

Received October 5, 2021, accepted October 31, 2021, date of publication November 4, 2021, date of current version November 11, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3125600

Nonlinear Data-Based Hydrodynamic Modeling of a Fixed Oscillating Water Column Wave Energy Device

MARCO ROSATI¹, THOMAS KELLY², AND JOHN V. RINGWOOD¹, (Senior Member, IEEE)

¹Centre for Ocean Energy Research (COER), Maynooth University, County Kildare, Maynooth, W23 F2H6 Ireland

²Centre for Renewable Energy at Dundalk IT (CREDIT), Dundalk Institute of Technology (DkIT), County Louth, Dundalk, A91 K584 Ireland

Corresponding author: Marco Rosati (marco.rosati.2021@mumail.ie)

This work was supported by the Science Foundation Ireland through the Marine and Renewable Energy Ireland (MaREI) Centre under Grant 12/RC/2302_P2 and Grant 13/IA/1886.

ABSTRACT System identification (SI) techniques represent an alternative strategy to provide the hydrodynamic model of oscillating water column (OWC) devices, compared to more traditional physics-based methods, such as linear potential theory (LPT) and computational fluid dynamics (CFD). With SI, the parameters of the model are obtained, by minimizing a model-related cost function, from input-output data. The main advantage of SI is its simplicity, as well as its potential validity range, where the dynamic model is valid over the full range for which the identification data was recorded. The paper clearly shows the value of a global nonlinear model, both in terms of accuracy and computational simplicity, over an equivalent multi-linear modelling solution. To this end, the validation performance of the nonlinear model is compared to the results provided by a range of linear models. Furthermore, in order to provide a more comprehensive comparative analysis, some practical aspects related to real-time implementation of multi-linear and nonlinear SI models are discussed. For the experimental campaign, real wave tank (RWT) data of a scaled OWC model are gathered from the narrow tank experimental facility at Dundalk Institute of Technology (DkIT). Particular attention is paid to the selection of suitable input signals for the experimental campaign, in order to ensure that the model is subjected to the entire range of equivalent frequencies, and amplitudes, over which model validity is required.

INDEX TERMS Data-based hydrodynamic modelling, linear ARX model, nonlinear KGP model, oscillating water column, real wave tank, system identification, wave energy.

I. INTRODUCTION

The global wave energy potential has been estimated by different authors [1]–[3], who report around 16000 – 18500 TWh/year, and a slow variation rate, of around 500 TWh/decade on average [3]. This makes wave energy a significant and relatively constant source of renewable energy.

Wave energy converters (WECs) harness wave power, and one of the most promising devices is the so-called oscillating water column (OWC), depicted in Fig. 1. Note that the scaled OWC tested in the current work, which is presented in Section II-A, is not equipped with turbine/generator group

and, hence, represents a simplified version of the device in Fig. 1. Its simplicity of operation and the presence of few moving components, all located above the water level, are the main advantages of OWCs. The motions of a water column, excited by incident waves, alternatively compress and decompress an air volume contained inside a pneumatic chamber. The chamber is connected to the atmosphere and, as a result, a bidirectional air flow is generated. A power-take-off (PTO) mechanism, namely a self-rectifying air turbine directly coupled with an electric generator, converts pneumatic energy to electricity, which is finally transferred to the power grid. A comprehensive review of OWC devices and air turbines is provided in [4]. Currently, in addition to the high production and operational costs, the commercial development of OWCs is also hindered because of the lack of efficient control systems and, in relation to this, it is imperative to

The associate editor coordinating the review of this manuscript and approving it for publication was Claudia Raibulet.

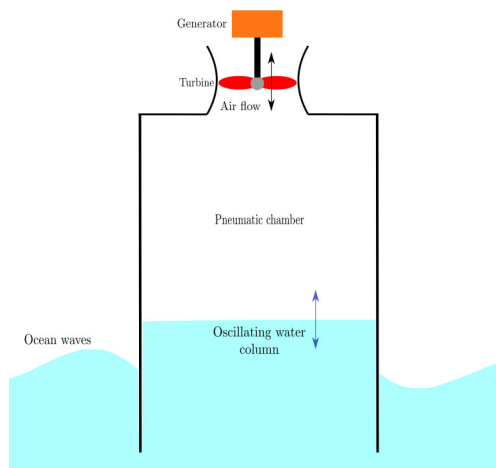


FIGURE 1. Schematic representation of a generic OWC device.

note that the control problem is intrinsically characterized by modelling-related assumptions and requirements [5].

System identification (SI) techniques [6] represent an alternative strategy to provide the hydrodynamic model of OWC devices [7], compared to the laborious process of model determination from first principles, in which, as a consequence of linear potential theory (LPT) assumptions, the model is linearised around its equilibrium point. In contrast to LPT, the main advantage of SI, besides its simplicity, is its validity range, meaning that the dynamic model is valid over the full range for which the identification data have been recorded. Ultimately, SI is a data-based modelling approach in which the parameters of the model are obtained from input-output data by minimizing a cost function, related to the model fidelity. SI models can be solely based on data (black-box models), or else can incorporate some physics-based information (grey-box models) [8]. In either case, a suitable model, and a fitting criterion, which is needed in order to evaluate the performance of the model, have to be chosen. The parameters of the model are then identified by feeding a numerical optimisation algorithm with training, or identification, data. Finally, the model is validated against a separate set of validation data, meaning that the capacity of the model to predict the behaviour of the device, while working with a different data set, is assessed. This procedure is summarized by the loop shown in Fig. 2. An essential aspect to keep in mind is that the validation data must be different from those employed during the identification. Moreover, it has to be noted that the choice of a suitable input signal is of primary importance for the parametric model to perform well in the real sea environment. The range of frequencies, and amplitudes, of the input signals employed in the training phase should correspond to the range of equivalent frequencies, and amplitudes, over which model validity is required.

Although the application of SI in wave energy is relatively recent, SI has already been successfully employed in WEC modelling [9]–[15]. In two joint publications [10], [11], SI is used in order to derive a force-to-position model,

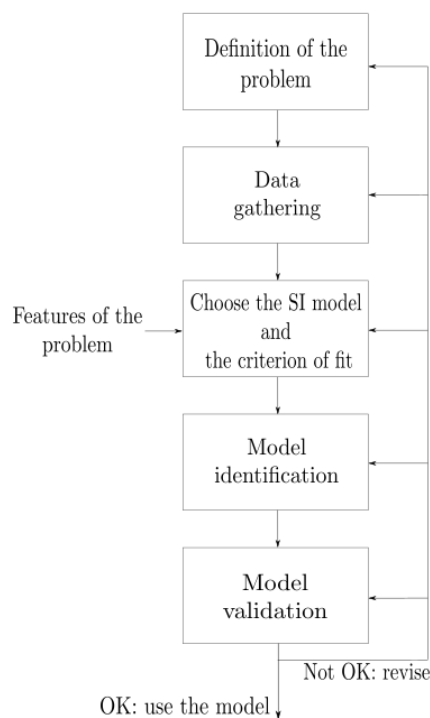


FIGURE 2. Schematic representation of the SI loop.

with a force being the PTO force, and a wave-to-position model of a WEC by using three discrete time parametric structures. Two of the three model structures, namely the autoregressive with exogenous input (ARX) model and the nonlinear Kolmogorov-Gabor polynomial (KGP) model, are *linear in the parameters*, whereas the multi-layer perceptron (MLP) artificial neural network (ANN), is nonlinear in the parameters, with a consequently greater challenge in optimal parameter identification. The data are gathered from numerical wave tank (NWT) experiments, which are built by means of an open source computational fluid dynamics (CFD) software. In [13], the wave-to-force and the force-to-position models of a WEC are identified, in the frequency domain, by using a linear black-box model, and a nonlinear Hammerstein-Wiener model. Note that the output in the wave-to-force model is the excitation force, whereas the input in the force-to-position model is the PTO force. In a recent paper [14], a linear force-to-position model for a three-body hinged-barge WEC is identified using real wave tank (RWT) data and, similarly to the previous cases, the input is the PTO force. The authors identify a parametric model for the transfer function (TF) of the system by following three different strategies. The first approach is carried out in the frequency domain, whereas the other two utilize time domain identification techniques. Finally, in [7], some linear wave-to-position models of a scaled OWC device, which is tested in irregular and regular waves, are identified. For the irregular wave case, the identification is carried out using linear ARX models whereas, for the regular wave case, an identification procedure is developed in the frequency domain.

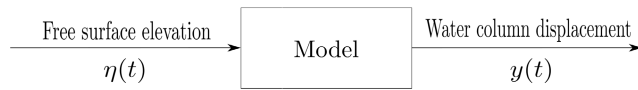


FIGURE 3. Schematic representation of the model considered in this work.

In this work, system identification is employed to derive linear and nonlinear black-box hydrodynamic models of a scaled fixed OWC device. The models are wave-to-position models of the type depicted in Fig. 3. On the left side of Fig. 3, the input of the model is the free surface elevation (FSE), namely the water elevation associated with an undisturbed (by the presence of a device) wave. On the right side of Fig. 3, the output of the model is the height of the water column measured at its centroid. Ultimately, the wave-to-position model in Fig. 3 is a mapping between FSE and the displacement of the water column. In order to develop the SI models, the time traces of the free surface elevation, $\eta(t)$, and the water column displacement, $y(t)$, are recorded and sampled during RWT experiments, where a scaled fixed OWC model is tested in irregular waves (IWs). The data sets are then employed in order to derive separate linear and nonlinear models of the type depicted in Fig. 3. Despite the fact that, normally, nonlinear models are significantly more complex than linear models, in this work, the nonlinear modelling procedure is somewhat a natural, and relatively straightforward, extension of the linear modelling approach, at least for the specific structure of the SI models sought here (linear ARX and nonlinear KGP models). Ultimately, the current paper focuses on assessing the value of a global nonlinear modelling approach over a multi-linear modelling solution for OWCs. To this end, the performance of a nonlinear KGP model is compared to the performance of a set of linear ARX models. Furthermore, in spite of the fact that 'linear' might superficially sound more attractive than 'nonlinear', the multi-linear approach considered here has some significant caveats. Therefore, in order to provide a more comprehensive and fair comparison, some practical aspects related to real-time implementation of multi-linear and nonlinear SI models are discussed. In the literature, hydrodynamic model for the characterization of OWC systems has already been the subject of a number of high-profile exercises [16]–[19]. These exercises typically adopt either linear BEM models or a CFD-based approach. However, CFD may be too complicated for many practical purposes, such as real-time control, whereas linear models may be inadequate for use in OWC modelling [20]. Our SI models, which are naturally validated against data gathered from the real system, offer the possibility of capturing nonlinear effects, with fast computation, without making assumptions on small motion around the equilibrium point. Therefore, by filling the void between simplicity linear model and computationally onerous CFD model, our SI models are ideal for hydrodynamic control of OWCs, which, to date, has received poor attention due to the absence of suitable control-oriented models.

The remaining of this paper is organised as follows. Section II contains details of the testing facility and the experimental campaign. In Section III, linear parametric models are identified and, in Section IV, the nonlinear data-based modelling procedure is carried out. In the first part of Section V, the validation performance of the models are reported and compared, while the second part of Section V is dedicated to some considerations concerning the real-time implementation of the aforementioned models. Finally, conclusions are drawn in Section V.

II. REAL WAVE TANK DATA GATHERING

A. TESTING FACILITY

The narrow tank at Dundalk Institute of Technology (DkIT), schematically illustrated in Fig. 4, is used in order to gather the RWT experimental data. On the left side of Fig. 4, the tank is equipped with a flap-type wave generator. On the right side, an absorption beach is positioned in order to minimise wave reflections from the beach towards the model. Since there is only one direction of wave propagation, the wave motion is basically two-dimensional. The tank has a length of 18 m, a width of 350 mm, a depth of 1 m, a freeboard of 200 mm and a probe-to-probe distance, d_{pp} (Fig. 4), of approximately 3 m, though its precise value is unknown. During an experimental run, the displacement of the water column, $y(t)$, is recorded at its centroid by means of a wave probe, as shown in Fig. 4. In order to record the FSE, $\eta(t)$, at the same location where $y(t)$ is measured, the experiment has to be identically repeated in the absence of the OWC model in the tank. Particular care has been taken in order to ensure that $\eta(t)$ (for the absent OWC case) and $y(t)$ are measured at the same location. Furthermore, it should be noted that the acquisition system and the wave maker are entirely isolated from each other, hence the data acquisition and the generation of excitation waves do not start at the same moment. Therefore, before using the data for OWC modelling, the time traces of the input need to be temporally aligned with those of the output. To this end, the FSE measurements gathered from the up-wave probe (Fig. 4) during the two experiments are cross-correlated in order to estimate the time delay. The capability of the tank to reproduce a pseudo-random wave elevation time series, i.e. the tank repeatability, has already been assessed in previous work [17]. The 1:5 scaled fixed OWC model is made in marine plywood and scaled in accordance with Froude scaling [21]. The internal width of the chamber, equal to the width of the water column, is 288 mm. An iris valve, which simulates the effect of a turbine by creating an orifice with a variable diameter, is mounted at the top of the pneumatic chamber, and the diameter of its pupil is set, mainly due to practical reasons, at 30 mm.

In general, in any OWC model, the air spring effect due to air compressibility does not scale correctly if the air volume is scaled geometrically, i.e. scaled by the cube of the scaling factor. This implies that, when the pneumatic chamber is relatively small, the effects of air compressibility are usually negligible. In [22], the effects of air compressibility in OWCs

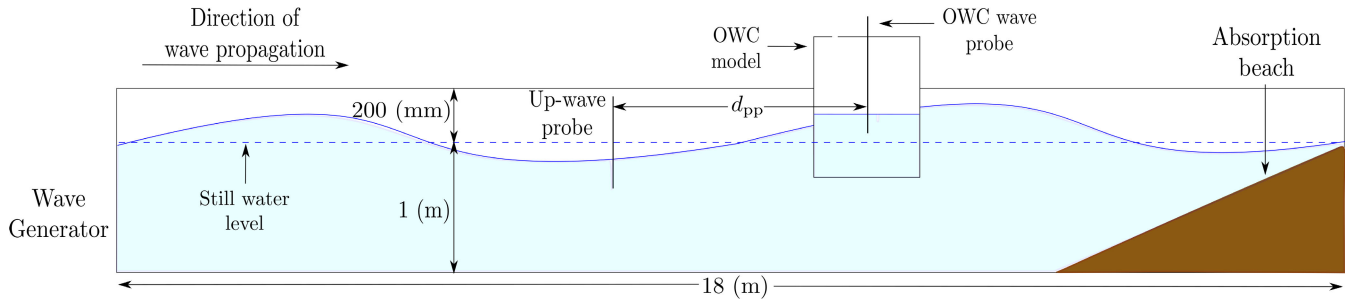


FIGURE 4. Schematic representation of the narrow tank at DkIT.

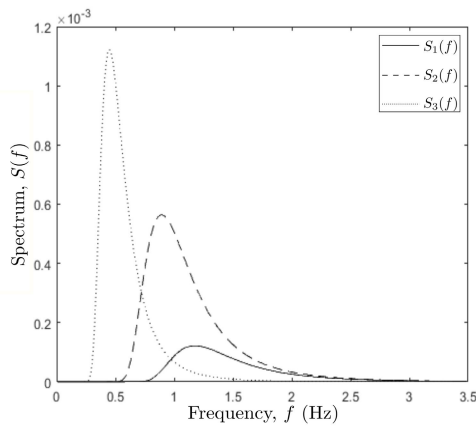


FIGURE 5. Bretschneider spectra.

TABLE 1. Bretschneider spectra.

$S_i(f)$	T_e (s)	H_s (mm)
$S_1(f)$	0.85	40
$S_2(f)$	1.12	75
$S_3(f)$	2.23	75

are discussed and three possible thermodynamic models are analysed and compared. Furthermore, despite the fact that the behaviour of an oscillating water column can be quite complex, the water column mainly displaces vertically as a piston. This piston-like mode of motion is known as the pumping mode, which is also the main mode responsible for power production in OWC devices. Ultimately, in this work, the effects of air compressibility are neglected because the air volume is relatively small and, moreover, the pumping mode is assumed to be the only active mode of the water column.

A more comprehensive description of the narrow tank and the OWC model can be found in [21].

B. EXPERIMENTAL CAMPAIGN

During the experimental campaign, the time traces of the water column displacement, $y(k)$, and the free surface elevation, $\eta(k)$, are sampled and recorded over time instants, $k = 1, 2, \dots, N$, by means of the National Instrument LabVIEW software [23].

The data collected during the experiments comprise three sets of input-output data, which have also been employed for a separate study [24]. The OWC model is tested under three sea states, which correspond to different Bretschneider spectra (Fig. 5), identified as $S_1(f)$, $S_2(f)$ and $S_3(f)$. The wave maker utilises an inbuilt function in order to generate a time sequence of waves, whose spectrum matches the target Bretschneider spectrum. A generic Bretschneider spectrum can be expressed as a function of the frequency, f , as:

$$S_i(f) = \frac{5}{16} H_s^2 \frac{f^{-5}}{T_e^4} \exp \left[-\frac{5}{4} (T_e f)^{-4} \right], \quad i = 1, 2, 3. \quad (1)$$

Table 1 reports the mean energy period, T_e , and the significant wave height, H_s , for the three wave spectra employed in this work. For each spectrum, $y(k)$ is recorded for 30 minutes and sampled at a frequency of acquisition, F_s , of 128 Hz. For all the frequency bands of the three Bretschneider spectra, the sampling frequency is comfortably in excess of the minimum sampling rate required by the Nyquist-Shannon sampling theorem [25]. Then, after removing the OWC model from the wave tank, the experiment is identically repeated in order to separately record $\eta(k)$ where the OWC was positioned. For the sake of simplicity, the three sets of input-output data, which are recorded during the experiments with spectra $S_1(f)$, $S_2(f)$ and $S_3(f)$, are respectively called D_1 , D_2 and D_3 . Note that each spectrum covers a different range of frequencies, and amplitudes, over which the OWC device is tested. The most significant spectrum is $S_3(f)$ since, in this case, the frequencies of the exciting waves lie within the resonant band of the pumping mode, which is the frequency band in which the uncontrolled device is designed to operate, in order to maximise power extraction [26], [27]. However, it is worth providing some insight into the behaviour of the device over the frequency range of $S_2(f)$ and $S_3(f)$, since a hypothetical control system may change the frequency band in which the device optimally captures wave energy. Indeed, a change in the PTO damping may significantly affect the hydrodynamic performance of an OWC [4], [28], [29]. For instance, a rotational speed control strategy changes the turbine damping which, in turns may affect the wave energy absorption process (hydrodynamic performance) of the OWC. The purpose of hydrodynamic control is to extend the bandwidth of the

system effectively increasing the range of frequencies at which the device can resonate [30].

III. LINEAR DATA-BASED MODELLING

A. METHODOLOGY

In the linear data-based modelling case, SI is carried out by following a somewhat standard approach, which is schematically represented in Fig. 8. A linear polynomial ARX model, whose unknown parameters are a_i and b_i , is chosen as the SI model:

$$\hat{y}(k) = - \sum_{i=1}^{n_a} a_i y(k-i) + \sum_{i=0}^{n_b} b_i u(k-n_d-i). \quad (2)$$

In equation (2), $\hat{y}(k)$ is the predicted model output at time k , $y(k)$ and $u(k)$ are the k -th samples of the output and input respectively, while n_a and n_b are the orders of the ARX model, with n_d the input delay. On the right side of equation (2), the two summation terms containing the outputs and the inputs are, respectively, the autoregressive (AR) part and the exogenous (X) part of the ARX model. In this work, the values of n_a , n_b and n_d are determined from a procedure described later on in this section. The number of unknowns in the ARX model of equation (2) is $n_a + n_b + 1$.

In order to compare the performance of the models in the identification and validation phases, the normalized root mean squared error (NRMSE) is chosen as the error metric:

$$NRMSE = \sqrt{\frac{\sum_k |y(k) - \hat{y}(k)|^2}{\sum_k |y(k)|^2}}. \quad (3)$$

In order to determine the orders n_a , n_b and n_d , a sequence of systematic model identification trials, with incremental changes in n_a , n_b and n_d , has been implemented. At the end of each trial, the performance of the model is evaluated on the basis of the NRMSE. Since the training NRMSE is not always a reliable estimator of the model performance, n_a , n_b and n_d should be selected on the basis of the validation NRMSE [6]. On the left side of Fig. 6, the NRMSE is computed for different combinations of n_a and n_b (at fixed n_d), whereas, on the right side of Fig. 6, a focus on the case of $n_b = 2$ is shown. Ultimately, the leading model order is n_a , while only minor changes in the value of the NRMSE are found by varying n_b and n_d . As already stated, the number of unknowns depends on the model orders and, generally, the loss function decreases as the number of model parameters increases. Therefore, a trade-off between model accuracy and model complexity has to be made; hence, to this end, the principle of parsimony is adopted. This principle potentially helps to set the variables in order to find a 'parsimonious model', which is a model whose complexity ideally matches the problem complexity. To this end, consideration will be given to the complexity of the solution (in terms of the number of model parameters), as well as the model error metric, in the spirit of [31]. Ultimately, in this work, $n_a = 6$ and $n_b = 2$ were chosen. Since the validation NRMSE has a minimum for $n_d = -4$ (as shown in Fig. 7), the value of the input delay

is set to -4 , noting that a negative value of n_d implies that the ARX model is noncausal, meaning that both past and future values of the input are utilized in order to predict the output value, $y(k)$. In contrast, if $n_d > 0$, the ARX model is causal, where the output prediction only relies on past values of the input. The choice of a negative value for n_d is explained as follows. In wave energy systems, where the FSE is used as the input signal, there is a noncausal relationship between the FSE and the motion (velocity/displacement) [32], as shown by the impulse response functions, for example in [33]. The noncausal input/output relationship is due to the fact that the FSE is not the real physical input that causes the output. Rather, the real input which provokes the device motion is the associated instantaneous pressure field, due to the exciting waves, which acts on the submerged surface of the device. In practice, an up-wave measurement, or forecast, is used to remove this noncausality [33].

The unknown parameters of the ARX model, a_i and b_i , are found by solving a relatively simple least-squares (LS) optimisation problem, which is constructed from equation (2). In particular, equation (2), i.e. the model predictor, can be rewritten in a more compact form by introducing two vectors, namely the regression vector, $\varphi(k)$:

$$\varphi(k) = [-y(k-1) \dots -y(k-n_a) \quad u(k-1+n_k) \dots u(k-n_b+n_k)]^T, \quad (4)$$

and the parameter vector, θ :

$$\theta = [a_1 \ a_2 \ \dots \ a_{n_a} \ b_0 \ \dots \ b_{n_b}]^T. \quad (5)$$

Therefore, combining equations (2), (4), and (5):

$$\hat{y}(k) = \theta^T \varphi(k) = \varphi(k)^T \theta. \quad (6)$$

Equation (6) is a simple linear regression which can be trivially extended for all time instants, $k = 1, 2, \dots, N$, as:

$$\hat{\mathbf{y}} = \mathbf{\Lambda} \theta, \quad (7)$$

where $\hat{\mathbf{y}} = [y(1) \ y(2) \ \dots \ y(N)]^T$ is the model output vector, and $\mathbf{\Lambda}$ is the regression matrix. The prediction error vector is defined as:

$$\boldsymbol{\epsilon} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{\Lambda} \theta, \quad (8)$$

and the LS optimization problem, i.e. find $\theta = \theta_{\text{opt}}$ such that the LS error is minimized, can be written as:

$$\theta_{\text{opt}} = \underset{\theta}{\text{argmin}} (\boldsymbol{\epsilon}^T \boldsymbol{\epsilon}), \quad (9)$$

with the well-known solution:

$$\theta_{\text{opt}} = (\mathbf{\Lambda}^T \mathbf{\Lambda})^{-1} \mathbf{\Lambda}^T \mathbf{y}. \quad (10)$$

The identification and validation phases are carried out as follows. The input-output time traces contained in D_1 , namely $\eta_1(k)$ and $y_1(k)$, are temporally split into two parts, D_{1_a} and D_{1_b} . Then, the data set D_{1_a} is used in order to identify an ARX model, M_{1_a} , which is finally validated against D_{1_b} . Similarly, M_{1_b} is identified by using D_{1_b} , and validated against D_{1_a} .

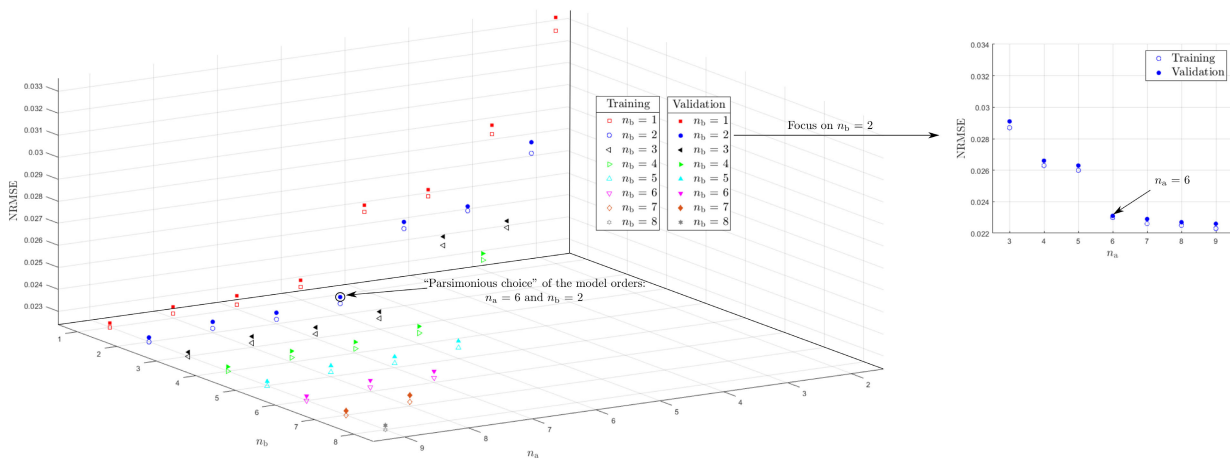


FIGURE 6. NRMSE of separate ARX models, identified from D_{3_b} and validated against D_{3_a} , using different combinations of n_a and n_b , with $n_d = -4$.

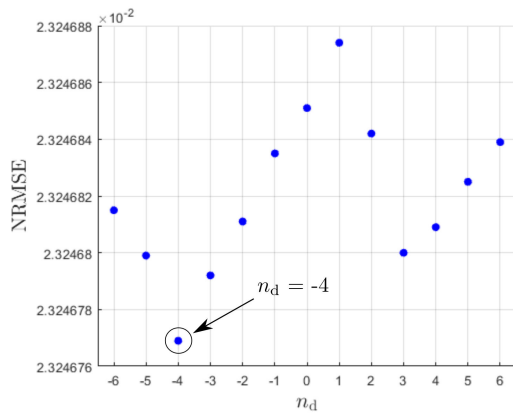


FIGURE 7. NRMSE resulting from the validation of different ARX models, identified from D_{3_b} and validated against D_{3_a} , with $n_a = 6$ and $n_b = 2$.



FIGURE 8. Linear data-based modelling approach for the data set D_1 .

This procedure is schematically shown in Fig. 8. The same approach is also repeated with D_2 and D_3 , and, ultimately, six ARX linear parametric models are identified and validated.

B. MODEL IDENTIFICATION

Table 2 reports the percentage values of the fidelity of the models, defined as $F_m = (1 - NRMSE)$, with respect to their corresponding identification data. These F_m values remain, as expected, relatively consistent in relation to the data set employed for the identification. For instance, the fidelity of models M_{1_a} and M_{1_b} , whose identification data are, respectively, D_{1_a} and D_{1_b} , is always about 94%. Moreover, it is interesting to note that the model fidelity broadly reflects the signal-to-noise ratio (S/N) of the identification data. As a matter of fact, the experimental data D_3 , which have the

TABLE 2. Fidelity values for ARX model training.

Model	M_{1_a}	M_{1_b}	M_{2_a}	M_{2_b}	M_{3_a}	M_{3_b}
F_m	94.43	94.24	96.56	96.26	97.68	97.89

highest S/N, produces the models with the highest fidelity, namely M_{3_a} and M_{3_b} . This consideration becomes intuitive when the values in Table 2 are compared to the peaks of the spectra in Fig. 5. The presence of relatively high levels of noise in experimental data can significantly affect the identification/validation performance of a model [7].

IV. NONLINEAR DATA-BASED MODELLING

A. METHODOLOGY

For the nonlinear modelling approach, a nonlinear KGP model, whose unknown parameters are $a_{i,j}$ and $b_{i,j}$, is employed:

$$\hat{y}(k) = \sum_{j=1}^p \left[- \sum_{i=1}^{n_a} a_{i,j} y^j(k-i) + \sum_{i=0}^{n_b} b_{i,j} u^j(k-n_d-i) \right], \tag{11}$$

where the polynomial order, p , is an additional model structural parameter to set. In principle, the KGP model may also contain some cross-product terms (or cross-product regressors), namely $c_{i,j} y^j(k-i) u^j(k-n_d-i)$ with $c_{i,j}$ being unknown parameters. However, since the cross-product terms may lead to instability problems, they have been discarded and, as a result, equation (11) does not employ any cross-product terms.

Note that, although the polynomial model in equation (11) is clearly nonlinear, the KGP model is, just like the ARX model, linear in the unknown parameters. Therefore, since the KGP model is a somewhat natural extension of the ARX model, some aspects of the methodology introduced in section III-A still apply. First of all, the aforementioned linearity in the parameters allows the unknown parameters

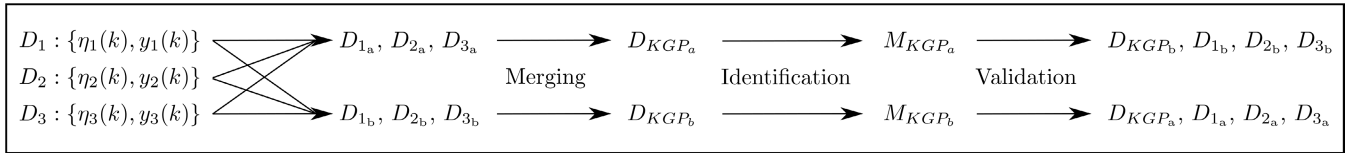


FIGURE 9. Schematic representation of the nonlinear data-based modelling approach.

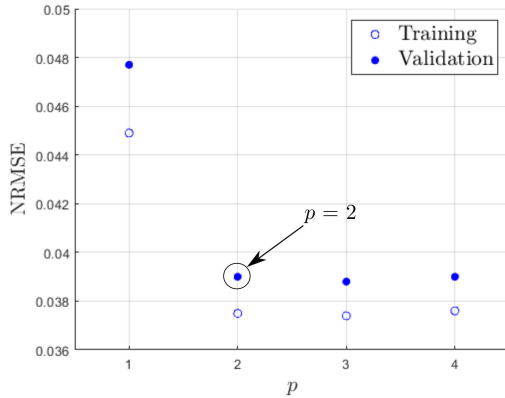


FIGURE 10. NRMSE of different KGP models, identified from D_{KGP_b} and validated against D_{KGP_a} , with $n_a = 6$, $n_b = 2$ and $n_d = -4$.

of the KGP model to be determined by solving a LS problem. Indeed, it is possible to express equation (11) as a linear regression, such as equation (6), and write the LS optimization problem, as shown in equation (9). Secondly, the orders of the KGP model, including p , can be determined by implementing a sequence of systematic trials, similar to that adopted in section III-A. In this work, since no significant improvements in the NRMSE are achieved for $p > 2$ (as shown in Fig. 10), p is set to 2, whereas n_a , n_b and n_d are identical to those used for the ARX models. The number of unknown parameters for the KGP model (equation (11)) is easily calculated as $p(n_a + n_b + 1)$.

The training and validation performance of the KGP models is assessed using the same error metric, namely F_m , adopted for the linear modelling approach.

The identification and validation phases are carried out as schematically shown in Fig. 9. Firstly, as for the ARX models, the input/output time traces of the data sets D_1 , D_2 and D_3 are split into two parts. Afterwards, the data sets of the a-group are merged together and, consequently, a new data set, D_{KGP_a} , is constructed. The same procedure is repeated for the b-group of data and, hence, another corresponding data set, D_{KGP_b} , is derived. The two data sets, D_{KGP_a} and D_{KGP_b} , are, respectively, used in order to train two KGP models, M_{KGP_a} and M_{KGP_b} . Finally, M_{KGP_a} is validated against data D_{KGP_b} , D_{1b} , D_{2b} and D_{3b} , whereas M_{KGP_b} is validated against D_{KGP_a} , D_{1a} , D_{2a} and D_{3a} .

B. MODEL IDENTIFICATION

The model fidelity F_m , already introduced in section III-B, with respect to the identification data for the KGP models

TABLE 3. Fidelity values for KGP model training.

Model	M_{KGP_a}	M_{KGP_b}
F_m	96.25	96.10

is reported in Table 3. The two models provide, as expected, similar results, in terms of fidelity, in the identification phase.

V. RESULTS AND COMPARISON

A. MODEL VALIDATION PERFORMANCE

Tables 4 and 5, respectively, report the validation performance of the six ARX and two KGP models, when validated against the designated set of validation data (VD). Each model is validated using three different output predictions, namely 1-step-ahead, 5-step-ahead and 10-step-ahead predictions. For clarity, in a generic k -step-ahead prediction, the i -th output value is computed by using previously measured outputs up to time instant $i - k$ and relevant inputs up to time instant i . For instance, Fig. 11 compares the predicted output, $\hat{y}_{ARX}(t)$, of the ARX model M_{3b} , the predicted output, $\hat{y}_{KGP}(t)$, of the KGP model M_{KGP_b} , and the measured output, $y(t)$, for the validation data set D_{3a} .

In table 4, the model performance remains consistent throughout the validation phase. First of all, the a-group and b-group models always provide similar (symmetric) results. For instance, in the 5-step-ahead prediction of M_{2a} and M_{2b} , the model fidelity is, respectively, 85.09% and 85.91%. Secondly, the model fidelity unsurprisingly decreases as the number of prediction steps ahead increases. Finally, the fidelity of M_{1a} (and M_{1b}) is always the lowest, whereas the fidelity of M_{3a} (and M_{3b}) is the best in all cases, due to superior S/N.

As in the case of the ARX models, the two KGP models also show strong consistency during the validation phase (table 5). Indeed, models M_{KGP_a} and M_{KGP_b} display, in the sense explained above for the ARX models, a similar (symmetric) fidelity. Furthermore, the validation of M_{KGP_a} (and M_{KGP_b}) against D_{KGP_b} (and D_{KGP_a}) provides a 'mean' fidelity. For instance, in the 5-step-ahead predictions of M_{KGP_a} against data sets D_{1b} , D_{2b} , D_{3b} and D_{KGP_b} the fidelities are, respectively, 81.61%, 85.05%, 92.95% and 86.54%. Note that $(81.61\% + 85.05\% + 92.95\%)/3 \approx 86.54\%$.

It should be clear from tables 4 and 5 that, if the same VD are considered across ARX and KGP models, the ARX models perform better than the KGP models. For instance, when VD are D_{3b} , the fidelity of the ARX model M_{3a} for the 1-step-ahead, 5-step-ahead and 10-step-ahead predictions is, respectively, 97.89%, 93.57% and 88.49%. For the nonlinear modelling, M_{KGP_a} has a fidelity of 97.71%, 92.95% and

TABLE 4. Fidelity values, F_m , for ARX model validation.

Model	VD	1-step-ahead	5-step-ahead	10-step-ahead
M_{1_a}	D_{1_b}	94.24	82.03	60.22
M_{1_b}	D_{1_a}	94.43	81.93	59.54
M_{2_a}	D_{2_b}	96.25	85.09	67.97
M_{2_b}	D_{2_a}	96.56	85.91	69.56
M_{3_a}	D_{3_b}	97.89	93.57	88.49
M_{3_b}	D_{3_a}	97.68	93.47	88.20

TABLE 5. Fidelity values, F_m , for KGP model validation.

Model	VD	1-step-ahead	5-step-ahead	10-step-ahead
M_{KGP_a}	D_{1_b}	94.25	81.61	57.43
M_{KGP_a}	D_{2_b}	96.33	85.05	66.69
M_{KGP_a}	D_{3_b}	97.71	92.95	84.58
M_{KGP_a}	D_{KGP_b}	96.10	86.54	69.57
M_{KGP_b}	D_{1_a}	94.48	81.73	57.55
M_{KGP_b}	D_{2_a}	96.58	85.82	68.99
M_{KGP_b}	D_{3_a}	97.70	92.98	84.64
M_{KGP_b}	D_{KGP_a}	96.25	86.84	70.39

84.58% for the 1-step-ahead, 5-step-ahead and 10-step-ahead predictions, respectively. The difference between the ARX and KGP models becomes more and more pronounced as the number of prediction steps increases. However, it is worth noting that, for the nonlinear case, a single KGP model (M_{KGP_a}) can provide good validation performance for different data (D_{1_b} , D_{2_b} and D_{3_b}). On the other hand, the ARX model M_{1_a} can provide good results only when the validation data are D_{1_b} . As a matter of fact, the cross-validation performance of the ARX models is relatively poor [7].

B. REAL-TIME IMPLEMENTATION

Despite the fact that the linear modelling approach may appear appealing and convenient, there are some practical aspects which should be discussed. Such aspects concern the necessary steps required to implement a multi-linear model in real-time (Fig. 12) to cover the complete operational range.

In sections III-B and V-A, multiple linear ARX models, corresponding to different (fixed) sea states, have been identified and validated. However, in order to employ linear ARX models in real-time, it is essential to derive a 'single' multi-linear model which can provide a system representation across all three sea states. To this end, a switching/interpolating mechanism for either the parameters, or the outputs, of the multiple linear models is required. Furthermore, since the switching/interpolating mechanism is based on knowledge of the specific sea state, it is imperative to identify the sea state itself, which in turns requires the fitting of spectral model to Fourier-transformed FSE data gathered from wave probes. It should be clear that the procedure illustrated in Fig. 12 has two critical interconnected items: the (explicit) switching/interpolating mechanism and the identification of the sea state. Indeed, the efficacy of

the switching/interpolating mechanism also depends on the correct identification of the sea state.

For a switching mechanism, one of the available linear models is selected to work with the corresponding identified sea state. If, instead, an interpolating mechanism is adopted, either the interpolated model parameters or the interpolated model outputs are chosen, utilising a set (at least 2) of parameters/outputs corresponding to the relevant set of sea states. It should be noted that, since the spectral models of the sea state are usually defined by two parameters (H_s and T_e), the interpolation is actually a two dimensional interpolation. In either model selection mechanism, the multi-linear model is expected to provide the best possible performance when the identified sea state is exactly one of the sea states originally employed during the experimental campaign (i.e. in the identification phase). In this peculiar situation, the fidelity of the multi-linear model, regardless of the model selection rationale (i.e. switching or interpolating mechanism), should be comparable to the values reported in table 4. Arguably, when the identified sea state does not perfectly coincide with one of those employed during the experimental campaign, a multi-linear model equipped with a switching mechanism is expected to provide poorer performance than a multi-linear model with a suitable interpolating mechanism.

The identification of the sea state can be a complex task to perform in real-time and in a real sea environment. Firstly, a suitable location for the wave probes has to be identified and, moreover, a spectral model for the sea state has to be chosen. In order to obtain an adequate statistical description, a sufficient number of FSE data has to be gathered from the wave probes; otherwise, the accuracy of the spectral model might be too poor and, consequently, the inferred multi-linear model could be far from being the best model to associate with the (true) actual sea state. Furthermore, regardless the spectral model accuracy, all the operations needed to identify the sea state clearly require some time. Therefore, delays in deducting the multi-linear model might severely affect the overall efficacy, and accuracy, of the multi-linear model itself. In the worst-case scenario, the multi-linear model interpolation/switching will not be able to keep up with the evolution of the sea state and, due to the poor cross validation performance of the linear ARX models [7], the fidelity of the multi-linear model may be poor.

The main advantage of the nonlinear model over the multi-linear alternative is its self-adapting capability, meaning that, since the KGP model is able to intrinsically perform the interpolation of the predicted output, the nonlinear modelling solution does not require either the identification of the sea state, nor an explicit switching/interpolating mechanism. Furthermore, the number of unknown parameters of the KGP model, p ($n_a + n_b + 1$), is likely lower than the unknown parameters of the multi-linear modelling solution, N_{LM} ($n_a + n_b + 1$), where N_{LM} is the number of multiple linear ARX models.

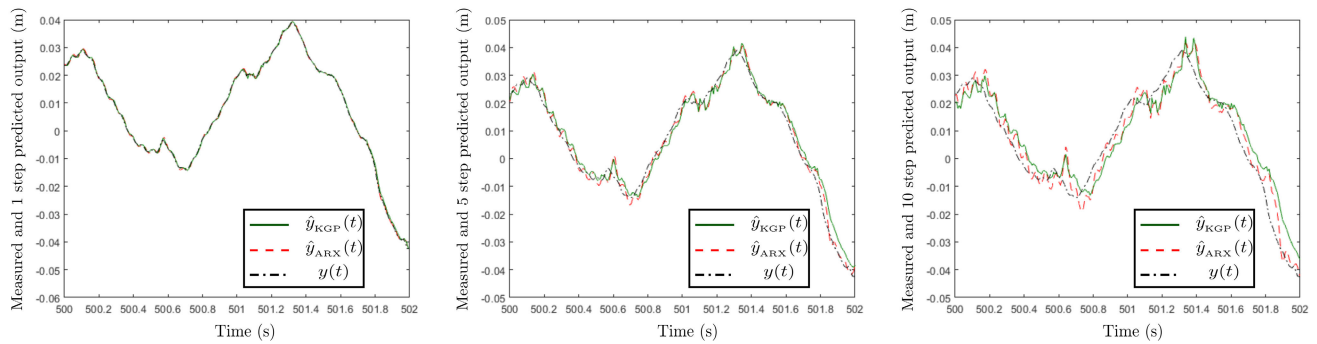


FIGURE 11. From left to right, comparison between the measured output, $y(t)$ (from D_{3a}), and the 1-step-ahead, 5-step-ahead and 10-step-ahead predicted outputs, $\hat{y}_{ARX}(t)$ and $\hat{y}_{KGP}(t)$, respectively related to the validation of the ARX model M_{3b} and the KGP model M_{KGP_b} .

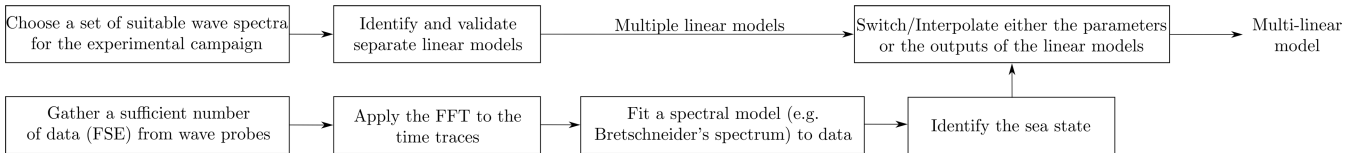


FIGURE 12. Schematic procedure related to the real-time implementation of a multi-linear ARX model.

VI. CONCLUSION

The current paper shows the value of a nonlinear data-based modelling approach over a linear, or multi-linear, modelling solution. Despite the fact that the multiple linear ARX models perform slightly better than the KGP models during the validation (tables 4 and 5) on their specific data sets, the procedure to implement in real-time (and in a real sea environment) a multi-linear ARX model, to cover the full range of operation, is found to be quite complex and time-consuming. Ultimately, since the KGP model is somewhat a natural and straightforward extension of the ARX model, not only is the nonlinear modelling approach relatively simple but, actually, requires even less effort than the multi-linear alternative! Furthermore, since wave probes for the identification of the sea state are not required in the nonlinear modelling solution, the capital and operational costs associated with such wave probes are nullified. Apart from the KGP model, there may be other nonlinear SI models, such as simple neural networks, which are suitable to model the hydrodynamics of an OWC. In order to assess the most suitable nonlinear SI model, in terms of accuracy/complexity, a comprehensive comparative study would be required. Here, we simply show the efficacy of a simple nonlinear model structure, which has a natural evolution from its linear ARX counterpart, and the validity of a nonlinear approach.

The main advantage of the KGP models is that the (implicit) interpolation and the sea state selection come 'free', as a direct consequence of using a nonlinear modelling approach. Arguably, it may be desirable to weight the optimization problem (equation (9)) depending on the likelihood of having certain sea states. In this way, the model parameters would be tuned in order to take into greater account the most likely sea states and, consequently, the model should perform better in those frequent sea states. The weighting operation could be done either automatically, by consulting

the sea state records collected at the considered deployment site (where frequent sea states will be well represented), or manually, by using a suitable weighting function during the identification phase. In either case, in order to select suitable training data, the historical records of the sea state of the deployment site are needed. Indeed, the most critical aspect of SI is the choice of suitable input/output data and, clearly, the identification data should be sufficient to ensure an adequate training phase.

We believe that our wave-to-displacement model is suitable for hydrodynamic control of OWCs, which to date has received poor attention due to the absence of suitable control-oriented model. Despite the fact that, to the best of authors' knowledge, relevant control strategies for OWCs currently focus on turbine speed control, there is also an opportunity to develop a twofold control strategy for both turbine speed and hydrodynamic control. For instance, in [34], aerodynamic control to maximize the performance of the air turbine is utilized together with peak power control to cope with potentially dangerous wave energy peaks. Moreover, in [35], a passive relief valve system is studied to counteract chamber pressure losses due to pressure skewness. Our model is ideal for model-based control applications, being both computationally and parametrically simple, and also captures the essential hydrodynamic nonlinearity in the system.

REFERENCES

- [1] G. Mork, S. Barstow, A. Kabuth, and M. T. Pontes, "Assessing the global wave energy potential," in *Proc. 29th Int. Conf. Ocean, Offshore Arctic Eng.*, vol. 3, Jan. 2010, pp. 447–454.
- [2] K. Gunn and C. Stock-Williams, "Quantifying the global wave power resource," *Renew. Energy*, vol. 44, pp. 296–304, Aug. 2012.
- [3] B. G. Reguero, I. J. Losada, and F. J. Méndez, "A global wave power resource and its seasonal, interannual and long-term variability," *Appl. Energy*, vol. 148, pp. 366–380, Jun. 2015.
- [4] A. F. O. Falcão and J. C. C. Henriques, "Oscillating-water-column wave energy converters and air turbines: A review," *Renew. Energy*, vol. 85, pp. 1391–1424, Jan. 2016.

- [5] N. Faedo, S. Olaya, and J. V. Ringwood, "Optimal control, MPC and MPC-like algorithms for wave energy systems: An overview," *IFAC J. Syst. Control*, vol. 1, pp. 37–56, Sep. 2017.
- [6] L. Ljung, *System Identification: Theory for the User*. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.
- [7] M. Rosati, T. Kelly, D. G. Violini, and J. V. Ringwood, "Data-based hydrodynamic modelling of a fixed OWC wave energy converter," in *Proc. 14th Eur. Wave Tidal Energy Conf. 5th–9th September 2021*, Plymouth, U.K., 2021, pp. 1–10.
- [8] L. Ljung, "Perspectives on system identification," *Annu. Rev. Control*, vol. 34, no. 1, pp. 1–12, 2010.
- [9] G. D. Gkikas and G. A. Athanassoulis, "Development of a novel nonlinear system identification scheme for the pressure fluctuation inside an oscillating water column-wave energy converter—Part I: Theoretical background and harmonic excitation case," *Ocean Eng.*, vol. 80, pp. 84–99, Apr. 2014.
- [10] J. Davidson, S. Giorgi, and J. V. Ringwood, "Identification of wave energy device models from numerical wave tank data—Part 1: Numerical wave tank identification tests," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 1012–1019, Jul. 2016.
- [11] S. Giorgi, J. Davidson, and J. V. Ringwood, "Identification of wave energy device models from numerical wave tank data—Part 2: Data-based model determination," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 1020–1027, Jul. 2016.
- [12] G. Bacelli, R. G. Coe, D. Patterson, and D. Wilson, "System identification of a heaving point absorber: Design of experiment and device modeling," *Energies*, vol. 10, no. 4, p. 472, Mar. 2017.
- [13] H. Cho, G. Bacelli, and R. G. Coe, *Linear and Nonlinear System Identification of a Wave Energy Converter*. Albuquerque, NM, USA: Sandia National Laboratories (SNL-NM), 2018.
- [14] F. Jaramillo-Lopez, B. Flannery, J. Murphy, and J. V. Ringwood, "Modelling of a three-body hinge-barge wave energy device using system identification techniques," *Energies*, vol. 13, no. 19, p. 5129, Oct. 2020.
- [15] R. G. Coe, G. Bacelli, and D. Forbush, "A practical approach to wave energy modeling and control," *Renew. Sustain. Energy Rev.*, vol. 142, May 2021, Art. no. 110791.
- [16] H. B. Bingham, Y.-H. Yu, K. Nielsen, T. T. Tran, K.-H. Kim, S. Park, K. Hong, H. A. Said, T. Kelly, J. V. Ringwood, R. W. Read, E. Ransley, S. Brown, and D. Greaves, "Ocean energy systems wave energy modeling task 10.4: Numerical modeling of a fixed oscillating water column," *Energies*, vol. 14, no. 6, p. 1718, Mar. 2021.
- [17] T. Kelly, T. Dooley, and J. Ringwood, "Experimental determination of the hydrodynamic parameters of an OWC," in *Proc. 12th Eur. Wave Tidal Energy Conf.*, Cork, Ireland, 2017, pp. 1–10.
- [18] D.-Z. Ning, R.-Q. Wang, Q.-P. Zou, and B. Teng, "An experimental investigation of hydrodynamics of a fixed OWC wave energy converter," *Appl. Energy*, vol. 168, pp. 636–648, Apr. 2016.
- [19] S. Dai, S. Day, Z. Yuan, and H. Wang, "Investigation on the hydrodynamic scaling effect of an OWC type wave energy device using experiment and CFD simulation," *Renew. Energy*, vol. 142, pp. 184–194, Nov. 2019.
- [20] M. Penalba, T. Kelly, and J. V. Ringwood, "Using NEMOH for modelling wave energy converters: A comparative study with WAMIT," in *Proc. 12th Eur. Wave Tidal Energy Conf. (EWTEC)*, Cork, Ireland, 2017.
- [21] T. Kelly, "Experimental and numerical modelling of a multiple oscillating water column structure," Ph.D. dissertation, Dept. Electron. Eng., Nat. Univ. Ireland Maynooth, Maynooth, Ireland, 2018.
- [22] A. F. O. Falcão and J. C. C. Henriques, "The spring-like air compressibility effect in OWC wave energy converters: Hydro-, thermo- and aerodynamic analyses," in *Proc. Honoring Symp. Professor Carlos Guedes Soares Mar. Technol. Ocean Eng.*, vol. 11A, Jun. 2018, Art. no. V11AT12A038.
- [23] *National Instruments: Test, Measurement and Embedded Systems*. Accessed: Nov. 1, 2016. [Online]. Available: <http://www.ni.com/>
- [24] S. Jacob, "Experimental and numerical modelling of an oscillating water column structure in irregular sea states," M.S. thesis, Dept. Electron. Mech. Eng., Dundalk Inst. Technol., Dundalk, Ireland, 2019.
- [25] A. J. Jerri, "The Shannon sampling theorem—Its various extensions and applications: A tutorial review," *Proc. IEEE*, vol. 65, no. 11, pp. 1565–1596, Nov. 1977.
- [26] J. Falnes, "Optimum control of oscillation of wave-energy converters," in *Proc. 11th Int. Offshore Polar Eng. Conf.*, Stavanger, Norway, Jun. 2001, pp. 1–12.
- [27] J. Falnes, *Ocean Waves and Oscillating Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [28] J. C. C. Henriques, J. C. C. Portillo, W. Sheng, L. M. C. Gato, and A. F. O. Falcão, "Dynamics and control of air turbines in oscillating-water-column wave energy converters: Analyses and case study," *Renew. Sustain. Energy Rev.*, vol. 112, pp. 571–589, Sep. 2019.
- [29] A. F. D. O. Falcão and P. A. P. Justino, "OWC wave energy devices with air flow control," *Ocean Eng.*, vol. 26, no. 12, pp. 1275–1295, Dec. 1999.
- [30] J. V. Ringwood, G. Bacelli, and F. Fusco, "Energy-maximizing control of wave-energy converters: The development of control system technology to optimize their operation," *IEEE Control Syst.*, vol. 34, no. 5, pp. 30–55, Oct. 2014.
- [31] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.
- [32] J. Falnes, "On non-causal impulse response functions related to propagating water waves," *Appl. Ocean Res.*, vol. 17, no. 6, pp. 379–389, Dec. 1995.
- [33] U. A. Korde and J. V. Ringwood, *Hydrodynamic Control Wave Energy Devices*. Cambridge, U.K.: Cambridge Univ. Press, 2016.
- [34] J. C. C. Henriques, R. P. F. Gomes, L. M. C. Gato, A. F. O. Falcão, E. Robles, and S. Ceballos, "Testing and control of a power take-off system for an oscillating-water-column wave energy converter," *Renew. Energy*, vol. 85, pp. 714–724, Jan. 2016.
- [35] K. Monk, V. Winands, and M. Lopes, "Chamber pressure skewness corrections using a passive relief valve system at the pico oscillating water column wave energy plant," *Renew. Energy*, vol. 128, pp. 230–240, Dec. 2018.



MARCO ROSATI received the B.S. degree in aerospace engineering and the M.S. degree in aeronautical engineering from the Politecnico di Milano, in 2017 and 2020, respectively. He is currently pursuing the Ph.D. degree with the Electronic Engineering Department, Maynooth University, Ireland. From 2019 to 2020, he was an Intern with the Institut National des Sciences Appliquées/Institut Clément Ader (INSA/ICA), Toulouse, France, where he worked on experimental microfluidics and developed a molecular tagging thermometry technique. He is also a member of the Centre for Ocean Energy Research, Maynooth University. His research interests include wave energy and, in particular, on oscillating water column devices.



THOMAS KELLY received the bachelor's degree in mechanical engineering from the Dublin Institute of Technology, in 1997, and the Ph.D. degree from Maynooth University, in 2018. His undergraduate dissertation focused on hybrid airships. Then, he worked for a number of years in the semiconductor industry before completing a master's degree in renewable energy systems with the Dundalk Institute of Technology, in 2010. Both his master's dissertation and a Ph.D. thesis focused on the modeling of water energy converters. He is currently a Lecturer at Dundalk IT, where he is also involved with research in association with the Centre for Renewable Energy. His research interests include the areas of renewable energy, with a particular focus on marine renewables.



JOHN V. RINGWOOD (Senior Member, IEEE) received the Diploma degree in electrical engineering from the Dublin Institute of Technology, Dublin, Ireland, in 1981, and the Ph.D. degree in control systems from the University of Strathclyde, Glasgow, U.K., in 1985. He is currently a Professor in electronic engineering with Maynooth University, Ireland, where he is also the Director of the Centre for Ocean Energy Research. He is also a Chartered Engineer and a fellow of Engineers Ireland. His current research interests include timeseries modeling, wave energy, and biomedical engineering. He serves on the editorial boards for the *Journal of Ocean Engineering and Marine Energy*, *IEEE TRANSACTIONS ON SUSTAINABLE ENERGY*, *IET Renewable Power Generation*, and *Energies*.