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

Yerai Peña-Sanchez[#1], John V. Ringwood[#2]

[#]Centre for Ocean Energy Research, Maynooth University

Maynooth, Co. Kildare, Ireland

[1]Yerai.Pena.2017@mumail.ie

[2]John.Ringwood@nuim.ie

Abstract—In order to extract as much energy as possible from ocean waves, an optimal control must be implemented in a wave energy converter (WEC), which requires the knowledge of the future incident waves (η). One of the most used methods to predict the future η, is to use a linear combination of past η values. Several models can be found in the literature, but only two of these models are compared in this paper, the autoregressive (AR) and autoregressive moving average (ARMA) models. 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 showing no improvement against the AR, and [2], which states that the ARMA model does improve the AR. The present paper shows that the two models achieve a similar performance for all the different conditions analysed.Thus, due to the simplicity and the lower computational requirement, the AR model is chosen as the best model for prediction.

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

I. INTRODUCTION

Maximising the extracted power is crucial in reducing the cost of wave energy [3]. To maximise the extracted energy, control can be inplemented in the wave energy converter (WEC), where only optimal control ensures that all the possible energy has been harvested from ocean waves [4].

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

Two methods can be used to forecast the free surface elevation. The first method 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 into account the multi-directional waves [9], the radiated and diffracted waves, and the non-linear propagation of the waves.
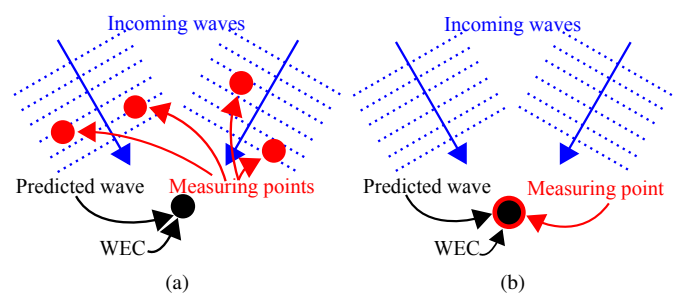

Fig. 1. Two main methods to forecast free surface elevation: space reconstruction of the free surface elevation using distant measurements (a) and forecasting based on the measurements of a single point (b)

The second forecasting method, illustrated in Fig. 1(b), is based on the measurements of a single point 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 in the literature: The autoregressive (AR) [1]- [11], the autoregressive moving average (ARMA) [1]- [2]- [12], the Kautz method [12], Neural Networks [1], etc. Among the mentioned methods, AR is one of the simplest and, at the same time, accurate [1] method. 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] where, compared to an AR model, achieves better results. Given the inconsistency of the results in the papers [1] and [2], the main purpose of this paper is to compare the two different forecasting models and determine the advantages and disadvantages of the two models under some conditions. This study uses real wave data from 3 different locations.

The remainder of the paper is organised as follows: in Section II the considerations taken into account to make a fair comparison are introduced. Section 2 describes the autoregressive model, while Section 10 presents the autoregressive moving average model. Section V shows the data used in the comparison. Finally, Section VI, compares the performances of the two forecasting models, and conclusions are drawn in Section VII.

II. COMMON FEATURES

Since the aim of this paper is to compare the AR and ARMA models under the same conditions, a number of processes must be carried out in the same way for the two models. For example, the way the model coefficients are identified. Different procedures can be found in the literature [13] such as *Yule-Walker*, *Burg*, *Hannan-Rissanen* (HR), etc methods. Unfortunately, no method can be used for both models. Therefore, a least square (LS) method is used for the AR model and a variation of the HR algorithm for the ARMA model in this paper. Coefficients of the ARMA model cannot be identified using LS, because the future values not only depend on past measurable values of η, but also on unobserved values (i.e. white noise). 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 method in the performance of the models.

Given that the coefficients are identified solving different LS problems, which minimizing 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 (η) and the predicted one ($\hat{\eta}$) is computed. Then, the error is normalised against η and the accuracy is showed in percentage:

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

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

The wave data records used in this paper are shorter than 30 minutes. Therefore, sea states have 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 η value based on past η, and is described as follows,

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

where η_k is the free surface elevation at time instant k, p the number of past values the model is based on, ϕ_i the regression

coefficients and w_k the white noise. Thus, in order to predict a new value of η, Equation (2) can be rewritten as,

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

where η_{k-1}^* is a vector of previous η values as as follows,

$$\eta_{k-1}^* = \begin{bmatrix} \eta_{k-1} & \eta_{k-2} & \cdots & \eta_{k-p} \end{bmatrix} \qquad (4)$$

and ϕ^* a vector containing all the regression coefficients as,

$$\phi^* = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_p \end{bmatrix}^T \qquad (5)$$

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

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

which, is a linear LS problem and where N is the number of past η values available. Coefficients from Equation 6 are identified as follows,

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

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

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

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

$$K_{p+1,N} = \begin{bmatrix} \eta_{p+1} & \eta_{p+2} & \cdots & \eta_{N-1} & \eta_N \end{bmatrix}^T \qquad (9)$$

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 coefficients where sea state changes must be implemented. In the literature several methods to update the coefficients can be found. One possible method is to vary slowly the parameters in each step with a Recursive Least Squares (RLS) [14]. Alternatively, coefficients can be recalculated every time the sea-state varies.

IV. AUTOREGRESSIVE MOVING AVERAGE MODEL

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

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

where q is the order of the moving average (MA) part and θ the coefficients to combine the past noise values. As stated in section II, ARMA coefficients can not be identified as a simple LS problem because the noise (w) is not measurable. However it is possible to identify the coefficients solving two LS problems [2].

The first LS problem is the same as in Equation (7) of the AR model. 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. The order of the definitive ϕ is defined as p. In [13], as part of the Hannnan-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 similarly in both models, the LS problem of Equation (??) is solved to get $\hat{\phi}$. Once parameters are determined, the noise w can be estimated as follows,

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

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

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

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

$$\beta^* = \begin{bmatrix} \phi_1 \cdots \phi_p & \theta_1 \cdots \theta_q \end{bmatrix}^T \tag{13}$$

and is given by,

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

where X is a matrix which contains past η and $\hat{w}$ values as,

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

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

Finally, the Hannan-Rissanen estimation of the noise variance ($\hat{Q}$) is obtained using the following expression,

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

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

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

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

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

$$G = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T \tag{19}$$

$$C = \begin{bmatrix} 1 & \theta_1 & \theta_2 & \cdots & \theta_{n-1} \end{bmatrix} \tag{20}$$

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

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

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

- Time update:

$$\hat{x}_k^- = A\hat{x}_{k-1} \tag{22}$$

- Measurement update:

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

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

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

V. WAVE DATA

Real sea data from three different locations 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.
- **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 1.28Hz frequency. This data has been obtained from the Irish Marine Institute.
- **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 component. Finally, Galway Bay wave data is mostly composed by wind waves, which can not be used to extract power by WECs, but has been selected in order to try the models in different conditions.

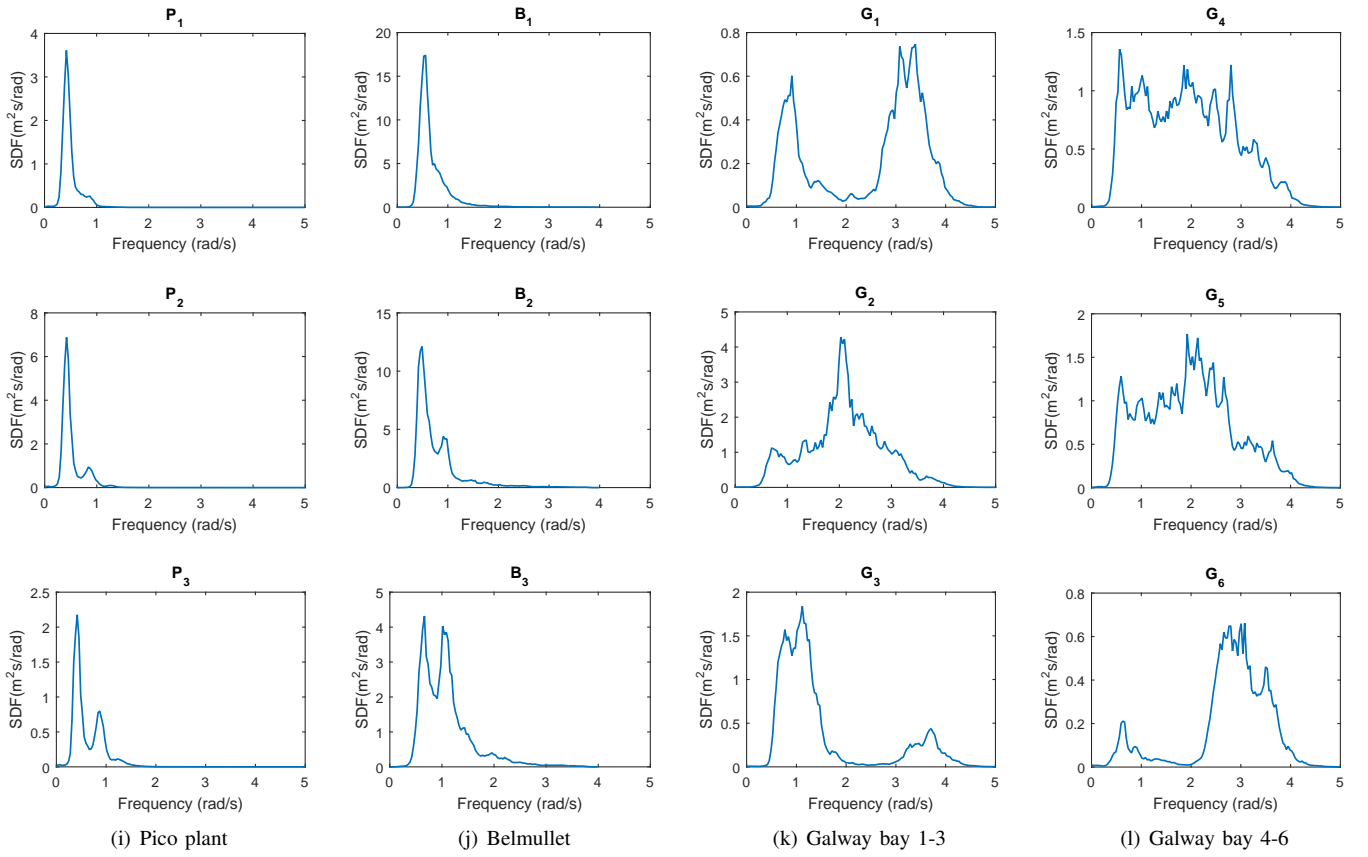To create the smoothed spectra showed in Fig. 2, the data have been processed as explained in [15].

(i) Pico plant   (j) Belmullet   (k) Galway bay 1-3   (l) Galway bay 4-6

Fig. 2. The spectra used for the comparison

VI. RESULTS

In [2], the prediction horizon (T_h) is set to $4s$, determined to be half of the typical wave period of the used 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.
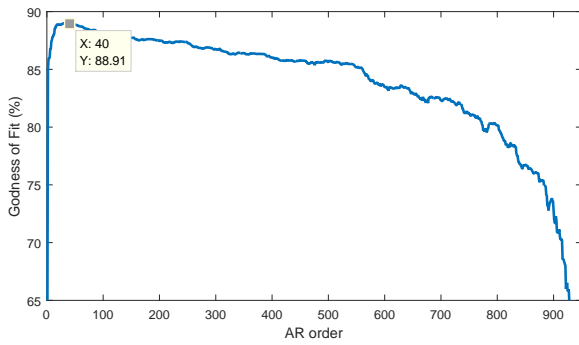

Fig. 3. The variation of the GoF of the AR model for the third Pico plant wave data set when the order of the model increases

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. Then, the model order for the highest accuracy is selected.

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

The impact of different combinations of p and q orders on the GoF of the ARMA model 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 simpler because the regression coefficients can be directly used on the past η values to predict new values while, in the ARMA model, a KF is needed to obtain the state x in order to predict new η values. In particular, for the first data set from Pico plant, where the order is similar for both models (89 for the AR and 58+29 for the ARMA), only 0.01s are needed by
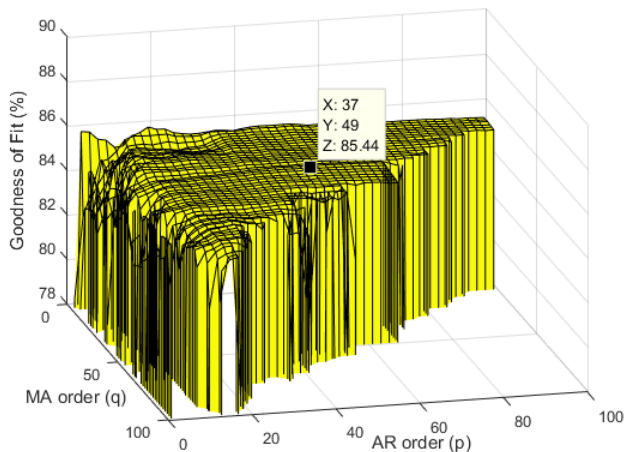
Fig. 4. The variation of the GoF of the ARMA model for the third Pico plant wave data set with different p and q order combinations

the AR to identify the parameters while, 0.25s are needed by the ARMA. Once the regression coefficients and the state x are identified, the ARMA model predicts a new η value using just one equation (Equation (24)), while the AR model needs a equation loop, whose size depends on T_h. However, as it is shown in Fig. 5, where the time needed to predict a new η value is shown, the ARMA model is not faster than the AR.
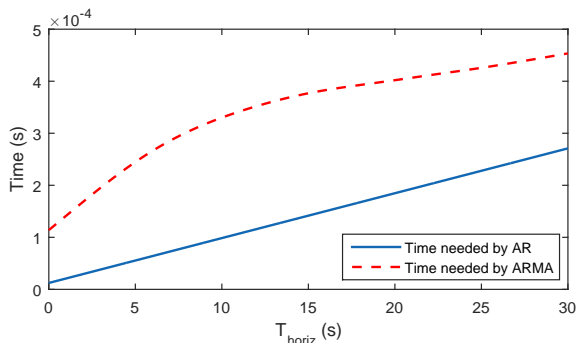

Fig. 5. The time needed for the prediction, for different T_h

One can notice that, while the AR order only depends on p, the ARMA model order depends on p and q which, as it is shown in Fig. 4, makes the 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 used 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. 6 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 an slightly better GoF comparing to the AR. In Belmullet data sets, AR achieves, generally, a slightly better GoF than the ARMA model. Finally, in the Galway Bay data sets, both models have a similar performance, where the AR model achieves slightly better predictions.

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], in case the process to obtain the regression coefficients explained in [2] is better than the one used in [1]. Nevertheless, it has been demonstrated 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 coefficients are updated in time, which may not be the best option due to the length of the used data sets. In contrast, the ARMA model in [2] keeps the parameters constant, so it is not fair to compare it to a model which updates the coefficients. Therefore, in this paper both models
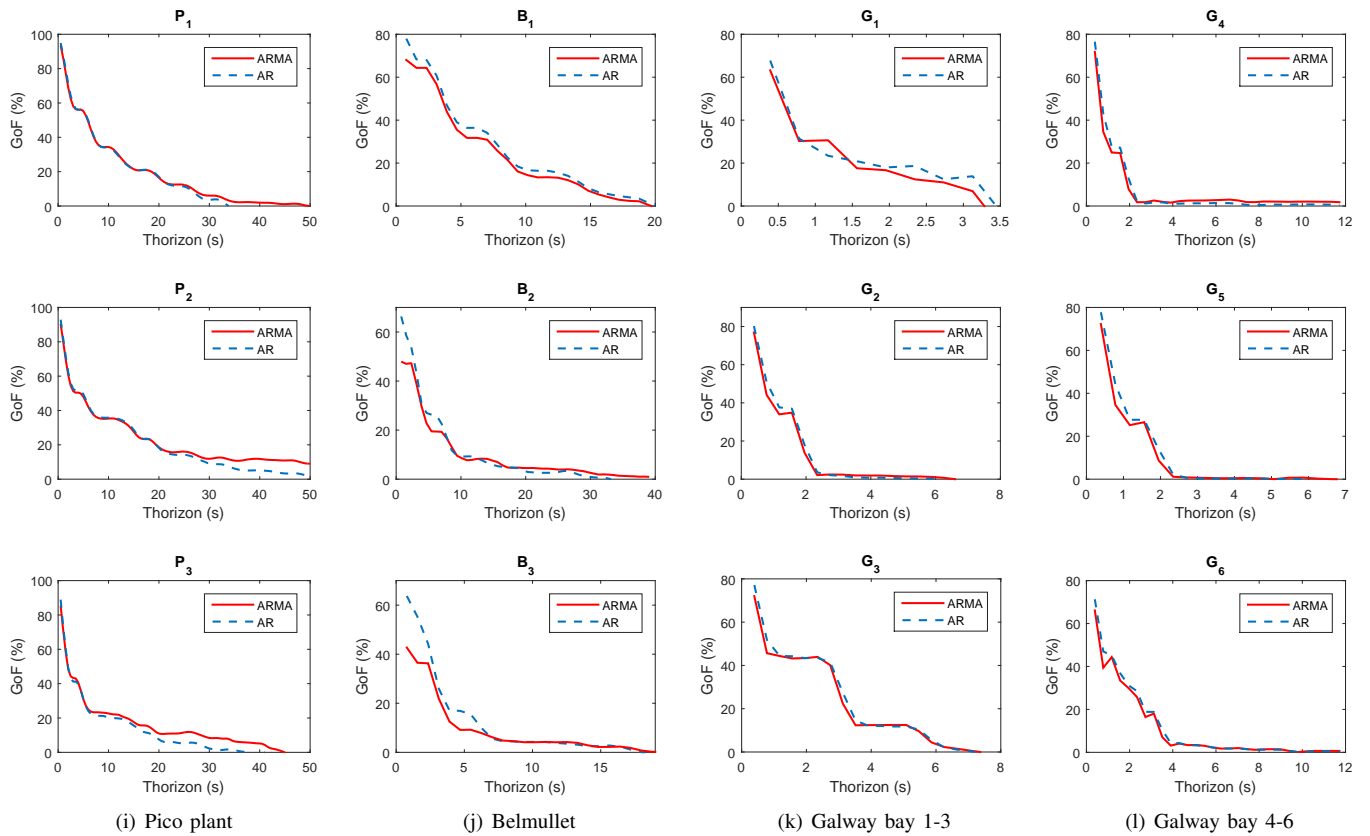
Fig. 6. The *GoF* for diferent T_h in all data sets

maintain the regression coefficients 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 Sections III and IV, that the ARMA model is more complex than the AR, while Section VI shows both obtain similar accuracies. Therefore, the conclusion presented in [1], saying the ARMA does not provide any improvement compared to the AR, is 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] M. Ge and E. C. Kerrigan, "Short-term ocean wave forecasting using an autoregressive moving average model," in *Control (CONTROL), 2016 UKACC 11th International Conference on*. IEEE, 2016, pp. 1–6.

[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. Scruggs, S. Lattanzio, A. Taflanidis, and I. Cassidy, "Optimal causal control of a wave energy converter in a random sea," *Applied Ocean Research*, vol. 42, pp. 1–15, 2013.

[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 devicesbenefits over independent control," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 4, pp. 1091–1099, 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.