Short-Term Forecasting of Sea Surface Elevation for Wave Energy Applications: The Autoregressive Model Revisited

Yerai Peña-Sanchez 🆔, Alexis Mérigaud 🆔, and John V. Ringwood 🆔

Abstract—For wave energy converter control applications, autoregressive (AR) models have been proposed as a simple wave forecasting method, solely based on measured or estimated values of the past wave elevation (or excitation force) signal. Using offline-filtered wave time series, AR models can achieve accurate forecasts several wave periods into the future. In this paper, the AR method is examined from the broader perspective of linear, Gaussian processes. In particular, assuming Gaussian waves and perfect knowledge of the wave spectrum, it is possible to derive a theoretically-optimal wave elevation predictor. It is shown that, in realistic situations, AR models can achieve a performance comparable to the theoretically optimal, spectrum-based predictor, both in simulated wave time series and using actual wave elevation records.

Index Terms—Autoregressive model (AR), filtering, forecasting, Gaussian process, time series, wave energy.

I. INTRODUCTION

TO MAKE wave energy profitable, wave energy converters (WECs) must convert the maximum possible amount of mechanical energy into electricity for a given device cost. To improve power conversion, a control strategy can be applied to the WEC, which optimizes the system loading [1]. In proposing causal WEC controllers [2], [3], some researchers avoid the need to predict the future wave elevation η, though the optimal control law is, in general, noncausal [1], [3]–[5], i.e., an optimal controller requires knowledge of the future η or excitation forces acting on the device. Although only prediction of sea surface elevation η is performed in this paper, the strategies introduced here can be applied, in an analogous mode, to the prediction of the excitation forces acting on a device.

Sea surface elevation forecasting differs from sea water level forecasting [6], significant waveheight forecasting [7], or more generally sea-state forecasting [8]. Sea-state forecasting consists of predicting wave statistics, for time horizons ranging from 1 to 48 h, as opposed to wave elevation forecasting that predicts
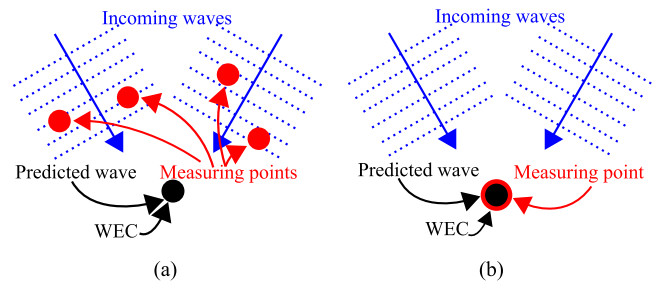
Fig. 1. Two main methods to forecast free surface elevation. (a) Using distant measurements. (b) Based on the measurements of a single point.

wave elevation a few seconds into the future. The reader interested in sea-state forecasting may refer to [8] for a comparison between methods based on meteorological models, and time-series-based approaches, for significant waveheight forecasting. In the literature, two main approaches are found for short-term wave elevation forecasting [9]. The first method predicts future η values at the WEC location, based on measurements taken at one or more points located at a certain distance from the WEC, as shown in Fig. 1(a). The number of measurement points, required to obtain an accurate prediction, depends mainly on the directional spreading of the incoming waves [10]. For a unidirectional sea-state, a single up-wave measuring point would lead to a deterministic prediction of η [10]. The second method, shown in Fig. 1(b), solely uses past measurements at the WEC location to predict η; that is, considering the measurements of η as a time series, future values can be predicted as a function of past values. From a hardware requirement point of view, the second method is simpler, mainly because it uses only measurements from the WEC itself, whereas the first method requires more equipment to measure η at different points. However, the knowledge of up-wave information may *improve* the prediction, compared to that achieved by using only past measurements taken at the WEC location. In this paper, only the approach of Fig. 1(b) is analyzed.

In the vast majority of cases, within the power production region of WECs, ocean waves can be described as a Gaussian, linear process, only excluding shallow water conditions (for a more detailed discussion, see [11, Ch. 9]). Additionally, for short periods of time (around 30 min), the process can also be assumed stationary [11]. Therefore, the use of a prediction model that exploits the stationary, linear statistical structure of

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

2   IEEE JOURNAL OF OCEANIC ENGINEERING

sea waves is pertinent. Several linear prediction models can be found in the literature. In [9], the simple AR model is found to achieve accurate predictions for the case of low-frequency swell waves, which are the most energetic waves and, thus, offer the most potential for wave energy conversion. In [10], a forecasting strategy is introduced which, assuming perfect knowledge of the wave spectrum, provides a theoretical upper bound for the prediction accuracy. An alternative linear prediction model to AR is the autoregressive moving average model which, for wave elevation prediction, achieves similar results to AR models [12].

Although the AR model introduced in [9] and the theoretical best predictor introduced in [10] have a different model structure, and their regression coefficients are identified using different techniques, both models are consistent with the linear, stationary assumptions of the ocean waves described and, therefore, both should lead to similar results. While the AR predictor structure uses a unique set of parameters (for the one-step-ahead prediction), and the multistep ahead prediction is carried out iteratively, the theoretically-optimal predictor has a different set of coefficients for each prediction horizon (T_{hor}) and, thus, directly predicts the multistep ahead wave elevation, with no need for iteration. In the following, the predictor structure with a different set of coefficients for each prediction horizon will be termed direct multistep (DMS). Regarding the identification method, while the AR model identifies the model coefficients from wave time series, the DMS prediction strategy introduced in [10] identifies the parameters using statistical information contained in the wave spectrum. For the sake of clarity, the different forecasting strategies used in this paper are termed using the predictor structure as name and the identification method in the subindex. For example, the forecasting strategy which has the DMS predictor structure and identifies the parameters from the wave spectrum is referred to as DMS_{Sp}.

Overall, the aim of this paper is to revisit the AR forecasting model by the following statements.

1) Comparing the predictions achieved by the AR model to those obtained by the DMS_{Sp} strategy which, in theory, yields the best achievable predictions.
2) Comparing two different identification methods for the AR predictor model. The first one is based on solving a linear least square (LLS) problem, termed AR_{LLS}. In contrast, a second forecasting strategy identifies the AR parameters through a non-LLSs problem, termed long-range predictive identification (LRPI) [9], [13] and, therefore, identified as AR_{LRPI}.
3) Additionally, the idea of using filtered η values for the prediction is investigated. In some studies [9], [13], it is shown that, when using an AR model, offline filtering can improve the forecasting accuracy. However, for real time operation, since online filters add a delay to the filtered signal (T_{delay}), online filtering may not improve the overall forecasting accuracy. This paper shows the difference between the prediction accuracy obtained using online and offline filtering of η.

A further forecasting strategy is introduced which, while having the same model structure as the theoretical best predictor, identifies the parameters from time series by solving an LLS problem, termed DMS_{LLS}. In summary, four different forecasting strategies are compared in this study (DMS_{Sp}, DMS_{LLS}, AR_{LLS}, and AR_{LRPI}). Under ideal conditions, all strategies should perform similarly. However, in real conditions, when the data available for identification are imperfect, and the sea-state is not perfectly stationary, the relative performance of the different strategies may differ. Therefore, the four strategies are compared using both simulated and real wave data.

The remainder of this paper is organized as follows. In Section II, the two AR-model-based forecasting strategies, namely AR_{LLS} and AR_{LRPI}, are introduced, along with the potential use of filtering. Section III describes the DMS_{Sp} and DMS_{LLS} prediction methods. In Section IV, all the methods analyzed in this paper are compared using numerically simulated wave time series. In Section V, the prediction methods are investigated using real wave data. Finally, conclusions are drawn in Section VI.

II. AR MODEL

A. AR Model

The AR model assumes that η_k, the value of η at time instant k, where k is the discrete time index ($t = kT_s$) and T_s the sampling time, depends linearly on its own previous values, through the parameters ϕ_i. Therefore, the p-step ahead predicted values of η ($\hat{\eta}_{k+p|k}$), where $p \in \{1 \ldots P\}$ and $T_{\text{hor}} = PT_s$, can be expressed as follows:

$$\hat{\eta}_{k+p|k} = \sum_{i=1}^{H} \phi_i \hat{\eta}_{k+p-i|k} \qquad (1)$$

where H is the order of the model and $\hat{\eta}_{k+p-i|k} = \eta_{k+p-i}$ when $k + p - i \leq k$, since the value of η_{k+p-i} is known and, hence, there is no need of prediction.

Several methods to determine the optimal order H of an AR model (e.g., Akaike's information theoretic criterion [14]) can be found in the literature. However, such methods are not discussed in this paper, since the aim is to show how the forecasting strategies perform *for different orders*, rather than how the best order can be obtained for each method.

B. AR Model Identification

Two different approaches for the identification of the ϕ_i in (1) are studied in this paper. The first approach minimizes the one-step-ahead prediction error. Therefore, given a set of training data of dimension N_{tr}, the coefficients are identified by minimizing the following cost function:

$$J_{\text{LLS}} = \sum_{k=H+1}^{N_{\text{tr}}} \left(\eta_k - \hat{\eta}_{k|k-1} \right)^2 \qquad (2)$$

which is an LLS problem. Therefore, the AR model based forecasting strategy, whose coefficients are identified by minimizing (2), is termed AR_{LLS}.

For longer prediction horizons, an identification method that minimizes the p-step ahead prediction error may be more pertinent. The coefficients of the AR_{LRPI} forecasting strategy are

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

PEÑA-SANCHEZ *et al.*: SHORT-TERM FORECASTING OF SEA SURFACE ELEVATION FOR WAVE ENERGY APPLICATIONS
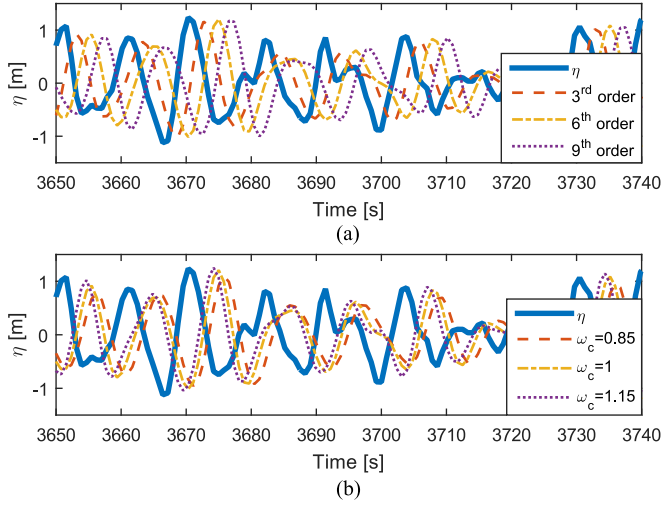
3

Fig. 2. (a) Section of η_f using Butterworth filters of 3rd, 6th, and 9th orders with $\omega_c = 1$ rad/s and (b) the same section of η_f filtered with a 6th order Butterworth filter with $\omega_c = 0.85$, 1, and 1.15 [rad/s].

identified by minimizing the following multistep ahead cost function:

$$J_{\mathrm{LRPI}} = \sum_{j=1}^{P} \sum_{k=H+1}^{N_{\mathrm{tr}}} (\eta_k - \hat{\eta}_{k|k-j})^2. \tag{3}$$

Since (3) represents a non-LSSs problem, its minimization is more complex problem than the minimization of (2). Thus, the time needed to identify the parameters of $\mathrm{AR}_{\mathrm{LRPI}}$ is significantly higher than the one needed for $\mathrm{AR}_{\mathrm{LLS}}$.

C. Wave Signal Filtering

It is well known that, if a signal is decomposed into one containing the low-frequency content and the other containing the high-frequency content, the signal with the higher frequencies is the less predictable one [15]. Since the low-frequency wave components are the most energetic [16], some studies [9], [13], lowpass filter η offline before the prediction process. However, for online filtering, the filtered signal η_f is delayed with respect to the original signal η [17]. Therefore, if a prediction of T_{hor} seconds is required, $T_{\mathrm{hor}} + T_{\mathrm{delay}}$ needs to be predicted.

T_{delay} depends mostly on the cutoff frequency ω_c and the order of the filter. By way of example, Fig. 2 shows how η_f is delayed with respect to η for different orders of a Butterworth filter [see Fig. 2(a)] and different ω_c [see Fig. 2(b)]. The simulated η used in this section is generated from a JONSWAP spectrum [18] with a peak period T_P of 10 s, a significant waveheight H_S of 2 m, and a peak enhancement factor γ of 3.3.

For the example shown in Fig. 2(a), η is filtered by using 3rd, 6th, and 9th order Butterworth filters with $\omega_c = 1$[rad/s], corresponding to a T_{delay} of, approximately, 2.3, 3.9, and 6.2 s, respectively. Additionally, Fig. 2(b) illustrates the differences between η and η_f using a 6th order Butterworth filter with various ω_c. It is clear that the delay added by the filter increases when the order increases or when ω_c decreases, i.e., when more high-frequency components are filtered out.
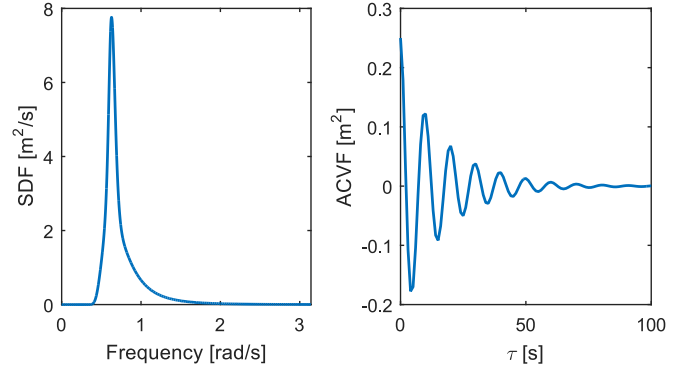

Fig. 3. SDF and ACVF for a JONSWAP spectrum with $T_P = 10$ s, $H_S = 2$ m, and $\gamma = 3.3$.

III. BROADER PERSPECTIVE—OCEAN WAVES AS A GAUSSIAN PROCESS

A. Theoretically Optimal Predictor

Ocean waves can, in most cases, be modeled as a zero-mean Gaussian process [11]. If a short duration time signal is considered, with respect to the typical time-length in which the sea condition changes (typically, 30 min to 3 h), the wave elevation process can be considered stationary. Finally, the wave elevation process is also considered to be ergodic, which means that the statistics of the process, obtained from time series, are equal to the ensemble statistics [11].

Thus, a stationary, ergodic Gaussian process is fully described by its mean, $\hat{\eta} = 0$ for the case of ocean waves, and its autocovariance function (ACVF), which is defined as follows:

$$R_{\eta\eta}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \eta_t \eta_{t+\tau} \, dt. \tag{4}$$

One can notice that $R_{\eta\eta}$ only depends on τ, due to the stationarity of the process. Additionally, the maximum of $R_{\eta\eta}(\tau)$ is at $\tau = 0$, where $R_{\eta\eta}(0) = \sigma^2$, the variance of the process (also equal to m_0, the zeroth order spectral moment).

The ACVF can be obtained directly from the spectral density function (SDF), since SDF and ACVF are a Fourier transform pair, according to the Wiener–Khintchine theorem. Therefore, the statistical properties of a stationary Gaussian sea are fully characterized by its ACVF or, equivalently, by its SDF [19]. As an example, Fig. 3 shows an SDF and the corresponding ACVF for a JONSWAP spectrum ($H_s = 2$ m, $T_p = 10$ s).

Once the statistical properties of ocean waves are well-defined, and assuming perfect knowledge of the SDF, it is possible to derive a theoretically optimal predictor, in a mean-square error sense. First, define the sequences of vectors $\mathbf{x}_k \in \mathbb{R}^H$ and $\mathbf{y}_k \in \mathbb{R}^P$, which contain, respectively, H past consecutive measurements and P consecutive forecast samples, as follows:

$$\mathbf{x}_k = \begin{bmatrix} \eta_k \\ \eta_{k-1} \\ \vdots \\ \eta_{k-H+1} \end{bmatrix} \quad \text{and} \quad \mathbf{y}_k = \begin{bmatrix} \eta_{k+P} \\ \vdots \\ \eta_{k+2} \\ \eta_{k+1} \end{bmatrix} \tag{5}$$

for any positive integer $k > H$. Let $\mathbf{v}_k \in \mathbb{R}^{H+P}$ be a vector containing $\mathbf{x}_k$ and $\mathbf{y}_k$ as follows:

$$\mathbf{v}_k = \begin{bmatrix} \mathbf{y_k} \\ \mathbf{x_k} \end{bmatrix}. \tag{6}$$

From the definition of a Gaussian process, $\mathbf{v}_k$ follows a multivariate normal distribution of dimension $H + P$, characterized by its mean, $\mu = 0_{\mathbb{R}^{H \times P}}$, and its variance-covariance matrix, $\Sigma_{vv} \in \mathbb{R}^{(H+P) \times (H+P)}$, whose components $\Sigma_{vv_{ij}}$ are defined as follows:

$$\Sigma_{vv_{ij}} = \mathbb{E}[\eta_{k+H+1-i} \eta_{k+H+1-j}] \tag{7}$$

where $\forall i, j \in [\![1; H + P]\!]^2$. All the diagonal elements of $\Sigma_{vv_{ij}}$ are equal to the variance of the process, σ^2. Denoting $r_i = R_{\eta\eta}(i\Delta t)$, for all $i \in \mathbb{N}$, each row of Σ_{vv} is a discretized and time-shifted version of the ACVF, [see (4)] as follows:

$$\Sigma_{vv} = \begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r_{H+P-1} \\ r_1 & r_0 & r_1 & \cdots & r_{H+P-2} \\ r_2 & r_1 & r_0 & & \vdots \\ \vdots & \vdots & & \ddots & r_1 \\ r_{H+P-1} & r_{H+P-1} & \cdots & r_1 & r_0 \end{bmatrix}. \tag{8}$$

Splitting Σ_{vv} into past-measurement and forecasting blocks

$$\Sigma_{vv} = \begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{xx} \end{bmatrix} \tag{9}$$

where $\Sigma_{yy} \in \mathbb{R}^{P \times P}$, $\Sigma_{xx} \in \mathbb{R}^{H \times H}$, $\Sigma_{yx} \in \mathbb{R}^{P \times H}$, and $\Sigma_{xy} \in \mathbb{R}^{H \times P}$ are defined as follows:

$$\Sigma_{yy} = \begin{bmatrix} r_0 & \cdots & r_{P-1} \\ \vdots & \ddots & \vdots \\ r_{P-1} & \cdots & r_0 \end{bmatrix}$$

$$\Sigma_{xx} = \begin{bmatrix} r_0 & \cdots & r_{H-1} \\ \vdots & \ddots & \vdots \\ r_{H-1} & \cdots & r_0 \end{bmatrix} \tag{10}$$

$$\Sigma_{yx} = \Sigma_{xy}^T = \begin{bmatrix} r_H & \cdots & r_{H+P-1} \\ \vdots & \ddots & \vdots \\ r_1 & \cdots & r_H \end{bmatrix}.$$

The conditional distribution of y and x, $y_k | x_k$, is a multivariate Gaussian process [20]. Using $\mu_x = 0_{\mathbb{R}^H}$ and $\mu_y = 0_{\mathbb{R}^P}$, the mean $\mu_{y_k|x_k}$ and variance $\Sigma_{y_k|x_k}$ are given as follows:

$$\mu_{y_k|x_k} = \Sigma_{yx} \Sigma_{xx}^{-1} x_k$$

$$\Sigma_{y_k|x_k} = \Sigma_{yy} - \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy}. \tag{11}$$

The best predictor of y is given by $\mu_{y_k|x_k}$ as

$$\hat{y}_k = \mu_{y_k|x_k} = \Sigma_{yx} \Sigma_{xx}^{-1} x_k \tag{12}$$

and the corresponding mean-square prediction error value, for each p, is given by the diagonal terms of $\Sigma_{y_k|x_k}$.

Explicitly, the set of optimal parameters θ, which gives the best linear combination of the measurements of the vector x_k, is obtained by solving the following system of linear equations:

$$\theta = \Sigma_{yx} \Sigma_{xx}^{-1}. \tag{13}$$

Unlike the AR model shown in (1), (12) has a set of H parameters for each prediction horizon p. Therefore, this predictor structure, which directly predicts the p-step-ahead values, with no need of iteration, is defined as follows:

$$\hat{\eta}_{k+p|k} = \sum_{i=1}^{H} \theta_{p,i} \eta_{k-i}. \tag{14}$$

As outlined in Section I, the predictor structure of (14) is termed DMS. The forecasting strategy which combines the DMS structure and the identification method which derives the forecasting coefficients using wave statistical information, as in (13), is referred to as DMS_{Sp}.

Given the Gaussian nature of the wave process, DMS_{Sp} can be considered as an upper bound for the prediction accuracy of any linear or nonlinear forecasting method which predict η by any function of the same set of previous values, provided that system (13) can be accurately solved.

B. Identifying a Direct Multistep Model Using Time Series

The DMS_{Sp} forecasting strategy can be implemented provided that spectral estimates are available. However, it is also possible to identify the parameters for a forecasting strategy with the DMS predictor structure introduced in (14) without using the wave spectrum, but rather using previous values of η. The identification of such a forecasting strategy is carried out by minimizing a separate cost function for each p defined as

$$J_{DMS_{LLS}}(p) = \sum_{k=H+1}^{N_{tr}} (\eta_{k+p} - \hat{\eta}_{k+p|k})^2 \tag{15}$$

which describes p separate LLS problems. To this end, the forecasting strategy which has the DMS predictor structure of (14) and whose parameters are identified by solving the p different LLS problems shown in (15) is termed DMS_{LLS}.

Note that the coefficients for the one-step-ahead prediction of the DMS_{LLS} strategy coincide with those identified for the AR_{LLS} forecasting strategy. Consequently, both strategies will lead to the same one-step-ahead prediction accuracy. Additionally, in theory, if the parameters of the DMS_{LLS} strategy are identified using a sufficiently long data set, and provided that system (13) is accurately solved to obtain the DMS_{Sp} parameters, both approaches should result in the same set of parameters $\theta_{p,i}$.

IV. NUMERICAL SIMULATION RESULTS

A. Performance Metrics

The metric used to describe the prediction accuracy, for each prediction horizon, is the goodness of fit (GoF). The p-step ahead GoF (GoF(p)) measures the quadratic error between the real value, η_{k+p}, and the predicted value, $\hat{\eta}_{k+p|k}$, and is formally

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

PEÑA-SANCHEZ *et al.*: SHORT-TERM FORECASTING OF SEA SURFACE ELEVATION FOR WAVE ENERGY APPLICATIONS 5

expressed as follows:

$$\text{GoF}(p) = 1 - \sqrt{\frac{\sum_{k=1}^{N} \left(\eta_{k+p} - \hat{\eta}_{k+p|k}\right)^2}{\sum_{k=1}^{N} \left(\eta_{k+p}\right)^2}} \qquad (16)$$

where N is the number of points used to evaluate the GoF.

Additionally, a multistep performance index (MSPI) is introduced, which takes into account all the prediction horizons considered between, say, 1 and P. Therefore, the MSPI provides an overview of the performance achieved by a given forecasting strategy for all time horizons at once. The MSPI is expressed as follows:

$$\text{MSPI} = 1 - \frac{1}{\sigma^2 P} \sum_{i=1}^{P} \left(\frac{1}{N} \sum_{k=1}^{N} (\eta_{k+i} - \hat{\eta}_{k+i|k})^2 \right). \qquad (17)$$

Some authors use correlation as a performance metric to quantify the accuracy of their forecasting strategies [21]. However, since correlation only represents the phase difference between two signals, the amplitude differences are not reflected in the obtained correlation measure and is therefore not used in the current study.

The two performance metrics used in this paper, GoF and MSPI, are based on the root-mean-square error between actual and predicted values of η. Therefore, they statistically measure the magnitude of the difference between actual and predicted η values, instead of measuring only the difference in amplitude or phase, giving a more realistic measure of the prediction error.

B. Wave Simulation

The simulated wave time series of this paper are generated from a JONSWAP spectrum with $H_S = 2$ m, $T_P = 10$ s, and $\gamma = 3.3$. The simulated data sets are 30 min long with a sampling time of $T_S = 1$ s. Both the identification of the parameters and the prediction are carried out using a complete 30-min data set ($N_{tr} = 1800$ points). The data set used for identification is different from the data set where the prediction is performed.

If the spectrum of a stationary process is zero for some nonzero frequency interval (*band-limited* spectrum), the process is *predictable* (or *deterministic*) [22], which means that, for any $\varepsilon > 0$, there exists a linear predictor whose one-step-ahead prediction error is less than ε. For such a process, the prediction error could be made arbitrarily small by increasing the order of the predictor [22]. This is clearly not a reasonable assumption for ocean waves recorded at a single location. Furthermore, any realistic wave elevation measurement or estimation method would carry some level of measurement noise. In view of those considerations, it is reasonable to assume that the process, as recorded by a measurement device, would present some degree of unpredictability.

Simulated JONSWAP spectral values are zero to machine accuracy for low frequencies (see Fig. 3), and should therefore be predictable (i.e., the one-step-ahead prediction error should tend to zero when increasing the predictor order). To make simulations more realistic, an arbitrarily small white-noise component ($1e^{-10}$) is added to the SDF used to generate η. Note that, when spectra are computed from real wave data, no such modification

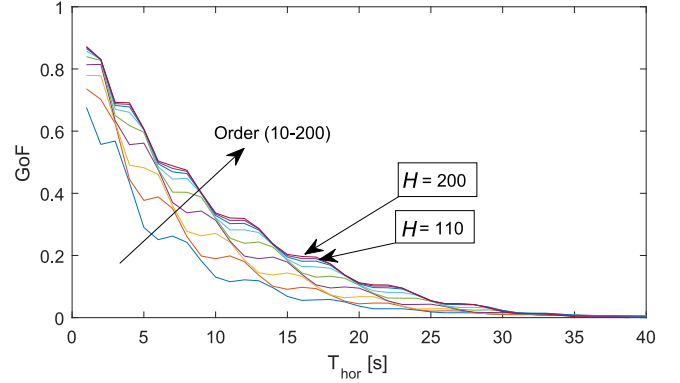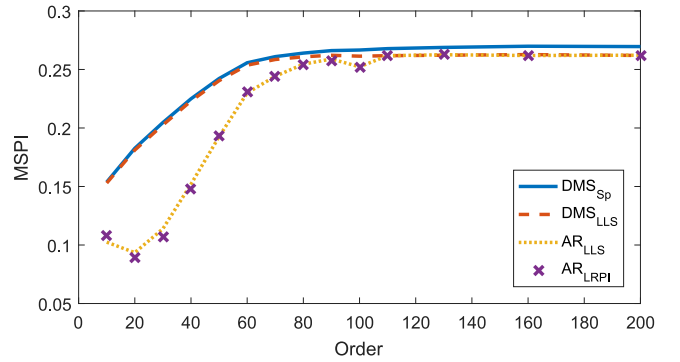Fig. 4. GoF of the prediction achieved by DMS$_{\text{LLS}}$ using different orders.

Fig. 5. MSPI obtained by the different forecasting methods using simulated wave data.

is necessary, because the recorded low-frequency spectral content is nonzero. Finally, the case where the simulated spectrum is band-limited is briefly commented in Section IV-E.

C. Comparison of Different Forecasting Strategies

The AR$_{\text{LLS}}$, AR$_{\text{LRPI}}$, and DMS$_{\text{LLS}}$ strategies require consistent identification of their prediction coefficients, using a sufficiently large learning data set (N_{tr} data points). Typically, regardless of the method, it is found that, to obtain accurate identification of the forecasting coefficients, the length of the training data should be 15 times the order of the model ($N_{tr} \approx 15H$) (in other words, using more learning data points does not bring any significant improvement to the prediction accuracy).

Unsurprisingly, for all the prediction methods analyzed, the achieved GoF increases with the model order, up to a given order, beyond which the GoF does not improve significantly. For example, Fig. 4 shows the GoF achieved by DMS$_{\text{LLS}}$ with orders from 10 to 200. For the chosen wave spectrum, the GoF achieved by DMS$_{\text{LLS}}$ does not significantly improve for orders larger than 80. In particular, it is shown that the GoF achieved by a DMS$_{\text{LLS}}$ of order $H = 110$, and a DMS$_{\text{LLS}}$ of order $H = 200$, is almost identical.

The order, beyond which the GoF improvement is not significant, is similar for all the strategies analyzed, as can be seen in Fig. 5, where the MSPI of the different strategies is illustrated. One can notice that the DMS$_{\text{Sp}}$ and DMS$_{\text{LLS}}$ strategies perform identically below $H \approx 80$, and outperform the two
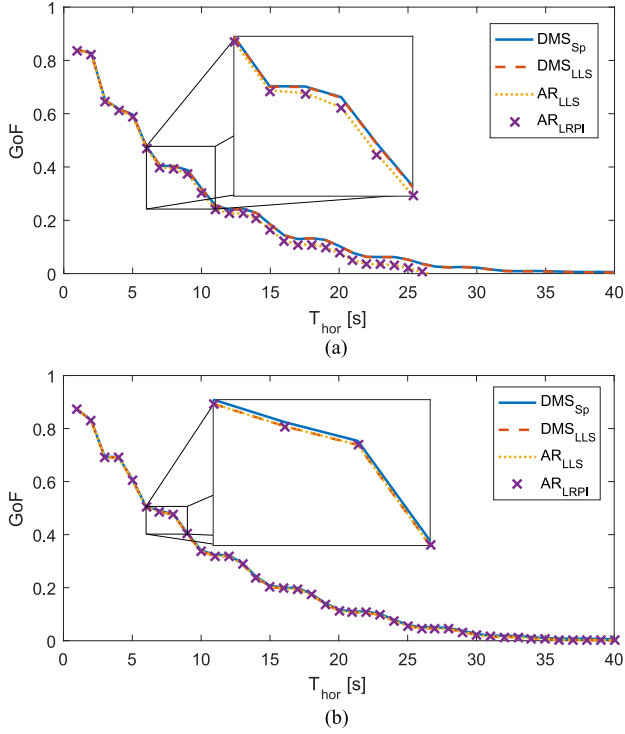
Fig. 6. GoF of the prediction using the different forecasting strategies with (a) $H = 50$, and (b) $H = 150$.

AR-model-based strategies. Moreover, for $H > 110$, the DMS$_{\text{LLS}}$, AR$_{\text{LLS}}$, and AR$_{\text{LRPI}}$ strategies achieve the same prediction accuracy, whereas the DMS$_{\text{Sp}}$ strategy performs slightly better than the other three.

In Fig. 6, the comparison between the forecasting strategies is further exemplified, in terms of GoF. Fig. 6(a) shows that, for $H = 50$, the DMS$_{\text{Sp}}$, and DMS$_{\text{LLS}}$ strategies perform identically and outperform, as shown before in Fig. 5, the AR$_{\text{LLS}}$ and AR$_{\text{LRPI}}$ strategies, which achieve the same GoF. Note that the difference between the DMS and AR predictors increases with T_{hor}, due to the error accumulated by the iterations required by AR models to obtain the p-step-ahead forecast. Additionally, Fig. 6(b) shows that, for $H = 150$, all strategies perform quasi-identically, although the DMS$_{\text{Sp}}$ strategy is slightly more accurate. It should also be noted that, the AR$_{\text{LRPI}}$ strategy achieves almost the same results as the AR$_{\text{LLS}}$, while requiring significantly more time for the identification (for example, for $H = 50$, solving the LRPI nonlinear problem is typically around 20 times slower than solving the LLS problem of AR$_{\text{LLS}}$, and the relative difference increases with the order).

The SDF used for the simulation is not predictable (as explained in Section IV-B), and the logarithm of its spectrum is integrable (Paley–Wiener condition [22]). Thus, the Kolmogorov–Szego's mean square error formula [22] can provide the theoretical lowest one-step-ahead prediction error, which would be obtained by increasing the predictor order to infinity, as follows:

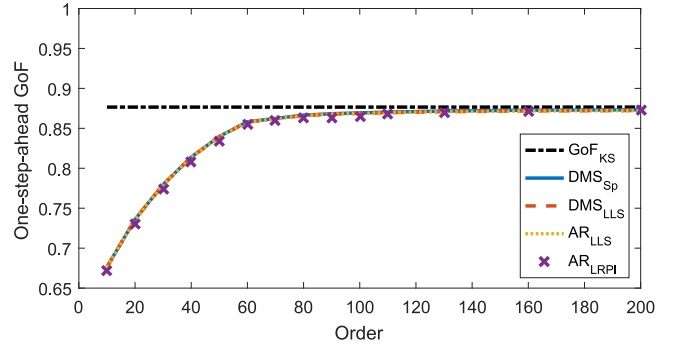$$\varepsilon_\infty^2 = \exp\left\{2\int_0^\pi \log \text{SDF}(\omega)d\omega\right\}. \tag{18}$$

Fig. 7. One-step-ahead prediction of the different forecasting strategies, for different orders, and the theoretical best one-step-ahead prediction associated to the SDF.

The one-step-ahead GoF corresponding to the Kolmogorov–Szego's (GoF$_{\text{KS}}$) mean square error formula [see (18)] can be derived as follows:

$$\text{GoF}_{\text{KS}} = 1 - \sqrt{\frac{\varepsilon_\infty^2}{\sigma^2}}. \tag{19}$$

Thus, by using (19), it is possible to compare the GoF$_{\text{KS}}$, related to the SDF used in this paper (see Fig. 3), to the one-step-ahead GoF obtained by the DMS$_{\text{Sp}}$, DMS$_{\text{LLS}}$, AR$_{\text{LLS}}$, and AR$_{\text{LRPI}}$ strategies. As shown in Fig. 7, the four strategies achieve almost identical one-step-ahead GoF (identical for the case of AR$_{\text{LLS}}$ and DMS$_{\text{LLS}}$). In fact, the difference between the GoF$_{\text{KS}}$ related to the SDF and the one-step-ahead GoF achieved by the DMS$_{\text{Sp}}$, DMS$_{\text{LLS}}$, AR$_{\text{LLS}}$, and AR$_{\text{LRPI}}$ strategies decreases asymptotically when the order increases.

In this section, identification and prediction are performed on different data sets, randomly generated from the same wave spectrum (see Fig. 3). Thus, both data sets share identical statistical properties, characterized by the wave spectrum. Therefore, the use of higher order models always results in a more accurate prediction, because more statistical information from the spectrum can be taken into account. However, real sea wave conditions, characterized by the wave spectrum, vary in time. Therefore, the forecasting strategies may be subject to some additional errors, because the wave condition evolves, between the period where the coefficients are identified, and the period over which the models are used to perform prediction. In particular, high-order models may be more affected by the inherent nonstationarity of ocean waves. However, as explained in Section V, for the DMS$_{\text{Sp}}$ strategy, the error related to the evolving wave condition can be assessed by the correct computation of the SDF. Overall, the impact of changing meteorological conditions is assessed in Section V, by using real wave data.

D. Use of Filtering in AR Model Prediction

As mentioned in Section II-C, low-frequency signals are more predictable than high-frequency signals. By way of example, Fig. 8(a) shows the GoF achieved by the AR$_{\text{LLS}}$ strategy, for the prediction of filtered wave signals, η_f, using offline Butterworth filters of various orders (see Section II-C). The prediction

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

PEÑA-SANCHEZ *et al.*: SHORT-TERM FORECASTING OF SEA SURFACE ELEVATION FOR WAVE ENERGY APPLICATIONS

7

Fig. 9. One-step-ahead GoF of the prediction achieved by DMS_{LLS} using different orders, when using wave data generated from a band-limited spectrum (solid blue) or from a nonband-limited spectrum (dashed red).
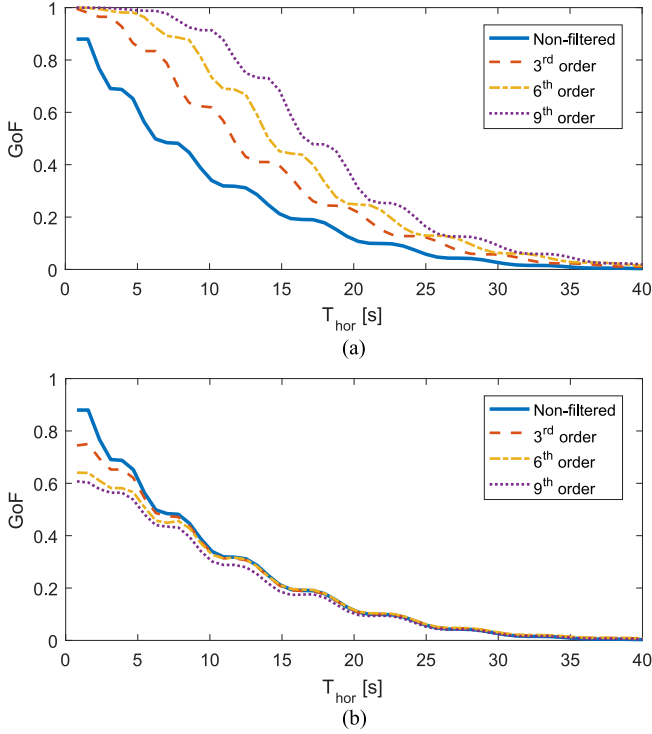

Fig. 8. GoF of the prediction computed with AR_{LLS} of $H = 100$ using η filtered (a) offline and (b) online with a Butterworth filter of different orders.

Fig. 10. MSPI obtained by the different forecasting methods using simulated wave data generated from a predictable spectrum.

achieved is significantly more accurate than the one achieved on nonfiltered η.

However, note that any realizable, online filter is necessarily causal, and therefore consists of some (usually linear) function of the previous measurements. Therefore, given the considerations presented in Section III-A, any predictor, based on online-filtered measurements, is also suboptimal with respect to DMS_{Sp} (provided that the same set of previous measurements is used, and that DMS_{Sp} avoids numerical inaccuracies). In view of the results shown in Section IV-C, showing the performance of the DMS_{Sp} strategy, it should not be possible to obtain results as favorable as in Fig. 8(a), when using an online filter.

As an illustration, Fig. 8(b) shows the performance of the AR_{LLS} forecasting strategy when using η_f, with the filtering carried out online. In Fig. 8(b), the GoF of the prediction, using η_f, drops below the performance of the prediction using nonfiltered η, which is due to the combination of two effects: first, the predicted η is compared to the actual (nonfiltered) wave elevation; second, as explained in Section II-C, the delay introduced by the filter must be taken into account.

The results shown in this section clearly illustrate how real-time filtering fails to improve the prediction of the nonfiltered wave signal. Although only Butterworth filters have been used here, the online implementation of any other filter would always lead to a prediction which is less accurate than that achieved using the nonfiltered signal directly. Indeed, the use of a filter, combined with an AR model, simply results in some other function of the past values, which is then necessarily suboptimal with respect to DMS_{Sp} using the same set of past measurements.
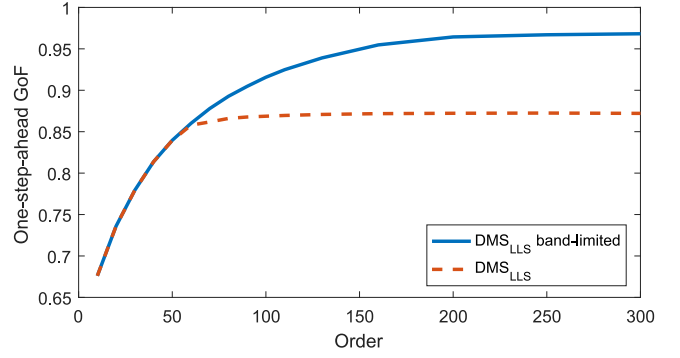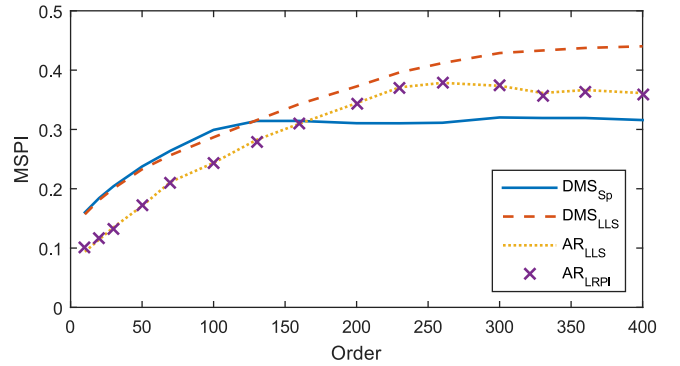
However, the above-mentioned considerations do not discard the idea of online filtering within the scope of real-time measurement or estimation of the wave elevation, especially in the presence of measurement noise.

E. Use of a Predictable Spectrum

The use of a band-limited simulated spectrum deserve some brief discussion. Here, as opposed to the previous sections, no white-noise component is added to the JONSWAP SDF, and therefore the process is band-limited—to machine accuracy. Fig. 9 shows the one-step-ahead prediction obtained by DMS_{LLS} of different orders, using wave data generated from both the noisy and band-limited JONSWAP spectra. It can be seen that, while for the nonpredictable spectrum the maximum one-step-ahead GoF is achieved with an order of $H \approx 60$, for a band-limited spectrum the one-step-ahead GoF continues increasing up to an order of $H \approx 250$.

Additionally, Fig. 10 shows the MSPI obtained by the different strategies, to show how their prediction accuracies are affected by the use of a simulated band-limited spectrum. The prediction accuracy achieved by the DMS_{LLS} strategy continue to increase up to orders larger than $H \approx 400$. Additionally, one can notice that the DMS_{Sp} strategy does not achieve accuracies as good as the other three strategies while, when using a nonpredictable spectrum, it is always the most accurate. The reason is that, when using a nonpredictable spectrum, the
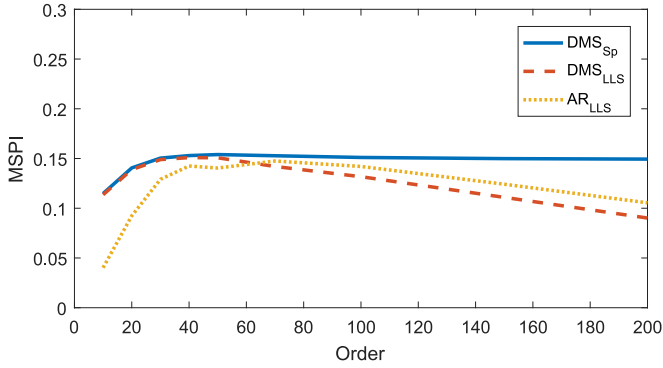
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

8   IEEE JOURNAL OF OCEANIC ENGINEERING

Fig. 11.   MSPI obtained by DMS_{Sp}, DMS_{LLS}, and AR_{LLS} for real wave data.

Fig. 12.   GoF of the prediction using DMS_{Sp}, DMS_{LLS}, and AR_{LLS} of $H = 50$.

result obtained from (13), although theoretically consistent with the assumption of stationary Gaussian waves, may be subject to numerical inaccuracies for large orders H; indeed, since the process is predictable, for large orders the covariance matrix in (13) is close to singular.

The simulated results obtained from a nonpredictable (noisy) spectrum depend, to some extent, on the level of white noise added to the JONSWAP formulation. However, as will be seen in Section V, the results obtained using real data are relatively similar to those obtained from the nonpredictable spectrum, in the sense that the prediction accuracies do not increase beyond orders in the region of 50–100 (consider, for example, Figs. 5 and 11). Therefore, in the opinion of the authors, the spectra used for simulations should be nonpredictable, to avoid unreasonably optimistic results.

V. RESULTS USING REAL WAVE DATA

The data used in this section are from Belmullet, Ireland, and have been provided by the Irish Marine Institute, Galway, Ireland.[1] Wave elevation time series have been recorded using a Datawell Waverider, Heerhugowaard, The Netherlands, [23] data buoy, and each data set is 30-min long, recorded at a 1.28 Hz sampling frequency. The results shown in this section are for December 2010.

In the remainder of this paper, for the sake of clarity, the results from the AR_{LRPI} strategy are not shown, because they are identical to AR_{LLS}, while requiring a computationally expensive identification procedure.

To derive the DMS_{Sp} parameters, it is important to use an appropriate procedure for wave SDF estimation. Assume, first, that the sea is purely stationary, i.e., that the wave spectrum does not evolve over time. Next, consider a 30-min wave signal window. This 30-min signal carries some statistical characteristics of the underlying wave spectrum, but also has some inherent randomness; in other words, even if the wave spectrum is stationary, two consecutive 30-min records will never be the same. Therefore, in theory, to obtain a perfect estimate of the spectrum, an infinite record would be necessary. However, if only a 30-min data set is available, applying signal processing techniques, such as windowing and smoothing, can help to separate the effects of the
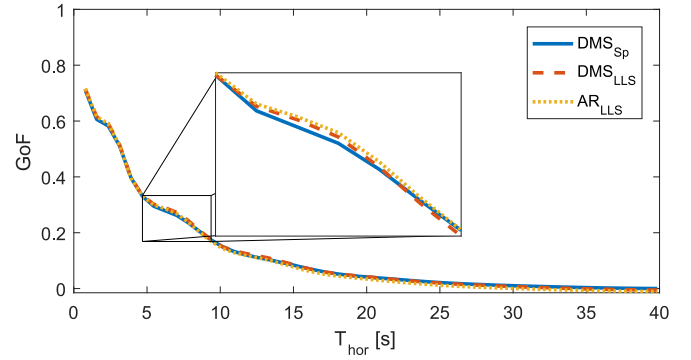
[1]https://www.marine.ie

short-term variability from the statistical information specific to the underlying spectrum [24].

For real sea records, it can be considered that the underlying wave spectrum slowly evolves over time, but such an underlying, slowly evolving spectrum must still be distinguished from the effect of short-term randomness. To compute the half-hourly SDF, the procedure retained is that used in [23]. Each 30-min data set is divided into l-second overlapping sections, with each section multiplied by a Tukey windowing function, before applying a fast Fourier transform to obtain an SDF estimate. Then, the half-hourly SDF is computed as the average of the SDFs of all sections. The resulting SDF is smoother when reducing l (and consequently increasing the number of overlapping sections). However, a smaller l also implies a lower frequency resolution for the SDF ($\Delta f = 1/l$). For this case, it has been found that $l = 5$ min gives the best estimate of the underlying wave spectrum.

The DMS_{Sp} parameters are derived from the SDF computed in the half-hourly data set, previous to the half-hourly data set where the parameters are used to predict η, so that, overall, only past information is used to forecast η. Similarly, for the strategies whose identification is based on a time series of η, the parameters are identified in the half-hourly data set previous to that used for wave prediction.

As shown in Figs. 11 and 12, the AR_{LLS}, DMS_{Sp}, and DMS_{LLS} strategies achieve very similar prediction accuracies. In particular, Fig. 11 indicates that, for an order of 50, the three methods perform almost identically, as further confirmed in Fig. 12.

For the strategies whereby the identification is based on time series (AR_{LLS} and DMS_{LLS}), there is an order H beyond which the model seems to "overlearn" statistical information due to the short-term randomness (as opposed to the underlying spectrum), thus marginally decreasing the prediction accuracy. As far as DMS_{Sp} is concerned, since the effect of short-term randomness is mitigated by the appropriate SDF computation methodology, the prediction accuracy does not decrease once the maximum accuracy is achieved, as shown in Fig. 11.

A significant difference between predicting simulated or real waves using a forecasting strategy identified using previous values of η (AR_{LLS} and DMS_{LLS}) is that, compared to the simulated waves case, the choice of the order of the model H can have a greater impact on the prediction performance for the case

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

PEÑA-SANCHEZ *et al.*: SHORT-TERM FORECASTING OF SEA SURFACE ELEVATION FOR WAVE ENERGY APPLICATIONS 9
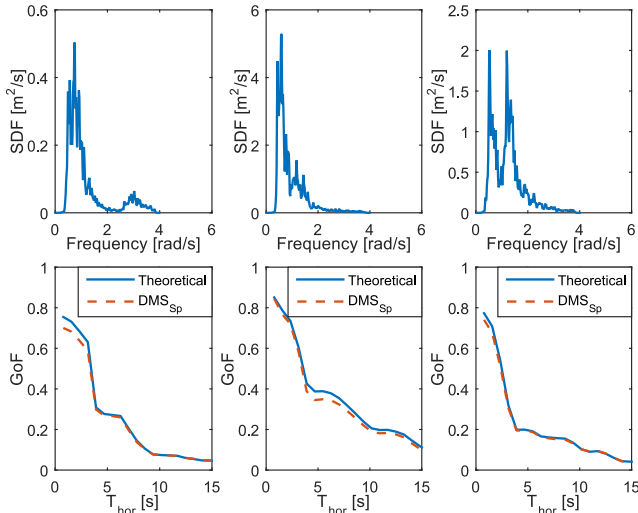

Fig. 13. Three different half-hourly SDFs (upper graphs), and the GoF achieved by DMS$_{Sp}$ and the theoretical best GoF (bottom graphs).

where real waves are used. In particular, for simulated waves, the performance is similar for any chosen order higher than the "optimal" order (H_{opt}) which, in this case, as shown in Fig. 5, is $H_{opt} \approx 80$. However, when using real waves, the performance decreases if the chosen H is greater than the "optimal" order which, in this case, as shown in Fig. 11, is $H_{opt} \approx 40$.

Overall, one can notice that the predictions obtained using real data are significantly less accurate than those achieved using simulated data, which may be due to several reasons.

1) The sea-state is nonstationary, which implies that the statistical characterization, learned in one given data set by any model is, in general, not the same in the next data set.
2) The SDFs of real wave data have significantly more high-frequency content than idealized sea-states, such as JON-SWAP spectrum (see the difference between the SDF of Fig. 3 and the SDFs of Fig. 13).
3) Finally, real wave data may be subject to measurement noise which, like the previous point, influences the frequency content of the signal.

To assess the relative importance of the above factors in the predictor performances, the following comparison is made (see bottom graphs of Fig. 13): The red dashed curve shows the empirical GoF achieved by DMS$_{Sp}$, identified using the SDF from the previous half-hourly data set, while the blue solid curve is the theoretical best GoF, associated with the SDF of the data set where the prediction is carried out [using (12)]. In other words, the blue curve indicates how the optimal predictor would perform, assuming perfect knowledge of the underlying spectrum. It can be appreciated that the difference between the two GoFs is minimal, thus suggesting that the difference, between simulated and actual prediction performances, is mainly because real wave spectra have more high-frequency content, rather than resulting from the nonstationarity of wave conditions.

Although the data used is from a single location, the results obtained in the different actual sea states confirm the conclusions obtained for the simulated wave data, where all the strategies perform similarly.

VI. DISCUSSION AND CONCLUSION

Overall, it is well-known that stationary AR models and stationary, discrete-time Gaussian processes are two appropriate alternative descriptions for nondeterministic, stationary, discrete-time linear processes; for any Gaussian linear process, there exists an AR(p) process such that the difference in the two ACVFs can be made arbitrarily small for all lags [25]. In fact, identifying a suitable AR model (either AR$_{LLS}$ or AR$_{LRPI}$), or adopting the perspective of a Gaussian vector (DMS$_{Sp}$ or DMS$_{LLS}$), implicitly assume the same underlying linear statistical properties. Therefore, under the assumption of stationary, Gaussian waves, and good identification, the various methods studied in this paper should perform similarly. The results shown in this paper, both from simulated and real wave data sets, indeed confirm that all approaches have almost identical accuracies.

The methods examined in this paper rely on the linearity of the wave process (which is a reasonable assumption within the power production region of WECs), and thus do not cover the cases where a Gaussian wave field cannot be assumed, which essentially correspond to shallow-water waves (with marked peak-to-trough asymmetry) or extreme sea states.

In particular, the results obtained by the DMS$_{Sp}$ strategy can reliably be considered as an upper bound for the achievable prediction accuracy (since the method is free of numerical issues). Therefore, if their parameters are correctly identified, the DMS$_{LLS}$, AR$_{LLS}$, and AR$_{LRPI}$ strategies yield results close to the best achievable prediction. From this point of view, all strategies studied here seem to be interesting candidate tools for short-term wave forecasting. However, the AR$_{LLS}$ strategy is particularly simple, and does not seem to be significantly outperformed by DMS$_{Sp}$.

Nevertheless, even the best achievable predictions remain relatively inaccurate, with a GoF lower than 50% for one wave period ahead in the favourable case of simulated data, and 20% in real sea-states. Such values are in contrast with the results of [9] and [13] (of the order of 90%–100% accuracy more than one wave period ahead), suggesting that the accuracies obtained from data filtered offline are unreasonably optimistic. Based on simple numerical examples, this paper shows that the use of online-filtered wave elevation values for the prediction always leads to inferior results than using the nonfiltered wave elevation.

Therefore, more accurate predictions may require the use of several measurement points in the vicinity of the WEC (such as, measurement buoys, other WECs, or remote sensing techniques). The DMS$_{Sp}$ strategy can be readily extended to such cases [10]. Similarly, autoregressive exogenous input models can provide an interesting extension of the AR model structure [13].

REFERENCES

[1] J. V. Ringwood, G. Bacelli, and F. Fusco, "Energy-maximizing control of wave-energy converters: The development of control system technology to optimize their operation," *IEEE Control Syst.*, vol. 34, no. 5, pp. 30–55, Oct. 2014.
[2] F. Fusco and J. V. Ringwood, "A simple and effective real-time controller for wave energy converters," *IEEE Trans. Sustain. Energy*, vol. 4, no. 1, pp. 21–30, Jan. 2013.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

10   IEEE JOURNAL OF OCEANIC ENGINEERING

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

[4] G. Bacelli, P. Balitsky, and J. V. Ringwood, "Coordinated control of arrays of wave energy devices benefits over independent control," *IEEE Trans. Sustain. Energy*, vol. 4, no. 4, pp. 1091–1099, Oct. 2013.

[5] 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.

[6] M. Vaziri, "Predicting Caspian Sea surface water level by ANN and ARIMA models," *J. Waterway, Port, Coastal, Ocean Eng.*, vol. 123, no. 4, pp. 158–162, 1997.

[7] A. Etemad-Shahidi and J. Mahjoobi, "Comparison between M5 model tree and neural networks for prediction of significant wave height in lake superior," *Ocean Eng.*, vol. 36, no. 15/16, pp. 1175–1181, 2009.

[8] G. Reikard, P. Pinson, and J.-R. Bidlot, "Forecasting ocean wave energy: The ECMWF wave model and time series methods," *Ocean Eng.*, vol. 38, no. 10, pp. 1089–1099, 2011.

[9] F. Fusco and J. Ringwood, "Short-term wave forecasting for real-time control of wave energy converters," *IEEE Trans. Sustain. Energy*, vol. 1, no. 2, pp. 99–106, Jul. 2010.

[10] A. Mérigaud and J. V. Ringwood, "Incorporating ocean wave spectrum information in short-term free-surface elevation forecasting," *IEEE J. Ocean. Eng.*, 2018, doi: 10.1109/JOE.2018.2822498.

[11] M. K. Ochi, *Ocean Waves: The Stochastic Approach*, vol. 6. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[12] Y. Peña-Sanchez and J. V. Ringwood, "A critical comparison of AR and ARMA models for short-term wave forecasting," in *Proc. 12th Eur. Wave Tidal Energy Conf.*, 2017, pp. 961.1–961.6.

[13] F. Paparella, K. Monk, V. Winands, M. F. P. Lopes, D. Conley, and J. V. Ringwood, "Up-wave and autoregressive methods for short-term wave forecasting for an oscillating water column," *IEEE Trans. Sutain. Energy*, vol. 6, no. 1, pp. 171–178, Jan. 2015.

[14] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, Dec. 1974.

[15] K. Sayood, *Introduction to Data Compression*. San Mateo, CA, USA: Morgan Kaufmann, 1996.

[16] M. K. Ochi, *Applied Probability and Stochastic Processes: In Engineering and Physical Sciences*. Hoboken, NJ, USA: Wiley 1990.

[17] A. V. Oppenheim, *Discrete-Time Signal Processing*. Noida, India: Pearson Educ., 1999.

[18] D. Hasselmann, M. Dunckel, and J. Ewing, "Directional wave spectra observed during JONSWAP 1973," *J. Phys. Oceanogr.*, vol. 10, no. 8, pp. 1264–1280, 1980.

[19] A. Mrigaud and J. V. Ringwood, "Free-surface time-series generation for wave energy applications," *IEEE J. Ocean. Eng.*, vol. 43, no. 1, pp. 19–35, Jan. 2018.

[20] M. L. Eaton, *Multivariate Statistics: A Vector Space Approach*. Hoboken, NJ, USA: Wiley, 1983.

[21] M. R. Belmont *et al.*, "An examination of the feasibility of linear deterministic sea wave prediction in multidirectional seas using wave profiling radar: Theory, simulation, and sea trials," *J. Atmos. Ocean. Technol.*, vol. 31, no. 7, pp. 1601–1614, 2014.

[22] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*. New York, NY, USA: McGraw-Hill, 2002.

[23] B. Datawell, *Datawell Waverider Installation Guide*, Datawell BV, Heerhugowaard, The Netherlands, 2012.

[24] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications*. Cambridge, U.K.: Cambridge Univ. Press, 1993.

[25] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*. New York, NY, USA: Springer, 2013.

Yerai Peña-Sanchez received the Diploma in renewable energies engineering from the University of the Basque Country, Leioa, Spain, in 2016. Since October 2016, he has been working toward the Ph.D. degree at the Centre for Ocean Energy Research, Maynooth University, Maynooth, Ireland.

He is currently working on estimation and forecasting of wave excitation forces.

Alexis Mérigaud received the Engineering degree in 2012 from ENSTA, Paris, France, where he specialized in energy management and economics. Since January 2015, he has been working toward the Ph.D. degree from the Centre for Ocean Energy Research, Maynooth University, Maynooth, Ireland.

He is working on numerical simulation and control for linear and nonlinear wave energy converter models, as well as real-time wave forecasting. After his engineering, he studied Public administration for two years at Sciences Po Paris. His previous experience in the field of energy includes economic modeling, as well as technical regulatory aspects, of energy networks and markets.

John V. Ringwood received the Diploma in electrical engineering from the Dublin Institute of Technology, Dublin, Ireland, in 1981, and the Ph.D. degree in control systems from Strathclyde University, Glasgow, U.K., in 1985.

He is currently a Professor of Electronic Engineering at the Maynooth University Ireland and the Director of the Centre for Ocean Energy Research. His research interests include time series modeling, wave energy, control of plasma processes, and biomedical engineering.

Dr. Ringwood is a Chartered Engineer and a Fellow of Engineers Ireland.