Reactive Control of a Wave Energy Converter using Artificial Neural Networks

E. Anderlini[a,c,d,*], D.I.M. Forehand[b], E. Bannon[c], M. Abusara[d]

[a] *Industrial Doctoral Centre for Offshore Renewable Energy, University of Edinburgh, Faraday Building, Colin Maclaurin Road, Edinburgh, EH9 3DW, UK*
[b] *Institute of Energy Systems, University of Edinburgh, Faraday Building, Colin Maclaurin Road, Edinburgh, EH9 3DW, UK*
[c] *Wave Energy Scotland, 10 Inverness Campus, Inverness, IV2 5NA, UK*
[d] *College of Engineering, Mathematics and Physical Sciences, University of Exeter, Penryn Campus, Penryn, TR10 9FE, UK*

Abstract

A model-free algorithm is developed for the reactive control of a wave energy converter. Artificial neural networks are used to map the significant wave height, wave energy period, and the power take-off damping and stiffness coefficients to the mean absorbed power and maximum displacement. These values are computed during a time horizon spanning multiple wave cycles, with data being collected throughout the lifetime of the device so as to train the networks off-line every 20 time horizons. Initially, random values are selected for the controller coefficients to achieve sufficient exploration. Afterwards, a Multistart optimization is employed, which uses the neural networks within the cost function. The aim of the optimization is to maximise energy absorption, whilst limiting the displacement to prevent failures. Numerical simulations of a heaving point absorber are used to analyse the behaviour of the algorithm in regular and irregular waves. Once training has occurred, the algorithm presents a similar power absorption to state-of-the-art reactive control. Furthermore, not only does dispensing with the model of the point-absorber dynamics remove its associated inaccuracies, but it also enables the controller to adapt to variations in the machine response caused by ageing.

Keywords: Wave Energy Converter (WEC), Artificial Neural Networks (ANNs), reactive control, Multistart optimization.

*Principal corresponding author
Email address: E.Anderlini@ed.ac.uk (E. Anderlini)

1. Introduction

With a possible resource of up to 2.1 TW of power worldwide [1], wave energy can become an important future energy resource, thus decreasing society's greenhouse gas emissions. At the moment, the wave energy industry is not mature yet: numerous wave energy converter (WEC) devices have been developed, but none has been established as the best design yet. Reference [2] provides a thorough review of some of the most promising recent technologies. Point absorbers are an established type of offshore WECs [2]. They comprise of a floating body, whose dimensions are small relative to the characteristic wavelength, excited by ocean waves that drive a power take-off (PTO) system, which absorbs energy. WECs are envisioned to be installed in groups, i.e. wave farms, so as to reap the benefits of economies of scale [3]. However, for simplicity we analyse a single, axisymmetric unit subject to motions in heave.

Over the years, various control schemes have been proposed for the maximization of energy absorption of WECs, with [4] and [5] presenting comprehensive reviews of the initial and recent studies in the field. In theory, optimal power generation can be obtained through complex-conjugate control, since it regulates the system so as to achieve resonance with the incoming waves [4]. Nevertheless, this is impractical in reality due to the associated large motions of, and loads on, the machine in extreme seas. Thus, alternative control strategies have been implemented, which consider physical constraints on the motions, forces and power rating of the WEC [3].

Latching, model-predictive and simple-but-effective control are real-time techniques for the control of WECs. With latching control, first developed by [6], there is an alternation over a wave cycle of stages when the device is linearly damped and locked in place by the PTO system. Resonance is achieved by regulating the duration of each phase [7]. Model predictive control computes at each time step the force that maximizes energy absorption during a future time horizon [8, 9]. Simple-but-effective control applies a force that is calculated by fitting a narrow-banded function to the wave excitation force [10]. While the scaling of latching control to wave farms poses serious problems, [11] have applied model predictive control to multi-body WECs, and [12, 13] to an array of three point absorbers. Although these methods include limits on the response and loading of WECs, their behaviour is strongly influenced by the quality of the forecast wave excitation force and of the model of the device dynamics [5]. In addition, model predictive control presents a very

high computational cost associated with the real-time optimization. Simple-but-effective control results in similar power generation to model predictive control, but presents a simpler implementation [5].

An alternative type of control strategies relies on time-averaged sea states, thus assuming stationary wave conditions over a prescribed time [3]. With reactive control, simulations are run to calculate the combination of PTO damping and stiffness coefficients that maximise the generated energy in each sea state. Resistive control represents a specific case, where the stiffness term is zero. Force and displacement constraints can be included within the numerical model and cost function, respectively. While this technique may be associated with lower energy extraction than on-line control strategies [13], it is less computationally intensive and presents a simple implementation. Furthermore, the control scheme can be easily extended to the treatment of wave farms, as considered by [3].

All aforementioned methods are strongly affected by the accuracy of the model of the body dynamics they use. For this reason, modelling errors can result in a drop in the generated power. Additionally, the control strategies cannot adapt to changes in the response of the WEC caused by its ageing, with marine biofouling playing a major role. Therefore, in a previous article the authors have developed an algorithm for resistive control based on reinforcement learning that learns the optimal PTO damping coefficient in every sea state directly from experience [14]. This work has been extended to the reactive control of a point absorber with a reaction plate in [15]. In contrast to resistive control, reactive control can lead to much higher efficiencies but requires an extension of the search space to two variables, namely the PTO damping and stiffness coefficients. For this reason, learning time in each sea state can become very long depending on the refinement of the discretization of the PTO coefficients. Furthermore, continuous values of the control parameters could result in higher efficiencies. Artificial Neural Networks (ANNs) represent an alternative set of machine learning algorithms which are popular in the computer science industry. They can yield smooth, non-linear function approximations [16] and therefore provide an elegant solution to the above two problems with reinforcement learning. ANNs have been used to provide real-time system identification for WEC dynamics by [17] and [18]. Furthermore, [17] have successfully applied the ANNs model to the control of the AWS Archimedes Swing WEC.

Here, ANNs will be applied for the first time to the reactive control of a point absorber. Hence, they are employed to map the sea state conditions

averaged over a time interval and the applied PTO coefficients to the mean power and maximum displacement that occur over the duration of the time interval. The resulting mapping will be used to select optimal PTO damping and stiffness coefficient at the start of each time interval, once learning has been completed. Numerical simulations are run in both regular and irregular waves to test the efficiency and convergence properties of the proposed control algorithm.

2. Reactive Control of a Point Absorber

2.1. System Description

A point-absorber with an electromechanical PTO is considered, as for example analysed by [19] or proposed by [20]. Removing the hydraulic stage in the power conversion process results in an increase in efficiency [20]. Furthermore, as opposed to direct-drive PTO, the use of smaller, cheaper rotating generators is still possible [20].

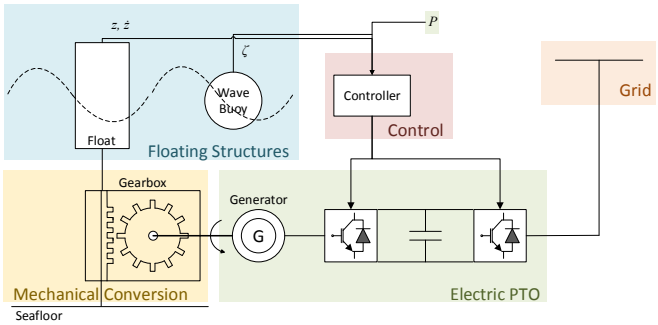


Figure 1: Scehmatic diagram of the WEC with its electromechanical PTO.

As shown in Figure 1, the movement of the float is converted into rotational motion through a mechanical stage. This mechanism drives a generator, which can be of a permanent magnet design as proposed by [20]. A variable-frequency converter delivers the generated power to the electrical grid at the requested frequency. The controller controls the generator through the machine-side converter in order to maximise energy absorption. The grid-side converter keeps a constant DC-link voltage and controls the active and reactive power transmitted to the network [21].

In order to select optimal control actions, the controller requires the heaving body displacement, z, and velocity, $\dot{z}$, as well as the wave elevation, ζ.

While the former two variables are inferred from on-board accelerometers, the latter is usually provided by an separate wave buoy for the whole wave farm. Furthermore, the generated power P is obtained from the electric PTO system.

2.2. Hydrodynamics Modelling

The hydrodynamic model has been obtained as in [14]. With the assumptions of small body motions and linear wave theory, it is possible to express the response of the point absorber through the superposition of inertial, hydrostatic, radiation, excitation and control forces [22]. Therefore, modelling the radiation force according to Cummins [23], it is possible to obtain the following time-domain equation of motion of the WEC [24]:

$$\left(M + A_{3,3}(\infty)\right)\ddot{z}(t) + \int_0^t K_{3,3}(t-\tau)\dot{z}(\tau)\mathrm{d}\tau + C_{3,3}z(t) = F_3(t) + F_{\mathrm{PTO}}(t), \quad (1)$$

with the index 3 expressing heaving motions. M is the float mass, $C_{3,3}$ the hydrostatic stiffness coefficient, $A_{3,3}(\infty)$ the added mass at infinite wave frequency, and $K_{3,3}(t)$ the radiation impulse response function. The panel-code WAMIT has been used for their determination. The right-hand side of (1) comprises of the sum of the PTO force, F_{PTO}, and the wave excitation force, F_3. The derivation of (1), as well as a more thorough explanation can be found in [25].

Equation (1) is shown graphically in Figure 2. The radiation force is approximated through a state-space system in order to speed up the simulations. The state-space matrices have been computed as described in [24].

2.3. Reactive Control

As can be seen in Figure 2, with reactive control the sum of a damping and a stiffness term yields the PTO force [3]:

$$F_{\mathrm{PTO}}(t) = B_{\mathrm{PTO}}\dot{z}(t) + C_{\mathrm{PTO}}z(t). \quad (2)$$

In electromechanical PTO units, variations in the generator excitation current or the power converter conduction angle result in changes in the PTO damping and stiffness coefficients [5], B_{PTO} and C_{PTO} respectively. As we deal with B_{PTO} and C_{PTO} directly in this article for simplicity, the method may in fact be applied to hydraulic or direct-drive PTO systems as well.

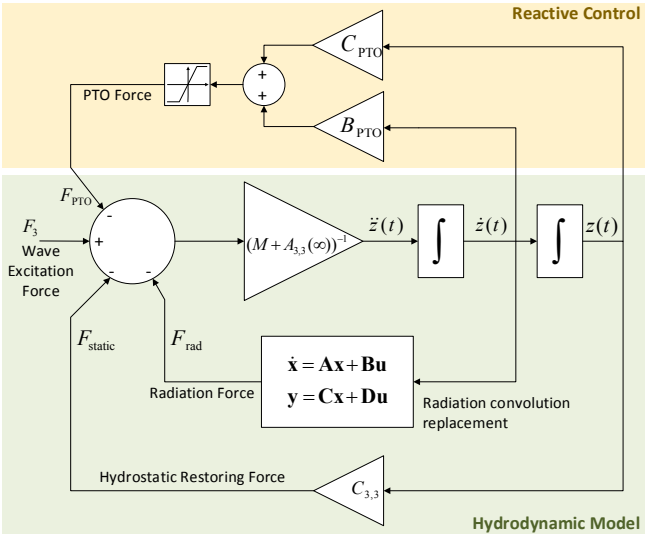


Figure 2: Block diagram employed in the computation of the float dynamics.

The control force is actually clipped at $\pm F_{\mathrm{Max}}$ owing to the physical limits of the PTO system. In Figure 2, this is represented through a saturation block. The calculation of the generated power with reactive control is standard [25], and given by [3]:

$$P(t) = \begin{cases} \eta F_{\mathrm{PTO}}(t)\dot{z}(t) & \text{if } F_{\mathrm{PTO}}(t)\dot{z}(t) > 0 \\ F_{\mathrm{PTO}}(t)\dot{z}(t)/\eta & \text{if } F_{\mathrm{PTO}}(t)\dot{z}(t) \leq 0 \end{cases}, \tag{3}$$

with η being the PTO efficiency.

Including the effects of force saturation, it is possible to maximize the energy extraction through the selection of suitable PTO damping and stiffness coefficients, which depend on the wave energy period, T_{e}, and the significant wave height, H_{s}. In regular waves, these correspond to the wave period and height respectively. Furthermore, the float displacement is to be bounded to $|z| < z_{\mathrm{Max}}$ so as to avoid structural damage in highly energetic waves. This means that the buoy should be prevented from reaching the end stops of the mechanical system, with the limits derived during the design stage. Hence, the optimal PTO damping and stiffness coefficients, $B_{\mathrm{PTO,opt}}$ and $C_{\mathrm{PTO,opt}}$ respectively, need to be chosen so that they maximise the generated power, while abiding by the displacement constraint.

At the moment, the state-of-the-art approach for reactive control is to pre-calculate $B_{\mathrm{PTO,opt}}$ and $C_{\mathrm{PTO,opt}}$ for a set of discrete sea states, generating a matrix. The simulations are run using a time-domain model similar to the one described here in order to account for the force saturation. Once the point absorber is at sea, the controller meets the PTO coefficients associated with the encountered sea state through the power electronics. However, this technique suffers from modelling errors, and it cannot recognized changes in the WEC dynamics caused by its ageing.

3. ANN-based Reactive Control of WECs

In order to obtain a model-free control with a continuous search space, the authors propose to use ANNs to learn from experience the mapping between the mean absorbed power and the maximum PTO displacement, and the sea state and the PTO damping and stiffness coefficients. This corresponds to system identification. However, rather than being on-line as in [17, 18], due to their statistical nature these parameters are computed from the data collected over a number of wave cycles based on the observation that the energy content of waves changes with wave groups [26]. In particular, the length of the time interval or time horizon is selected to be long enough to ensure the full decay of transient effects associated with a change in PTO coefficients. As a result, the coefficients from previous time intervals do not greatly affect the data in the current interval so that simpler feedforward ANNs can be used instead of autoregressive and local recurrent ANNs [16].

Hence, in order to train the ANNs, values of H_{s}, T_{e}, B_{PTO}, C_{PTO}, the mean absorbed power, P_{avg}, and $\max z$ are collected for each time horizon throughout the operation of the device, as entries of the training vector. The estimates for P_{avg} and $\max |z|$ can be expressed through the functions $f\left(H_{\mathrm{s}}, T_{\mathrm{e}}, B_{\mathrm{PTO}}, C_{\mathrm{PTO}}\right)$ and $g\left(H_{\mathrm{s}}, T_{\mathrm{e}}, B_{\mathrm{PTO}}, C_{\mathrm{PTO}}\right)$ respectively. The trained ANNs will then be fed to optimization functions in order to find the optimal PTO damping and stiffness coefficients for every new time horizon based on the forecast sea state conditions.

3.1. Application of ANNs to the Reactive Control of WECs

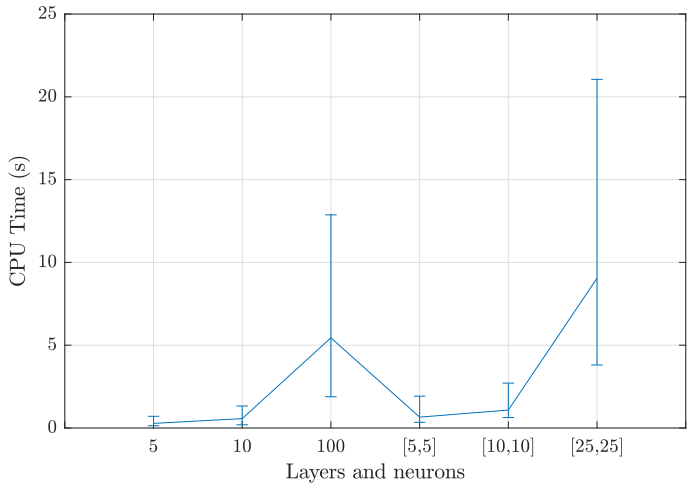
ANNs are a class of supervised learning algorithms [16]. Taking inspiration from their biological equivalent, they present a network of interconnected nodes, or neurons. Each neuron is a computational unit that maps input to output values. By combining multiple neurons in a number of layer, so that

the output of the neurons in one layer becomes the input to the neurons in the next layer, ANNs can be used to fit non-linear functions with a large number of input values.

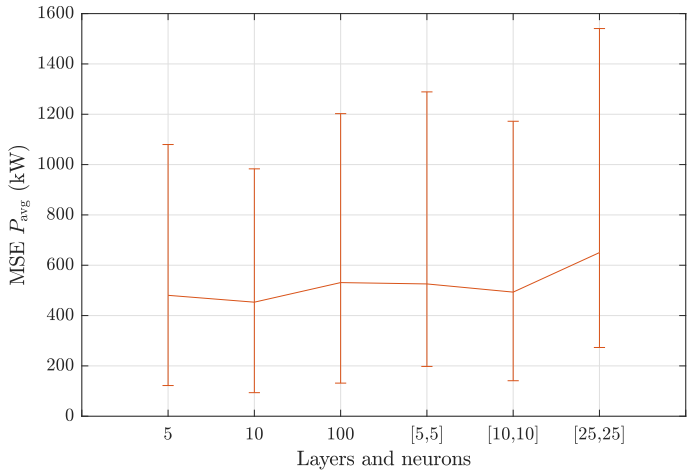
As aforementioned, in this work ANNs are employed in order to map the mean generated power and the maximum displacement at the PTO to H_{s}, T_{e}, B_{PTO} and C_{PTO}. This is achieved through a multi-layer, feedforward ANN with two output variables: P_{avg} and $\max|z|$.

In order to select a suitable size for the ANN, a preliminary study was conducted to assess the performance of possible network configurations in estimating the mean absorbed power (hence, ignoring $\max|z|$ and reducing the number of output variables to one). In particular, a single hidden layer with 5, 10, and 100 neurons, and two hidden layers with 5, 10 and 25 neurons each have been considered. For each configuration, 25 cases have been generated as the combination of 5 different random initializations of the weight matrices [27] and 5 training and test datasets. In fact, a single training dataset has been sampled from simulations in irregular waves for the sea states in Table 1, which has also been used to pre-initialize the ANN-based control in Section 4.3. According to standard practice with ANN training [27], the whole set has been subdivided into the five distinct training and test sets by randomly reordering it, and each time selecting the first 250 points for the test set (about 10%) and the remaining 2239 samples for the training set (approximately 90%). For each case, the ANN has been trained using the training samples, and then used to estimate P_{avg} for the test set. The mean square error between the prediction and the actual mean generated power value has been calculated, as well as the overall computing time required for the ANN implementation described below. Afterwards, the mean and standard deviation of these values have been computed for each network configuration, and plotted in Figure 3.

From Figure 3, it is clear that the decision on the size of the ANN should be based on a compromise between performance and accuracy. On the one hand, denser networks result in greater memory requirements and computational cost, as shown in Figure 3a. In particular, it is interesting to notice that the configuration with two hidden layers with 10 neurons each, which contains a total of 100 connections between the two hidden layers, presents a much lower computational cost than a single layer with 100 neurons, mainly due to implementation reasons. On the other hand, the deeper the network, the greater the number of features that can be matched from the original function; similarly, the greater the number of neurons, the more complex



(a)



(b)

Figure 3: Mean central processing unit (CPU) time (a) and mean square error (MSE) (b) associated with the prediction of the mean generated power for different ANN configurations in terms of hidden layers and neurons for 5 weight initializations and 5 training and test sets. The upper bar corresponds to the sum of the mean value and half the standard deviation, while the lower bar to the minimum value of all cases in order to prevent negative values.

the fitted function shape [16]. An example is the lower mean square error associated with the configurations with 10 neurons as compared with those with 5 in Figure 3b. Nevertheless, an excessive number of neurons can result

in overfitting the input data [27], i.e. fitting the random noise in addition to the underlying relationship, which is highly undesirable since the ANN is expected to generalise the shape of the P_{avg} and $\max|z|$ curves. In Figure 3b, this evidently occurs for a single hidden layer with 100 neurons and two hidden layers with 25 neurons each. Although a single hidden layer seems to perform best, this preliminary study has been carried out on a relatively small dataset, considering only a limited number of sea states. Therefore, it has been preferred to use a configuration with two hidden layers each with 10 neurons in order to represent the possible extra features associated with the additional sea states. Additionally, this results in only a minor increase in computational time. Similar results are obtained from the mapping of the maximum displacement.

A schematic diagram of the feedforward ANN can be seen in Figure 4. The network presents an input layer with 4 neurons (one for each input variable), two hidden layers with $m = 10$ and $n = 10$ neurons each, and an output layer with two output variables. Furthermore, it is possible to see that the input and hidden layers have an additional bias term, which is required to find the intercept of the fitted functions at each stage in the ANN [27]. Each layer l presents input and output variables, which are expressed as $\boldsymbol{x}_l$ and $\boldsymbol{y}_l$ respectively in vector notation. The input variables correspond to $\boldsymbol{y}_1$, while the output to $\boldsymbol{y}_4$. The signal between each two matrices is multiplied by weight matrices $\boldsymbol{W}_i$, with $i = 1, 2, 3$. The weight matrices for the bias terms are represented as $\boldsymbol{b}_i$.

Given the input data for one training example, $\boldsymbol{y}_1$, it is possible to obtain $\boldsymbol{y}_4$, the ANN estimate for the output by propagating the signal from one layer to the next one. Using this technique, known as forward propagation, the input and output vectors of each layer $l = 2, 3, 4$ can be computed for each training point in matrix notation, which is convenient from a programming perspective, as follows [27]:

$$\boldsymbol{x}_l = \boldsymbol{W}_{l-1}\boldsymbol{y}_{l-1} + \boldsymbol{b}_{l-1}, \tag{4}$$

$$\boldsymbol{y}_l = e_l(\boldsymbol{x}_l). \tag{5}$$

In (5), e_l denotes the activation function of the neurons in each layer. Figure 4 shows that the two hidden layers use the tanh activation function, while the output layer presents a linear activation function. The hyperbolic tangent is a standard smooth, non-linear activation function, which is superior to the sigmoid function, as its output is zero-centred [16]. As described in [16], for a

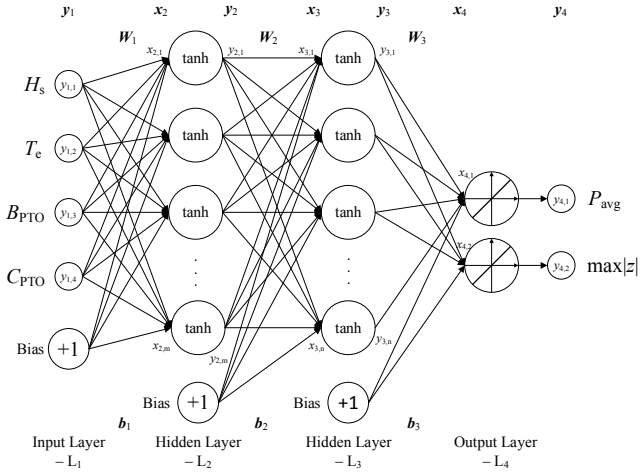


Figure 4: Schematic diagram of the feedforward ANN for the approximation of the mean generated power or maximum PTO displacement.

small number of layers, as in this case, tanh is preferred over rectified linear units, which are standard in deep learning. The linear activation function is employed in the output layer in order to return a real value, not bounded within ± 1 as would be the case if tanh had been used instead.

The mapping between input and output is dictated by the weights of the ANN [16]. Hence, learning can occur by tuning these parameters based on the training data so as to minimize an objective function, which is a measure of the error between actual and predicted output. In order to update the weight matrices, it is necessary to calculate a gradient matrix that indicates the change in the error due to a change in each weight. The gradient matrices are computed by propagating the error signal, or sensitivity, $\boldsymbol{s}_l$, from the output layer to the input layer in a process known as backpropagation [27]:

$$\begin{aligned} \boldsymbol{s}_4 &= -\left(\boldsymbol{y}_{\mathrm{tr}} - \boldsymbol{y}_4\right) \odot \dot{e}_4(\boldsymbol{x}_4) \\ \boldsymbol{s}_l &= \left(\boldsymbol{W}_l^T \boldsymbol{s}_{l+1}\right) \odot \dot{e}_l(\boldsymbol{x}_l) \end{aligned}, \tag{6}$$

where $\odot$ indicates the Hadamard, or element-wise, product, and $\dot{e}_l$ the first derivative of the activation function of each layer. $\boldsymbol{y}_{\mathrm{tr}}$ indicates the exact output of each training sample, i.e. the variables the ANNs should fit. Therefore,

$\boldsymbol{y}_{\text{tr}} = \begin{bmatrix} P_{\text{avg}} & \max|z| \end{bmatrix}^T$. The change in the weight matrices is given by [27]:

$$\Delta \boldsymbol{W}_l = \boldsymbol{s}_{l+1} \boldsymbol{y}_l^T, \tag{7}$$

$$\Delta \boldsymbol{b}_l = \boldsymbol{y}_l. \tag{8}$$

The equations above are used if the ANNs are trained using one training sample at a time, such as when the simple gradient descent scheme is applied [27]. Nevertheless, batch-mode training, i.e. employing multiple training samples at a time, is much more efficient. For this reason, the highly efficient Levenberg-Marquardt backpropagation training algorithm has been adopted instead [27]. A detailed explanation of the method, including the necessary extra equations, can be found in [28, 27] in matrix notation. The implementation within the Mathworks neural networks toolbox has been used in this work, with the default settings.

It is important to notice that the input variables, i.e. H_{s}, T_{e}, B_{PTO} and C_{PTO}, need to be normalized through their mean and standard deviation before being fed to the ANNs for training. Furthermore, the mean power values have also been normalized with respect to the maximum (for positive values) and minimum (for negative values). This has been necessary because the points lying on the $B_{\text{PTO}} = 0$ boundary of the search space presented excessively high negative power values that seriously affected the quality of the function fit.

3.2. Multistart Optimization

At the start of every new time horizon, the controller should select the PTO damping and stiffness coefficients that will result in maximum energy extraction for the predicted sea state during the horizon, in compliance with the constraint on the PTO displacement. This is clearly a non-linear optimization problem, since both P_{avg} and $\max|z|$ are non-linear functions of H_{s}, T_{e}, B_{PTO} and C_{PTO}. In addition, the values of the PTO damping and stiffness coefficients must be bounded within sensible values, so that the problem is constrained as well.

By removing the dependence on the significant wave height and wave energy period from functions f and g due to space limitations for display purposes, the cost function can be expressed at the start of each new time horizon h as follows:

$$c(h) = \begin{cases} -f\left(B_{\text{PTO}}, C_{\text{PTO}}\right) & \text{if } \left|g\left(B_{\text{PTO}}, C_{\text{PTO}}\right)\right| \leq z_{\text{Max}} \\ +1 & \text{if } \left|g\left(B_{\text{PTO}}, C_{\text{PTO}}\right)\right| > z_{\text{Max}} \end{cases} \tag{9}$$

subject to:

$$B_{\min} \leq B_{\mathrm{PTO}} \leq B_{\mathrm{Max}},\, C_{\min} \leq C_{\mathrm{PTO}} \leq C_{\mathrm{Max}}.$$

The values of the maximum and minimm allowable PTO damping and stiffness coefficients can be derived using accurate, non-linear models during the design stage in order to prevent damage to the generator in the most energetic sea states likely to be encountered, where the buoy velocity and displacement are highest.

Genetic and other nature-inspired algorithms have been extensively used recently for the solution of non-linear optimization problems that present multiple minima, as in this case [29]. Nevertheless, in this work, a strong emphasis is given to performance, since the optimization needs to be repeated at the start of each new time interval. For this reason, it has been preferred to use the Multistart algorithm [30]. This technique consists in generating a number of start points, sampled randomly within the $B_{\mathrm{PTO}}, C_{\mathrm{PTO}}$ search space. Although convergence is not assured, a large number of starting points greatly increase the chances. A value of 100 starting points has been selected for this reason. From each point, an optimization is run using a non-linear, constrained programming solver. In particular, the Mathworks functions *MultiStart* and *fmincon* have been used respectively. The main advantage of this technique over alternative methods, such as global search, is its simple parallel implementation, which can result in large savings in computational time. For instance, one Multistart optimization using the cost function in (9) takes 8.62 s on a quad-core, i7 computer with 16GB RAM, whereas a global search takes 29.20 s. A greater number of cores and an implementation in a lower-order language, such as C or Fortran, can result in even greater computational savings.

3.3. Algorithm

Figure 5 shows the algorithm for the ANN-based reactive control of the point absorber described in this article. As aforementioned, a time-averaged approach is used, where new values of B_{PTO} and C_{PTO} are selected at the start of every new time horizon h and applied throughout its duration $D(h)$. On the one hand, a longer duration is preferable for the power averaging and sea state statistical analysis so as to produce less noisy training data. On the other hand, a shorter time span can result in faster training. Furthermore, the controller would be able to track changes in the sea state on a smaller time scale, thus moving towards real-time control and possibly higher energy

extraction. For these reasons, $D(h) = 20T_e(h)$ has been chosen in both regular and irregular waves.

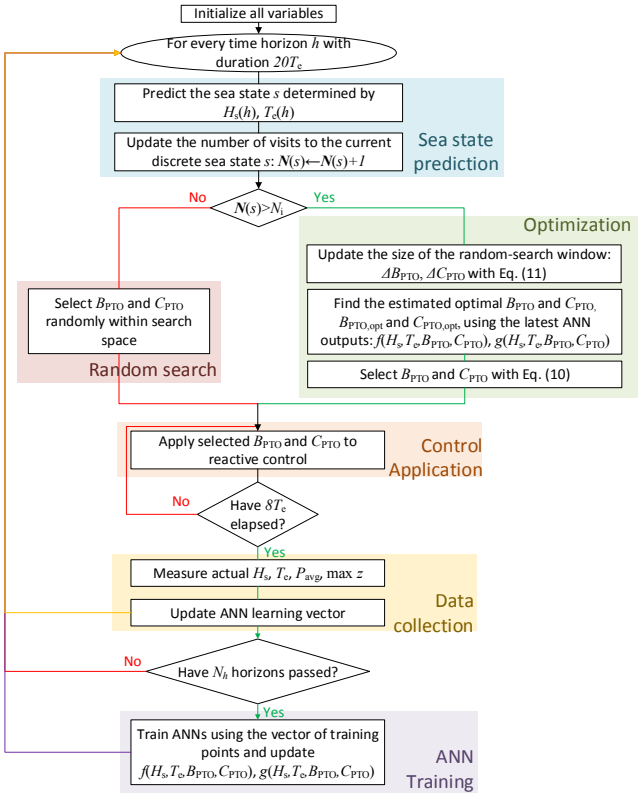


Figure 5: Flow chart of the ANN-based reactive control of a point absorber.

As can be seen from Figure 5, the first step in every time horizon is to predict the significant wave height and energy wave period during the time interval. Different approaches have been proposed for this problem, with example methods being Kalman filters, deterministic sea wave prediction [31], autoregressive models [32], and even ANNs [33]. Although these studies analyse the wave elevation, which is forecast with accuracy only 15 s into the future, it is assumed that similar strategies can be found for the forecast of the statistical wave conditions for one time horizon. For simplicity, in this initial work the actual values for H_s and T_e have been used, since the wave traces employed in the simulations are known in advance. $H_s(h)$ and $T_e(h)$ are then used to update the count of the number of observations in

the current discrete sea state, s. For this purpose a table, $\boldsymbol{N}$, is employed, with an entry for each discrete sea state with ranges of 1 m and 1 s for each dimension respectively.

During the first N_i visits to each discrete sea state, the values of the PTO damping and stiffness coefficients are selected randomly to ensure initial exploration. Once $\boldsymbol{N}(s) > N_\mathrm{i}$, the Multistart optimization can be run using the cost function in (9) in order to find the optimal coefficients, $B_{\mathrm{PTO,opt}}$ and $C_{\mathrm{PTO,opt}}$, for the forecast significant wave height and energy wave period. However, the ANN estimates f and g can be very inaccurate initially. For this reason, B_{PTO} and C_{PTO} are in fact selected randomly within a region around the optimum that shrinks with the number of data points collected in the sea state:

$$B_{\mathrm{PTO}} = B_{\mathrm{PTO,opt}} + \Delta B_{\mathrm{PTO}}, \tag{10}$$

where

$$\Delta B_{\mathrm{PTO}} = (r - 0.5) \cdot \mathrm{range}(B_{\mathrm{PTO}}) \cdot 0.9^{\boldsymbol{N}(s) - N_\mathrm{i}}, \tag{11}$$

with $r = [0, 1]$ a random number. The same applies to C_{PTO}. Upper and lower bounds are used to ensure the chosen values lie within the desired range. As more data points are collected in the optimal region, the accuracy of the ANN fit increases.

Once B_{PTO} and C_{PTO} are chosen and applied, measurements are employed to compute the mean absorbed power, maximum PTO displacement and actual $H_\mathrm{s}(h)$ and $T_\mathrm{e}(h)$ during the time interval. These values are in fact calculated using the data only after an initial time of $8T_\mathrm{e}(h)$ within the current horizon h in order to exclude the initial transient effects. This relatively long time also ensures that the time required for the Multistart optimization does not become an issue. Once the desired values are obtained, they are stored in memory as a data sample so that they can be used for training the ANN.

The ANN is trained every $N_h = 20$ time horizons, employing 90% of training points. The remaining 10% of the samples is used for validation and hence to check the quality of the fit. Each sample presents H_s, T_e, B_{PTO}, C_{PTO} as input, and P_{avg} and $\max|z|$ as output. The larger the number of training points, the less the risk of overfitting the data and the more accurate the estimates of the ANN. However, this will also cause an increase in training time and, more importantly, it may result in an excessive memory requirement. Therefore, for a practical application, it is expected that the number of training points will be limited to a large number, say 10^6. Care

will be needed in order to ensure that a similar number of data points is kept for each discrete sea state when overriding old data with new readings, as well as to explore a broad range of B_{PTO} and C_{PTO} values so as to aid the training of the ANN.

4. Simulation Results

4.1. Simulation System

The proposed algorithm has been tested using the same point absorber as in [22] and [14]: a floating vertical cylinder with 5-m radius and 8-m draught. Deep water is assumed in the determination of the hydrodynamic coefficients, with the radiation approximation state-space vector presenting five entries as in [14]. Similarly, the hydrodynamic model in Fig. 2 has been arranged in a state-space system and discretized using a first-order accurate Euler scheme, with a sampling time of 0.1 s. The same PTO force saturation and float displacement limits of 1 MN and ± 5 m respectively have also been adopted, as well as a PTO efficiency of 75%.

Figure 6 shows graphically the the program used for the simulation of the WEC. Instead of sensors installed on a wave buoy, in the simulations a wave model provides the wave elevation record as in [14], as can be seen in Figure 1. For irregular waves, the wave elevation is computed as the superposition of multiple individual wave components, whose amplitude is derived from the specified wave spectrum [26]. A value of 0.005 rad/s has been selected for the circular wave frequency step, since this value is smaller than the Nyquist frequency for a 15-minute window so as to prevent a repetition of the wave trace [34]. Therefore, each trace of irregular waves is generated as the combination of 15-minute-long time series, where the random number generator is initialized with a different seed for each component. In order to smooth the connection between the separate traces, a 20-point filter is employed over the last and first of each consecutive time series. The wave elevation time series has a dual purpose: on the one hand, it is used to establish H_{s} and T_{e} in each time horizon [26]; on the other hand, the convolution integral of the wave elevation and diffraction impulse response function produces the wave excitation force [35].

The search space has been limited to within $B_{\mathrm{min}} = 0$ and $B_{\mathrm{Max}} = 2$ MNs/m, and $C_{\mathrm{min}} = -1$ MN/m and $C_{\mathrm{Max}} = 0$ for the PTO damping and stiffness coefficients respectively. A wider search space has been selected for the PTO damping coefficient in order to prevent damage in large waves,

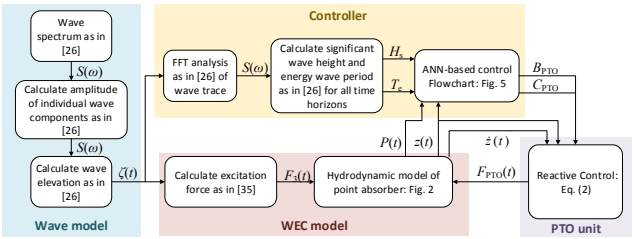


Figure 6: Flow chart of the program used in the simulations of the point absorber.

when greater damping and no stiffness are required. Nevertheless, the larger the search space, the longer the learning time; hence, an excessive search space needs to be avoided.

For the first 15 minutes of the simulations, no control force is applied in order to let the system dynamics settle. For this reason, all wave traces are in fact generated with an extra 15-minute interval at the start.

4.2. Results in Regular Waves

In regular waves, a 6-hour-long wave trace with unit amplitude and a wave period of 8 s has been analysed. As can be seen in Figure 7a and Figure 7b, the ANN-based algorithm learns successfully the optimal PTO damping and stiffness coefficients respectively. In particular, the optimal values (dotted lines) have been obtained with a Multistart optimization using a wave trace lasting 20 minutes and the analysed WEC model. Figure 7c shows the difference in the mean power generated with ANN-based control and state-of-the-art reactive control, where $P_{\mathrm{avg,opt}} = 176.24$ kW. A value of $N_{\mathrm{i}} = 40$ has been used.

4.3. Results in Irregular Waves

In irregular waves, even within a single sea state, the significant wave height and wave energy period do vary, if they are measured within a short time interval like $20T_{\mathrm{e}}$. Since reinforcement learning in [14] and [15] employs discrete states, it was possible to show the convergence behaviour of the algorithm in one sea state only. Conversely, the accuracy of ANNs is greatly improved and the effects of overfitting greatly reduced the wider the range of their samples [27] and thus the wider the range of sea conditions. For this reason, the proposed ANN-based reactive control algorithm is run for the 9 wave traces shown in Table 1. Each wave time series is generated

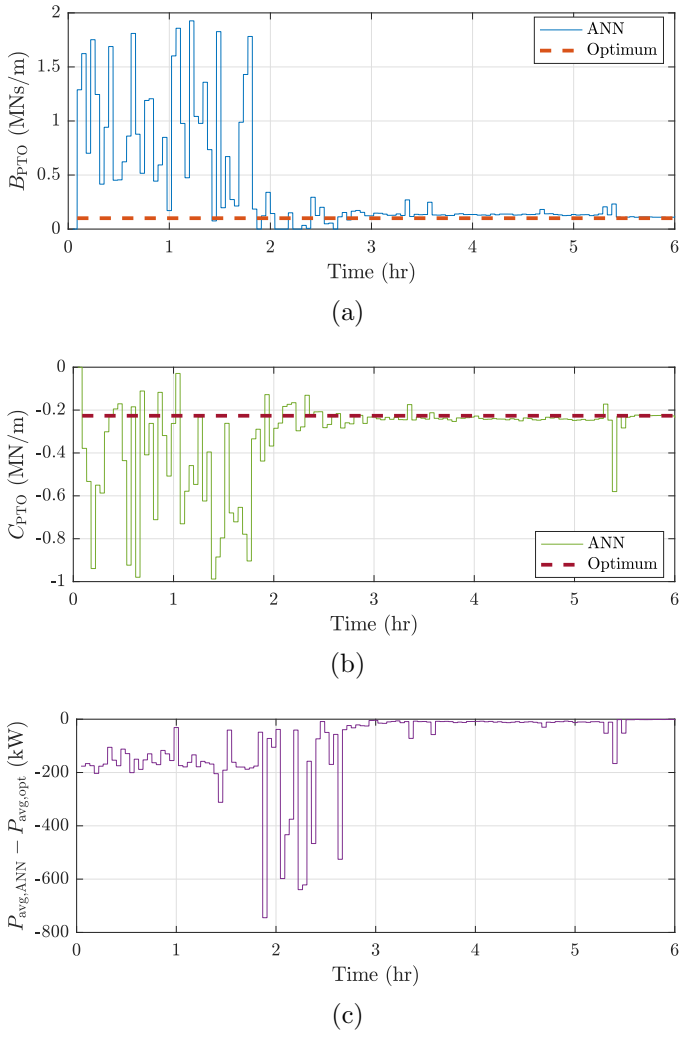


Figure 7: PTO damping (a) and stiffness (b) coefficients obtained from the ANN-based control as compared with the optimal value in regular waves with $H_{\mathrm{s}} = 2$ m and $T_{\mathrm{e}} = 8$ s. (c) shows the difference in the corresponding mean generated power.

with a Bretschneider spectrum (thus, broad-banded) [26] and lasts 3 hours. Although these wave traces have been simulated independently due to computational constraints, they should be treated as a continuous, time series where 9 independent sea states are observed in the order provide in Table 1, with a value of $N_{\mathrm{i}} = 120$ being used. In particular, for each wave trace the list of samples is initialized with the values observed in the previous runs. The

series of a sea states is repeated another time but with a different seed number to the random number generator for a total wave trace with an overall duration of 54 hours (excluding the 15 minutes required for the initialization of each wave trace).

Table 1: Significant wave height, energy wave period and duration of the wave traces used for the analysis of the ANN-based control in irregular waves.

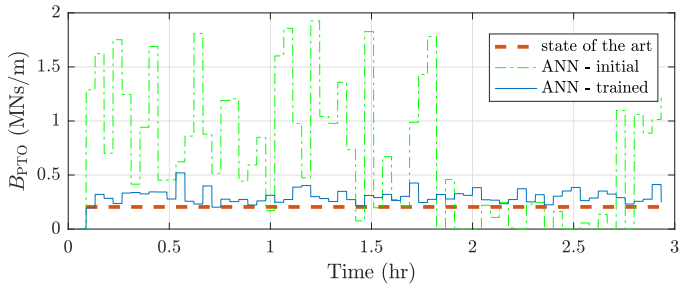
[!h] H_s (m)	2	1	1	1	2	2	3	3	3
T_e (s)	8	8	9	10	9	10	10	9	8
duration (hr)	3	3	3	3	3	3	3	3	3
no. repetitions	2	2	2	2	2	2	2	2	2

The learning behaviour of the proposed ANN-based reactive control algorithm in irregular waves is displayed in a compact way in Figure 8. The figure shows the controller performance for the first wave trace, i.e. $H_s = 2$ m and $T_e = 8$ s. In particular, the very first run (when the list of samples is empty at the start) is shown with dotted lines and labelled as "initial", since learning has just been initialized. The system is simulated in the same wave conditions again *after* the control has been applied for 54 hours in the wave traces shown in Table 1. The corresponding performance is shown with continuous lines in Figure 8 and labelled as "trained", since learning has completed by then with a large number of samples being available for the training of the ANN. Furthermore, in this case the exploration rate has almost fully decayed, as the discrete sea state has already been experienced for 6 hours. Additionally, the optimal value for the PTO coefficients and the corresponding absorbed energy is calculated running a MultiStart optimization of the WEC model in the same wave trace.

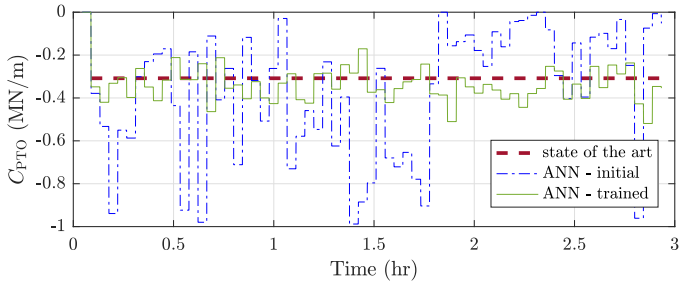
5. Discussion

5.1. Regular Waves

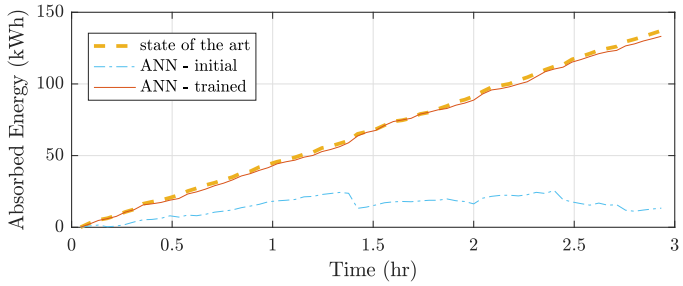
As shown in Figure 7, the ANN-based algorithm learns the optimal PTO damping and stiffness coefficients in regular waves within 4 hours after being randomly initialized. In the figures, it is possible to recognize three distinct regions: an initial region where completely random actions are selected ($\boldsymbol{N}(s) \leq N_i$), a section where random actions are taken around the expected



(a)



(b)



(c)

Figure 8: PTO damping (a) and stiffness (b) coefficients adopted by the ANN-based control at the start and after 54-hours of training in the wave conditions shown in Table 1 in irregular waves with $H_{\mathrm{s}} = 2$ m and $T_{\mathrm{e}} = 8$ s. Additionally, (a) and (b) display the results of state-of-the-art reactive control. The corresponding curves for the absorbed energy are plotted in (c).

optimum within a shrinking range (until $0.9^{\boldsymbol{N}(s)-N_i} \to 0$), and a final part where convergence has been reached. Within this last region, it is interesting to notice three random points (after approximately 5.5 hours). These are caused by the Multistart algorithm converging towards the wrong local optimum in the corresponding time horizons. This is a possibility that needs to

be taken into account when designing the control for an actual device, with its probability decreasing with the number of starting points. Nevertheless, the low computational cost means this optimization method is still preferred over global search or genetic algorithms. Oddly, the three random points also provide the ANNs with the missing training points for perfect convergence to the optimal PTO coefficients.

It should be noted that the ANN-based algorithm presents faster learning than reinforcement learning, which requires approximately 6 hours in regular waves with resistive control and 8 hours with reactive control in [14] and [15], respectively.

5.2. Irregular Waves

The convergence of the algorithm to the optimal PTO coefficients in irregular waves is shown by the "trained" lines in Figure 8. Oscillations in the values obtained with the ANN-based control are due to changes in wave conditions over the smaller time scale of $20T_\mathrm{e}$. The energy absorption is almost identical to state-of-the-art reactive control applied using the optimal coefficients for the WEC model in this wave trace.

In this case, the comparison in learning performance between reinforcement learning and ANNs is harder to understand. At first sight, 54 hours may seem like a very long learning time. However, this corresponds to 6 hours of learning time per discrete sea state, which is less than the 10 hours required by reinforcement learning in a single sea state of irregular waves for resistive control in [14]. Once a sufficient number of points is obtained, the ANN can generalise the information to unseen sea states, thus further reducing the learning time as compared with reinforcement learning with discrete states. In addition, the convergence time should be assessed in the context of the lifetime of a WEC, which is expected to be 20 to 25 years long [25].

In this work, discrete sea states have been analysed, each lasting 3 hours due to practical issues with the code implementation. In reality, the energy content in waves changes uniformly in time (hence, not through discrete sea states), with the duration of a typical sea state being 0.5 to 6 hours [26]. Since P_avg and $\max|z|$ can be considered to be purely dependent on the values of B_PTO, C_PTO, H_s and T_e in the *current* time interval, the samples of the ANN are independent of past data. Therefore, the algorithm can be safely applied to realistic, continuously varying wave conditions. In fact, the quality of the mapping provided by the ANN is expected to improve in continuously varying sea states, which result in a broader range of samples

[27]. Furthermore, under realistic wave conditions, the ANN-based reactive control is expected to result in higher energy absorption than state-of-the-art reactive control, since the latter uses a look-up table with discrete sea states, thus being less responsive to changes in wave energy over a shorter time scale. Additionally, the ANN-based method can adapt to changes in the device dynamics with time, e.g. due to marine growth.

5.3. Practical Considerations

Although ANNs are a supervised learning strategy, they are employed here in an approach reminiscent of reinforcement learning, which is unsupervised, and that entails exploration. This may result in damage to and even failure of the device if explorative negative actions are selected at the wrong time, e.g. a high PTO stiffness coefficient with low damping in highly energetic waves. Strategies that rely on explorations suffer from this problem, but the ANN-based control is more affected than reinforcement learning because:

- reinforcement learning [15] makes a step change in the PTO coefficients at the start of each interval. Hence, it is difficult to encounter highly negative situations, since the algorithm corrects the PTO coefficients as soon as it starts receiving negative feedback on the actions it has selected in that particular sea state. Conversely, the proposed ANN-based method is able to explore the whole search space during the first observations of a particular discrete sea state.

- the quality of the mapping produced by the ANN is improved for a wider range of samples. Hence, in order to improve the training process, the algorithm is incentivized to explore most combinations of the PTO coefficients in each sea state.

In order to prevent failure or damage to the WEC, two practical possibilities should be investigated:

- initializing the ANN with samples pre-generated using accurate, non-linear models of the WEC. In particular, simulations should be run using extreme values of the PTO coefficients so that once the algorithm is applied on the real-device, the Multi-Start optimization should home in onto the optimal conditions rather than risk selecting extreme control settings.

- an alternative approach consists in initially applying state-of-the-art reactive control with the look-up table approach and slowly changing the PTO coefficients in each sea state. The collected data will be then employed for the training of the ANN, and then the proposed algorithm will be applied. This process is designed to remove the exploration stage from the presented scheme, focusing only on the supervised nature of ANN algorithms.

At the moment, exact knowledge of the values of H_s and T_e during the following time horizon is assumed. In practice, errors will be associated with the estimation method [31, 32, 33]. Nevertheless, including information on the expected future wave excitation is a fundamental tool for the control of WECs in order to try to achieve optimal performance [5]. This is a further improvement over the reinforcement learning algorithm proposed in [15], since the study assumed the wave height and period to be identical between neighbouring horizons.

As compared with reinforcement learning, the selection of the control action with the proposed method requires greater computational power. Nevertheless, an on-line implementation is completely feasible with modern hardware and parallel computing. As described in Section 3.3, the control strategy consists in two main stages. On the one hand, the weight of the ANNs are updated every 20 algorithm steps, using all training points (at most, say, 10^6). This process occurs off-line, with the older weights not being overwritten on the memory until the new ones are ready, so that computing time is not an issue. On the other hand, at every time step, a new training point is collected (minimal computation effort), and a new action is selected through the Multistart optimization. This process is speeded up through parallel processing, and possibly an implementation in a low-order computational language, e.g. C. As described in Section 3.2, one optimization using the current simulation in Matlab and a quad-core i7 computer with 16 GB RAM takes less than 9 s. This period is less than 10% of the minimal expected time horizon duration, namely 100 s for a 5-s wave energy period, which is the smallest encountered in typical sea states [26]. Hence, computation efficiency is not critical in this case, since the rate of change of the plant is much slower than that of the control algorithm.

Real-time strategies are more efficient than time-averaged methods for the control of WECs [5]. Hence, although machine learning schemes are interesting due to their model-free approach of the WEC control problem, they will

need to be applied in real-time in order to compete with model-predictive control. The application of real-time system identification with ANNs to real-time strategies, such as model predictive control, will be investigated in the future.

6. Conclusion

In this article, an on-line, model-free strategy has been developed for the reactive control of WECs using ANNs. The aim is to maximise energy absorption, whilst limiting the PTO displacement to prevent failure in energetic sea conditions. A simple model of a point absorber has been employed to analyse the behaviour of the algorithm. Firstly, regular waves show that the strategy learns rapidly the optimal PTO damping and stiffness coefficients because of their periodicity. A longer convergence time is necessary in irregular waves, since the ANNs require a greater number of training samples in order to learn the mapping between the mean absorbed power and PTO displacement, and the significant wave height, wave energy period, and the PTO damping and stiffness coefficients. Nevertheless, this ensures the scheme can recognize variations in the wave conditions on a shorter time scale than state-of-the-art reactive control, which uses discrete sea states. Furthermore, implementation on a full-scale WEC is simple, as the technique is independent of models of the machine dynamics. More importantly, this method is able to treat changes in the device response as the structure is affected by marine biofouling.

Acknowledgment

The Industrial Doctoral Training Centre for Offshore Renewable Energy is a partnership of the universities of Edinburgh, Exeter and Strathclyde. This work was supported partly by the Energy Technologies Institute and the Research Councils Energy Programme (grant EP/J500847/), and partly by the Engineering and Physical Sciences Research Council (grant EP/J500847/1). Additionally, the first author's Eng.D. project is sponsored by Wave Energy Scotland.

Wave Energy Scotland is taking an innovative approach to supporting the development of wave energy technology by managing the most extensive technology programme of its kind in the sector, concentrating on key areas which have been identified as having the most potential impact on long term levellised cost of energy and improved commercial viability.

References

[1] K. Gunn, C. Stock-Williams, Quantifying the Potential Global Market for Wave Power, Proceedings of the 4th International Conference on Ocean Engineering (ICOE 2012) (2012) 1–7.

[2] A. F. D. O. Falcão, Wave energy utilization: A review of the technologies, Renewable and Sustainable Energy Reviews 14 (3) (2010) 899–918. doi:10.1016/j.rser.2009.11.003.

[3] A. J. Nambiar, D. I. M. Forehand, M. M. Kramer, R. H. Hansen, D. M. Ingram, Effects of hydrodynamic interactions and control within a point absorber array on electrical output, International Journal of Marine Energy 9 (2015) 20–40. doi:10.1016/j.ijome.2014.11.002.

[4] S. H. Salter, J. R. M. Taylor, N. J. Caldwell, Power conversion mechanisms for wave energy, Proceedings of the I MECH E Part M 216 (1) (2002) 1–27. doi:10.1243/147509002320382112.

[5] J. V. Ringwood, G. Bacelli, F. Fusco, Energy-Maximizing Control of Wave-Energy Converters: The Development of Control System Technology to Optimize Their Operation, IEEE Control Systems Magazine 34 (5) (2014) 30–55.

[6] K. Budal, J. Falnes, Optimum Operation of Wave Power Converter, Marine Science Communications 3 (2) (1977) 133–150.

[7] A. Babarit, A. H. Clément, Optimal latching control of a wave energy device in regular and irregular waves, Applied Ocean Research 28 (2) (2006) 77–91. doi:10.1016/j.apor.2006.05.002.

[8] J. Hals, J. Falnes, T. Moan, Constrained Optimal Control of a Heaving Buoy Wave-Energy Converter, Journal of Offshore Mechanics and Arctic Engineering 133 (1) (2011) 011401. doi:10.1115/1.4001431.

[9] T. K. A. Brekken, On Model Predictive Control for a point absorber Wave Energy Converter, Proceedings of the IEEE Trondheim PowerTech (2011) 1–8doi:10.1109/PTC.2011.6019367.

[10] F. Fusco, J. V. Ringwood, A simple and effective real-time controller for wave energy converters, IEEE Transactions on Sustainable Energy 4 (1) (2013) 21–30. doi:10.1109/TSTE.2012.2196717.

[11] K. U. Amann, M. E. Magaña, S. Member, O. Sawodny, Model Predictive Control of a Nonlinear 2-Body Point Absorber Wave Energy Converter With Estimated State Feedback, Sustainable Energy, IEEE Transactions on 6 (2) (2015) 336–345.

[12] D. Oetinger, M. E. Magaña, S. Member, O. Sawodny, Decentralized Model Predictive Control for Wave Energy Converter Arrays, Sustainable Energy, IEEE Transactions on 5 (4) (2014) 1099–1107.

[13] O. Sawodny, D. Oetinger, M. E. Magaña, Centralised model predictive controller design for wave energy converter arrays, IET Renewable Power Generation 9 (2) (2015) 142–153. doi:10.1049/iet-rpg.2013.0300.

[14] E. Anderlini, D. I. M. Forehand, P. Stansell, Q. Xiao, M. Abusara, Control of a Point Absorber using Reinforcement Learning, Transactions on Sustainable Energy 7 (4) (2016) 1681–1690.

[15] E. Anderlini, D. I. M. Forehand, E. Bannon, Q. Xiao, M. Abusara, Reactive Control of a Two-Body Point Absorber using Reinforcement Learning, Ocean Engineering (IDCORE Special Issue) (2017) in press.

[16] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444. arXiv:arXiv:1312.6184v5, doi:10.1038/nature14539.

[17] D. Valério, M. J. G. C. Mendes, P. Beirão, J. Sá da Costa, Identification and control of the AWS using neural network models, Applied Ocean Research 30 (3) (2008) 178–188. doi:10.1016/j.apor.2008.11.002.

[18] S. Giorgi, J. Davidson, J. V. Ringwood, Identification of Wave Energy Device Models From Numerical Wave Tank DataPart 2: Data-Based Model Determination, IEEE Transactions on Sustainable Energy 7 (3) (2016) 1020–1027. doi:10.1109/TSTE.2016.2515500.

[19] E. Tedeschi, M. Carraro, M. Molinas, P. Mattavelli, Effect of control strategies and power take-off efficiency on the power capture from sea waves, IEEE Transactions on Energy Conversion 26 (4) (2011) 1088–1098. doi:10.1109/TEC.2011.2164798.

[20] L. Castellini, M. D. Andrea, N. Borgarelli, Analysis and Design of a Recipracating Linear Generator for a PTO, in: International Symposium on Power Electronics, Electrical Drives, Automation and Motion Analysis, 2014, pp. 1373–1379.

[21] C. Wei, Z. Zhang, W. Qiao, L. Qu, An Adaptive Network-Based Reinforcement Learning Method for MPPT Control of PMSG Wind Energy Conversion Systems, IEEE Transactions on Power Electronics 8993 (c) (2016) 1–1. doi:10.1109/TPEL.2016.2514370.

[22] J. N. Newman, Marine Hydrodynamics, MIT Press, 1977.

[23] W. E. Cummins, The impulse response function and ship motions, Schiffstechnik 47 (9) (1962) 101–109.

[24] D. Forehand, A. E. Kiprakis, A. Nambiar, R. Wallace, A Bidirectional Wave-to-Wire Model of an Array of Wave Energy Converters, IEEE Transactions on Sustainable Energy 7 (1) (2016) 118–128. doi:10.1109/TSTE.2015.2476960.

[25] J. Cruz, OceanWave Energy, Springer-Verlag, 2008.

[26] L. H. Holthuijsen, Waves in Oceanic and Coastal Waters, Cambridge University Press, 2007.

[27] M. T. Hagan, H. B. Demuth, M. H. Beale, O. De Jesús, Neural Network Design, 2nd Edition, PWS Publishing, 1996.

[28] M. T. Hagan, M. B. Menhaj, Training Feedforward Networks with the Marquardt Algorithm, IEEE Transactions on Neural Networks 5 (6) (1994) 989–993. doi:10.1109/72.329697.

[29] J. S. Arora, Introduction to Optimum Design, 3rd Edition, Academic Press, 2012. arXiv:arXiv:1011.1669v3, doi:10.1016/B978-0-12-381375-6.00009-7.

[30] Z. Ugray, L. Lasdon, J. Plummer, F. Glover, Scatter Search and Local NLP Solvers : A Multistart Framework for Global Optimization, Information Systems 19 (May) (2006) 328–340. doi:10.1287/ijoc.1060.0175.

[31] G. Li, G. Weiss, M. Mueller, S. Townley, M. R. Belmont, Wave energy converter control by wave prediction and dynamic programming, Renewable Energy 48 (2012) 392–403.

[32] F. Fusco, J. V. Ringwood, Short-term wave forecasting with ar models in real-time optimal control of wave energy converters,

IEEE International Symposium on Industrial Electronics (2010) 2475–2480doi:10.1109/ISIE.2010.5637714.

[33] H. E. Shoori J., B. Ling, B. A. Batten, Use of artificial neural networks for real-time prediction of heave displacement in ocean buoys, 3rd International Conference on Renewable Energy Research and Applications, ICRERA 2014 (2015) 907–912doi:10.1109/ICRERA.2014.7016517.

[34] G. F. Franklin, J. D. Powell, A. Emami-Naeini, Feedback Control of Dynamic Systems, 6th Edition, Pearson, 2008.

[35] J. Falnes, Ocean waves and Oscillating systems, paperback Edition, Cambridge University Press, 2005. doi:10.1016/S0029-8018(02)00070-7.