

# A Critical Comparison of AR and ARMA Models for Short-term Wave Forecasting

Yerai Peña-Sánchez<sup>#1</sup>, John V. Ringwood<sup>#2</sup>

<sup>#</sup>Centre for Ocean Energy Research, Maynooth University

Maynooth, Co. Kildare, Ireland

<sup>1</sup>Yerai.Pena.2017@mumail.ie

<sup>2</sup>John.Ringwood@nuim.ie

**Abstract**—In order to extract as much energy as possible from ocean waves, an optimal control needs to be implemented in a Wave Energy Converter (WEC), which requires knowledge of the elevation of the future incident waves ( $\eta$ ). One of the most used methods to know future  $\eta$  is to use a linear combination of past  $\eta$  values, for which more than one model can be found in the literature. This paper compares two of these models, Auto Regressive (AR) and Auto Regressive Moving Average (ARMA). Real wave data from different locations is used to determine which model is the best and in which scenario. This comparison addresses the discrepancies between [1], where the ARMA model is discarded for not showing any improvement against AR, and [2], which states that the ARMA model is better than the AR. The paper shows that the two models achieve a similar performance for all the different conditions analysed in this paper. Due to the simplicity and the lower computational requirement, the AR model is the best one.

**Index Terms**—Wave energy, free surface elevation forecasting, autoregressive model, autoregressive moving average model, optimal control

## I. INTRODUCTION

Maximizing the extracted power is instrumental in reducing the cost of wave energy [3]. In order to achieve that maximization of the extracted energy, one of the goals to meet is to apply an optimal control to the WEC [4].

Some the control algorithms for WECs avoid the need for predicting the free surface elevation ( $\eta$ ) or the excitation force ( $F_{ex}$ ) [5]- [6]. However, in general, due to the non-causality of the optimal PTO force, knowledge of future  $\eta$  or  $F_{ex}$  is necessary to implement optimal control to the device [4]- [7]- [8]. In the literature  $F_{ex}$  is usually used as an input to the controller of the WEC due to the simplicity of the problem and the fact that, commonly,  $F_{ex}$  is computed linearly, which makes it predictable. However, in the case of non-linear WECs,  $F_{ex}$  can not be predicted based on its past values so the input to control the WEC must be  $\eta$ . Thus, since  $\eta$  is the only variable which can be used as input to control the WEC in linear and non-linear systems, this paper focuses on the prediction of  $\eta$ .

Two methods are mainly used to forecast the free surface elevation. The first one reconstructs the free surface using one or more measurements taken close to the device [3]- [9]- [10], as shown in Fig. 1(a). Hence, to measure the free surface elevation, more devices may be needed, which means a higher structural cost and a more complex model that takes

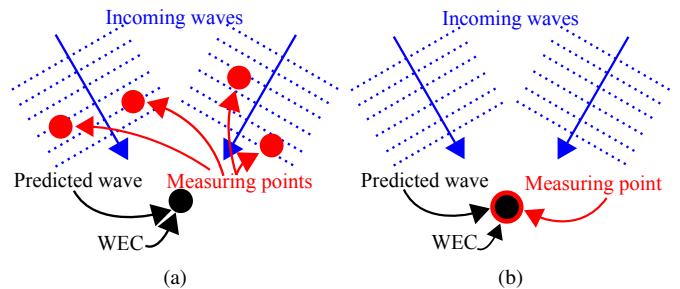


Fig. 1. Two main methods to forecast free surface elevation: space reconstruction of the free surface elevation using distant measurements (a) and forecasting based on the measurements of a single point (b)

into account the multi-directional waves [9], the radiated and diffracted waves, and the non-linear propagation of the waves.

The second forecasting method, illustrated in Fig. 1(b), is based on the measurements of a single point, the position of the device, to forecast the future values of the free surface elevation treating the past values as time series. This second method, does not need additional devices to measure the free surface elevation and does not have to take into account the radiated and diffracted waves or multi-directional waves, which makes it simpler than the first method. A number of wave forecasting models have been developed: Autoregressive (AR) [1]- [11], Autoregressive Moving Average (ARMA) [1]- [2]- [12], Kautz [12], Neural Networks [1] etc. Among the mentioned methods, AR is one of the simplest and, at the same time, accurate [1], in particular for the prediction of the low-frequency waves (swell), which are the most energetic waves.

The ARMA model is tested in [1], and is discarded for showing no improvement compared to the AR method. Conversely an ARMA model is developed in [2] which, compared to an AR model, achieves better results. The main purpose of this paper is compare the two different forecasting models and determine the reason of the conflict between [1] and [2]. This study uses real wave data from 3 different locations in order to determine which of the two methods is the best, or performs better for a given sea state.

## II. COMMON FEATURES

Since the aim of this paper is to compare the AR and ARMA models under the same conditions, there are some processes that are carried out in the same way for the two models, although may not be the most adequate when just one model is to be analysed. An example is the way to identify the coefficients of the models, for which different procedures can be found in the literature [13] as *Yule-Walker*, *Burg*, *Hannan-Rissanen* (HR), etc. There is no method which can be used for both models. Therefore, in this paper, a least square (LS) method is used for the AR model and a variation of the HR algorithm for the ARMA model. Coefficients of the ARMA model cannot be identified using LS, because the future values not only depend on past measurable values, but also on unobserved values. However, as shown in [2], it is possible to use a variation of the HR method based on two LS problems. Thus, coefficients are identified similarly for both models, which enables a fair comparison, minimising the impact of the identification methods in the performance of the models.

Given that the coefficients will be identified solving different LS problems, which minimizes the quadratic error between predicted and real values, the way to measure the accuracy of the prediction should have the same criteria. Thus, a Goodness of Fit (*GoF*) is used to describe the accuracy of the models where the quadratic error between the real value ( $\eta$ ) and the predicted one ( $\hat{\eta}$ ) is computed. Then, the error is normalised against  $\eta$  and the accuracy is showed in percentage:

$$GoF = \left( 1 - \frac{\sqrt{\sum (\eta_{k+T_h} - \hat{\eta}_{k+T_h|k})^2}}{\sqrt{\sum (\eta_{k+T_h})^2}} \right) 100 \quad (1)$$

where  $T_h$  is the prediction horizon and  $\hat{\eta}_{k+T_h|k}$  is the predicted value of  $\eta$  at  $k + T_h$  from the time instant  $k$ .

As explained in Section V, the used wave data records are shorter than 30 minutes. Therefore, sea state has been considered stationary and, thus, the coefficients of the AR and ARMA models are assumed constant.

## III. AUTOREGRESSIVE MODEL

The AR model expresses the value of a new  $\eta$  value based on past  $\eta$  values and is described as follows,

$$\eta_k = \sum_{i=1}^p (\phi_i \eta_{k-i}) + w_k \quad (2)$$

where  $\eta_k$  is the free surface elevation at time instant  $k$ ,  $p$  the number of past values the model is based on,  $\phi_i$  are the autoregression coefficients, or regressors, and  $w_k$  is white noise which is not taken into account. Thus, in order to predict a new value of  $\eta$ , Equation (2) can be rewritten as,

$$\hat{\eta}_{k|k-1} = \eta_{k-1}^* \phi^* \quad (3)$$

where  $\eta_{k-1}^*$  is a vector of previous  $\eta$ ,

$$\eta_{k-1}^* = [\eta_{k-1} \quad \eta_{k-2} \quad \cdots \quad \eta_{k-p}] \quad (4)$$

and  $\phi^*$  is a vector with all the regressors,

$$\phi^* = [\phi_1 \quad \phi_2 \quad \cdots \quad \phi_p]^T \quad (5)$$

The regressors of the AR are identified minimizing the following cost function,

$$J_{\phi^*} = \sum_{i=p+1}^N (\eta_i - \hat{\eta}_{i|i-1})^2 \quad (6)$$

which, is a linear Least Square (LS) problem and where  $N$  is the number of past  $\eta$  values available. The way to identify the coefficients from Equation 6 is,

$$\phi^* = (Z_{1,N-1}^T Z_{1,N-1})^{-1} (Z_{1,N-1}^T K_{p+1,N}) \quad (7)$$

where  $Z_{1,N-1} \in \mathbb{R}^{(N-p-1) \times p}$  is defined as,

$$Z_{1,N-1} = \begin{bmatrix} \eta_p & \eta_{p-1} & \cdots & \eta_2 & \eta_1 \\ \eta_{p+1} & \eta_p & \cdots & \eta_3 & \eta_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \eta_{N-1} & \eta_{N-2} & \cdots & \eta_{N-p+1} & \eta_{N-p} \end{bmatrix} \quad (8)$$

and  $K_{p+1,N} \in \mathbb{R}^{(N-p-1) \times 1}$  as,

$$K_{p+1,N} = [\eta_{p+1} \quad \eta_{p+2} \quad \cdots \quad \eta_{N-1} \quad \eta_N]^T \quad (9)$$

As stated in Section II, the AR coefficients are constant. However, for a longer data set or the real life, a model which updates the AR regressors while sea state changes must be applied. In the literature more than one method to update the coefficients can be found. One possible way is to vary slowly the parameters in each step with a Recursive Least Squares (RLS) as shown in [14]. Another way would be to recalculate the coefficients after some time, which will depend on how fast the sea state changes.

## IV. AUTOREGRESSIVE MOVING AVERAGE MODEL

The ARMA model is not only based on past  $\eta$  values to express a new  $\eta$  value, it also takes into account the past values of the noise as,

$$\eta_k = \sum_{i=1}^p \phi_i \eta_{k-i} + \sum_{i=1}^q \theta_i w_{k-i} + w_k \quad (10)$$

where  $q$  is the order of the Moving Average (MA) part and  $\theta$  are the MA coefficients. As stated in section II, ARMA coefficients can not be identified as a simple LS problem because the noise ( $w$ ) is not measurable. However, as in [2], it is possible to identify the coefficients solving two LS problems.

The first LS problem is the same of the AR model in Equation (7). Since the coefficients obtained from this first LS problem are a first estimation, coefficients are denoted as  $\hat{\phi}$  and the order of the model is  $h$  instead of  $p$ , which is the order of the definitive  $\phi$ . In [13], for the Hannan-Rissanen algorithm, the estimation of  $\hat{\phi}$  is computed using the Yule-Walker algorithm. However, as explained in section II, in order to identify the parameters as similar as possible for

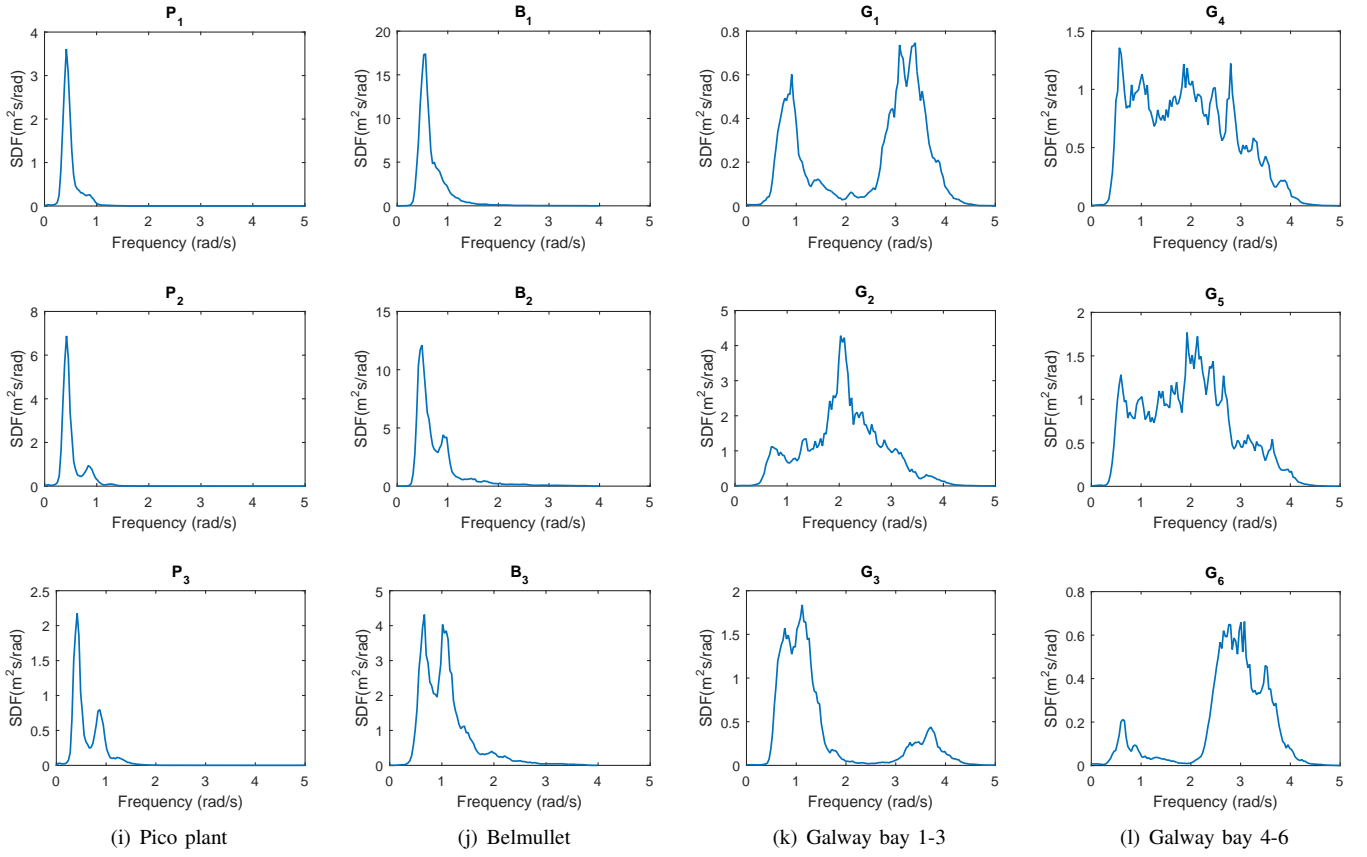


Fig. 2. The spectra used for the comparison

AR and ARMA, the LS problem of Equation (7) is solved to get  $\hat{\phi}$ . Having the parameters, the noise  $w$  can be estimated as follows,

$$\hat{w}_k = \eta_k - \sum_{i=1}^h \eta_{k-i} \hat{\phi}_i \quad (11)$$

Once  $w$  is estimated ( $\hat{w}$ ), the AR and MA parameters are estimated with a second LS problem, minimizing the following cost function,

$$J_{\beta^*} = \sum_{i=r+1}^N (\eta_i - \hat{\eta}_{i|1})^2 \quad (12)$$

where  $r = \max(q, p)$ . Thus,  $\beta^*$ , contains the parameters as,

$$\beta^* = [\phi_1 \cdots \phi_p \quad \theta_1 \cdots \theta_q]^T \quad (13)$$

and is given by,

$$\beta^* = (X^T X)^{-1} (X^T K_{h+r+1, N}) \quad (14)$$

where  $X$  is a matrix which contains  $\eta$  and  $w$  values as,

$$X = \begin{bmatrix} \eta_{h+r} & \cdots & \eta_{h+r-p+1} & \hat{w}_{h+r} & \cdots & \hat{w}_{h+r-q+1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \eta_{N-1} & \cdots & \eta_{N-p} & \hat{w}_{N-1} & \cdots & \hat{w}_{N-q} \end{bmatrix} \quad (15)$$

and  $K_{h+r+1, N}$  is a vector which contains values of  $\eta$  from  $h+r+1$  to  $N$ .

Once the coefficients  $\phi$  and  $\theta$  are determined, as in [2], the Hannan-Rissanen estimation of the noise variance ( $\hat{Q}$ ) is obtained using the following expression,

$$\hat{Q} = \frac{J_{\beta^*}}{N - h - r} \quad (16)$$

In order to predict new  $\eta$  values, a steady state Kalman Filter (KF) is used to obtain the values of the state  $(x_k)_{k=1}^N$ , as in [2]. The state space representation of the ARMA model is,

$$\begin{aligned} \hat{x}_{k+1} &= A\hat{x}_k + Gw_{k+1} \\ y_k &= C\hat{x}_k \end{aligned} \quad (17)$$

where  $A \in \mathbb{R}^{r \times r}$  is the matrix which contains the AR coefficients,  $G \in \mathbb{R}^{r \times 1}$  is the weighting matrix of the noise,  $y_k$  is the output of the system and  $C \in \mathbb{R}^{1 \times r}$  is the matrix which contains the MA coefficients.  $A$ ,  $G$  and  $C$  matrices are,

$$A = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{r-1} & \phi_r \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \quad (18)$$

$$G = [1 \quad 0 \quad \cdots \quad 0]^T \quad (19)$$

$$C = [1 \quad \theta_1 \quad \theta_2 \quad \cdots \quad \theta_{n-1}] \quad (20)$$

where  $\phi_i = 0$  for  $i > p$  and  $\theta_i = 0$  for  $i > q$ . As in [2], the error covariance  $P_\infty$  and the Kalman gain  $L_\infty$ , which are constant, are obtained as,

$$\begin{aligned} P_\infty &= AP_\infty A^T + GQG^T - AP_\infty C^T (CP_\infty C^T)^{-1} CP_\infty A^T \\ L_\infty &= (P_\infty C^T) (CP_\infty C^T)^{-1} \end{aligned} \quad (21)$$

Thus, the KF Time Update (TU) and Measurement Update (MU) are composed by two equations. In TU the state is estimated and when a new measurement is available, in MU, the state is updated using  $L_\infty$  as,

- Time Update:

$$\hat{x}_k^- = A\hat{x}_{k-1} \quad (22)$$

- Measurement Update:

$$\hat{x}_k = \hat{x}_{k-1}^- + L_\infty(y_k - C\hat{x}_{k-1}^-) \quad (23)$$

Once the state  $\hat{x}$  is determined, it is possible to recursively predict future observations  $\bar{\eta}_{k+T_h|k}$  as,

$$\bar{\eta}_{k+T_h|k} = CA^{T_h}\hat{x}_k \quad (24)$$

## V. WAVE DATA

Real sea data from three different places have been used in this paper:

- **Pico plant**, which is located in Pico island, approximately in (38.56, -28.45), in the Azores. This data has been recorded using an Aquadopp 1001 with a frequency of 2Hz. Data consist of 30 minutes sets recorded continually. This data has been obtained from the Irish Marine Institute.
- **Belmullet** is located in the west coast of Ireland, approximately (54.27, -10.28). Data consist of 30 minutes sets recorded continually with a Waverider buoy with a 2.56Hz frequency.
- **Galway Bay**, west coast of Ireland, data has been obtained from the Irish Marine Institute. This data consist of 20-minute records for each hour measuring values with a frequency of 2.56Hz recorder with a Waverider buoy. The buoy is located at, approximately (53.23, -9.26), where the water depth is, approximately, 20m.

The data fragments have been selected attempting to an implementation of different sea states, for which Fig. 2 shows the spectra. The spectra of Pico plant wave data are narrow banded at low-frequencies, showing that is mainly swell wave, with no high-frequency component. The spectra of the data from Belmullet, are also narrow banded at low frequencies, but they also have high-frequency components. Finally, Galway Bay wave data is mostly composed by wind waves, which can not be used to extract power by WECs. However, this data has been selected in order to try the models in different scenarios.

To create the smoothed spectra showed in Fig. 2, the data have been processed as explained in [15].

## VI. RESULTS

In [2] the prediction horizon ( $T_h$ ) used is 4s due to the fact of using half of the typical wave period of the spectra. However, as data with different spectra is used in this paper, the typical wave period is not the same for all spectra, thus the results are showed as a function of  $T_h$ .

To determine the order of the models, different combinations are compared as in Fig. 3-4, where the accuracies of one step ahead predictions are compared and, then, the model order for the highest accuracy is selected.

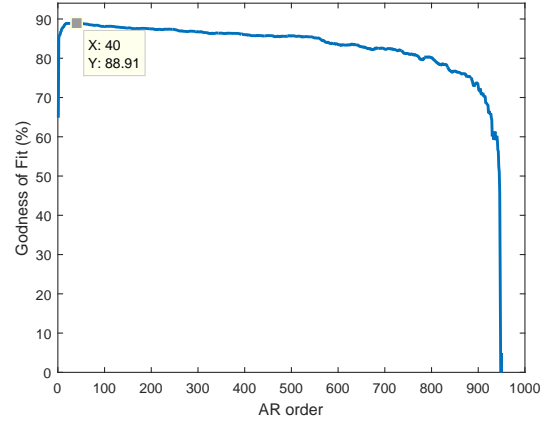


Fig. 3. The variation of the  $GoF$  of the AR model for the third Pico plant wave data set when the order of the model increases

In Fig. 3 it is shown how the  $GoF$  of an AR model changes when the order of the model increases. As explained in Equation 6, the coefficients are identified minimizing the error of the prediction in one step. Thus, as Fig. 3 illustrates, the model order has also been chosen comparing one step ahead predictions'  $GoF$ .

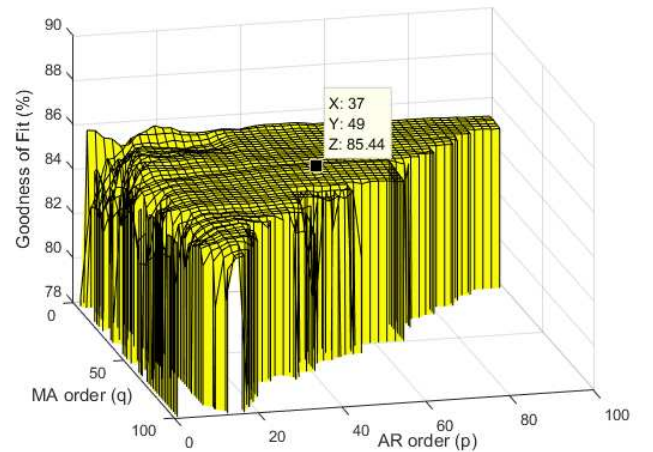


Fig. 4. The variation of the  $GoF$  of the ARMA model for the third Pico plant wave data set with different  $p$  and  $q$  order combinations

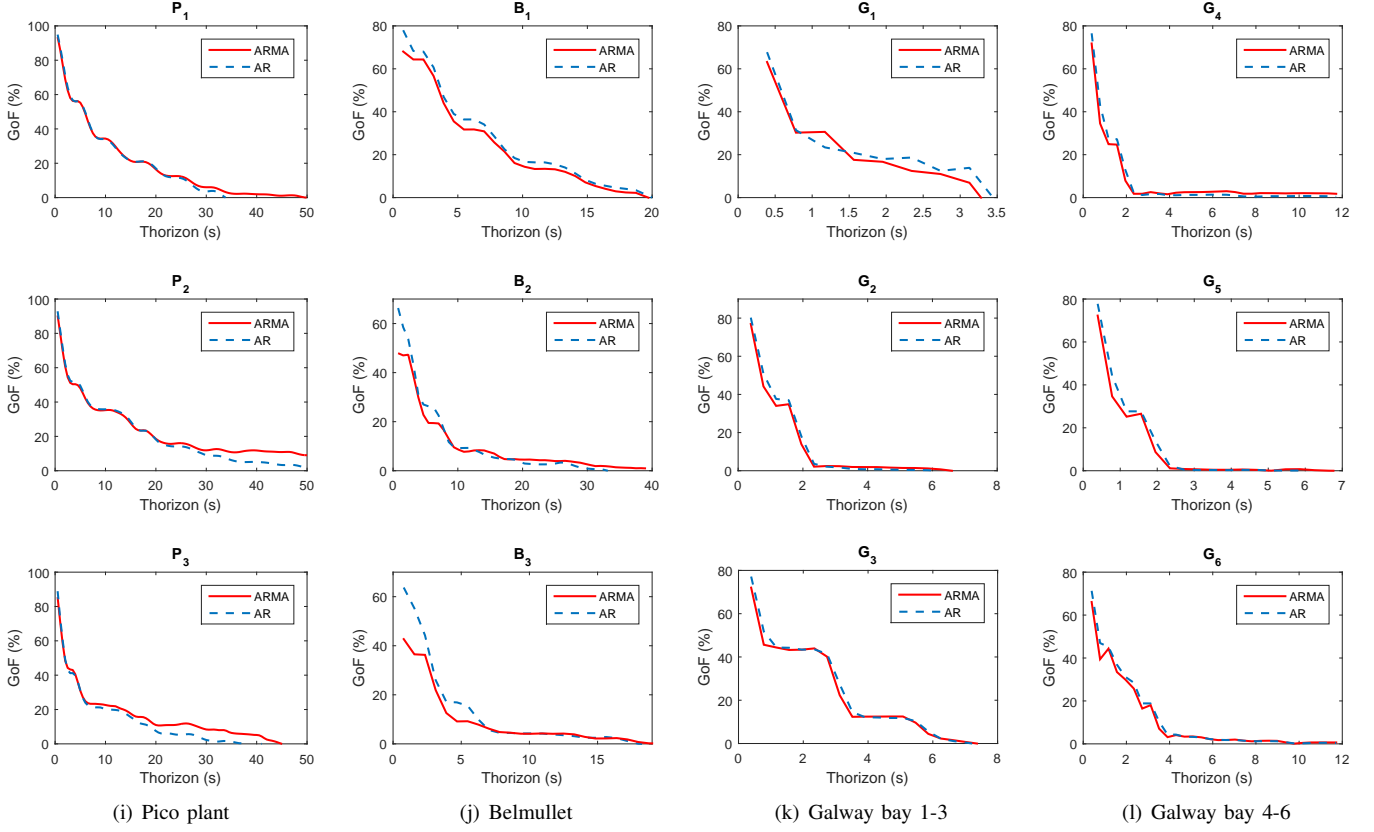


Fig. 5. The  $GoF$  for different  $T_h$  in all data sets

The impact of different combinations of  $p$  and  $q$  orders in the ARMA model on the  $GoF$  is shown in Fig. 4. As explained for the AR model, the order has been chosen comparing the  $GoF$  of the one step ahead prediction of different  $p$  and  $q$  order combinations. The order of the first AR estimation ( $h$ ) must be higher than the order of  $r$  [13], and it has been proven that the variation of  $h$  does not affect the  $GoF$  (as long as it is higher than  $r$ ), so  $h = 110$  have been used for all the data sets.

Since just one LS problem is solved and the order of the model is usually lower, the process of identifying the parameters is faster in the AR model than in the ARMA model. Furthermore, the prediction process of the AR model is more simple because the regressors can be directly used on the past  $\eta$  values to predict new values while in the ARMA model a KF is needed to obtain the state  $x$  in order to predict new  $\eta$  values. However, once the regressors and the state  $x$  are identified, the ARMA model predicts a new  $\eta$  value using just one equation (Equation (24)), while the AR model needs a loop whose size depends on  $T_h$ . Thus, while the time needed by the AR model to predict a new value increases when  $T_h$  increases, as shown in Fig. 6, the variation of  $T_h$  does not affect considerably on the time needed by the ARMA model to predict a new value. Fig. 6 illustrates the time needed for each model to predict  $\eta$  for the third data set of Pico plant with a model order of  $p = 40$  in the case of the AR and  $p = 49$  and  $q = 37$  for the ARMA.

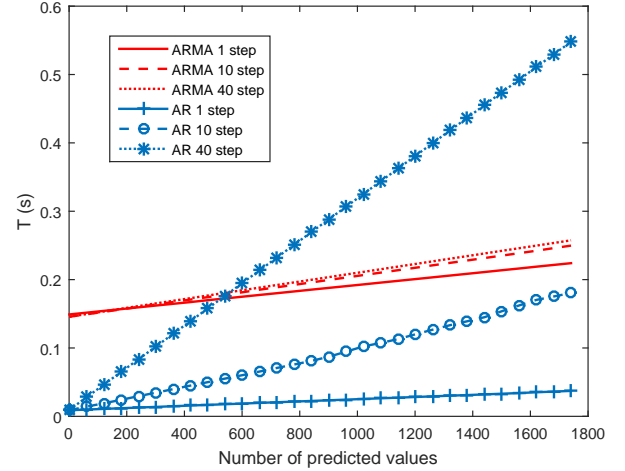


Fig. 6. The time needed to predict, when  $T_h$  changes

One can notice that, while AR only has  $p$ , the ARMA model combines  $p$  and  $q$  which, as it is shown in Fig. 4, makes the model order selection more complex than in the case of the AR model. As it can be concluded from Table I, the AR model order is up to 100, while the ARMA model order is between 1 and 60 for  $p$  and 1 and 40 for  $q$ . Apart from the number of possible combinations, as explained before in this section, the ARMA model needs more time to identify the parameters, which need to be identified for each different combination.

Hence, while approximately 3 seconds are required to generate a graph like Fig. 3, whose  $p$  goes from 1 to 100, around 12 minutes are needed to obtain a graph like Fig. 4, whose  $p$  goes from 1 to 60 and  $q$  goes from 1 to 40.

The AR and ARMA model orders and the one step ahead  $GoF$  are shown in Table I, where the one step prediction is shown to be always more accurate using AR than ARMA.

TABLE I  
MODEL ORDER AND ONE STEP AHEAD PREDICTION  $GoF$  OF ALL DATA SETS

		AR		ARMA			
		$p$	$GoF$	$h$	$p$	$q$	$GoF$
Pico	1	89	94.94	110	58	29	93.47
	2	40	92.86	110	14	21	90.42
	3	40	88.92	110	49	37	85.44
Belmullet	1	62	77.98	110	46	10	68.06
	2	66	66.34	110	22	16	48.39
	3	55	63.73	110	16	7	43.37
Galway	1	7	67.76	110	7	4	63.21
	2	38	80.25	110	7	16	76.92
	3	41	77.16	110	28	37	72.05
	4	42	73.40	110	19	25	71.87
	5	26	77.12	110	19	16	72.32
	6	31	71.21	110	16	13	66.24

Fig. 5 shows how the  $GoF$  of each data set changes when  $T_h$  increases. One can notice that the two models achieve a similar performance. In the Pico plant data sets, both models achieve the same performance when  $T_h < 20s$ . However, for  $T_h > 20s$ , ARMA model achieves a slightly better  $GoF$  comparing to the AR. In the data sets from Belmullet, AR achieves, generally, a slightly better  $GoF$  than the ARMA model. And finally, in the Galway Bay data sets, both models have a similar performance, where the AR model achieves a light improvement on the prediction.

Generally, it is fair to say that both models achieve a similar performance, which is not surprising due to the fact that the ARMA model is a sum of AR model and a noise which is estimated using another AR model.

## VII. CONCLUSIONS

The main purpose of this paper is to solve the conflict between [1], which discards the ARMA model for not showing any improvement against AR, and [2], which states that the ARMA model achieves a more accurate prediction than AR. It has been deemed convenient to try the ARMA model explained in [2] because, the process to obtain the regressors is not described in [1], maybe the process explained in [2] is better than the one used in [1]. Nevertheless, it has been proven that the ARMA model explained in [2] does not show any improvement comparing to an AR model of similar characteristics.

The conclusion obtained in [2] is probably due to the fact that the AR model used updates its coefficients in time, which may not be the best option due to the length of the used data sets. Furthermore, the ARMA model used in [2] keeps the parameters constant, so it is not fair to compare it to a model which updates them. Because of that, in this paper both models maintain the regressors constant. To ensure that the ARMA model is the same as in [2], the initial idea of this paper was using the same data used in [2] and compare the results but, unfortunately, that data was not available.

One can notice from Section III and IV, that the ARMA model is more complex than the AR and, from Section VI, that the accuracies obtained using ARMA are not better than the ones obtained from AR. Therefore, the conclusion of [1], which says that the ARMA does not achieve any improvement comparing to the AR, can be confirmed.

## REFERENCES

- [1] F. Fusco and J. Ringwood, "Short-Term Wave Forecasting for Real-Time Control of Wave Energy Converters," *IEEE Transactions on Sustainable Energy*, vol. 1, no. 2, pp. 99–106, 2010.
- [2] E. C. Kerrigan, "Short-term Ocean Wave Forecasting Using an Autoregressive Moving Average Model," 2008.
- [3] K. Budal and J. Falnes, "Wave power conversion by point absorbers: A Norwegian project," *International Journal of Ambient Energy*, vol. 3, no. 2, pp. 59–67, 1982.
- [4] J. V. Ringwood, G. Bacelli, and F. Fusco, "Energy-maximizing control of wave-energy converters: The development of control system technology to optimize their operation," *IEEE Control Systems*, vol. 34, no. 5, pp. 30–55, 2014.
- [5] F. Fusco and J. V. Ringwood, "A simple and effective real-time controller for wave energy converters," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 1, pp. 21–30, 2013.
- [6] J. T. Scruggs and S. Lattanzio, "Optimal Causal Control of an Ocean Wave Energy Converter in Stochastic Waves," *9th EWTEC -European Wave and Tidal Energy Conference*, p. CDR0M paper #414, 2011.
- [7] G. Li, G. Weiss, M. Mueller, S. Townley, and M. R. Belmont, "Wave energy converter control by wave prediction and dynamic programming," *Renewable Energy*, vol. 48, pp. 392–403, 2012.
- [8] G. Bacelli, P. Balitsky, and J. V. Ringwood, "Coordinated Control of Arrays of Wave Energy Devices Benefits Over Independent Control," *IEEE Transactions on Sustainable Energy*, pp. 1–9, 2013.
- [9] M. R. Belmont, J. M. K. Horwood, R. W. F. Thurley, and J. Baker, "Filters for linear sea-wave prediction," *Ocean Engineering*, vol. 33, no. 17-18, pp. 2332–2351, 2006.
- [10] J. Tedd and P. Frigaard, "Short term wave forecasting, using digital filters, for improved control of Wave Energy Converters," in *Proceedings of the International Offshore and Polar Engineering Conference*, 2007, pp. 388–394.
- [11] F. Paparella, K. Monk, V. Winands, M. F. P. Lopes, D. Conley, and J. V. Ringwood, "Benefits of up-wave measurements in linear short-term wave forecasting for wave energy applications," *2014 IEEE Conference on Control Applications, CCA 2014*, pp. 2048–2053, 2014.
- [12] M. P. Schoen, J. Hals, and T. Moan, "Wave prediction and robust control of heaving wave energy devices for irregular waves," *IEEE Transactions on Energy Conversion*, vol. 26, no. 2, pp. 627–638, 2011.
- [13] P. J. Brockwell and R. a. Davis, *Introduction to Time Series and Forecasting*, Second Edition, 2002.
- [14] F. Fusco, "Short-term Wave Forecasting as a Univariate Time Series Problem," no. December, 2009.
- [15] "Datawell Waverider Installation Guide Datawell BV oceanographic instruments," pp. 1–8, 2012.