F = \Delta v(-0.0026)(248.9) = -0.65\Delta v \quad (19)$$

The average power consumed in kW over each Δt_1 in hours for a Δv in $\frac{m}{s}$ can be calculated using (19), the total energy required to fill the tanks with water in kWh , and Δt_1 :

$$P_{ballast}^{HD} = \begin{cases} 0, & \Delta v < 0 \\ \frac{(14.02)(0.65\Delta v)}{\Delta t_1}, & \Delta v > 0 \end{cases} \quad (20)$$

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 (17):

$$\Delta F_F = \Delta z \frac{dF_F}{dz} = -0.0026\Delta z \quad (21)$$

The average power in kW consumed over Δt_2 in hours for Δz in meters can be calculated using (21), the total energy used to fill the tanks with water, and Δt_2 in hours:

$$P_{ballast}^{CD} = \begin{cases} 0, & \Delta z > 0 \\ \frac{(14.02)(-0.0026\Delta z)}{\Delta t_2}, & \Delta z < 0 \end{cases} \quad (22)$$

Therefore, based on (10), (11), (20), and (22), the average net power of the OCT system is calculated as:

$$P_{net} = \frac{1}{2} \times 1030 \times 100\pi \times 0.415v^3 - P_{ballast}^{HD} - P_{ballast}^{CD} \quad (23)$$

III. PROPOSED METHODOLOGY

A. Problem Formulation and Optimization Framework

Schematic diagram of the studied problem for maximizing OCT generated power through a spatiotemporal approach is shown in Fig. 2. According to the predicted ocean velocity, the OCT system location should be determined each Δt_2 to maximize the output power.

To address the complexities of this problem, a novel hierarchical spatiotemporal optimization and control framework, as shown in Fig. 3, is developed. The ocean velocity is

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

Algorithm 1 Upper-level Spatiotemporal Optimization

- 1: Initialize n data sample $v(z, t)$ and optimization method parameters (MPC and RL methods)
- 2: **for** each sampling time Δt_2 **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: **end for**
- 8: Output optimal depth z^*
- 9: **end for**

modeled with GP method, and the forecasted velocity is used to calculate the output power of 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 adjust detailed turbine control dynamics using the prescribed set-points from the upper-level. On the lower-level, 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} . The control inputs to the electrical controller (i.e., generator controller) include the electromechanical torque T_{em} , generator current I , and rotor angular speed ω . This paper will focus on the upper-level spatiotemporal optimization, and the pseudocode is presented in Algorithm 1. Two approaches including RL algorithm and MPC method are considered in this paper.

B. MPC-based Optimization Design

MPC is a powerful method for optimizing the objective function by predicting the future actions. The MPC-based algorithm was used to maximize the output power of an airborne system [12], maximize the output power of an autonomous underwater vehicle [21], and minimize the operational cost of wind turbine considering the fault-tolerant method [11]. In this paper, MPC 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.

Fig. 3. Proposed hierarchical spatiotemporal optimization and control.

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 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:

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

subject to

$$40 \leq z(i|p) \leq 200 \quad (25)$$

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

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 $\sigma^2(z_*, t_*)$, and β_1 and β_2 are the gains of two terms. The exploitation term $J_{\text{exploitation}}$ is defined as:

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

where $E(P(z(i|p)))$ shows the expected power according to the predicted ocean velocities, and P_{rated} is the rated power calculated by (23). 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 $J_{\text{exploration}}$ is defined as:

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

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 (2), 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 sampling time using (9). The objective function is calculated over the prediction horizon, while optimal depth is determined.

Algorithm 2 Proposed MPC-based Design

- 1: Initialize n data sample $v(z, t)$, $i = 0$, Δt_1 , Δt_2 , and T
- 2: **for** each sampling time Δt_2 **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 (25) and obtain optimal depth trajectory $z(p)$;
- 7: **end for**
- 8: Output optimal depth z^* ;
- 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 the RL approach, the set of actions 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 this 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 (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. Transitions between different water depths 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 , 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 a reasonable prediction horizon in order 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., 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:

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

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

Algorithm 3 Proposed RL-based Design

```

1: Initialize  $n$  data sample  $v(z, t)$ ,  $\Delta t_1$ ,  $\Delta t_2$  and  $T$ 
2: for each sampling time  $\Delta t_2$  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;
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:         $Q_{t+1}(s_t, a_t) \leftarrow Q(s_t, a_t) +$ 
14:         $\alpha[R_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)]$ 
15:         $s \leftarrow s_{t+1}$ 
16:      end for
17:    end for
18:  end for
19: Output optimal depth  $z^*$ ;
20: end for

```

depth change) should be done in order to increase the net power at each time step. The reward function is defined as:

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

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.

Q-learning algorithm is used 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.

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

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

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

TABLE II. Optimization parameters.

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

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

To show the efficiency of the proposed method, it is tested with real ocean velocity data and the example OCT discussed in Section II. Recorded ocean velocities V are modeled by GP model in (9). After modeling ocean velocities, 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 optimization are presented in Table II.

The spatiotemporal optimization for maximizing output power of the OCT system is compared with three approaches: **Case without Spatiotemporal Optimization:** In this case, the OCT system is located in a fixed 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 operating 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 $J_{exploration}$.

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 Δz) should be determined to maximize the total power. The algorithm is trained continuously while the system is online, improving performance all the time.

B. Comparative Results

The obtained results are shown in Fig. 4. The control action is defined as the depth change in this 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 presented in Fig. 4 (a), the

Fig. 4. Comparing optimal depth and velocity over 200 hours. (a) Optimal depth; (b) Optimal velocity.

Fig. 5. Comparing optimal power under MPC based optimization, RL based optimization and case without spatiotemporal optimization.

optimal depth are similar for both MPC-based and RL-based algorithms. However, 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. 4 (b)), which verifies that the two algorithms are both making optimal selections.

In Fig. 5, the harvested power of the OCT system is presented, which shows that the output power of OCT increases when 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. In addition, the average power of the OCT system in terms of P_{OCT} , $P_{ballast}^{HD}$, $P_{ballast}^{CD}$, and P_{net} for three different operating depths are presented in Table III.

C. Discussions

RL algorithm as a 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 feasible for maximizing the output power of the OCT system. Another important feature of the presented method is its robustness to errors and uncertainties. In MPC algorithm,

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

z (Time = 0) m	P_{OCT} kW	$P_{ballast}^{HD}$ kW	$P_{ballast}^{CD}$ kW	P_{net} 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

TABLE IV. Comparing the robustness according to the percent decrease in cumulative energy using accurate GP vs inaccurate GP.

MPC Algorithm		RL Algorithm	
MWh	%	MWh	%
Accurate GP Model			
39.653	-	40.366	-
Inaccurate GP Model			
38.393	3.18	39.636	1.8
36.847	7.08	37.901	6.1

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 IV, in which 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 over the MPC-based 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. 6. 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 the two compared optimization approaches. RL algorithm outperforms MPC method, and 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.

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

Fig. 6. Comparing cumulative energy under MPC based optimization, RL based optimization and case without spatiotemporal optimization.

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. 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 lower-level mechanical controller (i.e., flight controller) and electrical controller (i.e., generator controller) will be thoroughly designed in the next work, while their interactions with the upper-level (i.e., spatiotemporal optimization) is going to be fully investigated. 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 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] 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] T. Wilberforce, Z. El Hassan, D. A., J. Thompson, B. Soudan, and A. Olabi, "Overview of ocean power technology," *Energy*, vol. 175, pp. 165–181, may 2019.
- [6] 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, March 2018.
- [7] 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.
- [8] Y. Saidi, A. Mezouar, Y. Miloud, K. Djamel Eddine Kerrouche, B. Brahmi, 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.
- [9] 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.
- [10] D. R. Song, Q. A. Li, Z. Cai, L. Li, J. Yang, M. Su, and Y. H. Joo, "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, 2019.
- [11] T. Jain and J. Julien Yame, "Fault-tolerant economic model predictive control for wind turbines," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 4, pp. 1696–1704, October 2019.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] 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.
- [19] 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.
- [20] 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.
- [21] 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.
- [22] 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.
- [23] 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 25th International Ocean and Polar Engineering Conference*. Int. Society of Offshore and Polar Engineers, 2015.
- [24] 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.
- [25] 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.
- [26] 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.
- [27] A. Hasankhani, J. VanZwieten, Y. Tang, B. Dunlap, 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.
- [28] 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.
- [29] 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.
- [30] 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.
- [31] The fleet type submarine online submarine trim and drain systems (chapter 2). [Online]. Available: <https://maritime.org/doc/fleetsub/trim/chap2.htm>