

# *IET Renewable Power Generation*

## Special Issue Call for Papers

---

**Be Seen. Be Cited.  
Submit your work to a new  
IET special issue**

Connect with researchers and  
experts in your field and  
share knowledge.

Be part of the latest research  
trends, faster.





[Read more](#)



The Institution of  
Engineering and Technology

## CASE STUDY

# Air pressure forecasting for the Mutriku oscillating-water-column wave power plant: Review and case study

Jorge Marques Silva<sup>1,2</sup>  | Susana M. Vieira<sup>1</sup>  | Duarte Valério<sup>1</sup>  |  
João C. C. Henriques<sup>1</sup>  | Paul D. Sclavounos<sup>2</sup>

<sup>1</sup> IDMEC, Instituto Superior Técnico, Universidade de Lisboa, Portugal

<sup>2</sup> Laboratory for Ship and Platform Flows, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

## Correspondence

João C. C. Henriques, IDMEC, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal.  
Email: joaochenriques@tecnico.ulisboa.pt

## Funding information

Fundação para a Ciência e a Tecnologia, Grant/Award Numbers: UIDB/50022/2020, SFRH/BD/136521/2018; Fulbright Research Grant with the support of FCT; MIT Portugal Seed Funding 2019

## Abstract

The high variability and unpredictability of renewable energy resources require optimization of the energy extraction, by operating at the best efficiency point, which can be achieved through optimal control strategies. In particular, wave forecasting models can be valuable for control strategies in wave energy converter devices. This work intends to exploit the short-term wave forecasting potential on an oscillating water column equipped with the innovative biradial turbine. A Least Squares Support Vector Machine (LS-SVM) algorithm was developed to predict the air chamber pressure and compare it to the real signal. Regressive linear algorithms were executed for reference. The experimental data was obtained at the Mutriku wave power plant in the Basque Country, Spain. Results have shown LS-SVM prediction errors varying from 9% to 25%, for horizons ranging from 1 to 3 s in the future. There is no need for extensive training data sets for which computational effort is higher. However, best results were obtained for models with a relatively small number of LS-SVM features. Regressive models have shown slightly better performance (8–22%) at a significantly lower computational cost. Ultimately, these research findings may play an essential role in model predictive control strategies for the wave power plant.

## 1 | INTRODUCTION

The ocean waves are a relevant renewable energy resource that, if largely exploited, can play a significant role in the clean energy supply of countries with sea coasts [1]. Wave energy is clean, abundant and extraordinarily powerful, especially offshore. In comparison to wind power, wave energy is significantly less variable, allowing for the accurate forecasting of the supply of electricity in advance, which is essential for energy planning strategies [2].

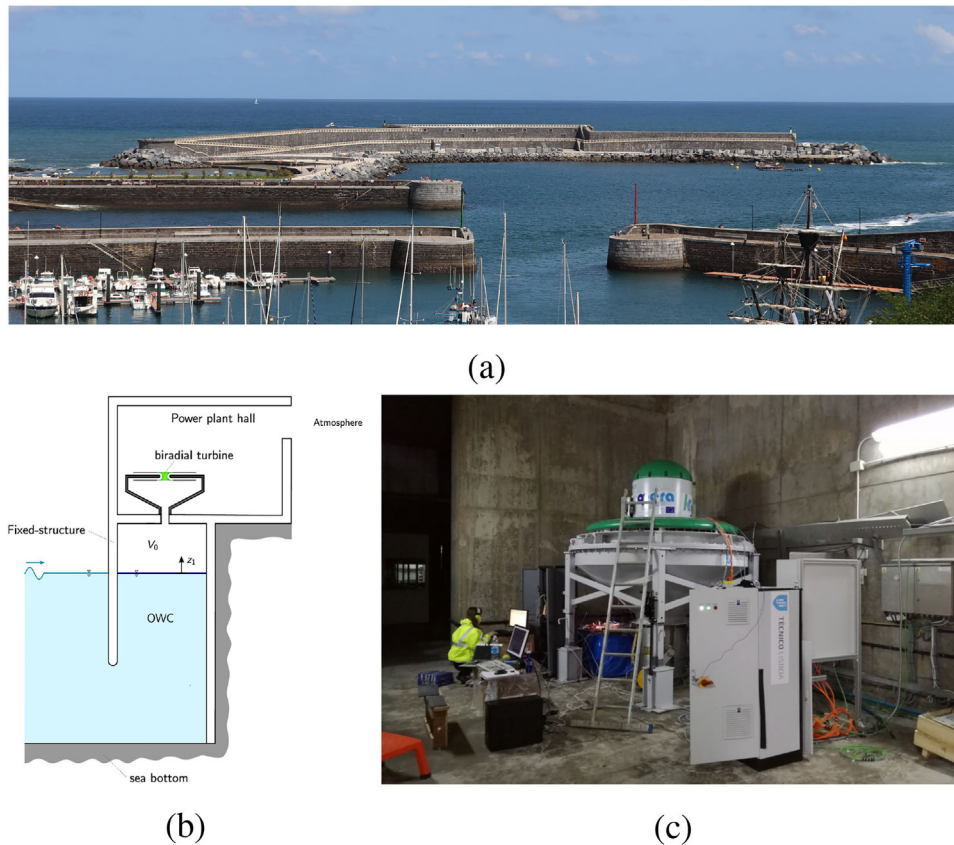
The oscillating water column (OWC) is widely regarded as the simplest and most reliable type of wave energy converter (WEC) [1]. An OWC power plant consists of a fixed or floating hollow structure with its bottom surface open below the mean water level. The upper part of the structure forms an air chamber where a turbine is installed and connected to the outer atmosphere through a duct. The incident waves induces the motion of the water column which compresses and expands

the air inside the chamber. This creates a pressure difference with the atmosphere that drives a self-rectifying air turbine coupled to an electrical generator through a shaft, as part of the power take-off (PTO) system.

Three main types of self-rectifying air turbines have been proposed for this application: the Wells turbine, the axial impulse turbine and the novel biradial turbine [3]. A comparison reveals that the biradial turbine has the highest peak efficiency and the wider range of operation of the three types of turbines [4]. Moreover, the biradial turbine damping is almost insensitive to the rotational speed, given its near-quadratic relationship between the flow rate and the pressure head. This feature can be especially important for the pressure forecasting, as the control action will not have a major impact on future values of pressure in the air chamber. Besides, the air pressure signal is mostly clean and easily measurable, as opposed to wave elevation or excitation force. Under these assumptions, the pressure forecasting is relevant for the subsequent model predictive control (MPC)

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *IET Renewable Power Generation* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology



**FIGURE 1** (a) Shoreline view of the Mutriku wave power plant; (b) the biradial turbine during the commissioning tests of the EU H2020 OPERA project [6] and (c) schematic cross section view of the Mutriku power plant [5]

of the wave power plant. For these reasons, the present work is focused on short-term forecasting of the air chamber pressure on an OWC operating with the biradial turbine.

As a case study, the Mutriku power plant is here presented. Located in the Basque Country, Spain, the Mutriku power plant was commissioned in 2011 and it consists of 16 fixed OWCs (see Figure 1). Initially, each OWC was equipped with a Wells turbine and an electrical generator with 18.5 kW of rated power. In June 2017, a biradial turbine was installed on top of one OWC's air chamber with a 30-kW generator (see Figures 1b and 1c), within the framework of the EU H2020 OPERA project (<http://opera-h2020.eu/>). A more detailed description of the Mutriku power plant and performance comparison between the Wells and biradial turbines can be found on [5]. The experimental data used in the current work was measured and obtained in situ, and it is available at <https://doi.org/10.5281/zenodo.4926029>.

This paper is subdivided in the following sections: the current state of the art and the interest of this paper's developments are discussed in Section 2; two linear regressive algorithms are detailed in Section 3 to be subsequently examined for comparison purposes; Section 4 describes support vector machines (SVM) as the primary forecasting method; Section 5 presents the PTO dynamics of the biradial turbine; Section 6 explains the forecasting simulation results; lastly,

Sections 7 and 8 allow space for final remarks and future course of work.

## 2 | LITERATURE REVIEW

The use of wave forecasting models is valuable for constructing improved control strategies for WEC devices and can be found in published literature for this purpose since 1998. It is useful to classify forecast attempts in three different ways:

- As to the duration of the forecast; Mérigaud et al. have classified ocean forecasting for wave energy production [7]: short-term (in the order of seconds); medium-term (hourly and daily basis), useful for the wave energy market; and long-term (several years), used to evaluate the viability of a WEC project;
- As to the purpose of the forecast, there are distinctions between the evolution with time of variables (e.g. wave elevation, or excitation force acting upon a particular WEC), and the evolution with time of statistical parameters of the variables (e.g. significant height of the waves, peak period or power available for extraction). Understandably, the first option is usually connected with short-term forecasts, and the latter with medium and long-term forecasts;



**TABLE 1** Wave energy forecasting published literature: Short term

Year	Authors	Topic	Data	Resolution	Source	Location	Techniques
2019	Skene et al. [8]	Short-term forecasting	Wave elevation	0.031 s	Simulated	-	FFT
2018	Fernandes et al. [9]	Short-term forecasting	Pressure; Turbine speed; Wave elevation	0.5 s	Real	Mutriku, Basque Country, Spain	ANN
2018	Peña-Sanchez et al. [10]	Short-term forecasting	Wave elevation	1 s (simulated); 0.78 s (real)	Real & simulated	Belmullet, Ireland	AR; GP
2018	Shi et al. [12]	Short-term forecasting	Wave elevation	0.78 s (real)	Real	Pacific Marine Energy Center	ANN; AR; GP
2017	Garcia-Abril et al. [13]	Short-term forecasting	Excitation force	0.1 s	Simulated	-	AR; KF
2017	Peña-Sanchez & Ringwood [11]	Short-term forecasting	Wave elevation	0.39 s; 0.5 s; 0.78 s	Real	Belmullet & Galway Bay, Ireland; Pico, Azores, Portugal	AR; ARMA
2015	Monk et al. [14]	Short-term forecasting	Wave elevation; Wave power	0.5 s	Real	Pico, Azores, Portugal	AR
2014	Paparella et al. [15]	Short-term forecasting	Wave elevation	0.5 s	Real	Pico, Azores, Portugal	AR; ARX; FIR
2014	Paparella et al. [16]	Short-term forecasting	Wave elevation	0.5 s	Real	Pico, Azores, Portugal	ANN; AR; ARX; FIR
2011	Sheng & Lewis [18]	Short-term forecasting	Airflow; OWC motions: Surge, Sway, Heave, Roll, Pitch and Yaw	-	Real	Hydraulics & Maritime Research Centre, Ireland	ANN
2010	Fusco & Ringwood [19]	Short-term forecasting	Wave elevation	0.39 s; 0.78 s	Real	Galway Bay, Ireland; Pico, Azores, Portugal	ANN; AR
2010	Fusco & Ringwood [20]	Short-term forecasting	Wave elevation	0.39 s; 0.78 s	Real	Galway Bay, Ireland; Pico, Azores, Portugal	AR
2009	Fusco [21]	Short-term forecasting	Wave elevation	0.39 s; 0.78 s	Real	Galway Bay, Ireland; Pico, Azores, Portugal	ANN; AR; EKF

- As to the forecast technique, there are distinctions between linear and nonlinear methods; among the latter, the SVM technique is particularly popular and especially appealing for WEC applications.

Published literature on this subject is presented in a condensed form in Tables 1 and 2, using the classification paradigms just presented, and described in detail in what follows.

## 2.1 | Short-term forecasting

Skene et al. have recently developed an algorithm that uses the fast Fourier transform (FFT) to predict wave elevation in real time, using simulated data [8]. The algorithm was capable of predicting times in the order of the peak wave period and distances in the order of the peak wavelength.

Several other studies have implemented artificial neural networks (ANN) and other solutions, such as autoregressive (AR) method and some of its variations: autoregressive exogenous (ARX), autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA). In 2018, Fernandes et al. formulated an ANN model to predict the air chamber

pressure of an OWC [9]. This study used real data from the biradial turbine developed at Instituto Superior Técnico, Lisbon. The model has shown good prediction accuracy for horizons ranging from 1 to 3 steps ahead and worse above this threshold, for a sampling period of 0.5 s.

Peña-Sanchez et al. have revisited the AR model and the Gaussian process (GP) techniques to forecast sea surface elevation, using simulated and real data [10]. Although both techniques presented identical accuracies, the AR model was simpler and delivered similar performance. Overall, predictions were inaccurate and particularly worse for real data. The authors also experimented the ARMA model which offered similar accuracies [11]. However since the ARMA algorithm was more complex, the AR model was preferred. Shi et al. employed AR, ANN and GP methods to forecast and optimally control a WEC [12]. On the one hand, a similar conclusion was drawn—the AR solution was simpler and presented similar prediction accuracy. On the other hand, ANN models presented flexible learning structures. Garcia-Abril et al. have advocated that accurate excitation force prediction was required for control strategies to optimize energy capture using AR models [13]. Relevant features that would contribute to the improvement of the excitation force predictions were: the prediction method order, the time step, the forecast horizon and potentially the application of a low-pass

**TABLE 2** Wave energy forecasting published literature: Medium term

Year	Authors	Topic	Data	Resolution	Source	Location	Techniques
2018	James et al. [24]	Medium-term forecasting	Significant wave height; Wave period	3 h	Real	California, United States	ANN; SVM; SWAN
2017	Akhil P. & Deka [25]	Medium-term forecasting	Significant wave height	-	-	-	ANN; ANFIS; ELM; GA; SVR
2017	Mérigaud et al. [7]	Long, medium and short-term forecasting	Wave elevation; Wave power	30 min (medium)	Real & simulated	Galway Bay, Ireland; Pico Island, Azores, Portugal	AR; ARX
2017	Zheng et al. [52]	Long, medium forecasting	Significant wave height; Swell energy; Swell index; Wave power density	-	Reviewed	Indian Ocean; Java & Australia; Clipperton Island, Pacific Ocean	ANFIS; ANN; ARIMA; ECMWF; SVM; SWAN
2016	Gopinath & Dwarakish [45]	Medium and short-term forecasting	Significant wave height	3 h	Real	New Mangalore Port, India	ANN; FFBB; PSO
2015	Reikard et al. [48]	Medium-term forecasting	Significant wave height; Wave period; Wave power	1 h	Real	Washington & Oregon, United States	AR; ECMWF
2015	Reikard et al. [47]	Medium-term forecasting	Significant wave height; Wave period; Wave energy flux; Wave power	1 h	Real	British Columbia, Canada	ARIMA; SWAN
2014	Hadadpour et al. [50]	Medium-term forecasting	Significant wave height; Wave period; Wave energy flux	1 to 12 h	Real	Anzali, Caspian Sea, Iran	ANN
2011	Reikard et al. [46]	Medium-term forecasting	Significant wave height; Wave period; Wave energy flux	1 h	Real	National Data Buoy Centre, United States; Integrated Science Data Management, Canada	ANN; AR; ECMWF
2009	Reikard [49]	Medium-term forecasting	Significant wave height; Wave period; Wave energy flux; Wind speed	1 h; 1 d	Real	United States	ANN; AR; KF; SW
1998	Deo & Sridhar Naidu [51]	Medium-term forecasting	Significant wave height	1 h; 3 h	Real	Yanam, India	ANN; AR

filter. In the previously mentioned study, Mérigaud et al. have tested short-term forecasts time series [7]: a simple AR model offered an accurate prediction for up to two typical wave periods into the future (data from Galway Bay and Pico Island). No real benefit in using ANN models, unless highly nonlinear sea states were encountered, which were unlikely to pertain the power production operational region of WECs. Spatial results have shown little accuracy advantage in making an up-wave measurement (Pico Island).

Earlier in 2015, Monk et al. focused on optimizing Pico's OWC pneumatic power (in Azores archipelago, Portugal) by controlling the pressure valve which regulates pressure variance, based on AR forecasting [14]. An increase of 15% in power production as well as significant reductions in the frequency of the severest turbine stalls were observed, compared to the previous basic control strategy. Performance enhancements could be achieved using short-term wave forecasting with basic equipment and minimal investment. Paparella et al. have published two studies on up-wave measurements to evaluate its relevance on short-term wave forecasting [7, 15, 16]. A finite impulse

response (FIR) model was designed to study the up-wave measurements. An ARX model was executed to predict the wave elevation in the air chamber, combining its past values and the up-wave elevation measurements. Equivalently, the ARX model did not reasonably improve the AR model accuracy and, due to the added complexity of ANNs, linear models were preferable. In the cited studies, the authors initially concluded there was not much benefit in making up-wave measurements and the prediction of the excitation force should be considered. However, Peña-Sanchez et al. confirmed that previous conclusions were related with the use of filtered wave elevation signals, and up-wave measurements could in fact improve the prediction accuracy [10]. Finally, this idea was reinforced in a recent study by Mérigaud and Ringwood [17].

In 2011, an ANN model was applied for short-term prediction of relevant OWC variables: the airflow and OWC motions (surge, sway, heave, roll, pitch and yaw) [18]. Sheng and Lewis have trained a 3-layer ANN which presented very good predictions. There was no need for large ANN perceptrons to provide fast convergence. The network parameters

could work over a broad data range for a good prediction. Finally, Fusco and Ringwood used a simple AR model for wave elevation forecasting for real-time optimal control [19, 20]. Research findings showed that the AR algorithm was able to capture the cyclical wave behavior and offered a very accurate prediction of the low-frequency swell waves, for up to two wave periods. Once again, there was no real benefit in using ANN models, unless on highly nonlinear sea states. As suggested, the WEC controller may require the prediction of the wave excitation force, which was practically a low-pass filtered form of the wave elevation. A previous report by one of the authors additionally considered an extended Kalman filter (EKF): a computationally light and simple method has proven to be significantly effective for sea states with narrow band (5- to 10-s predictions) but entirely ineffective for wider-banded wave systems [21].

Regarding the relevance of prediction horizon, Faedo et al. stated in [22] that state-of-the-art forecasting algorithms are only able to provide an accurate prediction for horizons of a couple seconds. As a contrast, Henriques et al. tested latching control on a OWC [23]—significant gains were only found for receding horizon time intervals greater than the typical energy period of the waves. Thus, the optimal prediction horizon depends on the type of WEC device and the applied control strategy.

## 2.2 | Medium-term forecasting

Other studies have been developed for medium-term forecasting. This approach aims to forecast wave statistical parameters such as the significant wave height  $H_s$  or the wave peak period  $T_p$ . This can be useful to estimate the wave energy flux later. In 2018, James et al. have compared the physics-based simulating waves nearshore (SWAN) model to machine learning (ML) techniques such as ANN and SVM to forecast  $H_s$  and  $T_p$  [24]. ML models were an accurate and computationally efficient surrogate for the SWAN model, running over 4000 times faster. The predicted wave conditions could be used to estimate the power-generation potential of WECs. The authors considered that a convolutional neural network deep learning model with additional data (such as latitude, longitude and bathymetric depth) should be implemented.

The application of ML techniques in wave height forecasting was reviewed in 2017 by Akhil P. and Deka [25]. The authors covered a broad range of methods, and here are a few take-aways:

- The predictive efficiency of an ML approach depended upon the size and quality of the available data set;
- The ANN model took more computational time [26–31] and hybrid models provided better results [32–36];
- The following features had most to least impact on wave height prediction: wind speed, air temperature, sea surface temperature and wind direction [37];
- The adaptive neuro-fuzzy inference system (ANFIS) model was adaptable to optimization, computationally efficient and

fast [37]. Hybrid fuzzy inference system and ANFIS could remove the non-stationary character of wind-wave data [38];

- A genetic algorithm (GA) could work with discontinuous data but required large data support and longer running time [39–42];
- Extreme learning machine (ELM) was a highly fast-training method with an excellent quality-based result [43];
- A support vector regression (SVR) algorithm removed calibration need and could predict wave height accurately [44].

According to Mérigaud et al., medium-term forecasts were useful for wave energy production for market participation [7]. Furthermore, basic sea state parameters were predicted: wave energy flux, significant wave height and wave period. Anyhow, more research was needed. Gopinath and Dwarakish used ANN models trained by particle swarm optimization (PSO) for real-time prediction of waves [45]. With conventional feed-forward back propagation (FFBP) network, the accuracy depended on data length used as input. In the case of the PSO-ANN network, accurate training was possible with smaller input length. Finally, the PSO-ANN network outperformed the conventional FFBP network.

Reikard et al. have presented studies on wave energy medium-term forecasting using time series, with an ARIMA model, and a physics-based model—the European Centre for Medium-range Weather Forecasts (ECMWF) [46, 47]. The authors have also combined wave with wind and solar resources [48]. According to these studies, the ECMWF model was well-suited to wave farm simulation and was able to capture seasonality. The time series models have predicted more accurately over short horizons (1–5 h, higher frequency signals) while the physics models have forecasted more accurately over longer horizons (above 6 h, lower frequency signals)—combining the two methods leads to increased accuracy. Although wave energy flux could be volatile, the wave farm power was smoother, more predictable and easier to forecast when compared to wind or solar. Also, wave reserves were far less expensive than others. As a conclusion, the findings argued strongly for the development of wave energy at coastal locations. On a previous publication, Reikard has tested time-series models such as ANN, AR, Kalman filter (KF) and sliding window (SW) [49]. Regarding the forecasting run frequency (hourly and daily), the higher the frequency, the higher the forecast error of the wave energy flux; time-varying parameter regression and hybrid models were preferred for experiments at hourly frequency. For the ANN model, the forecast accuracy was significantly influenced by the length of the training data set. With high forecast errors, utilities would need to use conventional power sources as a backup, in order to compensate for the inherent variability of waves.

Hadadpour et al. used ANNs for wave energy forecasting in a real study case in the Caspian Sea [50] (2014)—a sigmoid transfer function and 1 hidden layer with 5 neurons was employed using a lag time of 3 h. Increasing the forecasting horizon led to a decrease of the predictive accuracy: average errors of 11% for 1 h forecasting and 41% for 12 h. Also, forecasting the separate components performed better than a direct forecast of the wave energy flux. Models used to

forecast the significant wave and the peak wave period using its past values were more appropriate than others. Lastly, Deo and Sridhar Naidu [51] acknowledged the advantages of using the ANN technique over typical deterministic or statistical models. The cascade correlation scheme was the best training algorithm in terms of convergence speed. The AR results have shown lower covariation of the wave height predictions than the ANN models.

## 2.3 | Long-term forecasting

The assessment of future wave conditions in a time scale of several years is essential for resource evaluation, site selection, feasibility studies and project design. However, no actual meteorological forecast can be reasonably expected for such a long horizon, which is relevant for WEC control. Instead, physical models use past input data from a certain number of years, assuming it is representative of the expectations in the following years. Such studies are specifically referred to as hindcasts rather than forecasts.

A number of hindcast studies have been carried out, and wave hindcasting already presents a reasonable degree of standardization. Mérigaud et al. reported that hindcasts tests that were carried out in different wave climates offered good results [7]—the methodology was robust and coherent. Zheng et al. presented an overview of a medium- to long-term predictions of global wave energy resources [52]. The authors concluded that long-term climate trends and predictability of wave energy still remained scarce. Also, the overall capture efficiency and the WEC's lifespan had a relevant impact on the system's stability [52–55]. The utilization factor of wave energy was closely related with the significant wave height [53, 56], whereas energy quality was determined by the energy level occurrence [57]. Ultimately, Zheng et al. suggested several improvement opportunities for the wave energy prediction capacity, such as the incorporation of climate prediction methods and the application of numerical simulation methods and swell propagation characteristics.

## 2.4 | Concluding remarks

As reviewed previously in this section, a significant part of short-to-medium-term forecasting studies has focused on wave elevation, significant wave height or the wave period. For this purpose, a considerable extent of approaches was put forward, with the major spotlight not only on AR models (and its different forms) and ANNs but also on physics-based models such as ECMWF and SWAN. The majority of the short-term forecasting studies favor the use of regressive algorithms over nonlinear approaches, due to their low complexity and relatively similar performances. As for medium-term, the inclusion of more complex methods in combination with simpler ones can lead to better results. Hindcasting is not especially important for the control of WEC devices, since it is more related with the structural design of wave power plants.

Among ML models, SVM solutions are attracting increasing attention, mostly because they require less configuration complexity than ANN (e.g. the number of hidden layers, the number of hidden nodes, the learning rate, among others [58]) while presenting a robust procedure for solving problems in nonlinear classification, function estimation and density estimation. As an example, the Least Squares Support Vector Machine (LS-SVM) technique only involves two parameters [58]. Moreover, both LS-SVM and SVR algorithms for dynamic models are capable of outperforming ANN [58, 59]. LS-SVM has also been used in recent studies for renewable energy forecasting, namely wind power [60–63].

## 3 | LINEAR MODELS

### 3.1 | Autoregressive

An AR model is used to predict future values of a determined time-variant signal, based on past behavior. Given the correlation between the signal values, the forecasting model comprises a linear regression of its data, that is, a linear combination of the past values (called lags), hence the name autoregressive—the prefix auto stands for "self" [64]. The output of an AR model of order  $p$ ,  $AR(p)$ , at instant  $t$  can be expressed as

$$y_t(\boldsymbol{\phi}) = c + \sum_{i=1}^p \phi_i y_{t-i} + \varepsilon_t, \quad (1)$$

where  $c$  is a constant,  $\boldsymbol{\phi} = (\phi_1, \dots, \phi_p)^T$  is the model parametric vector and  $\varepsilon_t$  is the white noise (or randomness). The constant  $c$  is defined by the mean value of the time series  $\mu$ ,

$$c = \left(1 - \sum_{i=1}^p \phi_i\right) \mu. \quad (2)$$

AR models are fairly flexible at handling a wide range of different time series patterns. Moreover, they consist of a stochastic process with degrees of uncertainty or randomness. Therefore, one can obtain high but not perfect prediction accuracy (lower than 100%) [64, 65].

### 3.2 | Moving average

Among the modeling approaches for time series forecasting, the moving average (MA) model focuses on past prediction errors rather than past data values. Each predicted value is seen as a weighted average of the previous errors [65]. The predicted value of an MA model of order  $q$ ,  $MA(q)$ , is described as

$$y_t(\boldsymbol{\theta}) = \mu + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t, \quad (3)$$

where  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)^T$  is the model parametric vector.



For each forecasting cycle, the average is re-calculated to predict future values—thus the term “moving average”. This technique is especially useful for forecasting long-term trends [66].

### 3.3 | Autoregressive moving average

In the statistical analysis of a times series, an ARMA model presents a combination of AR and MA polynomials and describes a weakly stationary stochastic process. The ARMA model comprises two order dimensions,  $p$  and  $q$ , for the AR and MA models, respectively [67]. Accordingly, ARMA( $p, q$ ) yields

$$y_t(\phi, \theta) = c + \varepsilon_t + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i}. \quad (4)$$

## 4 | SUPPORT VECTOR MACHINES

### 4.1 | Overview

ML has experienced explosive growth in recent years fueled by the emergence of advanced algorithms capable of identifying complex patterns in data from a limited number of samples. Among other ML techniques, SVM is a powerful methodology for solving problems in nonlinear classification, function estimation and density estimation. Originally proposed by Vapnik [68, 69], SVM is gaining popularity in the ML realm, due to its attractive features and promising empirical performance [70].

### 4.2 | Linear classification

At its core, SVM primarily performs classification tasks by constructing hyperplanes in a multidimensional space that separates cases of different class labels. For the example of a linear SVM problem, let us assume a training data set of  $M$  points of the form  $(\mathbf{x}_i, y_i)$  where  $i = 1, \dots, M$ ,  $\mathbf{x}_i \in \mathbb{R}^N$  and  $y_i = \pm 1$ . In other words, each input vector  $\mathbf{x}_i$  with  $N$  dimensions is related with the respective binary output  $y_i$ , depending on the class  $\mathbf{x}_i$  belongs. The algorithm intends to find the “maximum-margin hyperplane” that separates the group of points  $\mathbf{x}_i$  for which the output is 1 from its counterpart group where the output equals  $-1$ . The hyperplane is defined so that the distance between itself and the nearest point  $\mathbf{x}_i$  from either group is maximized. It is a set of points  $\mathbf{x}$  satisfying

$$\mathbf{w} \cdot \mathbf{x} + b = 0, \quad (5)$$

where  $\mathbf{w}$  is the normal vector to the hyperplane. The parameter  $b/\|\mathbf{w}\|^{-1}$  determines the offset of the hyperplane from the origin along the normal vector  $\mathbf{w}$ .

### 4.3 | Nonlinear classification

The original hyperplane algorithm proposed by Vapnik consisted of a linear classifier. Although a linear problem may

seem simple to solve, most of the real-world data do not show a linear relation, which makes it hard to separate different classes. One way to solve a 2-dimensional nonlinear problem would be to map the data to a 3-dimensional space. However, if a nonlinear problem involves a considerable amount of dimensions, the algorithm computation also becomes more expensive.

Bernhard Boser et al. suggested a way to create nonlinear classifiers by applying the kernel trick [71]. The kernel trick avoids the explicit mapping to higher dimensions, allowing to operate in a transformed feature space with reduced complexity—every dot product is replaced by a nonlinear kernel function and the resulting algorithm is formally similar. As the example described in [72], let us consider two data points in a 3D space:

$$\mathbf{x}_a = (x_{a_1}, x_{a_2}, x_{a_3})^T, \quad (6)$$

$$\mathbf{x}_b = (x_{b_1}, x_{b_2}, x_{b_3})^T. \quad (7)$$

Assuming there is need to map this problem to a 9-dimensional space, the resulting transforming functions, denoted as  $\phi(\cdot)$ , would be:

$$\phi(\mathbf{x}_a) = (x_{a_1}^2, x_{a_1}x_{a_2}, x_{a_1}x_{a_3}, x_{a_2}x_{a_1}, x_{a_2}^2, \quad (8)$$

$$x_{a_2}x_{a_3}, x_{a_3}x_{a_1}, x_{a_3}x_{a_2}, x_{a_3}^2)^T,$$

$$\phi(\mathbf{x}_b) = (x_{b_1}^2, x_{b_1}x_{b_2}, x_{b_1}x_{b_3}, x_{b_2}x_{b_1}, x_{b_2}^2, \quad (9)$$

$$x_{b_2}x_{b_3}, x_{b_3}x_{b_1}, x_{b_3}x_{b_2}, x_{b_3}^2)^T.$$

The algorithm output is just a scalar determined by

$$\phi(\mathbf{x}_a)^T \phi(\mathbf{x}_b) = \sum_{i,j=1}^3 x_{a_i}x_{a_j}x_{b_i}x_{b_j}, \quad (10)$$

with a computational complexity of  $\mathcal{O}(N^2)$ . Using the kernel function  $k(\mathbf{x}_a, \mathbf{x}_b)$  instead, it is possible to reach the same result within the same space order, by

$$\begin{aligned} k(\mathbf{x}_a, \mathbf{x}_b) &= (\mathbf{x}_a^T \mathbf{x}_b)^2 = (x_{a_1}x_{b_1} + x_{a_2}x_{b_2} + x_{a_3}x_{b_3})^2 \\ &= \sum_{i,j=1}^3 x_{a_i}x_{a_j}x_{b_i}x_{b_j}. \end{aligned} \quad (11)$$

The computational complexity is reduced to  $\mathcal{O}(N)$ . Some common kernel functions include:

- Polynomial:

$$k(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j)^n. \quad (12)$$



- Gaussian radial basis function (RBF):

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2), \quad (13)$$

with  $\gamma > 0$ . Sometimes parametrized using  $\gamma = (2\sigma^2)^{-1}$ .

- Hyperbolic tangent:

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\kappa \mathbf{x}_i \cdot \mathbf{x}_j + c), \quad (14)$$

for some (not every)  $\kappa > 0$  and  $c < 0$ .

For further details, see [71, 72].

## 4.4 | Support vector regression

### 4.4.1 | Overview

As mentioned previously, SVM has been widely applied in classification problems (e.g. text and image classification or pattern recognition), by formulating them as convex optimization problems [73]. Drucker et al. have proposed a version of SVM for regression [74], commonly known as SVR. This method is a generalization of the classification problem, in which the model returns a continuous-valued output (a model function), as opposed to an output from a finite set [75]. The purpose is to find the function  $f(\mathbf{x})$  that best describes the relationship between the features and the target [68], while balancing complexity and prediction error. This is done by first defining a convex cost function to be minimized and finding an optimal region around the function that contains most of the training instances, known as  $\epsilon$ -tube [75]. In recent years, the SVR technique has proven successful for prediction tasks with promising results [76, 77].

### 4.4.2 | Definition

Considering a training data set of  $M$  points  $(\mathbf{x}_i, y_i)$  where  $i = 1, \dots, M$ ,  $\mathbf{x}_i \in \mathbb{R}^N$  and  $y_i \in \mathbb{R}$ , where the  $N$ -dimensional vector  $\mathbf{x}_i$  is the  $i$ th sample of the features and  $y_i$  is the corresponding target. The regression function  $f$ , prospectively used for prediction, relates with the SVM classification components: a hyperplane vector  $\mathbf{w}$ , a feature-space transformation  $\phi(\cdot)$  and a bias parameter  $b$ ,

$$f(\mathbf{x}_i) = \mathbf{w}^T \phi(\mathbf{x}_i) + b. \quad (15)$$

Depending on the problem, an appropriate kernel function is chosen for  $\phi(\cdot)$  (see Section 4.3).

In a linear regression, the objective is to minimize a  $\lambda$ -regularized error function [78]:

$$\frac{1}{2} \sum_{i=1}^M (f(\mathbf{x}_i) - y_i)^2 + \frac{1}{2} \lambda \|\mathbf{w}\|^2, \quad (16)$$

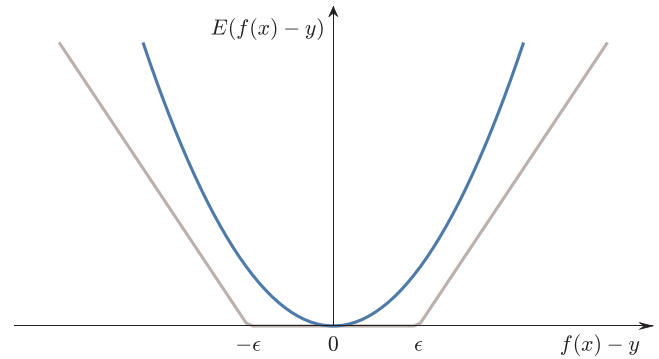


FIGURE 2  $\epsilon$ -insensitive error function

with  $\lambda > 0$ . The quadratic error function is replaced by an  $\epsilon$ -insensitive error function, in order to obtain sparse solutions [68], with  $\epsilon > 0$ —if the absolute difference between the target and the prediction is less than  $\epsilon$  the function returns zero. An example of an  $\epsilon$ -insensitive error function is given by [78]

$$E_\epsilon(f(\mathbf{x}_i) - y_i) = \begin{cases} 0, & \text{if } |f(\mathbf{x}_i) - y_i| < \epsilon \\ |f(\mathbf{x}_i) - y_i| - \epsilon, & \text{otherwise} \end{cases}, \quad (17)$$

and is illustrated in Figure 2.

Introducing the  $\epsilon$ -insensitive error function, Equation (16) becomes

$$C \sum_{i=1}^M E_\epsilon(f(\mathbf{x}_i) - y_i) + \frac{1}{2} \|\mathbf{w}\|^2, \quad (18)$$

where  $\lambda$  is replaced by an inverse parameter  $C$ , with  $C > 0$ , associated with the error term. Two slack variables are introduced into the optimization problem, for each data point  $\mathbf{x}_i$ :

$$\begin{cases} \xi_i > 0 \wedge \hat{\xi}_i = 0, & \text{if } y_i > f(\mathbf{x}_i) + \epsilon \\ \xi_i = 0 \wedge \hat{\xi}_i > 0, & \text{if } y_i < f(\mathbf{x}_i) - \epsilon, \\ \xi_i = \hat{\xi}_i = 0, & \text{otherwise} \end{cases}, \quad (19)$$

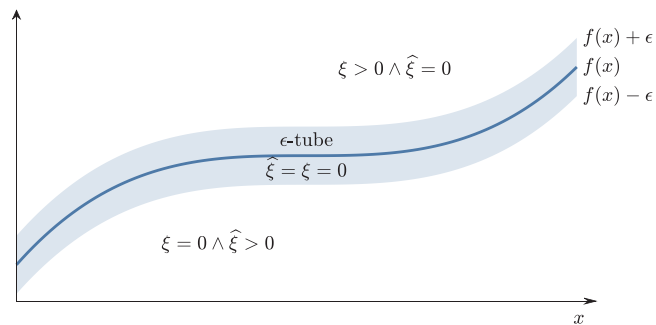
as illustrated in Figure 3. By virtue of the slack variables, only points that lie outside the  $\epsilon$ -tube will be considered for the optimization, provided the following conditions are verified:

$$\xi_i \geq 0, \quad (20)$$

$$\hat{\xi}_i \geq 0, \quad (21)$$

$$y_i \leq f(\mathbf{x}_i) + \epsilon + \xi_i, \quad (22)$$

$$y_i \geq f(\mathbf{x}_i) - \epsilon - \hat{\xi}_i. \quad (23)$$



**FIGURE 3** Illustration of an SVR problem, showing the regression function together with the  $\epsilon$ -tube

Excluding data points inside the  $\epsilon$ -tube, Equation (18) can be re-expressed as

$$C \sum_{i=1}^M (\xi_i + \hat{\xi}_i) + \frac{1}{2} \|\mathbf{w}\|^2, \quad (24)$$

which must be minimized subject to the constraints (20) to (23). This can be achieved with a Lagrangian optimization function [78],

$$\begin{aligned} L(\mathbf{a}, \hat{\mathbf{a}}, \boldsymbol{\mu}, \hat{\boldsymbol{\mu}}) &= C \sum_{i=1}^M (\xi_i + \hat{\xi}_i) + \frac{1}{2} \|\mathbf{w}\|^2 \\ &\quad - \sum_{i=1}^M (\mu_i \xi_i + \hat{\mu}_i \hat{\xi}_i) \\ &\quad - \sum_{i=1}^M a_i (\epsilon + \xi_i + f(\mathbf{x}_i) - y_i) \\ &\quad - \sum_{i=1}^M \hat{a}_i (\epsilon + \hat{\xi}_i - f(\mathbf{x}_i) + y_i), \quad (25) \end{aligned}$$

with the non-negative Lagrange multipliers  $a_i, \hat{a}_i, \mu_i$  and  $\hat{\mu}_i$ . Substituting  $f(\mathbf{x}_i)$  using Equation (15) and setting the Lagrangian derivatives with respect to  $\mathbf{w}, b, \xi_i$  and  $\hat{\xi}_i$  to zero,

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \Rightarrow \mathbf{w} = \sum_{i=1}^M (a_i - \hat{a}_i) \boldsymbol{\phi}(\mathbf{x}_i), \quad (26)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^M (a_i - \hat{a}_i) = 0, \quad (27)$$

$$\frac{\partial L}{\partial \xi_i} = 0 \Rightarrow a_i + \mu_i = C, \quad (28)$$

$$\frac{\partial L}{\partial \hat{\xi}_i} = 0 \Rightarrow \hat{a}_i + \hat{\mu}_i = C. \quad (29)$$

Applying these optimal solutions into Equation (25) to eliminate  $\mathbf{w}, b, \xi_i$  and  $\hat{\xi}_i$  and obtaining the dual Lagrangian function  $\tilde{L}$  [78], it follows

$$\begin{aligned} \tilde{L}(\mathbf{a}, \hat{\mathbf{a}}) &= -\frac{1}{2} \sum_{i,j=1}^M (a_i - \hat{a}_i)(a_j - \hat{a}_j) k(\mathbf{x}_i, \mathbf{x}_j) \\ &\quad - \epsilon \sum_{i=1}^M (a_i + \hat{a}_i) + \sum_{i=1}^M (a_i - \hat{a}_i) y_i, \quad (30) \end{aligned}$$

where the vectors  $\mathbf{a} = (a_1, \dots, a_M)^T$  and  $\hat{\mathbf{a}} = (\hat{a}_1, \dots, \hat{a}_M)^T$  collect the Lagrangian multipliers. The kernel function,

$$k(\mathbf{x}_i, \mathbf{x}_j) = \boldsymbol{\phi}(\mathbf{x}_i)^T \boldsymbol{\phi}(\mathbf{x}_j), \quad (31)$$

is here introduced as defined in Section 4.3.

The dual Lagrangian function needs to be maximized to attain the optimal multipliers, under the following constraints: Equation (27); since  $\mu_i$  and  $\hat{\mu}_i$  are non-negative, Equations (28) and (29) imply that

$$0 \leq a_i \leq C, \quad (32)$$

$$0 \leq \hat{a}_i \leq C. \quad (33)$$

Many quadratic problem solvers can handle this problem. Usually, only certain parts of the samples can satisfy the property  $a_i - \hat{a}_i \neq 0$ , which are the sparse solutions. They are called support vectors since only this subset of samples contribute to the model [79].

Finally, considering Equations (26) and (31) into Equation (15), the predictions for new inputs are determined by [78]

$$f(\mathbf{x}) = \sum_{i=1}^M (a_i - \hat{a}_i) k(\mathbf{x}, \mathbf{x}_i) + b. \quad (34)$$

The Karush–Kuhn–Tucker conditions, which state that at the solution the product of the dual variables  $a_i, \hat{a}_i, \xi_i$  and  $\hat{\xi}_i$  and the constraints (22), (23), (28) and (29) must vanish [78], are given by

$$a_i (\epsilon + \xi_i + f(\mathbf{x}_i) - y_i) = 0, \quad (35)$$

$$\hat{a}_i (\epsilon + \hat{\xi}_i - f(\mathbf{x}_i) + y_i) = 0, \quad (36)$$

$$(C - a_i) \xi_i = 0, \quad (37)$$

$$(C - \hat{a}_i) \hat{\xi}_i = 0. \quad (38)$$

From condition (35),  $a_i$  is non-zero if  $y_i = f(\mathbf{x}_i) + \epsilon + \xi_i$ , which means the data point  $y_i$  either lies on the upper boundary of the  $\epsilon$ -tube ( $\xi_i = 0$ ) or above ( $\xi_i > 0$ ). Analogously from

condition (36), if  $\hat{a}_i \neq 0$  implies  $y_i = f(\mathbf{x}_i) - \epsilon - \hat{\xi}_i$ , such points must lie either on ( $\hat{\xi}_i = 0$ ) or below ( $\hat{\xi}_i > 0$ ) the lower boundary of the  $\epsilon$ -tube.

Furthermore, since  $\epsilon$  is strictly positive, the two constraints  $y_i = f(\mathbf{x}_i) + \epsilon + \hat{\xi}_i$  and  $y_i = f(\mathbf{x}_i) - \epsilon - \hat{\xi}_i$  are incompatible, that is, a point  $y_i$  cannot be simultaneously on both upper boundary (or above) and lower boundary (or below). This means for every data point  $\mathbf{x}_i$ , either  $a_i$  or  $\hat{a}_i$  (or both) are zero.

The bias parameter  $b$  can be determined using a data point which satisfies  $0 < a_i < C$ : the condition (37) requires  $\hat{\xi}_i = 0$ ; and from condition (35), the outcome  $y_i = f(\mathbf{x}_i) + \epsilon$  is observed. Solving Equation (15) for  $b$  and using Equation (26),

$$b = y_i - \epsilon - \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_i) = y_i - \epsilon - \sum_{j=1}^M (a_j - \hat{a}_j) \kappa(\mathbf{x}_i, \mathbf{x}_j). \quad (39)$$

## 4.5 | Least squares support vector machines

### 4.5.1 | Overview

SVM algorithms have been introduced within the context of statistical learning theory and structural risk minimization. The high computational complexity is among the main drawbacks of standard SVM. Therefore a new technique—the LS-SVM—has been developed by Suykens and Vandewalle [80, 81]. This approach consists of a set of supervised learning methods that analyze data and recognize patterns for classification and regression problems. An additional advantage of the LS-SVM solution is that model optimization can be performed relatively fast. Finally, it is closely associated to regularization networks [82] and Gaussian processes [83] but additionally include primal-dual interpretations.

Some basic definitions about the LS-SVM method are presented here—further details in [84, 85]. Furthermore, the LS-SVM learning process is presented in the flowchart in Figure 4.

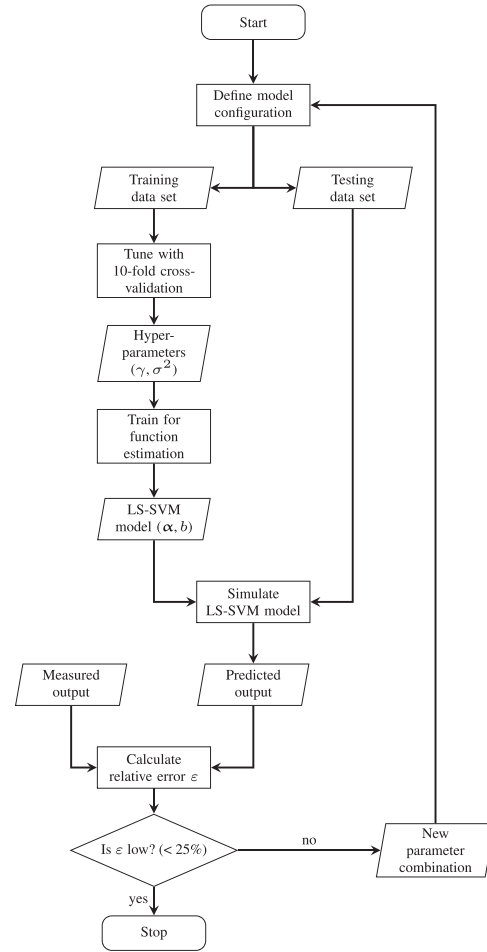
### 4.5.2 | Definition

Similarly to the example in Equation (16) in Section 4.4.2, the LS-SVM error function presented in [86] requires minimization,

$$\frac{1}{2} \gamma \sum_{i=1}^M e_i^2 + \frac{1}{2} \|\mathbf{w}\|^2, \quad (40)$$

subject to

$$y_i = f(\mathbf{x}_i) + e_i, \quad (41)$$



**FIGURE 4** LS-SVM learning process. In this flowchart, rectangles, parallelograms and diamonds represent processes, input/outputs and decisions, respectively. The model configuration entails the training data size, the feature vector size and the forecast horizon. The relative error  $\epsilon$  is the dimensionless root mean squared error, and it is considered low for values under 25%

and where the regularization constant satisfies  $\gamma > 0$ . The Lagrangian optimization function [86] is defined as

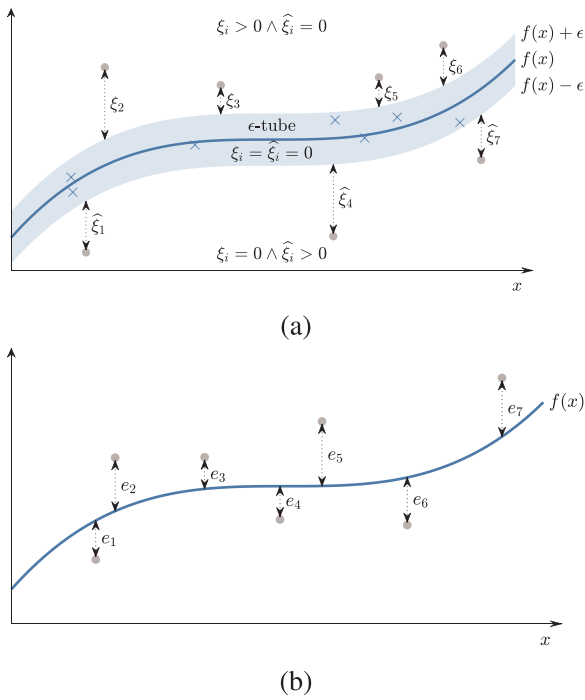
$$L(\boldsymbol{\alpha}) = \frac{1}{2} \gamma \sum_{i=1}^M e_i^2 + \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^M \alpha_i (e_i + f(\mathbf{x}_i) - y_i), \quad (42)$$

with the non-negative Lagrange multipliers  $\alpha_i$ , also known as support values. Substituting  $f(\mathbf{x}_i)$  using Equation (15) and setting the Lagrangian derivatives with respect to  $\mathbf{w}$ ,  $b$  and  $e_i$  to zero,

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \Rightarrow \mathbf{w} = \sum_{i=1}^M \alpha_i \boldsymbol{\phi}(\mathbf{x}_i), \quad (43)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^M \alpha_i = 0, \quad (44)$$

$$\frac{\partial L}{\partial e_i} = 0 \Rightarrow \alpha_i = \gamma e_i, \quad (45)$$



**FIGURE 5** Sparseness comparison between regression methods: (a) SVR, with the  $\epsilon$ -tube and (b) LS-SVM, with the values  $e_i$

$$\frac{\partial L}{\partial \alpha_j} = 0 \Rightarrow y_j = f(\mathbf{x}_j) + e_j. \quad (46)$$

These optimal conditions are similar to its SVR counterparts, shown in Equations (26)–(29), except for the condition (45), where the sparseness property is not applied, in the LS-SVM model.

Merging Equations (43)–(45) with Equation (46), the optimal support vector  $\alpha$  and parameter bias  $b$  are found by solving the following equation system [86]:

$$\begin{bmatrix} \mathbf{K} + \gamma^{-1} \mathbf{I}_M & \mathbf{1}_{M \times 1} \\ \mathbf{1}_{1 \times M} & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ b \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}, \quad (47)$$

where the kernel matrix  $\mathbf{K} \in \mathbb{R}^{M \times M}$  has the elements  $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$ . The target vector is  $\mathbf{y} = (y_1, \dots, y_M)^T$ , the vectors  $\mathbf{1}_{M \times 1} = \mathbf{1}_{1 \times M}^T = (1, \dots, 1)^T$  and  $\mathbf{I}_M$  is the identity matrix.

The resulting LS-SVM regression model for prediction, from Equations (43), (31) and (15), becomes

$$f(\mathbf{x}) = \sum_{i=1}^M \alpha_i k(\mathbf{x}, \mathbf{x}_i) + b. \quad (48)$$

Comparing Equations (40) and (24), Wang et al. conclude that LS-SVM is a reformulation of the SVR principle, using equality constraints [86]. Furthermore, LS-SVM training uses the least squares instead of the  $\epsilon$ -insensitive loss function, as illustrated in Figure 5: Figure 5a presents SVR with a  $\epsilon$ -tube (excluding interior points) and the slack variables ( $\xi_i$  and  $\hat{\xi}_i$ , corresponding to

the support vectors for sparse solutions); while in the LS-SVM model (Figure 5b) those features are replaced by error variables  $e_i$  (distances from each point to the regression function)—all the samples in LS-SVM are support vectors, meaning that all training data is used for regression [86].

Nevertheless, despite lack of sparseness, the LS-SVM algorithm has a great benefit: unlike the complex SVR problem which requires a computationally hard quadratic programming solver, the LS-SVM model is determined from a linear system solution (see Equation (47)), making it easier and faster to optimize [86]. Therefore LS-SVM model is preferred for this study.

## 5 | POWER TAKE-OFF SYSTEM

The biradial turbine was installed at the Mutriku power plant in June 2017 and operated until July 2018 (see Figures 1b and 1c), as a de-risk phase before the installation in the IDOM Marmok-A5 OWC spar buoy that was deployed at the BiMEP test site (Basque Country, Spain) in October 2018. The performance characteristics of the Mutriku’s biradial turbine are usually presented in dimensionless form, where

$$\Psi = \frac{p^* p_{at}}{\rho_{in} \Omega^2 d^2}, \quad (49)$$

$$\Phi = \frac{Q_{turb}}{\Omega d^3}, \quad (50)$$

$$\Pi = \frac{T_{turb}}{\rho_{in} \Omega^2 d^5}, \quad (51)$$

$$\eta_{turb} = \frac{\Pi}{\Phi \Psi}, \quad (52)$$

are the dimensionless pressure head, the dimensionless flow rate, the dimensionless power coefficient and turbine efficiency. In Equations (49), (50) and (51),  $p^*$  is the dimensionless air chamber pressure relative to the atmosphere, defined as

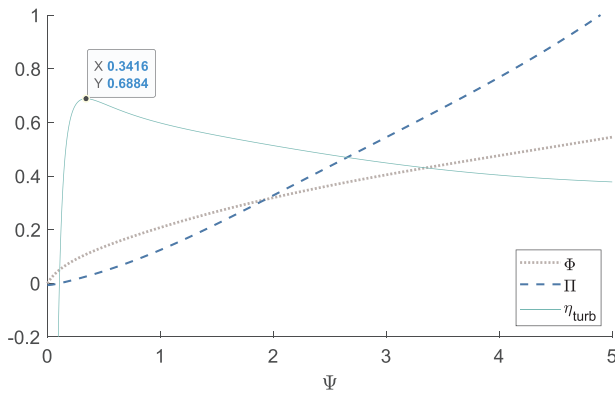
$$p^* = \frac{p}{p_{at}} - 1, \quad (53)$$

where  $p$  is the instantaneous pressure inside the air chamber, and  $p_{at}$  is the atmospheric pressure. The turbine rotational speed, rotor diameter, volumetric flow rate and turbine aerodynamic torque are denoted by  $\Omega$  (in radians per unit time),  $d$ ,  $Q_{turb}$ , and  $T_{turb}$ , respectively. The reference density  $\rho_{in}$  is defined in stagnation conditions at the turbine entrance.

For large Reynolds numbers ( $Re_{tip} > 10^6$ ) and low Mach numbers ( $Ma_{tip} < 0.3$ ), based on tip blade speed, the dimensionless variables  $\Phi$ ,  $\Pi$  and  $\eta_{turb}$  can be plotted as simple curves as function of  $\Psi$ , as shown in Figure 6. The curves for  $\Phi$  and  $\Pi$  were estimated by best fitting the turbine’s measured data points.  $\eta_{turb}$  follows Equation (52).

Axial impulse turbines and the biradial turbine are theoretical quadratic turbines. However, experimental tests of the biradial turbine tested at Mutriku showed a relationship almost quadratic





**FIGURE 6** Turbine curves: dimensionless flow rate  $\Phi$ , dimensionless power coefficient  $\Pi$  and efficiency  $\eta_{\text{turb}}$ , as functions of the dimensionless pressure head  $\Psi$ , for the biradial turbine used in the numerical simulations, based on [4]

between the flow rate coefficient and the dimensionless pressure head, approximately defined as

$$\Phi^{5/3} = \kappa \Psi, \quad (54)$$

where  $\kappa$  is a dimensionless constant. Replacing Equations (50) and (49), Equation (54) evolves into

$$Q_{\text{turb}} = \left( \frac{\kappa p^* p_{\text{at}}}{\rho_{\text{in}}} \right)^{3/5} \frac{d^{9/5}}{\Omega^{1/5}}, \quad (55)$$

after algebraic manipulation. Equation (55) shows that the flow rate is weakly affected by typical changes of the rotational speed since  $Q_{\text{turb}}$  is proportional to  $\Omega^{-1/5}$ .

As introduced in the Section 1, the PTO system for the OWC consists of a biradial turbine and a generator [5]. The dynamics of the turbine/generator set is described by

$$I\dot{\Omega} = T_{\text{turb}}(p^*, \Omega) - T_{\text{gen}}(\Omega), \quad (56)$$

where  $T_{\text{gen}}$  is the generator electromagnetic torque imposed by the rotational speed control law.  $I$  represents the moment of inertia of the turbine/generator set. The turbine aerodynamic torque is computed from Equation (51),

$$T_{\text{turb}} = \rho_{\text{in}} \Omega^2 d^5 \Pi(\Psi). \quad (57)$$

As a conclusion, the opportunity to explore the biradial turbine seems to be beneficial: the high-efficiency range and the near-quadratic behavior that makes the hydrodynamics almost decoupled from the turbine control. Consequently, the pressure is a satisfying and important element for the MPC system to be implemented at the wave power plant.

## 6 | RESULTS

An LS-SVM algorithm was developed for the forecast of the pressure signal  $p^*$  inside the air chamber of the Mutriku WEC

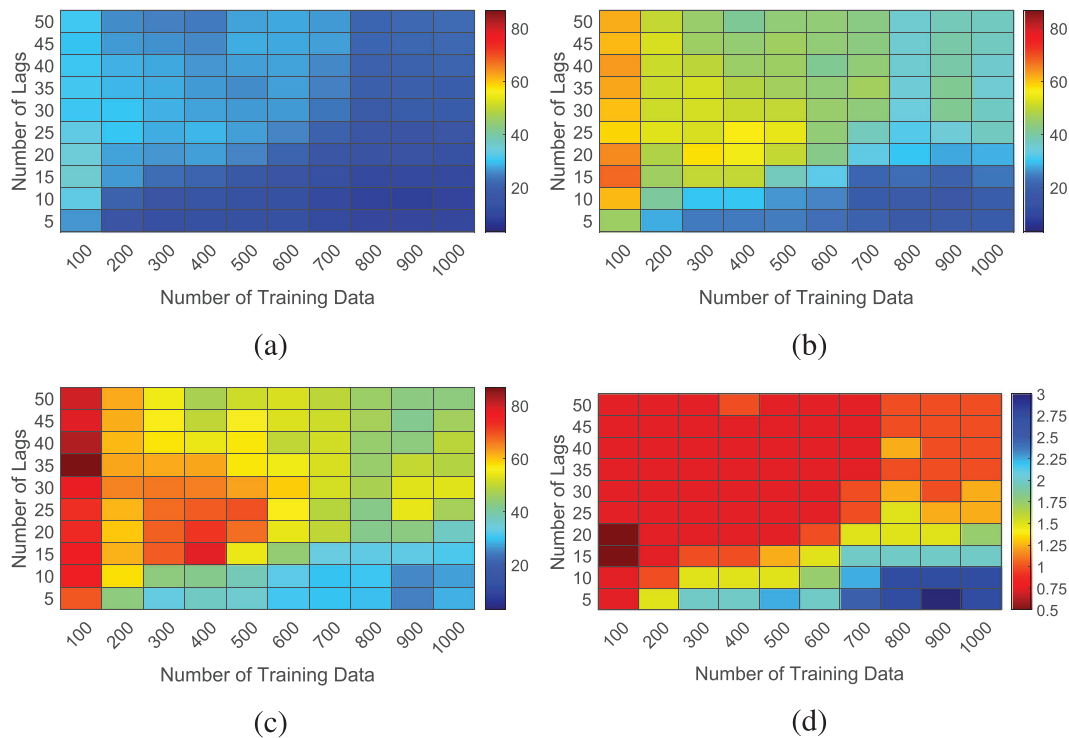
for subsequent use as input in a future MPC controller. As mentioned in Section 1, the pressure data used for training this algorithm was measured during sea trials of the biradial turbine at the Mutriku power plant. However, there was no onographic buoy up-wave of the Mutriku power plant, thus it was not possible to measure wave elevation. With no sea state data available and because the OWC is nonlinear—due to the air compressibility in the chamber and the air turbine behavior—the choice of a nonlinear technique such as LS-SVM, presents the benefit of describing both linear and nonlinear systems. Still, for comparison purposes, both AR and ARMA linear models were developed and the respective forecasts were carried out in parallel.

The design of the LS-SVM forecasting model formulation encompassed the definition of the training data set, the feature vector (or number of lags) and the forecast horizon. For each forecasting cycle, the model was tuned and trained before being tested—this impacted the computational effort, but delivered full prediction capability. Subsequently, the model hyperparameters were tuned. These are the regularization parameter  $\gamma$ , which determines the trade-off between error minimization and smoothness, and the squared bandwidth  $\sigma^2$  of the Gaussian RBF kernel. The simplex method was used to optimize, and the performance was estimated with 10-fold cross validation, using mean squared error as a cost measure. For each forecasting horizon window, the model training employed a data set of previously measured pressure values—in the form of a Hankel matrix, defined by the numbers of training data and features—to calculate the support values  $\alpha_i$  and the bias term  $b$  of the LS-SVM model. Finally, the model was tested using recently measured data as input (distinct from the training set), resulting on the forecast of the upcoming pressure values, within the defined time window.

An exhaustive sensitivity study was carried out to evaluate the best LS-SVM model configuration. The training data set size, the feature vector size and forecast horizon were varied. The same procedure was adopted to run the AR and ARMA algorithms. For the specific case of the ARMA model, an additional variable is introduced—the MA order, corresponding to the lagged error values (see Section 3.2). Each model configuration was ran for 5 min, using a time step of 0.25 s. The model performance was evaluated based on the relative prediction error. This error was determined by the root mean squared error (RMSE) with  $M$  points, and the original signal standard deviation  $\sigma$ ,

$$\begin{aligned} \varepsilon[\%] &= \frac{1}{\sigma} \text{RMSE}(p^*, \hat{p}^*) \times 100\% \\ &= \frac{1}{\sigma} \sqrt{\frac{1}{M} \sum_{i=1}^M (p_i^* - \hat{p}_i^*)^2} \times 100\%. \end{aligned} \quad (58)$$

Figure 7 show calculated prediction errors for forecasts of 1, 2 and 3 s ahead. The number of training data ranged from a hundred to a thousand data points, and the number of lags varied from 5 to 50 elements. Figure 7d maps the maximum forecast horizon possible to attain with a prediction error smaller than 25%.



**FIGURE 7** LS-SVM relative prediction error (%) of pressure using a forecast horizon of: (a) 1, (b) 2 and (c) 3 s; (d) maximum LS-SVM forecast horizon (in seconds) with relative error under 25%

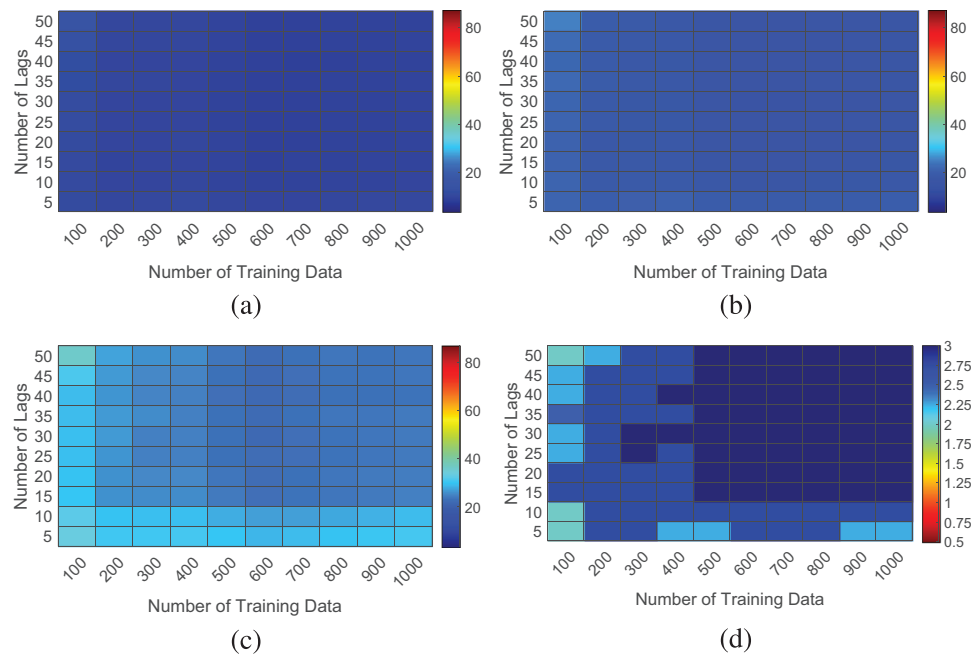
In regard to the LS-SVM parameter sweep, there is a significant increase in the relative prediction error as the forecast horizon is increased, as expected. Interestingly, there generally seems to be no need for large training data sets to achieve accurate forecasts. However, the features size plays a relevant role—the more features, the less accurate model is possible to obtain. The best models offer errors of 8.74%, 17.27% and 24.98% (average of 17%), depending on the forecasting horizon. The maximum forecast horizon map shows a similar trend, where the smaller horizons (from 0.5 to 1.75 s) represent the worst cases. Figure 7d also shows a curious trend on high training data configurations: the more features the smaller accurate horizons. This is probably a consequence of unnecessary past data training.

Analogously, Figures 8 and 9 show the outcomes of both AR and ARMA parameter sweeps, respectively. As for the ARMA sweep, the additional variable of MA lags was equally varied from 5 to 50 elements—along with this range and for each combination of training data and AR lags, the lowest error configurations were selected for graphical representation. The black squares in Figure 9 captioned as not a number (NaN) represent models that failed to simulate due to reduced training data size for high AR and MA lags. Compared to LS-SVM, AR and ARMA maps show more of a bluer trend, meaning there are more AR and ARMA configuration models which would deliver a lower prediction error and longer accurate horizons. This suggests that both the variation of the training data and features sizes would not impact the AR and ARMA outcomes as much

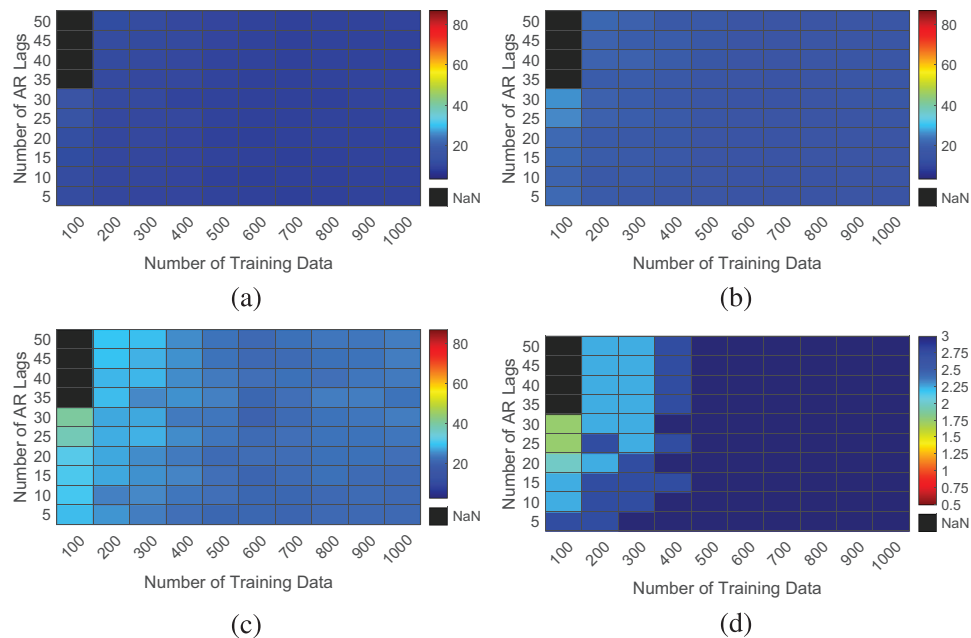
as it would in the case of LS-SVM solution. The regressive models outperformed the LS-SVM approach, with best model errors of 8.51%, 15.57% and 22.01% for AR (average of 15.36%) and 8.37%, 15.1% and 20.81% for ARMA (average of 14.76%), depending on the forecasting horizon.

All three forecasting algorithms were reproduced in more detail in order to assess their predicting potential. Considering models with the best performances picked from Figures 7 to 9, forecasts of the real pressure signal were put forward, considering horizons ranging from 1 to 3 s ahead. The lowest prediction errors defined the best LS-SVM performances (900 for training and 5–10 feature data points) and its AR and ARMA counterparts (training points: 600–700 for both AR and ARMA models; feature data points: 30–40 for AR models; 5 and 30 for ARMA models). For the purpose of these forecasts, the original signal was subdivided into 10 distinct segments, each segment with 5 peak periods of the signal (see Figure 10a). The peak period of the signal (12.5 s) was estimated from a power spectral density analysis of the original signal. Consequently, the total run time of the forecasts was 625 s. The dimensionless relative pressure signal fluctuated between the extreme values of  $-0.0489$  and  $0.0738$  ( $-4957 \text{ Pa} \leq (p - p_{at}) \leq 7481 \text{ Pa}$ ), with a mean value of  $0.0005849$  ( $\bar{p} - p_{at} = 59 \text{ Pa}$ ) and a standard deviation of  $0.0192$  ( $\sigma_{p-p_{at}} = 1948 \text{ Pa}$ ).

The prediction error  $\varepsilon$  was computed for each signal segment, and the respective values are mapped in Figure 10, considering different horizons. Real and forecasted pressure signals are shown in Figure 11—segment #2 from Figure 10a was



**FIGURE 8** AR relative prediction error (%) of pressure using a forecast horizon of: (a) 1, (b) 2 and (c) 3 s; (d) maximum AR forecast horizon (in seconds) with relative error under 25%

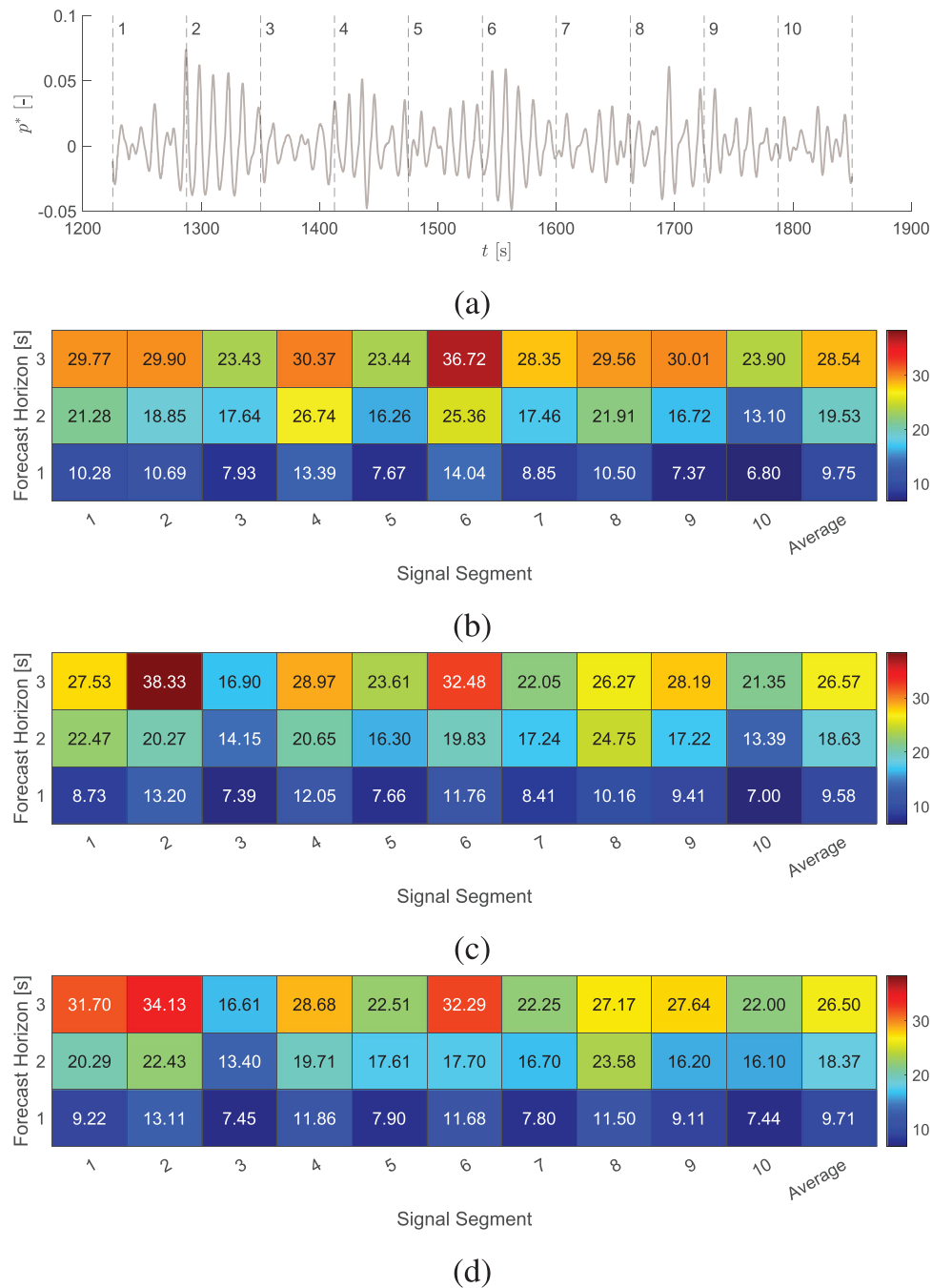


**FIGURE 9** ARMA relative prediction error (%) of pressure using a forecast horizon of: (a) 1, (b) 2 and (c) 3 s; (d) maximum ARMA forecast horizon (in seconds) with relative error under 25%

chosen to observe the forecasting performances. With respect to the forecast horizon from 1 to 3 s, LS-SVM results exhibit prediction errors ranging from 9.75% to 28.54%, while AR and ARMA vary from 9.58% to 26.57%. The results on Figures 10 and 11 indicate an overall better performance for AR and ARMA models, when compared to the LS-SVM model: for

shorter horizons AR and ARMA performances are only slightly improved, whereas AR and ARMA behave significantly better for longer ones.

In Figure 11, one can observe several jumps in the pressure signal, more noticeably in Figure 11 c. These jumps are not related to the measured pressure signal, they occur due to the



**FIGURE 10** (a) Real pressure signal subdivided into 10 distinct segments. Best prediction relative error [%]: (b) LS-SVM, (c) AR and (d) ARMA performances

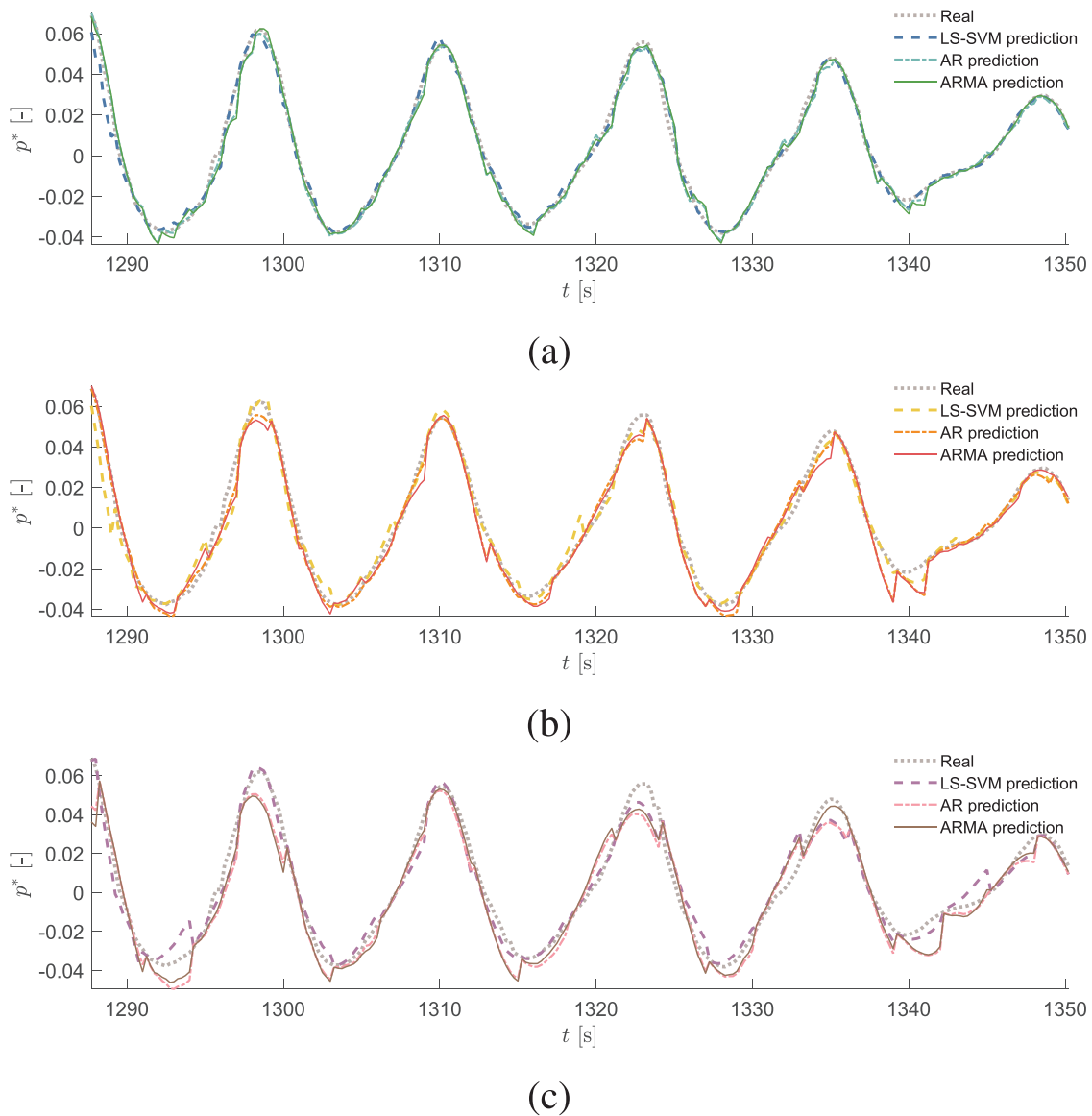
change of forecasting horizon window. Since the error increases with the horizon, we might observe jumps between the last and the first predicted data points, belonging to different forecasting cycles.

Figure 12 reports the computational effort demanded by the different algorithms, in the form of the run time ratios, that is, the elapsed run time divided by the real signal time span. LS-SVM is by far the slowest option, taking a run time over 70 times longer than the signal time span. This is due to the retuning and

retraining of the model every forecasting cycle, which could be simplified. As far as AR and ARMA are concerned, the computational effort can be feasible and is sensible to forecast horizon: the longer the horizon the less the computational expense. This makes sense, since longer horizons imply wider time windows, and consequently, the algorithm would need to run less cycles. Categorically, ARMA models are found to be the most efficient.

The presented results seem to favor AR and ARMA models for a future MPC control strategy.





**FIGURE 11** Pressure signals (segment #2): real, (a) 1-s, (b) 2-s and (c) 3-s ahead predictions

Forecast Horizon [s]	LS-SVM	AR	ARMA
3	73.7	1.23	0.72
2	72.89	0.95	0.44
1	73.41	3.77	1.32

**FIGURE 12** Run time ratio: the elapsed run time divided by the signal time span

## 7 | CONCLUSION

This study was focused on the problem of short-term wave forecasting, which is a key topic in the wave energy field, in

order to improve the economic viability of a WEC. Three algorithms were developed to estimate future values of pressure inside an OWC's air chamber equipped with a biradial turbine, at the Mutriku power plant, namely: LS-SVM, AR and ARMA. The best LS-SVM performing models include fewer features—consequently, with lower computational effort, there are no critical deviations from the real time series, with average errors in the order of 19%. For the cases of AR and ARMA, the prediction error is not significantly impacted by the inclusion of many features, with average errors around 18%. With such low levels of prediction error, these models suggest being valuable for control formulation in the future. Nevertheless, the worst LS-SVM models reveal a significant error increase slightly above 80% when extending the forecasting horizon, due to the scarcity of data for training. As for the signal comparison, prediction errors range from 6.8% to 38.33% and are sensible to the forecast horizon. These show how unreliable the forecast can be

if the model is not carefully tuned and trained. Yet, this is not the case for both AR and ARMA models—the worst models present considerably low errors, in the same order as the best ones.

Hence, in terms of the prediction capability, all three algorithms seem to be good candidates for forecasting in the short term. Still, AR and ARMA deliver better results for both lower and longer forecast horizons. Therefore, the regressive models are preferable for an MPC strategy in a real wave power plant. Ultimately, when compared to LS-SVM, these models are simpler and require less computational effort, with ARMA being the fastest alternative.

## 8 | FUTURE WORK

A robust strategy for this WEC control system is the main forthcoming objective. As the controller will be implemented in the generator in order to adjust the turbine rotational speed to its optimal, there is a need to evaluate the generator efficiency as well. This is a necessary step to quantify the real prediction potential of the OWC predictive controller, in terms of the prediction error, forecasting horizon and computational effort.

Preliminary tests implementing the described models are taking place to assess the improvement that can be achieved. The tests suggest a potential 5% increase in mean turbine power over time, with the biradial turbine. This is possible because the controller acts in advance towards optimal turbine speed, varying the generator power between positive and negative values in order to efficiently decelerate or accelerate the turbine, respectively.

Besides the biradial turbine, further work includes extending the pressure forecasting models to the Wells turbine. In this case, the rotational speed changes the flow rate and, as such, the chamber pressure is affected by the control algorithm—this is not the case for the biradial turbine, introduced in this paper.

## ACKNOWLEDGEMENTS

The first author was supported by (a) Fundação para a Ciência e Tecnologia (FCT) through a PhD Scholarship (SFRH/BD/136521/2018) and (b) a Fulbright Research Grant with the support of FCT for a research stay at the Laboratory for Ship and Platform Flows, at the Massachusetts Institute of Technology (MIT). The research groups at MIT and at Instituto de Engenharia Mecânica/Instituto Superior Técnico (IDMEC/IST) have been awarded an MIT Portugal Seed Funding in 2019 for a joint collaboration, offered by the MIT Portugal Partnership 2030. The IDMEC/IST team was also funded by FCT through IDMEC, under LAETA, project UIDB/50022/2020.

## CONFLICT OF INTEREST

The authors declare no conflict of interests.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the repository: “Air pressure forecasting for the Mutriku

oscillating-water-column wave power plant: Review and case study” at <https://doi.org/10.5281/zenodo.4926029>.

## ORCID

Jorge Marques Silva  <https://orcid.org/0000-0002-0092-0697>

Susana M. Vieira  <https://orcid.org/0000-0001-7961-1004>

Duarte Valério  <https://orcid.org/0000-0001-9388-4308>

João C. C. Henriques  <https://orcid.org/0000-0001-5850-9641>

## REFERENCES

- Falcão, A.F.O., Henriques, J.C.C.: Oscillating-water-column wave energy converters and air turbines: A review. *Renew. Energy* 85, 1391–1424 (2016)
- Chozas, J.F., Kofoed, J.P., Sørensen, H.C.: Predictability and variability of wave and wind. [https://vbn.aau.dk/ws/portalfiles/portal/75674382/Predictability\\_and\\_Variability\\_of\\_Wave\\_and\\_Wind.pdf](https://vbn.aau.dk/ws/portalfiles/portal/75674382/Predictability_and_Variability_of_Wave_and_Wind.pdf). Accessed 15 Nov 2020
- Falcão, A.F.O., Henriques, J.C.C., Gato, L.M.C.: Self-rectifying air turbines for wave energy conversion: A comparative analysis. *Renew. Sustain. Energy Rev.* 91, 1231–1241 (2018)
- Falcão, A.F.O., Gato, L.M.C., Nunes, E.P.A.S.: A novel radial self-rectifying air turbine for use in wave energy converters. Part 2. Results from model testing. *Renew. Energy* 53, 159–164 (2013)
- Henriques, J.C.C., Portillo, J.C.C., Sheng, W., Gato, L.M.C., Falcão, A.F.O.: Dynamics and control of air turbines in oscillating-water-column wave energy converters: Analyses and case study. *Renew. Sustain. Energy Rev.* 112, 571–589 (2019)
- Henriques, J.C.C., Lemos, J.M., Eça, L., Gato, L.M.C., Falcão, A.F.O.: A high-order Discontinuous Galerkin Method with mesh refinement for optimal control. *Automatica* 85, 70–82 (2017)
- Mérigaud, A., Ramos, V., Paparella, F., Ringwood, J.V.: Ocean forecasting for wave energy production. *J. Mar. Res.* 75(3), 459–505 (2017)
- Skene, D., Wolgamot, H., Geldard, J., Taylor, P., Draper, S.: Real-time prediction of unidirectional irregular waves. In: *The 34th International Workshop on Water Waves and Floating Bodies*. University of Bristol, Bristol (2019)
- Fernandes, M.P., Vieira, S.M., Henriques, J.C.C., Valério, D., Gato, L.M.C.: Short-term prediction in an oscillating water column using artificial neural networks. In: *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7. IEEE, Piscataway (2018)
- Peña.Sanchez, Y., Mérigaud, A., Ringwood, J.V.: Short-term forecasting of sea surface elevation for wave energy applications: The autoregressive model revisited. *IEEE J. Oceanic Eng.* 45(2), 462–471 (2020)
- Peña.Sanchez, Y., Ringwood, J.V.: A critical comparison of AR and ARMA models for short-term wave forecasting. In: *Proceedings of the 12th European Wave and Tidal Energy Conference*, pp. 1–6. Wave Energy Centre, Uppsala (2017)
- Shi, S., Patton, R.J., Liu, Y.: Short-term wave forecasting using gaussian process for optimal control of wave energy converters. In: *IFAC-PapersOnLine* 51(29), 44–49 (2018)
- Garcia.Abril, M., Paparella, F., Ringwood, J.V.: Excitation force estimation and forecasting for wave energy applications. *20th IFAC World Congr.* 50(1), 14692–14697 (2017)
- Monk, K., Conley, D., Winands, V., Lopes, M., Zou, Q., Greaves, D.: Simulations and field tests of pneumatic power regulation by valve control using short-term forecasting at the Pico OWC. In: *Proceedings of the 11th European Wave and Tidal Energy Conference*. Wave Energy Centre, Uppsala (2015)
- Paparella, F., Monk, K., Winands, V., Lopes, M.F.P., Conley, D., Ringwood, J.V.: Benefits of up-wave measurements in linear short-term wave forecasting for wave energy applications. In: *2014 IEEE Conference on Control Applications (CCA)*, pp. 2048–2053. IEEE, Piscataway (2014)
- Paparella, F., Monk, K., Winands, V., Lopes, M.F.P., Conley, D., Ringwood, J.V.: Up-wave and autoregressive methods for short-term wave

- forecasting for an oscillating water column. *IEEE Trans. Sustainable Energy* 6(1), 171–178 (2015)
17. Mérigaud, A., Ringwood, J.V.: Incorporating ocean wave spectrum information in short-term free-surface elevation forecasting. *IEEE J. Oceanic Eng.* 44(2), 401–414 (2019)
  18. Sheng, W., Lewis, A.: Short-term prediction of an artificial neural network in an oscillating water column. *Int. Soc. Offshore Polar Eng.* 21(04), 8 (2011)
  19. Fusco, F., Ringwood, J.V.: Short-term wave forecasting for real-time control of wave energy converters. *IEEE Trans. Sustain. Energy* 1(2), 99–106 (2010)
  20. Fusco, F., Ringwood, J.V.: Short-term wave forecasting with AR models in real-time optimal control of wave energy converters. In: 2010 IEEE International Symposium on Industrial Electronics, pp. 2475–2480. IEEE, Piscataway (2010)
  21. Fusco, F.: Short-term wave forecasting as a univariate time series problem. Maynooth University. <https://mural.maynoothuniversity.ie/2282/>. Accessed 15 Nov 2020.
  22. Faedo, N., Peña.Sanchez, Y., Ringwood, J.V.: Receding-Horizon Energy-Maximising Optimal Control of Wave Energy Systems Based on Moments. *IEEE Trans. Sustain. Energy* 12(1), 378–386 (2021)
  23. Henriques, J.C.C., Gato, L.M.C., Falcão, A.F.O., Robles, E., Fay, F.X.: Latching control of a floating oscillating-water-column wave energy converter. *Renew. Energy* 90, 229–241 (2016)
  24. James, S.C., Zhang, Y., O'Donncha, F.: A machine learning framework to forecast wave conditions. *Coastal Eng.* 137, 1–10 (2018)
  25. P, A.C., Deka, P.C.: Application of machine learning techniques in wave height forecasting in marine environment – A review. *Int. J. Innovative Res. Sci., Eng. Technol.* 6(5), 9649–9654 (2017)
  26. Deo, M.C., Kumar, N.K.: Interpolation of wave heights. *Ocean Eng.* 27(9), 907–919 (2000)
  27. Deo, M.C., Jha, A., Chaphekar, A.S., Ravikant, K.: Neural networks for wave forecasting. *Ocean Eng.* 28(7), 889–898 (2001)
  28. Makarynsky, O.: Improving wave predictions with artificial neural networks. *Ocean Eng.* 31(5), 709–724 (2004)
  29. Kalra, R., Deo, M.C., Kumar, R., Agarwal, V.K.: Artificial neural network to translate offshore satellite wave data to coastal locations. *Ocean Eng.* 32(16), 1917–1932 (2005)
  30. Makarynsky, O., Pires.Silva, A.A., Makarynska, D., Ventura.Soares, C.: Artificial neural networks in wave predictions at the west coast of Portugal. *Comput. & Geosci.* 31(4), 415–424 (2005)
  31. Mandal, S., Prabakaran, N.: Ocean wave forecasting using recurrent neural networks. *Ocean Eng.* 33(10), 1401–1410 (2006)
  32. Chen, B.F., Wang, H.D., Chu, C.C.: Wavelet and artificial neural network analyses of tide forecasting and supplement of tides around Taiwan and South China Sea. *Ocean Eng.* 34(16), 2161–2175 (2007)
  33. Deka, P.C., Mandal, S., Prahlada, B.R.: Multiresolution wavelet-ANN model for significant wave height forecasting. In: Proceedings of National Conference on Hydraulics, Water Resources, Coastal and Environmental Engineering - HYDRO 2010, pp. 230–235. Indian Institute of Technology, Mumbai (2010)
  34. Deka, P.C., Prahlada, R.: Discrete wavelet neural network approach in significant wave height forecasting for multistep lead time. *Ocean Eng.* 43, 32–42 (2012)
  35. Dixit, P., Londhe, S., Dandawate, Y.: Removing prediction lag in wave height forecasting using Neuro-Wavelet modeling technique. *Ocean Eng.* 93, 74–83 (2015)
  36. Dixit, P., Londhe, S.: Prediction of extreme wave heights using neuro wavelet technique. *Appl. Ocean Res.* 58, 241–252 (2016)
  37. Hashim, R., Roy, C., Motamedi, S., Shamshirband, S., Petković, D.: Selection of climatic parameters affecting wave height prediction using an enhanced Takagi-Sugeno-based fuzzy methodology. *Renew. Sustain. Energy Rev.* 60, 246–257 (2016)
  38. Stefanakos, C.: Fuzzy time series forecasting of nonstationary wind and wave data. *Ocean Eng.* 121, 1–12 (2016)
  39. Buckles, B.P., Petry, F.: Genetic algorithms. In: IEEE Computer Society Press Technology Series. IEEE Computer Society Press, Washington, DC (1992)
  40. Altunkaynak, A.: Adaptive estimation of wave parameters by Geno-Kalman filtering. *Ocean Eng.* 35(11), 1245–1251 (2008)
  41. Altunkaynak, A.: Sediment load prediction by genetic algorithms. *Adv. Eng. Software* 40(9), 928–934 (2009)
  42. Altunkaynak, A.: Prediction of significant wave height using genomultilayer perceptron. *Ocean Eng.* 58, 144–153 (2013)
  43. Cornejo.Bueno, L., Nieto.Borge, J.C., García.Díaz, P., Rodríguez, G., Salcedo.Sanz, S.: Significant wave height and energy flux prediction for marine energy applications: A grouping genetic algorithm – Extreme Learning Machine approach. *Renew. Energy* 97, 380–389 (2016)
  44. Salcedo.Sanz, S., Nieto.Borge, J.C., Carro.Calvo, L., Cuadra, L., Hesser, K., Alexandre, E.: Significant wave height estimation using SVR algorithms and shadowing information from simulated and real measured X-band radar images of the sea surface. *Ocean Eng.* 101, 244–253 (2015)
  45. Gopinath, D.I., Dwarakish, G.S.: Real-time prediction of waves using neural networks trained by particle swarm optimization. *Int. J. Ocean Clim. Syst.* 7(2), 70–79 (2016)
  46. Reikard, G., Pinson, P., Bidlot, J.R.: Forecasting ocean wave energy: The ECMWF wave model and time series methods. *Ocean Eng.* 38(10), 1089–1099 (2011)
  47. Reikard, G., Robertson, B., Buckham, B., Bidlot, J.R., Hiles, C.: Simulating and forecasting ocean wave energy in Western Canada. *Ocean Eng.* 103, 223–236 (2015)
  48. Reikard, G., Robertson, B., Bidlot, J.R.: Combining wave energy with wind and solar: Short-term forecasting. *Renew. Energy* 81, 442–456 (2015)
  49. Reikard, G.: Forecasting ocean wave energy: Tests of time-series models. *Ocean Eng.* 36(5), 348–356 (2009)
  50. Hadadpour, S., Etemad.Shahidi, A., Kamranzad, B.: Wave energy forecasting using artificial neural networks in the Caspian Sea. *Proc. Inst. Civ. Eng. - Marit. Eng.* 167(1), 42–52 (2014)
  51. Deo, M.C., Sridhar.Naidu, C.: Real time wave forecasting using neural networks. *Ocean Eng.* 26(3), 191–203 (1998)
  52. Zheng, C.W., Wang, Q., Li, C.Y.: An overview of medium- to long-term predictions of global wave energy resources. *Renew. Sustain. Energy Rev.* 79, 1492–1502 (2017)
  53. Zheng, C.W., Gao, Z.S., Liao, Q.F., Pan, J.: Status and prospect of the evaluation of the global wave energy resource. *Recent Pat. Eng.* 10(2), 98–110 (2016)
  54. Zheng, C.W., Pan, J.: Assessment of the global ocean wind energy resource. *Renew. Sustain. Energy Rev.* 33, 382–391 (2014)
  55. Justin, T.T., Barve, K.H., Ranganath, L.R., Dwarakish, G.S.: Assessment of wave energy potential along South Maharashtra coast. *Int. J. Earth Sci. Eng.* 9, 26–31 (2016)
  56. Zheng, C.W., Pan, J., Li, J.x.: Assessing the China Sea wind energy and wave energy resources from 1988 to 2009. *Ocean Eng.* 65, 39–48 (2013)
  57. Zheng, C.W., Zhuang, H., Li, X., Li, X.: Wind energy and wave energy resources assessment in the East China Sea and South China Sea. *Sci. China Technol. Sci.* 55(1), 163–173 (2012)
  58. Samui, P., Kothari, D.P.: Utilization of a least square support vector machine (LSSVM) for slope stability analysis. *Scientia Iranica* 18(1), 53–58 (2011)
  59. Acuña, G., Curilem, M.: Comparison of neural networks and support vector machine dynamic models for state estimation in semiautogenous mills. In: Aguirre, A.H., Borja, R.M., García, C.A.R. (eds.) MICAI 2009: Advances in Artificial Intelligence, pp. 478–487. Springer, Berlin Heidelberg (2009)
  60. Jiang, Y., Liu, S., Zhao, N., Xin, J., Wu, B.: Short-term wind speed prediction using time varying filter-based empirical mode decomposition and group method of data handling-based hybrid model. *Energy Convers. Manage.* 220, 113076 (2020)
  61. Jiang, Y., Zhao, N., Peng, L., Liu, S.: A new hybrid framework for probabilistic wind speed prediction using deep feature selection and multi-error modification. *Energy Convers. Manage.* 199, 111981 (2019)
  62. Mi, X., Liu, H., Li, Y.: Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine. *Energy Convers. Manage.* 180, 196–205 (2019)

63. Yu, C., Li, Y., Bao, Y., Tang, H., Zhai, G.: A novel framework for wind speed prediction based on recurrent neural networks and support vector machine. *Energy Convers. Manage.* 178, 137–145 (2018)
64. Glen, S.: Autoregressive model: Definition & the AR process. <https://www.statisticshowto.com/autoregressive-model/>. Accessed 15 Nov 2020
65. Hyndman, R.J., Athanasopoulos, G.: *Forecasting: Principles and Practice*, 2nd ed. OTexts, Melbourne (2018) <https://otexts.com/fpp2/>
66. Glen, S.: Moving average: What it is and how to calculate it. <https://www.statisticshowto.com/moving-average/>. Accessed 15 Nov 2020
67. Glen, S.: ARMA model. <https://www.statisticshowto.com/arma-model/>. Accessed 15 Nov 2020
68. Vapnik, V.N.: *The Nature of Statistical Learning Theory*. Springer, New York (1995)
69. Cortes, C., Vapnik, V.: Support-vector networks. *Mach. Learn.* 20(3), 273–297 (1995)
70. Hou, L., Yang, Q., An, J.: An improved LSSVM regression algorithm. In: 2009 International Conference on Computational Intelligence and Natural Computing, vol. 2, pp. 138–140. Springer, Cham (2009)
71. Boser, B.E., Guyon, I.M., Vapnik, V.N.: A training algorithm for optimal margin classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory. COLT '92, pp. 144–152. Association for Computing Machinery, Pittsburgh (1992)
72. Zhang, G.: What is the kernel trick? Why is it important? <https://medium.com/@zxr.nju/what-is-the-kernel-trick-why-is-it-important-98a98db0961d>. Accessed 15 Nov 2020
73. Vapnik, V.N.: *Statistical Learning Theory: Adaptive and Learning Systems for Signal Processing, Communications, and Control*. Wiley, New York (1998)
74. Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A., Vapnik, V.: Support vector regression machines. In: Mozer, M.C., Jordan, M., Petsche, T. (eds.) *Advances in Neural Information Processing Systems*, vol. 9, pp. 155–161. MIT Press, Cambridge (1997)
75. Awad, M., Khanna, R.: *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. Apress, Berkeley (2015)
76. Taghavifar, H., Mardani, A.: A comparative trend in forecasting ability of artificial neural networks and regressive support vector machine methodologies for energy dissipation modeling of off-road vehicles. *Energy* 66, 569–576 (2014)
77. Zhao, H.X., Magoulès, F.: Parallel support vector machines applied to the prediction of multiple buildings energy consumption. *J. Algorithms Comput. Technol.* 4(2), 231–249 (2010)
78. Bishop, C.M.: *Pattern recognition and machine learning*. Information Science and Statistics. Springer, New York (2006)
79. Zhao, H.X., Magoulès, F.: Feature selection for support vector regression in the application of building energy prediction. In: 2011 IEEE 9th International Symposium on Applied Machine Intelligence and Informatics (SAMII), pp. 219–223. IEEE, Piscataway (2011)
80. Suykens, J.A.K., Vandewalle, J.: Least squares support vector machine classifiers. *Neural Process Lett.* 9(3), 293–300 (1999)
81. Van Gestel, T., Suykens, J.A.K., Baesens, B., Viaene, S., Vanthienen, J., Dedene, G., et al.: Benchmarking least squares support vector machine classifiers. *Mach. Learn.* 54(1), 5–32 (2004)
82. Evgeniou, T., Pontil, M., Poggio, T.: Regularization networks and support vector machines. *Adv. Comput. Math.* 13(1), 1 (2000)
83. Wahba, G.: *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, Philadelphia (1990)
84. Suykens, J.A.K.: Nonlinear modelling and support vector machines. In: IMTC 2001. IN: Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference. Rediscovering Measurement in the Age of Informatics (Cat. No.01CH 37188). vol. 1, pp. 287–294. IEEE, Piscataway (2001)
85. Vallyon, J., Horváth, G.: Extended Least Squares LS-SVM. *Int. J. Comput. Inf. Eng.* 1(12), 3864–3872 (2007)
86. Wang, H., Hu, D.: Comparison of SVM and LS-SVM for regression. In: 2005 International Conference on Neural Networks and Brain, vol. 1, pp. 279–283. IEEE, Piscataway (2005)

**How to cite this article:** Marques Silva, J., Vieira, S.M., Valério, D., Henriques, J.C.C., Sclavounos, P.D.: Air pressure forecasting for the Mutriku oscillating-water-column wave power plant: Review and case study. *IET Renew. Power Gener.* 15, 3485–3503 (2021). <https://doi.org/10.1049/rpg.2.12289>