Physics-guided vector autoregressive model: application on short term wave forecasting for real-time safety control of an oscillating water column device

Miguel M. Ribeiro
miguelmribeiro@tecnico.ulisboa.pt
Instituto Superior Técnico, Universidade de Lisboa, Portugal
December 2021

Abstract

High energy sea states are valuable for wave energy converters (WEC). However, they pose a significant risk of wrecking devices. To assure safe operations, one requires knowledge of the WEC's future state. This paper introduces the *physics-guided vector autoregressive*(PGVAR) for nonlinear multivariate forecasting problems with known physics. This *physics-guided machine learning*(PGML) model estimates the oscillating water column device's rotation speed, acceleration, and relative chamber pressure purely based on their past history. The model is tested for Biradial and Wells turbines in data recorded during normal and extreme sea states. Other models such as *autoregressive*(AR) and *vector autoregressive*(VAR) are also presented and analysed. Results show an increasing accuracy from AR, VAR until PGVAR, all without needing filters for data. Finally, the hybrid PGVAR model forecasted real observations with enough time and accuracy for safety measures to be implemented in emergencies. Future applications on predictive control require the PGVAR to improve the turbine's efficiency.

Keywords: wave energy, oscillating water column, Wells turbine, Biradial turbine, turbine speed, forecasting, vector autoregressive model, physics-guided vector autoregressive model, physics-guided machine learning

1. Introduction

Wave energy converters (WECs) can bring relevant contributions for a more sustainable energy system around the world [1–3]. Power extraction technologies of ocean waves continue to be an active field of research and development [4]. Furthermore, interest is growing in integrating WECs in synergetic technologies to increment their performance and their economic and technical viability [5].

WECs' expanding applications demand more efficient models. Therefore, intelligent control algorithms emerge with strategies such as controlling oscillating WECs to tune their oscillations with the incident wave elevation [6]. Another strategy is to run the device in energetic sea states with an algorithm that automatically protects the device from hazards. Highly energetic seas might threaten WECs' integrity; hence, survivability strategies need to be applied to minimise structural damage. Author [7] provides insights on survival strategies as, for example, locking the power take-off system (PTO) to cease motion and electricity generation.

1.1. Motivation

The present work arises due to the need of improving WECs' capacity factors. The aim is to assure safe generation of electricity on high energy sea states. This is achieved with Artificial Intelligence (AI) algorithms which can keep devices performing in near limit conditions by automatically controlling safety valves on a wave-by-wave basis. This prolonged activity unlocks opportunities for installing turbines in places where the total annual wave energy is concentrated in short periods. For instance, the Mediterranean Sea exhibits great potential for low-priced Wells turbines when maintaining its services in harsher situations. Although less frequent, energetic waves contain the central portion of the power density in the Mediterranean Sea. As a consequence, it becomes essential to take full advantage of them.

Predictions are required by AI algorithms to regulate the devices. Depending on their purpose, predictions claim different time scales [8]: yearly, daily, hourly ([9]) and wave-by-wave forecasts. Namely, controlling safety valves demands horizons on the order of 1-2 seconds.

Accordingly, the necessity arises for short-term forecasting of a WEC's state in high energy sea states.

1.2. Background

1.2.1. Literature Review

Concerning the model's choice, one could argue that harmonic models would be suitable to forecast waves since these explicitly represent the signal as a sum of sines and cosines. Nevertheless, author [6] tested these and faced many problems arising from the high complexity of the resultant models. His models adapted online to wave's amplitudes and phases while having pre-fixed wave frequencies. Thus, the model's accuracy became highly connected with its capacity to cover the typical range of frequencies as much as possible. The pre-choice of frequencies was a challenging task and therefore did not fit the purpose. Even though harmonic models with variable frequency were tested, these still showed no promising results [6].

Neural Networks (NNs) were tried by, for example, [6, 10, 11]. However, even though NNs generally have a more flexible learning structure than linear models, these models showed no significant improvement when confronted with the added complexity. Plus, NNs have the disadvantage of working as a black-box model and hence cannot provide insights into the characteristics of the modelled process itself.

Another alternative could be Gaussian Processes for their ability to calculate forecasting uncertainties automatically [10]. Still, running the model requires the choice of permanent frequencies. As it was for cyclical models, the choice of fixed frequencies is non-trivial and not practical. Consequently, Gaussian Processes will not be considered.

On the contrary, linear functions of past observations are easier to develop and give intuitive parameters for the model's analysis, albeit they may be unsuitable for deeply nonlinear and complex processes. Moreover, linear models commonly consume less time and data to run when compared to more complex models. Both [6] and [10] propose Autoregressive (AR) models as simple solutions for their practical way of implicitly representing cyclical models. These linear models come in handy by automatically choosing the frequencies and having the online adaptive phase and amplitude capacity, all effortlessly. AR models come from a more prominent family of Autoregressive-Moving-Average (ARMA) models. These have become popular since the 1970s ([12, pp. 277]) as Box-Jenkins [13] methodology often showed better forecast performance than more complex, structural models. [6, 8] tested AR models achieving results with an accuracy of 90%-100% for more than one period ahead by using offline low pass filters for wave elevation forecasting. However, these results are not practical since online filters need to be applied in real life. Later, [14] repeated the experiments and clarified their accuracy was over-optimistic and concluded that non-filtered wave forecasting with AR yields better results than using online filters with AR. The same author also investigated ARMA models in [15]. These offered a comparable accuracy to AR, hence not compensating for the added complexity. AR models were used to control a pressure valve in real-time, optimizing Pico's OWC (in Azores archipelago, Portugal). Consequently, fewer occurrences of turbine stalls and an increase of 15% in power production were perceived when compared with the previous basic control strategy [16].

1.2.2. Mutriku OWC

Oscillating water column device (OWC) is broadly considered as the simplest, most reliable and developed type of WEC [2]. The OWC WEC can be understood as the interaction between the OWC hydrodynamics excited by waves, the air chamber and the PTO sub-system, where the PTO comprises both the turbine and the generator.

As a case study, this paper presents Mutriku's OWC-breakwater WEC. As the first multi-OWC constructed in Europe, it was inaugurated in July 2011 and built in the province of Gipuzkoa, northern Spain. Mutriku's wave power plant contains 16 air chambers that are 4.5 m wide, 4.3 m long and 7.45 m high (above the maximum Equinoctial Spring Tide Low Water). Each chamber is equipped with a Wells turbine of 18.5 kW rated power, summing up to a total installed capacity of 296 kW [17]. This is fed into the electrical grid powering an estimate of 100 homes. The Wells turbine is equipped in most OWC prototypes, and it is the most famous self-rectifying air turbine for wave energy applications [4]. Nonetheless, recently, *H2020 OPERA* project [4] installed a Biradial turbine developed at Instituto Superior Técnico, Lisbon [18, 19]. The Biradial turbine is a novel impulse turbine with peak and time-average efficiencies higher than the Wells turbines [20]. This paper explores the Wells and the Biradial air turbines. Both have a valve installed in series with the rotor for control purposes: that is, Wells has a butterfly valve placed in the duct that unites the turbine to the air chamber; Biradial has a simple axially sliding built-in high-speed safety valve [4]. The turbines' physical model are in 2.3.1.

1.3. Objectives

The objectives of the paper are:

- to introduce the hybrid physics-AI PGVAR model for nonlinear multivariate forecasting problems with known physics;
- to introduce the choice of a signal originated from a system with an inertial response, in this paper, the turbine's speed;
- to forecast the turbine's speed for safety control of oscillating water column wave energy converters (OWC WECs);
- to evaluate the accuracy and field applicability of PGVAR compared to Naive, AR and VAR models;
- to examine the previous models' performance
- to evaluate the use of linear filters;

1.4. Paper Outline

The paper is organized as follows. First, state of the art is presented in the Introduction 1. The following section, Forecasting models 2, comprises a description of the Autoregressive (AR), the Vector Autoregressive models (VAR) and the Physics-Guided Vector Autoregressive models (PGVAR), as well as the tools to operate them. Afterwards, available data are described in Section 3 and the model's implementation procedures in Section 4. Section 5 compares AR, VAR and PGVAR for Biradial turbine operating on the Mutriku wave power plant under highly energetic sea states. Results are reported and discussed. Finally, Section 6 compiles the main results and conclusions are drawn. Further work directions are proposed.

2. Forecasting Models

2.1. AR Model

Autoregressive models (AR) detect patterns on sequential data (such as time series) and estimate future values. It forecasts a variable of interest, assuming it can be obtained by a linear combination of past values of the variable. The autoregressive term refers to the regression of a variable against itself. Namely, the model will forecast 1 step the variable x based on its m number of past values.

$$x_{t+1} = \sum_{i=1}^{m} a_i x_{t-i+1} + \varepsilon_{t+1} \tag{1}$$

where the residual term is given by ε_{t+1} and the a_i coefficients are the model's weights for the linear combination of past values.

Following equation (1), if the a_i coefficients are estimated and residuals are Gaussian and white, the best L-step ahead forecast $\hat{x}(t+L|t)$ at time t is given by the equation:

$$\hat{x}_{t+L|t} = \sum_{i=1}^{m} a_i \hat{x}_{t-i+L|t} \tag{2}$$

where $\hat{x}_{t+L-i|t} \equiv x_{t+L-i}$ if $t+L-i \leq t$ since there is no need to estimate already known information from past values. Equation 2 is implemented running the model L times while using forecasted values as inputs to obtain the succeeding ones.

The performance of each model is measured by the goodness of fit index (GoF). GoF is a classic and intuitive tool for results' comparison. After running the model on test data, GoF outputs the prediction accuracy L time steps ahead as follows:

$$GoF(L) = \left(1 - \frac{\sqrt{\sum_t [x(t+L) - \hat{x}(t+L \mid t)]^2}}{\sqrt{\sum_t x(t+L)^2}}\right) \cdot 100 \tag{3}$$

2.1.1. Confidence intervals

The model's forecasts alone give incomplete information about future events, as they are certainly affected by errors. Hence, it becomes indispensable to estimate the error and confidence intervals to accompany the prediction results. After checking the gaussianity of the errors, the l-step ahead prediction error is given by a Gaussian distribution

$$\hat{\varepsilon}(t+l \mid t) = x(t+l) - \hat{x}(t+l \mid t) \approx \aleph\left(0, \sigma_l^2\right) \tag{4}$$

defined by its variance σ_l^2 or standard deviation σ_l. The variance σ_l^2 is easily estimated from the past errors of N observations through the equation:

$$\hat{\sigma}_l^2 = \frac{1}{N-1}\sum_{t=1}^{N}\hat{\varepsilon}(t+l \mid t)^2 \tag{5}$$

2.2. VAR Model

Vector Autoregressive model (VAR) is a multivariate forecasting model which assumes a single variable can obtained by linearly combining the past values of a group of variables. VAR models resemble AR models with the difference that their variables are vectors and their coefficients are matrices. The model's equation is as follows:

$$X_{t+1} = A_1 X_{t-1+1} + \ldots + A_m X_{t-m+1} + u_{t+1} \tag{6}$$

where X is a vector containing observations of k variables: $X = (x_1, x_2, \ldots, x_k)'$ and A_i is the $(k \times k)$ coefficient matrix that multiplies to the i^{th} lagged value of X, for $i = (1, 2, ..., m)$, in which m is the model's order, i.e., the number of past values required to forecast. Finally, u_{t+1} is the residual vector at the future time instance $(t+1)$ which consists of the difference between the observed X_{t+1} and the estimated $\hat{X}_{t+1}$ [21] .

The estimated L-step ahead forecast of the future $\hat{X}(t+L|t)$ at time t is given by the equation:

$$\hat{X}_{t+L|t} = A_1 X_{t-1+L|t} + \ldots + A_m X_{t-m+L|t} \tag{7}$$

where $\hat{X}_{t+L-i|t} \equiv X_{t+L-i}$ if $t + L - i \leq t$ since past information is already known. Equation 7 is implemented running the model L times while using the forecasted values as inputs to obtain the succeeding ones.

VAR is suitable for multivariate time series as it accounts for the interactions of all model inputs throughout time. Note that, similarly to AR, this method is only used under the assumption of stationary data.

2.3. PGVAR Model

The Physics-Guided Vector Autoregressive model (PGVAR) is a type of physics-guided machine learning (PGML) model used to forecast multivariate series. The PGVAR can obtain a variable by linearly combining the past values of a set of variables and correcting the result through physical equations. This hybrid model is composed of an ML-based VAR stage which feeds its output to a physics-based correction stage, mathematically:

$$\begin{cases} X_{t+1}^{\text{non-pg}} = A_1 X_{t-1+1} + \ldots + A_m X_{t-m+1} + u_{t+1} \\ X_{t+1} = F_{\text{pg}}(X_{t+1}^{\text{non-pg}}) \end{cases} \tag{8}$$

in which X is a vector containing observations of k variables: $X = (x_1, x_2, \ldots, x_k)'$ and A_i is the $(k \times k)$ coefficient matrix that multiplies to the i^{th} lagged value of X, for $i = (1, 2, ..., m)$, where m is the model's order, i.e., the number of past values required to forecast. The vector u_{t+1} is the residual vector at the future time instance $(t+1)$ which consists of the difference between the observed X_{t+1} and the estimated $\hat{X}_{t+1}$. Finally, F_{pg} denotes for the physics-guided function which converts the non-physics-guided forecasted vector $X_{t+1}^{\text{non-pg}}$ into its corrected version X_{t+1}.

The estimated L-step ahead forecast of the future $\hat{X}(t+L|t)$ at time t is given by the equation:

$$\begin{cases} \hat{X}_{t+L|t}^{\text{non-pg}} = A_1 X_{t-1+L|t} + \ldots + A_m X_{t-m+L|t} \\ \hat{X}_{t+L|t} = F_{\text{pg}}(\hat{X}_{t+L|t}^{\text{non-pg}}) \end{cases} \tag{9}$$

where $\hat{X}_{t+L-i|t} \equiv X_{t+L-i}$ if $t+L-i \leq t$ since the past values are already known. Equation 9 is implemented running the model L times while using the corrected forecasted values as inputs to obtain the succeeding ones (see fig 1).

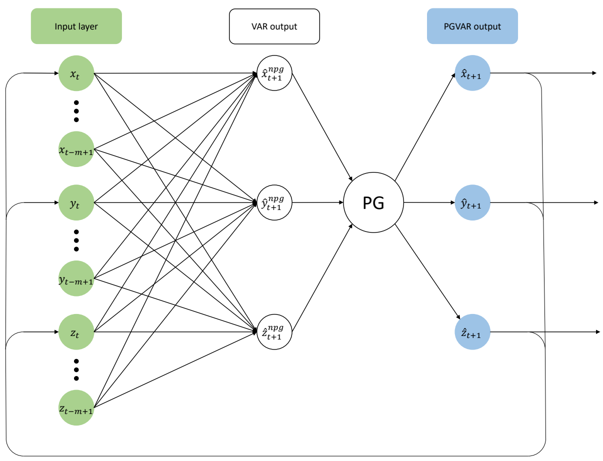

Figure 1: PGVAR(m) scheme

PGVAR model was built as an interpretable grey box model in which all nodes have physical meaning, and their connections resemble real-life connections to take leverage over prior knowledge of physical laws. As the model runs, the PG stage repeatedly corrects and constrains the outputs with predefined linear and nonlinear equations between variables. Consequently, results become more coherent with science while less dependent on training data. Ultimately, the PG stage adds robustness and generalizability to out-of-sample cases.

2.3.1. Physical model utilized

The PG stage implemented is governed by an OWC WEC system taken from [4].

To better understand the power flow from waves to the electrical grid, one can visualize Fig.2.

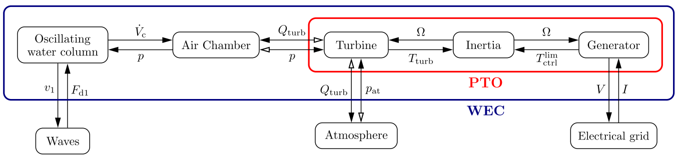

Figure 2: Wave-to-wire power-flow on an OWC wave energy converter. The bidirectional power-flow between the air chamber, the turbine and the atmosphere is represented by double arrows. In the figure, V, I and T_{ctrl}^{lim} stand for voltage, electrical current and generator electromagnetic torque, respectively.

The necessity of forecasting the WEC's state in high energy sea states (mentioned in Section1) demands choosing a variable from Fig.2 to represent the wave-WEC's behaviour. For example, chamber pressure could be chosen, for it is a filtered version of the wave elevation, in which the OWC serves as a low-pass filter. Then, following the same reasoning, the turbine filters the pressure signal and converts it to spinning motion. Consequently, this paper introduces the prediction of the turbine's speed to define the WEC's state.

Finally, the information provided by the forecast is sent to the OWC system's supervisor, which is programmed to make the turbine's rotation speed stay within safety limits. The supervisor can increase the generator torque or close the safety valve in less than 2 seconds for Wells (0.2 seconds for Biradial), thus assuring the turbine's safe operation. Note that air compressibility grants the possibility of closing the high-speed safety valve whenever instructed by the supervisor, but its insights are out of the scope of the paper.

Turbine-generator set. The equation employed in the PG stage supra 9 sets as variables of interest the chamber pressure, turbine speed and turbine acceleration, respectively $X = (p, \Omega, \dot{\Omega})$.

The PGVAR makes use of the VAR model capability to linearly combine and forecast variables of interest. However, VAR poorly predicts the turbine acceleration. Therefore, and also because p and Ω have an associated acceleration state, it was found beneficial to implement the highly nonlinear function that takes as inputs the estimates of $\hat{p}_{t+1}$ and $\hat{\Omega}_{t+1}$ and outputs $\hat{\dot{\Omega}}_{t+1}$:

$$F_{pg} = \begin{cases} \hat{p}_{t+1} = \hat{p}_{t+1}^{\text{non-pg}} \\ \hat{\Omega}_{t+1} = \hat{\Omega}_{t+1}^{\text{non-pg}} \\ \hat{\dot{\Omega}}_{t+1} = F_{pg}^{(3)}(\ \hat{p}_{t+1}^{\text{non-pg}},\ \hat{\Omega}_{t+1}^{\text{non-pg}}) \end{cases} \tag{10}$$

PGVAR model combined equations 9 and 10 where the 1^{st} and 2^{nd} variables are maintained and the 3^{rd} variable is corrected with a physics-guided function $F_{pg}^{(3)}$ (11) composed of several equations taken from [4].

Turbine acceleration is obtained by the torque balance equation between the turbine and the generator:

$$\dot{\Omega} = F_{pg}^{(3)}(p, \Omega) = (T_{turb} - T_{gen}^{em})/I, \tag{11}$$

in which I, T_{turb} and T_{gen}^{em} are respectively the moment of inertia of the rotating parts, the instantaneous turbine aerodynamic torque (13), and the instantaneous generator electromagnetic torque imposed to control the rotational speed (16). The Wells and Biradial turbines installed at Mutriku power plant have a correspondent diameter of 0.75 m and 0.50m, with their in [4].

Turbine system. Following [4], the turbine aerodynamic power P_{turb} can be presented in dimensionless form using the power coefficient Π as function of the dimensionless pressure head Ψ detailed in [4]. Here, p is the stagnation pressure head between the air chamber and the atmosphere, ρ_{in} is the turbine inlet density at stagnation conditions, Ω is rotational speed of the turbine, and D is the rotor diameter.

The reference density, ρ_{in}, is a function of the pressure difference between the atmosphere and the air chamber. The processes of inhalation and exhalation require different approaches to the reference density, as shown below:

$$\rho_{\text{in}} = \begin{cases} (p/p_{at}+1)^{1/\gamma} & \text{if } p \geq 0 \text{ (exhalation)} \\ \varrho_{\text{at}} \,, & \text{if } p < 0 \text{ (inhalation)} \end{cases} \tag{12}$$

Finally, given $P = T \cdot \Omega$, one can obtain the turbine aerodynamic torque:

$$T_{\text{turb}} = \varrho_{\text{in}}\, \Omega^2 d^5 \Pi(\Psi) \tag{13}$$

Generator system. While producing electricity, the generator acts as a controller applying resisting torque. It's power is proportional to the turbine speed (14).

$$P_{\text{gen}}^{\text{opt}} = a\Omega^b \tag{14}$$

In addition, the generator rated power must not be exceeded,

$$P_{\text{gen}}^{\text{em}} = \min\left(P_{\text{gen}}^{\text{opt}}, P_{\text{gen}}^{\text{rated}}\right) \tag{15}$$

Lastly, the electromagnetic torque applied by the generator is trivially obtained,

$$T_{\text{gen}}^{\text{em}} = P_{\text{gen}}^{\text{em}}\, \Omega^{-1} \tag{16}$$

3. Available Data

Even though a generic model was implemented in 2.3.1 for the OWC, for the data, a specific site location had to be chosen. Therefore, the observations were taken from Mutriku's OWC WEC described in Section 1. Among 10 months of data starting on *2017-07-19 18:07:05* and ending on *2018-05-21 07:14:20*, one data sets was chosen:

1. The state is characterized as the most energetic sea state in which the plant was still running. The storm hit the station on *2018-01-17* at around 5am with a significant wave height of $H_s = 5.28\,\text{m}$ and a wave period of $T_e = 13.97\,\text{s}$. Under these conditions, the OWC had a theoretical maximum absorption power of $\hat{P}_{\text{wave}} * L_{\text{w}} = 858.762\,\text{kW}$.This set is of most interest as it is an observed case where the turbine's speed reaches its limits (250 rad/s) without external manipulations such as closing the safety valve. The interaction between the waves and the OWC was recorded at a sampling time of $T_s = 0.2568\,\text{s}$. In total, one hour and 10 minutes (19000 data points) were extracted as the most prolonged consecutive interval of energy generation.

Prepossessing data sets included splitting the data into training data (80%) and testing data (20%), as the machine learning community considers it an optimal proportion.

Raw data was exploited for all results (Section 5) except when it explicitly states the data was filtered. In those cases, Butterworth filters were used to prove the vain application of linear filters. These low-pass filters have a predefined cut-off frequency, for which any frequency above it is mitigated and eventually suppressed from the signal. The filter details are out of the scope of the paper. Nevertheless, the reader is referred to [22].

Waves typical period lies between 7-15 seconds and, since one needs a period of measure of at least 1/20 of the smallest significant period, a minimum frequency of 2,9 Hz is advised for the measurements. Each data set has a generous sampling time T_s of 0.2568 seconds ($\approx$ 4 Hz) so that no information is lost in high sampling times.

The data sets that support the findings of this study was provided within the *EU H2020 OPERA* [4] project and are available in GitHub [23]

4. Implementation

The model was built using python's open source package: Statsmodels [24]. Namely, AR and VAR models developed in [25].

The AR model's regression coefficients are estimated by training the model for a set of data. Firstly, the AR equation is modelled using a state-space approach, and then parameters are estimated via maximum likelihood [26]. This method estimates the parameters combination that yields the maximum likelihood for a 1-step-ahead forecast [25, pp. 33–35].

Python scripts and descriptions can be accessed from the Github repository [23]

5. Results

5.1. Turbine's speed

The $x(t+l)$ is the signal to be forecasted for which x can be turbine's speed (Ω), turbine's acceleration ($\dot{\Omega}$) and pressure (p). And $\hat{x}(t+l/t)$ is the l steps ahead prediction based on information up to instant t. As expected on this type of signal, the predictability decreases over time since uncertainties are always present. A goodness of fit (GoF 3) value of 100%

would mean the signal is perfectly modelled, which is not possible for any prediction into the future ($l > 0$).

The first thing that emerges from Fig.3 is that Ω predictions are significantly better than other variable's predictions. Previous papers in this field focused solely on forecasting wave elevation or wave pressure signals. As far as the author is aware, only [11] used Ω to forecast the chamber pressure, yet the author did not forecast Ω itself. Nevertheless, this paper innovates by introducing a more predictable variable such as the turbine speed to ensure the device's safety. The turbine's inertia works similarly to a low pass filter that mitigates high-frequency disturbances and has a certain delay compared to the pressure signal. Its inertia increases Ω's dependence on its past values, smoothing the signal and thus making it good for forecasting. Additionally, one can take advantage of the fact that the pressure signal, in principle, contains information about the future Ω signal since the latter one is a delayed consequence of the first one (see Fig.2). This is done by using the VAR stage that linearly combines values of pressure to forecast Ω and by the PG stage that non-linearly combines values of pressure and Ω to obtain $\dot{\Omega}$. The acceleration contains information about the future value of the signal and, if integrated, it returns future values of Ω. To extract that information, one uses the linear combination of Ω and $\dot{\Omega}$ values on the VAR stage, which can be thought of as a discrete integration over time for a certain sampling frequency. All these advantages make Ω predictions far better than pressure ones (respectively $85\% >> 45\%$, even for the elementary AR-mle models). Negative values of GoF were converted to $GoF = 0\%$ for their inadequacy. Further analysis of the GoF of $\dot{\Omega}$ suggests it is not optimal. Although it depends on a precise signal (Ω), it is also affected by the uncertainty present in the pressure signal. Pressure's accuracy could rise with a different PGVAR, but focusing on pressure is out of the scope of this paper. Still, Section 6 proposes other works to be done in that direction.

5.2. Models

The following 5 models are compared: Naive, AR with maximum likelihood estimation (AR-mle), AR with ordinary least squares (AR-ols), VAR and PG-VAR models. Firstly, one needs a baseline with an elementary model such as the Naive one. This method simply assumes the future value is the same as the present one $\hat{x}(t+l/t) = x(t)$. Its accuracy rapidly de-

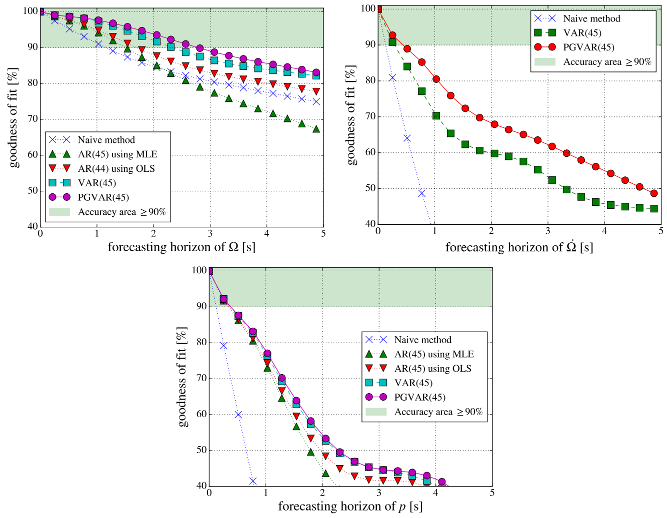

Figure 3: Goodness of fit of different forecasting models on the highly energetic sea state, forecasting: turbine's speed Ω, turbine's acceleration $\dot{\Omega}$ and pressure p.

cays on pressure and $\dot{\Omega}$ since these oscillate relatively quickly and with high amplitudes. The obtained GoF values below 40% were omitted as they mean the prediction diverges. Whereas for Ω predictions, Naive model gives reasonable results ($GoF = 85\%$ for 2 seconds forecast) due to the low amplitude variation of Ω values. This model is handy to create a baseline comparison between Ω forecasts. By outperforming the Naive baseline in Fig.3, the more sophisticated techniques are disclosing casual relationships on the signal. It happens to all models, except for the AR-mle model, which yields worse results than the Naive method for a horizon larger than 2 seconds. Above that, the AR using maximum likelihood estimation to train its parameters (AR-mle) takes significantly more time to run and more data when compared with the other methods in this paper. On the one hand, AR-ols, VAR and PGVAR take some insignificant 3 seconds to train with 1 hour of data and only 6 seconds to forecast 15 minutes of data. On the other hand, AR-mle took 170 and 1.5 seconds, respectively (times were obtained using a 2.80-GHz, 8-core Intel processor). Reversely, AR with ordinary least squares (AR-ols) is a simple model that can get accurate results with less running time and training data compared with AR-mle. It also proves that combining just one variable's past values already gives most of the information the model demands for accurate results ($GoF = 87.6\%$ for 2 seconds forecast of Ω on the energetic sea state). One can notice AR-ols is the only model using 44 past inputs instead of 45. This happened because the software was programmed to automatically choose the number of past values that yielded the best AIC criteria. Although AR-ols has shown to be good, there is a need to have a GoF high enough (say $GoF \geq 90\%$ as a reasonable reference) to have more reliable forecasts. Therefore, instead of just one variable, the VAR model linearly combines different variables to share information between them and, with it, manages to reduce about one-third of Ω's error (from $GoF = 87.6\%$ to 91.7% at 2 seconds of Ω on the energetic sea state). It is also seen that pressure's forecasts become better than AR-ols ones. VAR already showed good results (with Ω's $GoF \geq 90\%$), but one can take a step further and leverage the prior knowledge of physics to correct the predicted variables. This is the property of PGVAR used to improve its accuracy compared with the VAR model (for example, from $GoF = 91.7\%$ to 93.5% at 2 seconds of Ω on the energetic sea state). After having the VAR stage predictions, the following PG stage combines its variables according to known physics. Overall, there was a significant improvement from the common AR-mle to the novel PGVAR model (halving the error on Ω from $GoF = 85.1\%$ to $GoF = 93.5\%$). $\dot{\Omega}$ was needed for VAR and PGVAR models. In Fig.3 there is visible a gain of 10% on GoF from VAR to the PGVAR. In contrast, pressure does not show improvements from VAR to the PGVAR model. This was expected since the PGVAR model was mainly built for Ω.

5.3. Filters

Errors might come from uncertain measures, and thus one could argue that an online filter would solve the problem. However, despite increasing the initial accuracy, online filters cause a delay in the signal.Indeed, applying an AR model to a linearly filtered signal simply yields a different combination of past values when compared to applying an AR to the unfiltered signal. Thus, as it was concluded on [14], in real-life applications, forecasting unfiltered signals with AR models is better than forecasting linearly filtered delayed data. The same is said for the VAR and PGVAR models since they also take their unfiltered inputs and operate some linear combination of past values.

5.4. Discussion

In all experiments, the risk of operating at high energy sea states was undertaken by predicting WEC's reactions to incident waves. For it, the state of the WEC was measured through different variables. Among them, Ω proved to be a better alternative to the classic pressure or wave elevation. The turbine's inertia results in a smooth signal with gradual variations. The models were capable of defining an upper bound of confidence based on the residuals. That bound gives the maximum future values of Ω with an associated certainty. This ended up accomplishing the purpose of indicating if the turbine is at risk of crossing its maximum speed. Biradial turbines fit well for the purpose of forecasting as the regulation of Ω weakly affects the OWC hydrodynamics while still improving the power performance of the PTO (more details in [4]). As for the AR models, training via OLS and MLE yielded different outcomes. The first showed to be much better than the latter as the MLE required much more data and running time to reach the same accuracy as OLS. No benefits were proven upon using MLE. It should also be emphasized that AR models do not necessitate linear filters. This artificial intelligence algorithm is capable of learning from data with errors. Particularly, AR already combines the inputs as a linear filter but without causing any delay. Filters were discarded since online filtering causes a delay on the signal that works in contradiction with forecasts, and their added accuracy does not overcome their delay. AR's simplicity when compared with nonlinear models, along with its considerable accuracy, suggested AR models would fit the purpose (Section 1). Nevertheless, these models are limited to the information of the signal contained within itself. The solution came with VAR models. Following the physical model, several variables were tested. Upon analyzing their contributions to each other's forecasts, it was seen that only Ω, pressure and $\dot{\Omega}$ were influencing each other through the VAR equation's parameters. These weights managed to extract and combine different information present in the three variables. As a result, they worked to improve the accuracy while purely relying on statistical learning. As in most cases physical models exist, a step further was taken to take advantage of that knowledge. PGVAR utilized a physics-guided architecture falling under the category of *PG machine learning* (PGML) models. The PGVAR model was created as the final and most accurate model in this paper. The new PG stage does not need to train its nonlinear relationships; rather, it is predefined by the user based on theoretical models. Consequently, it wasn't trained with extensive data nor much operation time. The

PG stage does not depend on the available data. This feature adds robustness and generalizability to out-of-sample cases. Unlike classical black-box ML algorithms, a physical meaning is ascribed to all nodes on PGVAR, converting the model into a fully interpretable one. One can even explore the parameters of the VAR equation to understand from which lags and variables is the model getting information. To the best of the author's knowledge, the PGVAR model is original and outperforms the state-of-the-art models reducing the error by $\geq 20\%$ for VAR and $\geq 50\%$ for AR and, consequently, shrinking the size of the confidence intervals. Overall, the goodness-of-fit at 2 seconds forecast increased from $\leq 79\%$ in real wave elevation forecasting of previous articles [10, 11, 14], to $\geq 93\%$ in real Ω forecasting while still defining the WEC system's state. Note that other papers that use offline filters for the data have more accuracy than those presented here. However, those ideal filters cannot be applied on the field, and thus their results are not considered. Performance enhancements are achieved in short-term wave forecasting with PGVAR, while still requiring minimal investment in basic equipment, whether for measuring data or for running the model. The hybrid model managed to calculate the turbine's speed 2 seconds ahead with very high precision. These 2 seconds are of utmost importance since the safety valve installed in series with the turbine takes less time to close and protect the turbine from incident waves. Thus, overall, PGVAR emerged among other models as the solution to safely operate a turbine in high energy sea states.

6. Conclusions

This paper introduces the forecasting model PGVAR for multivariate nonlinear problems with known physical phenomenons. PGVAR model takes the artificial intelligence's capability of detecting patterns (VAR) and integrates it with a physics-guided model to account for the known nonlinear relationships (PG). As proof of concept, the PGVAR is evaluated by predicting two data sets with two types of turbines. The objective is to forecast a turbine's speed to assure it does not exceed its limits while relying solely on the system's past history.

6.1. Achievements

Experiments show that VAR and PGVAR perform satisfactorily in problems where only AR was originally tested. As opposed to the classic forecast of wave elevation or wave energy converter's chamber pressure, this paper predicts a much more stable variable, the turbine's speed, while still ensuring the turbine's integrity. To do this, the model outputs a confidence interval containing the future values of the turbine's speed. With it, the system has enough time to close its safety valves in case of an emergency. Furthermore, the PGVAR model demonstrated improvements compared to VAR ($\geq 20\%$ error reduction) and AR ($\geq 50\%$ error reduction) models justifying the added complexity upon building it. PGVAR turned out to be an interpretable model claiming effortless training, low computational costs and low-priced measuring equipment. Additionally, it was concluded that no real benefit could be expected in using linear filters since the AR components of the models already act as filters, enabling uncertainty measurements, yet without the delay caused by classical online filtering. By combining the empirical evidence presented, it can be claimed that PGVAR has higher accuracy than VAR, which in turn performs better than AR in multivariate forecasting problems with known physics. Thus, PGVAR models are a relevant contribution to the state-of-the-art physics-guided machine learning field (PGML). Moreover, the PGVAR is essential for future applications on predictive control to increase the turbine's efficiency.

6.2. Future work

There are several possible directions for further work. Regarding pressure's accuracy, one could increase it by adding up-wave measurements [27] or by replicating what was done in this paper for the turbine's speed. That means building a PGVAR with pressure's derivative obtained from the physical model in [4]. These new variables will upgrade the PGVAR overall performance and optimize the power output [4, 28]. The PG stage could correct the VAR outputs, calculate their derivatives and even compute other variables based on those outputs. Consequently, and as a closing mark, the introduced PG stage can increase the model's interpretability, accuracy, robustness and generalizability to out of sample cases.

Acknowledgements

This work was part of a Master's thesis.

References

[1] A. F. de O. Falcão, Wave energy utilization: A review of the technologies, Renewable and Sustainable Energy Reviews 14 (3) (2010) 899–918. doi:https://doi.org/10.1016/j.rser.2009.11.003. URL https://www.sciencedirect.com/science/article/pii/S1364032109002652

[2] A. F. O. Falcão, J. C. C. Henriques, Oscillating-water-column wave energy converters and air turbines: A review, Renewable Energy 85 (2016) 1391–1424. doi:10.1016/J.RENENE.2015.07.086.

[3] F. Fusco, J. V. Ringwood, A study of the prediction requirements in real-time control of wave energy converters, IEEE Transactions on Sustainable Energy 3 (1) (2012) 176–184. doi:10.1109/TSTE.2011.2170226.

[4] J. C. Henriques, J. C. Portillo, W. Sheng, L. M. Gato, A. F. Falcão, Dynamics and control of air turbines in oscillating-water-column wave energy converters: Analyses and case study, Renewable and Sustainable Energy Reviews 112 (2019) 571–589. doi:10.1016/j.rser.2019.05.010.

[5] D. Clemente, P. Rosa-Santos, F. Taveira-Pinto, On the potential synergies and applications of wave energy converters: A review, Renewable and Sustainable Energy Reviews 135 (2021) 110162. doi:10.1016/J.RSER.2020.110162.

[6] F. Fusco, J. V. Ringwood, Short-term wave forecasting for real-time control of wave energy converters, IEEE Transactions on Sustainable Energy 1 (2) (2010) 99–106. doi:10.1109/TSTE.2010.2047414.

[7] R. G. Coe, Y.-H. Yu, J. V. Rij, A Survey of WEC Reliability, Survival and Design Practices, Energies 2018, Vol. 11, Page 4 11 (1) (2017) 4. doi:10.3390/EN11010004. URL https://www.mdpi.com/1996-1073/11/1/4/htmhttps://www.mdpi.com/1996-1073/11/1/4

[8] A. Mérigaud, V. Ramos, F. Paparella, J. V. Ringwood, Ocean forecasting for wave energy production, THE SEA: THE SCIENCE OF OCEAN PREDICTION Journal of Marine Research 75 (2017) 459–505.

[9] G. Reikard, B. Robertson, J. R. Bidlot, Combining wave energy with wind and solar: Short-term forecasting, Renewable Energy 81 (2015) 442–456. doi:10.1016/J.RENENE.2015.03.032.

[10] S. Shi, R. J. Patton, Y. Liu, Short-term Wave Forecasting using Gaussian Process for Optimal Control of Wave Energy Converters, IFAC-PapersOnLine 51 (29) (2018) 44–49. doi:10.1016/j.ifacol.2018.09.467.

[11] M. P. Fernandes, S. M. Vieira, J. C. C. Henriques, D. Valerio, L. M. Gato, Short-term prediction in an Oscillating Water Column using Artificial Neural Networks, in: Proceedings of the International Joint Conference on Neural Networks, Vol. 2018-July, Institute of Electrical and Electronics Engineers Inc., 2018. doi:10.1109/IJCNN.2018.8489571.

[12] H. Tong, E. J. Hannan, M. Deistler, The Statistical Theory of Linear Systems., Vol. 152, John Wiley and Sons, 1988. doi:10.2307/2983139.

[13] G. Box, G. Jenkins, G. Reinsel, G. Ljung, Time series analysis: forecasting and control, 5th Edition, John Wiley \& Sons, 2015.

[14] Y. Pena-Sanchez, A. Merigaud, J. V. Ringwood, Short-Term Forecasting of Sea Surface Elevation for Wave Energy Applications: The Autoregressive Model Revisited, IEEE Journal of Oceanic Engineering 45 (2) (2018) 462–471. doi:10.1109/JOE.2018.2875575.

[15] Y. Peña-Sanchez, J. V. Ringwood, A Critical Comparison of AR and ARMA Models for Short-term Wave Forecasting, in: Proceedings of the 12th European Wave and Tidal Energy Conference. European Wave and Tidal Energy Conference Series, 2017, pp. 1–6.

[16] K. Monk, Q. Zou, D. Conley, V. Winands, M. Lopes, D. Greaves, Coastal Morphodyanmics View project Tide Structure Interaction View project, in: Proceedings of the 11th European 877 Wave and Tidal Energy Conference. European Wave and Tidal Energy Conference 878 Series, 2015. URL https://www.researchgate.net/publication/281931710

[17] Marine Energies - Ente Vasco de la Energía. URL https://www.eve.eus/Actuaciones/Actuaciones/Marina.aspx

[18] A. F. O. Falcão, L. M. C. Gato, E. P. A. S. Nunes, A novel radial self-rectifying air turbine for use in wave energy converters, Renewable Energy 50 (2013) 289 – 298. doi:10.1016/j.renene.2012.06.050.

[19] A. F. O. Falcão, L. M. C. Gato, E. P. A. S. Nunes, A novel radial self-rectifying air turbine for use in wave energy converters. Part 2. Results from model testing , Renewable Energy 53 (2013) 159 – 164. doi:10.1016/j.renene.2012.11.018.

[20] A. F. O. Falcão, J. C. C. Henriques, L. M. C. Gato, Self-rectifying air turbines for wave energy conversion: A comparative analysis, Renewable and Sustainable Energy Reviews 91 (2018) 1231–1241. doi:10.1016/j.rser.2018.04.019.

[21] H. Lütkepohl, New introduction to multiple time series analysis, Springer Science \& Business Media, 2005. URL https://books.google.com/books?hl=pt-PT{&}lr={&}id=COUFCAAAQBAJ{&}oi=fnd{&}pg=PR4{&}ots=wG3D9sYLHo{&}sig=FQaP5b2WCtE8Dk4sqmV3EC3Ohe8

[22] W. Chen, The electrical engineering handbook, Elsevier, 2004.

[23] joaochenriques (João C. C. Henriques) · GitHub. URL https://github.com/joaochenriques

[24] S. Seabold, J. Perktold, Statsmodels: Econometric and statistical modeling with python, in: 9th Python in Science Conference, 2010. URL https://pypi.org/project/statsmodels/

[25] C. Fulton, Estimating time series models by state space methods in Python : DismalPy (2015).

[26] J. Durbin, S. Koopman, Maximum likelihood estimation of parameters, in: Time Series Analysis by State Space Methods, Oxford University Press, 2013, pp. 170–189. doi:10.1093/acprof:oso/9780199641178.003.0007.

[27] A. Merigaud, J. V. Ringwood, Incorporating Ocean Wave Spectrum Information in Short-Term Free-Surface Elevation Forecasting, IEEE Journal of Oceanic Engineering 44 (2) (2019) 401–414. doi:10.1109/JOE.2018.2822498.

[28] J. C. Henriques, L. M. Gato, A. F. Falcão, E. Robles, F. X. Faÿ, Latching control of a floating oscillating-water-column wave energy converter, Renewable Energy 90 (2016) 229–241. doi:10.1016/J.RENENE.2015.12.065.