A Study on Short-Term Sea Profile Prediction for Wave Energy Applications

F. Fusco[1] and J. Ringwood[2]

[1]Electronic Engineering Department,
National University of Ireland Maynooth,
Maynooth, Co. Kildare, Ireland
E-mail: francesco.fusco@eeng.nuim.ie

[2]E-mail: john.ringwood@eeng.nuim.ie

Abstract

Control of wave energy converters requires knowledge of some seconds of the future behavior of certain physical quantities, in order to approach optimality. That is why short time prediction of the oncoming waves is a crucial problem in the field of wave energy, whose solution could bring great benefits to the effectiveness of the devices and to their economical viability.

This study is proposed as a preliminary approach to cope with this necessity, where wave forecasts are computed on the basis of past observations collected at the prediction site itself. Working on single point measurements allows the treatment of the wave elevation as a pure time series, so that a wide range of well established techniques from the stochastic time series modelling and forecasting field may be exploited. Among the proposed solutions there are some cyclical models, based on an explicit representation of the *a priori* knowledge about the real process. It is then shown how a lot simpler and more effective solution can be obtained through classical AR models, which are shown to be able to implicitly represent the cyclical behavior of real waves. As a comparison with AR models some results obtained with neural networks are also provided.

Keywords: wave energy, control of wave energy converters, wave forecasting, time series

Nomenclature

WEC	=	Wave Energy Converter
H_s	=	Significant wave height
η	=	Wave elevation
R^2	=	Predictability index
$AR(n)$	=	Auto Regressive model of order n
NN	=	Neural network
DHR	=	Dynamic Harmonic Regression
ω_c	=	Cut-off frequency
$\hat{x}(l/k)$	=	prediction of $x(l)$ based on information up to k

1 Introduction

Different approaches to wave energy extraction, either in the operating principle (oscillating bodies, oscillating water columns, etc...) and in the control technique require knowledge of some seconds of the future behavior of certain physical quantities, in order to approach optimality. These quantities may be the wave excitation force or the oscillation velocity in the case of oscillating bodies, excitation volume flux or air chamber pressure in the case of oscillating water columns, overtopping water flow in the case of an overtopping device, and so on [1],[2]. They are all strictly dependant (in some cases through a non causal transformation) on the incident wave on the device [3]. That is why short time prediction of the oncoming waves is a crucial problem in the field of wave energy, whose solution could bring great benefits to the effectiveness of the Wave Eenergy Converters (WECs) and consequently to their economical convenience.

The first approach that may be found in literature, to the best of the author's knowledge, was provided by Budal and Falnes [4] and utilises the Kalman filter to adaptively estimate the frequency, phase and amplitude of the wave excitation force acting on a heaving body, on the basis of distant pressure measurements. Very strict simplifying assumptions, requiring simple sinusoidal behavior of the excitation force and mono-directionality of wave propagation, are applied and their validity in real sea conditions is not tested. More recent solutions proposed a wave prediction computed by means linear digital filters where the inputs are either distant pressure measurements [2] or distant wave elevation [5]. Whereas the former [2] gives a very interesting approach to the design of the predicting filters (unfortunately not very relevant results were provided and hypothesis of mono-directionality is made), the study presented in [5] has the peculiarity of dealing with multi-directionality, but introduces a very high numerical complexity in the model.

Finally, it is worth mentioning the preliminary study by Voronovich [6], where the spatial prediction of wave elevation is provided by fitting an harmonic model to the distant observations, but only results obtained in simplified conditions were provided.

The approach that will be followed in this study is slightly different and was firstly presented by these authors in [7]. In particular, focus is put on forecasting the wave elevation based on past measurement collected at the prediction site itself, so that the wave elevation is treated as a univariate time series. The only study in literature following the same approach, at the best of the authors knowledge, was presented in [8], where the interest is however restricted to the prediction of some key characteristics of the wave excitation force (e.g. time until the next peak) in order to improve latching control performance, and results were presented only on simulated data.

Working on single point measurements (in particular at the location of the wave absorber) allows for some significant simplifications:

- multi-directionality does not need to be taken into account (provided that a non-directional absorber is assumed);
- wave spatial propagation laws have no effect and no hypotheses need to be made about them;
- there is no need to separate the incident wave (which is the one of interest) from the radiated wave due to the device motion.

There are some drawbacks, however, and in particular:

- the approach is not valid for directional devices (or non-directional absorbers arrays);
- it is still not clear how accurately the wave elevation can be measured at the device location.

This study deals with real sea wave observations and in section 2 an analysis of the available real data is presented, particularly the energy distribution at different frequencies, a quantification of the possible non-linearities and a quite interesting measure of predictability. Then sections 3 and 4 propose some possible forecasting models and compare the result achieved with them. Conclusions are finally outlined in section 5.

2 Analysis

The data available for this study was provided by the Irish Marine Institute and comes from a data buoy located in Galway Bay, on the West Coast of Ireland (at approximately $53^o 13' N, 9^o 18' W$). The data consists of 20 minute records sets for each hour, collected at a sampling frequency of $2.56 Hz$, for parts of years 2007 and 2008. The location is sheltered from the Atlantic Ocean so that the wave height magnitude is generally small, which makes it an ideal site for $1/4$ scale WEC prototypes.

An overall understanding of the main characteristics and properties of the waves at the observed location is provided in this section, so that the motivations for certain choices of the forecasting algorithms proposed in section 3 will be much clearer. There was no data available from different offshore locations so that no general conclusions may be drawn, but it is the opinion of these authors that the tools and the considerations in the following will be very valuable when dealing with any kind of wave elevation time series collected at any site.

In section 2.1, some overall statistics of the waves at the considered location are shown, and their properties are discussed relatively to the overall problem of wave forecasting which this paper is focused on. Then, as the water depth at the location is nearly $20 m$, the quantification of possible non-linearities that may arise due to relatively big waves or to irregularities in wind waves is discussed in section 2.2. As the aim of the study is wave forecasting treated as a time series problem, an interesting general theory about predictability is presented in section 2.3 where the feasibility of the problem is assessed without referring to any particular solution but just to the data itself.

2.1 Fourier analysis

The main tool for a first analysis of the waves is their spectral distribution, the *wave spectrum*, which shows how much energy is distributed at different frequency components of the wave, which are supposed to be completely independent of each other. Although offering a limited time-averaged information (a Wavelet transform would offer a more complete information in the time domain [9]) it is still very valuable in order to provide some overall characteristics of the sea conditions in different situations.

A first analysis, which is interesting to carry out, over the available hourly data sets, concerns the distribution of the significant wave height H_s and the peak and mean radian frequency of the spectrum, respectively ω_{peak} and ω_{mean}, and to assess if their behaviors are correlated to each other in some way. The significant wave height is a measure of the mean energy contained in the wave, while the ω_{peak} and the ω_{mean} can be a way to represent where the spectrum (and so the energy) of the wave is more concentrated. From Fig.1, it is clear how high energy wave systems show a much lower spread of the spectrum, centered at a low frequency (about $1 \, rad/s$), consisting of a well defined narrow peak (swell). The lower the energy, on the other hand, the more the distance between the peak and mean frequency, which denotes a much flatter spectrum where the high frequency wind waves have a similar energy content to the low frequency swell. The two sample spectra of Fig. 2 are particularly illustrative in this respect.

2.2 Non-linearity analysis

Ocean waves, like most of the systems in the real world, are not linear, and it would be helpful and valu-

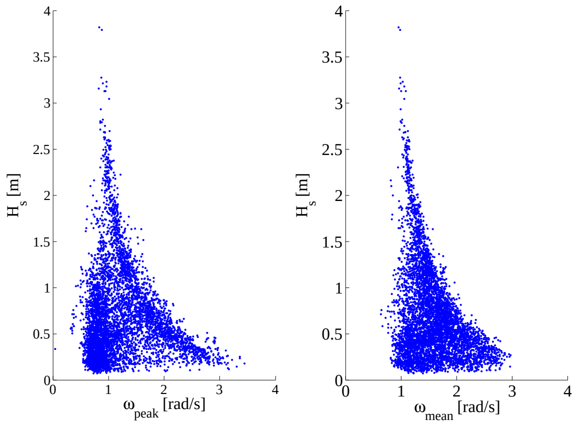

Figure 1: Big waves systems (high H_s) present well defined low frequency swells, while in low energy waves the high frequencies (wind waves) are also quite significant and can contain more energy then the swell.

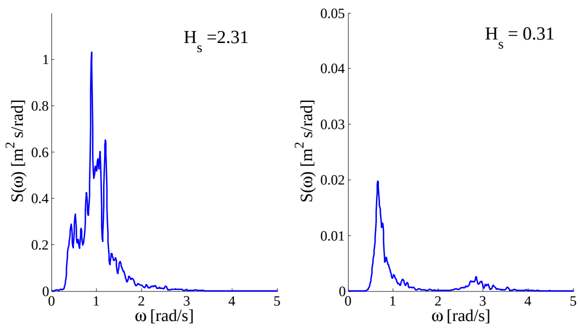

Figure 2: Sample high and low energy spectra, respectively occurred on the 5^{th} of January and on the 1^{st} of May 2007, at the Galway Bay location.

able to quantify how far from linearity they are so that, in the particular case of wave forecasting, an appropriate model can be chosen. Linearity in the case of waves implies linear superposition of harmonic components (sines and cosines), so that the distribution of the wave elevation results to be perfectly symmetric with respect to zero. While in the case of deep water this assumption is reasonably valid (the wave elevation distribution approximates a Gaussian [10]), in shallow water locations (wave length comparable to the water depth) higher order terms should be taken into account (refer to Stokes [11]) and their effect is to produce higher and narrower peaks then troughs, so that the distribution in not Gaussian any more. A statistical analysis of higher order momentum [12] (note that a Gaussian distribution is completely defined up to the second order momentum) can be utilised to detect the significance of such a non-linearity. Fig. 3 shows the skewness and kurtosis indices computed for each available data set and measuring, respectively the asymmetry of the distribution (null skewness denotes perfect symmetry [12]) and the peakedness of the distribution (a Gaussian distribution has a kurtosis equal to 3 [12]).

There is another possible non-linearity to take into account, which unfortunately is less quantifiable and can only be analysed through visual inspection. This is due to the interactions occurring between different harmonic components of the wave system, which are neglected in classical linear wave theory and in Fourier-Wavelet analysis. An higher order spectral analysis through the bispectrum [10] revealed to be quite effective in order to detect these interactions, but as previously stated, a real quantification would be hard to carry out and probably not really significant. This non-linearity is known to be more present in wind waves, that is at high frequency and low energy, that are less interesting from a wave energy point of view. A low-pass filtering of the wave elevation time series, in particular, may help to reduce their effect so that they should not be taken into account in the forecasting model. As an example compare Fig. 4 and Fig. 5, where the bispectrum is shown for two sample data sets. Relevant portions in off-diagonal parts of the graphs indicate interactions between different frequencies, and it is clear how in the low energy system very significant energy exchanges appear between high and low frequency wave systems. In the case of an high energy wave, on the other hand, most of the bispectrum is concentrated near the 45 degrees line, so that no relevant exchanges of this sort are expected.

It is important to say, however, that both the effects are not expected to be as relevant in deep water off-shore locations, where wave energy devices would most likely be located.

2.3 Predictability

As the focus of this study is on the multi-step-ahead prediction of the wave elevation time series, it would be a very valuable information to analyse the *predictability* of the time series, that is how accurately it can be predicted

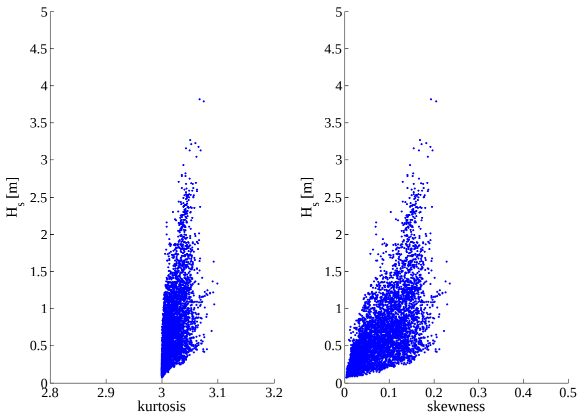

Figure 3: Third and fourth order statistical analysis, respectively through indices of kurtosis and skewness, for all the available data sets.

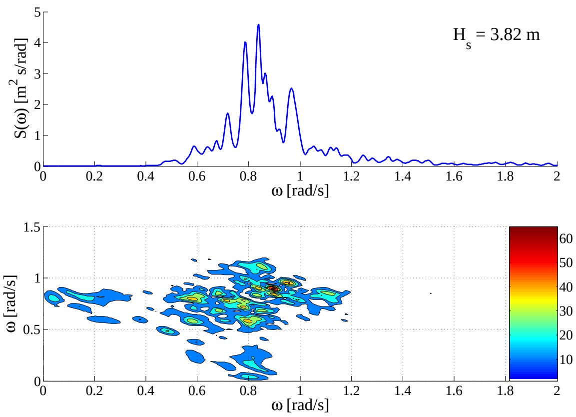

Figure 4: Bispectrum for an high energy wave system mostly concentrated around the 45 degrees line, apart from some small interactions revealed between the peak frequency and very low frequencies ($0.2-0.3$ rad/s).

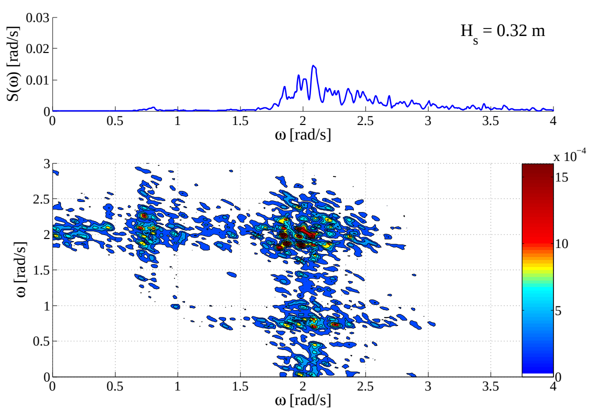

Figure 5: Bispectrum in the case of small waves shows significant interactions between high and low frequency components.

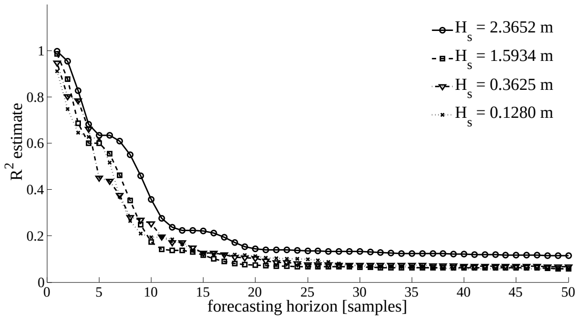

Figure 6: Predictability indices $R(k)^2$ estimated for data sets with different energy. It dies down very quickly in any case, with the highest energy wave system showing a slightly better predictability.

based only on past values, without making any assumptions on the actual model behind it or the forecasting techinque which will eventually be adopted. Such a measure is based, in the most general case, on the amount of information that past behavior of the signal contains about its future values. A simpler measure of predictability then the very general approach proposed in literature (based on the mutual information notion [13]) will be adopted here, which supposes that a linear relationship exists that relates the future values of the wave elevation to the past. This is, of course, a limiting assumption but it is still effective, as it will be shown, to provide at least some qualitative deductions. In particular, a predictability index $R^2(k)$ is estimated, defined as the ratio of the variance of the optimal k-step-ahead prediction, $\hat{\eta}(t+k/t)$, to the variance of the real wave elevation, $\eta(t)$:

$$R^2(k) \triangleq \frac{E\{\hat{\eta}(t+k/t)^2\}}{E\{\eta(t)^2\}} = 1 - \frac{\hat{\sigma}_k^2}{E\{\eta(t)^2\}} \qquad (1)$$

where it is supposed that the wave elevation $\eta(k)$ has a zero mean and, in the second formulation, the optimal k-step-ahead prediction error variance, $\hat{\sigma}_k^2 \triangleq E\{\hat{e}(t+k/t)^2\}$, is introduced. A very efficient algorithm for the estimation of $R(k)^2$, under the assumption of a linear univariate time series, was proposed in [14] and it is adopted here for the analysis of the available wave data. Fig. 6 shows the estimated predictability index $R^2(k)$, for a forecasting horizon of almost 20 seconds (exactly 50 samples), of four wave systems of different energy, which is expressed in terms of the significant wave height H_s. As expected from any real world time series, it is a non-increasing function of the prediction horizon. All the wave systems show a relatively poor predictability, which dies out very quickly after $2-4$ seconds ($5-10$ samples), with a slightly better behavior of the highest energy waves. This would be expected because of the better regularity of the high energy wave components, which corresponds to low frequency waves (as seen in section 2.1), while low energy and high frequency waves are more affected by non-linearities and irregularities (strong energy exchanges with low frequency swell, from the bispectrum of Fig. 5).

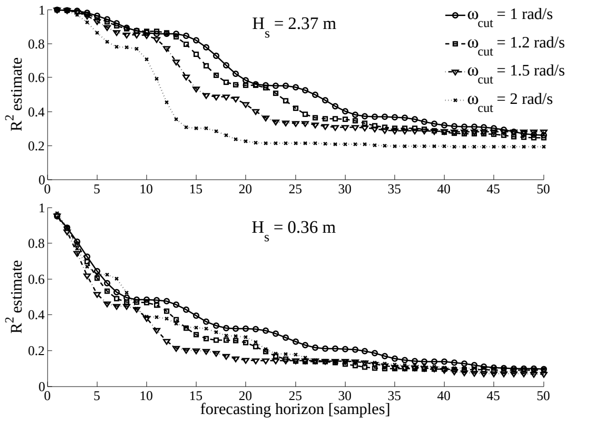

Figure 7: Predictability of a high energy and a low energy data set when low-pass filtering with different cut-off frequencies ω_c is applied. The better predictability of low frequency components is clear in the case of high energy systems.

In a wave energy context, however, one might be interested in forecasting only the high energy components, so that a low pass filter can be applied to the time series and a focus would be put exclusively on the low frequency components. In Fig.7, the estimated predictability index $R(k)^2$ is shown for the pre-filtered wave system with the highest energy ($H_s = 2.3652\,m$) when different cut-off frequencies ω_c are applied. It is clear how the overall predictability significantly improves with respect to the non-filtered waves. Moreover, the smaller the cut-off frequency, i.e. the lower the frequencies we limit the analysis to, the better the predictability of the time series, so that more accurate predictions, and further in the future, should be expected. The same improvement is not shown in the case of low energy waves, and this may be explained with the fact that low-pass filtering cuts out most of the energy of the signal, so that the harmonic components left in it have relatively small amplitude. This, however, might not be a problem in a wave energy context, as the actual energy that is lost, although being a great part of all the available energy, might still represent a reasonable and negligible loss compared to high energy wave systems.

3 Models

3.1 Cyclical models

From linear wave theory [1], a real ocean sea state may be modelled as a linear superposition of waves with different frequencies and propagating in different directions:

$$\eta(x,y,t) = \int_0^{+\infty} d\omega \int_{-\pi}^{+\pi} A(\omega,\beta)\cos(\omega t - kx\cos\beta - ky\sin\beta + \varphi_i(\omega))d\beta \tag{2}$$

where k is the wave number and β represents the direction of propagation in the x-y plane. If a specific location (x_0, y_0) is considered, the following simplified expression can then be obtained:

$$\eta(x_0,y_0,t) = \int_0^{+\infty} d\omega \int_{-\pi}^{+\pi} A(\omega,\beta)\cos(\omega t + \phi(\omega,\beta)) \tag{3}$$

where the directionality information is obviously lost and the constant terms $kx_0\cos\beta$ and $ky_0\sin\beta$ are included in the phase $\phi(\omega,\beta)$.

From this knowledge about the real process it is quite straightforward to choose, as a forecasting model for the wave elevation, a simple cyclical model, as it was also presented in [7], where the frequency domain is of course discretised:

$$\eta(t) = \sum_{i=1}^{m} a_i\cos(\omega_i t) + b_i\sin(\omega_i t) + \zeta(t) \tag{4}$$

An error $\zeta(t)$ has been introduced and the phase and amplitude information for each harmonic component is now contained in the parameters a_i and b_i.

The model (4) is completely characterised by the parameters a_i,b_i and by the frequencies ω_i. It could then be fitted to the data through some non-linear estimation procedure (the model is non-linear in the frequencies in particular) and utilised to predict the future behavior of the wave elevation time series. It needs, however, to be adapted to the time variations of the wave spectrum, which is non-constant at all, so that the first approach [7] has been to choose the frequencies in the model design phase and to keep them constant during its utilisation and estimation. In this way the model becomes perfectly linear in the parameters a_i,b_i and can be easily estimated and on-line adapted to the spectral variations of the sea.

The problem of choosing the frequencies can be divided in two sub-problems:

1. **Choice of the range:** This is a quite easy matter, as statistical information about the location can be utilised to properly define an upper and lower bound for the range. At this point, one may decide to include the range of higher frequencies where the low energy wind waves are, or to simply consider a narrower range including only the swell.

2. **Distribution of the frequencies in the range:** A robust choice would be a constant spacing between the frequencies over all the range, but a more efficient non-homogeneous distribution was also proposed in [7]. The latter however suffer from the problem of specificity, so that if the wave spectrum changes the frequencies might not be appropriate any more. If the frequencies are kept constant, then it would not be a proper choice.

Once the frequencies are determined, a model for the amplitudes has to be chosen. In [7], it was pointed out how they have to be adaptive to the wave, as constant amplitudes gave very poor results. Two adaptive models are proposed here, in particular:

Structural model: based on Harvey's structural model [15], the model (4) is expressed in the following discrete time form:

$$\eta(k) = \sum_{i=1}^{m} \psi_i(k) + \zeta(k) \tag{5}$$

$$\begin{bmatrix} \psi_i(k+1) \\ \psi_i^*(k+1) \end{bmatrix} = \\ = \begin{bmatrix} \cos(\omega_i T_s) & \sin(\omega_i T_s) \\ -\sin(\omega_i T_s) & \cos(\omega_i T_s) \end{bmatrix} \begin{bmatrix} \psi_i(k) \\ \psi_i^*(k) \end{bmatrix} + \\ + \begin{bmatrix} w_i(k) \\ w_i^*(k) \end{bmatrix}, \quad i = 1, \dots m \tag{6}$$

where it can be verified that $\psi_i(0) = a_i$ and $\psi_i^*(0) = b_i$. From equation (6), then, the following state space form, which is more familiar to work with, is easily derived:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{w}(k) \\ \eta(k) &= \mathbf{C}\mathbf{x}(k) + \zeta(k) \end{aligned} \tag{7}$$

where

$$\mathbf{x}(k) \triangleq [\psi_1(k) \quad \psi_1^*(k) \quad \dots \psi_m(k) \quad \psi_m^*(k)]^T \tag{8}$$

$$\mathbf{w}(k) \triangleq [w_1(k) \quad w_1^*(k) \quad \dots w_m(k) \quad w_m^*(k)]^T \tag{9}$$

$$\mathbf{A} \triangleq diag \left\{ \begin{bmatrix} \cos(\omega_i T_s) & \sin(\omega_i T_s) \\ -\sin(\omega_i T_s) & \cos(\omega_i T_s) \end{bmatrix} \right\} \tag{10}$$

$$\mathbf{C} \triangleq [1 \quad 0 \quad 1 \quad 0 \quad \dots \quad 1 \quad 0] \in \Re^{1 \times 2m} \tag{11}$$

Dynamic Harmonic Regression (DHR): Introduced by Young [16], it expresses a cyclical model of the type of eq. (4), where the a_i and b_i parameters evolve according to a Generalised Random Walk:

$$\begin{bmatrix} x_i(k+1) \\ x_i^*(k+1) \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ 0 & \gamma \end{bmatrix} \begin{bmatrix} x_i(k) \\ x_i^*(k) \end{bmatrix} + \begin{bmatrix} \delta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_i(k) \\ \varepsilon_i^*(k) \end{bmatrix} \\ x_i = a_i \quad \text{for } i = 1, \dots m \\ x_{i-m} = b_i \quad \text{for } i = m+1, \dots 2m \tag{12}$$

where x_i^* models a slope for the evolution of each parameter x_i. The disturbance terms ε_i and ε_i^* are still assumed to be Gaussian noises and introduce the variability in the model. A particular form of (12) was implemented in this study where the dynamic matrices are chosen in order to represent Harvey's local linear trend [15]:

$$\begin{bmatrix} x_i(k+1) \\ x_i^*(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_i(k) \\ x_i^*(k) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_i(k) \\ \varepsilon_i^*(k) \end{bmatrix} \tag{13}$$

for $i = 1, \quad 2, \quad \dots \quad 2m$. A state space form, then, can easily be derived, resulting in the following model:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \varepsilon(k) \\ \eta(k) &= \mathbf{C}(k)\mathbf{x}(k) + \zeta(k) \end{aligned} \tag{14}$$

where

$$\mathbf{x}(k) \triangleq [x_1(k) \quad x_1^*(k) \quad \dots x_{2m}(k) \quad x_{2m}^*(k)]^T \tag{15}$$

$$\varepsilon(k) \triangleq [\varepsilon_1(k) \quad \varepsilon_1^*(k) \quad \dots \varepsilon_{2m}(k) \quad \varepsilon_{2m}^*(k)]^T \tag{16}$$

$$\mathbf{A} \triangleq diag \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right\} \in \Re^{4m \times 4m} \tag{17}$$

$$\mathbf{C}(k) \triangleq [\cos(\omega_1 T_s) \quad 0 \quad \dots \quad \cos(\omega_m T_s) \quad 0 \\ \sin(\omega_1 T_s) \quad 0 \quad \dots \quad \sin(\omega_m T_s) \quad 0] \tag{18}$$

Both the models have the advantage of a state space representation, which is particularly suited to the application of the Kalman filter for a recursive on-line adaption. The initialisation is provided through means of regular least squares on a number of past observations and then the Kalman filter is applied on-line, once a proper covariance matrix for the state and output disturbances is provided. When the estimate of the model's parameters, $\hat{x}(k/k)$, is available at any instant k, the l-steps-ahead prediction $\hat{\eta}(k+l/k)$, based on the information up to k, is obtained through the free evolution of the model:

$$\hat{\eta}(k+l/k) = C(k+l)A^l \hat{x}(k/k) \tag{19}$$

There are, however, some strong limitations to this approach with cyclical models, that also emerged in [7], and that will be highlighted also in the results, section 4:

- The use of constant frequencies requires, for the sake of robustness, a dense and complete set, which adds considerable complexity to the model, and
- It is not clear how to choose the covariance matrices for the Kalman filter implementation

In the next section 3.2, it will be shown how AR models implicitly overcome these difficulties in a very effective, and simple, way.

3.2 Auto Regressive (AR) models

As a pure time series problem is under study, there is the advantage of the existence of a well established theory, from the time series field, which it is possible to utilise as well. As a comparison with the cyclical models, where the a priori knowledge that we have about the real system is explicitly taken into account, it is particularly interesting to analyse the properties of classical AR models.

The wave elevation $\eta(k)$ is supposed to be linearly dependant on a number n of its past values:

$$\eta(k) = \sum_{i=1}^{n} a_i \eta(k-i) + \zeta(k) \tag{20}$$

where a disturbance term $\zeta(k)$ has been also included. If the parameters a_i are estimated and the noise is supposed to be Gaussian and white, the best prediction of the future wave elevation $\hat{\eta}(k+l/k)$ at instant k is then given by:

$$\hat{\eta}(k+l/k) = \sum_{i=1}^{n} \hat{a}_i(k)\hat{\eta}(k+l-i/k) \tag{21}$$

where, obviously, $\hat{\eta}(k+l-i/k) \equiv \eta(k)$ if $k+l-i \leq k$ (i.e. the information is already acquired and there is no need of prediction).

The properties of such a very simple forecasting model become clearer if an explicit solution of the difference equation (21) is provided [17]:

$$\hat{\eta}(k+l/k) = \sum_{i=1}^{n} b_i(k) f_i(l) \quad (22)$$

Here, the coefficients $b_i(k)$ depend only on the forecasting origin (so they stay constant at each instant for the complete prediction time horizon) and are function of the initial conditions (the past n observations), whereas $f_i(l)$ are functions of the lead time l and, in general, they include damped exponential and damped sinusoidal terms completely determined by the roots p_i of the transfer function $\varphi(z)$ describing eq. (20) in the Z-domain:

$$\eta(z) = \frac{\zeta(z)}{\varphi(z)} \triangleq \frac{\zeta(z)}{\prod_{i=1}^{n}(z-p_i)} \quad (23)$$

The general shape of the prediction function is therefore completely determined by the poles, p_i, while the particular realisation of this general structure is determined, at each sampling instant, by the past values of the time series. It is particularly interesting to analyse the shape of the forecasting function (22) in the case of $m/2$ (when m is even) couples of complex-conjugate poles p_i and p_i^*:

$$\hat{\eta}(k+l/k) = \sum_{i=1}^{m/2} c_i(k)|p_i|^l \sin(\angle p_i k + \varphi_i(k)) \quad (24)$$

An AR model with only complex-conjugate poles is *implicitly* a cyclical model, where the frequencies are related to the phase, $\angle p_i$, of each pole and the amplitude and phase of each harmonic component are related to the last n observations of each time instant k, so that they adapt to the observations.

Note, then, that an adaptivity mechanism is already present even if the AR model is only estimated once on a batch data set. Only the frequencies are fixed, while amplitudes and phases are automatically updated on the basis of the recent past information.

A further degree of adaptivity can be introduced with an on-line estimation of the AR model parameters, a_i, which would introduce an on-line adaptivity of the frequencies as well. This is not considered in this paper, however, as it would go beyond its main scope.

3.3 Neural networks

It was shown in section 2.2 how the non-linearities appearing in the big low frequency waves, due to the relatively small water depth, are not really relevant. The study may therefore end with the cyclical and AR models provided through sections 3.1 and 3.2, particularly if the high frequency components are filtered out. It is however interesting, in the authors opinion, looking at a comparison with a most widespread tool for time series modelling and forecasting such as neural networks.

For the problem under study, a non-linear relationship of the following type is created through a multilayer perceptron [18]:

$$\eta(k) = \mathbf{NN}(\eta(k-1), \eta(k-2), \ldots \eta(k-n)) \quad (25)$$

so that the dependance between the current wave elevation and n past values is realised. The model is then trained through the back propagation algorithm on a set of batch data and utilised for multi-step-ahead prediction.

This is, of course, not the only possibility and many others could be considered. For example, a priori knowledge about the process (which would always be a more appropriate approach) may be included and a non-linear relationship of the following type may be considered instead:

$$\eta(k) = \mathbf{NN}(\cos(\omega_1 T_s k + \varphi_1), \ldots \cos(\omega_n T_s k + \varphi_n)) \quad (26)$$

but some of the limitations outlined in section 3.1, when cyclical models where considered, due to an appropriate choice of the frequencies, are still present. Here, there is a greater problem of how to consider the initial phases φ_i of the input harmonics, so the possibility was discarded.

In section 4, results will be shown and compared with the cyclical and AR models, for different neural network topologies, with two hidden layers and different numbers of inputs (regression order n).

4 Results

The possible forecasting models proposed in section 3 were tested on a significant sample data set, appropriately chosen among all those available (refer to section 2), representing different sea conditions. In particular, a single 20 minute data set for the 5^{th} of January 2007 and the 1^{st} of May 2007 have been chosen to provide differences between a high energy situation with a well defined swell and a low energy case, where the small swell is comparable with the high frequency wind waves (refer to Fig. 2).

The prediction accuracy is measured with the following goodness-of-fit index, which depends on the forecasting horizon l:

$$fit(l) = \left(1 - \frac{||\eta(k+l) - \hat{\eta}(k+l/k)||_2}{||\eta(k)||_2}\right) \cdot 100 \quad (27)$$

Here $||\cdot||_2$ is the Euclidean norm operator (root sum squared) over all the sampling instants k of the simulation (a 20 minutes data set), $\eta(k+l)$ is the wave elevation and $\hat{\eta}(k+l/k)$ is its prediction based on the information up to instant k. A 100% value for $fit(l)$ means that the wave elevation time series is perfectly predicted l steps in the future.

Starting from the cyclical models outlined in section 3.1, Table 1 shows their ability to predict the January data set with different choices of the frequency spacing $d\omega$ and for both the Harvey and the DHR model. The cases when almost all the spectrum is considered

$H_s = 2.31m$		$\omega_c = 2rad/s$	$\omega_c = 1.2rad/s$	
$model$	$d\omega$	$fit(5)$	$fit(5)$	$fit(10)$
Harvey	0.1	12.54%	37.13%	25.49%
	0.05	19.89%	41.91%	30.94%
	0.01	24.97%	48.54%	31.28%
DHR	0.1	−59.87%	−47.43%	−88.70%
	0.05	−59.01%	−46.59%	−86.29%

Table 1: Cyclical models on January high energy data set

$H_s = 0.31m$		$\omega_c = 3.5rad/s$	$\omega_c = 1.2rad/s$	
$model$	$d\omega$	$fit(5)$	$fit(5)$	$fit(10)$
Harvey	0.1	−10.27%	39.47%	23.44%
	0.05	−4.90%	45.78%	31.55%
	0.01	4.66%	55.82%	42.63%
DHR	0.1	−40.58%	−35.25%	−100.63%
	0.05	−40.19%	−33.64%	−96.67%

Table 2: Cyclical models on May low energy data set

(cut-off frequency $\omega_c = 2rad/s$) or when part of the lower energy spectrum is filtered out ($\omega_c = 1.2rad/s$) are shown. It is clear how the performance is quite poor when a wider range of frequencies is considered, while better prediction is achieved when focusing only on the low frequencies (together with a reduced complexity due to the narrower range considered). The accuracy, however, never goes past 50% for 5 sample-ahead predictions, corresponding to nearly 2 seconds. Moreover, the complexity of the models can be very high, e.g. a spacing of $d\omega = 0.01$ with a range $[0.3, 1.2]\,rad/s$ generates a state space model of order 182 for the Harvey's cyclical model and 364 for the DHR! Consequently, it can be concluded that, although different adjustments of the estimation procedure (particularly the initial choices for the covariance matrices required by the Kalman filter) may lead to better results, the solution does not seem to be very valuable.

Moving to AR models, then, Table 3 and Table 4 show a far better accuracy (a comparison with cyclical models is depicted in Figure 9), which improves with the regression order and if only low frequencies are considered. In particular, acceptable predictions may be obtained up to 30 samples in the future (nearly 12 seconds) when only the low frequencies are considered, particularly with regression orders 16 and 32 (higher orders did not show any significant improvement), so that the model remains computationally light with respect to the cyclical models. In the case of the low energy wave system the accuracy is even better then for high energy data (January) if only the low frequencies are considered. This is probably due to the non-linearities appearing in big waves at the considered location, as explained in section 2.2.

It is worth noting that these results have been obtained with *static* AR models, estimated only from a batch time series with regular least squares, and no adaptivity or recursive on-line estimation has been implemented (which was fundamental for the cyclical models). Here, in fact, as mentioned when describing the AR model in section 3.2, it was pointed out how also a static

$H_s = 2.31m$	$\omega_c = 2rad/s$	$\omega_c = 1.2rad/s$	
$model$	$fit(5)$	$fit(5)$	$fit(30)$
8	92.31%	97.94%	0.57%
16	98.31%	97.96%	53.38%
32	98.31%	97.96%	67.94%

Table 3: static AR models on January high energy data set

$H_s = 0.31m$	$\omega_c = 3.5rad/s$	$\omega_c = 1.2rad/s$	
$model$	$fit(5)$	$fit(5)$	$fit(30)$
8	40.47%	98.97%	24.95%
16	80.15%	99.00%	77.77%
32	98.53%	99.00%	81.96%

Table 4: static AR models on May low energy data set

AR(n) model provides cyclical components with amplitudes and phases time-varying on the basis of the last n observations.

It is particularly interesting also to show, Fig. 8, for the AR(16) model estimated on the January data set when $\omega_c = 1.2rad/s$, how its poles all lie approximately on the unit circle in the Z-plane and the corresponding frequencies are all contained in the significant part of the spectrum (in the interval $[0.3\,,\,1.2]\,rad/s$). This occurs in every situation as soon as the AR order is kept reasonably low. For order 32, for example, some of the frequencies, depending on the data set, are estimated out of the expected bound and, more in particular, close to the maximum $\pi * f_s$.

The focus can now be moved to the results obtained with neural networks, to see if they can better the performance of the simple AR models. Some structures where trained and the results obtained are shown in Table 5 and Table 6, respectively, for the January and May sample data sets. A comparison of these results with Table 3 and Table 4 regarding AR models highlights how neural networks offer great accuracy over short forecasting horizon (5 samples), but they reveal problems when predicting further in the future, as it can be clearly seen in Fig. 10, where 20-step-ahead forecasts obtained with an AR(16) model are also plotted as a comparison. The main reason for this behavior lies in the fact that the AR model implicitly takes into account of the real process cyclical dynamics and it is easier to choose the structure with reference to these characteristics, while neural networks are

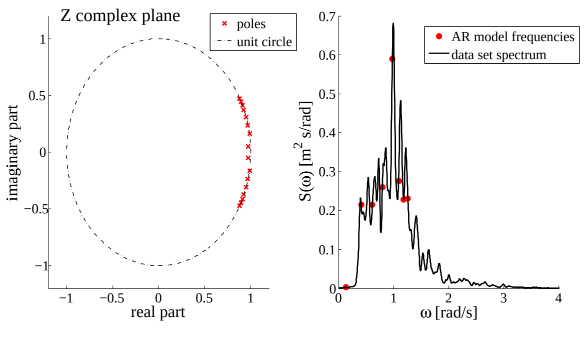

Figure 8: Poles and corresponding frequencies of the AR(16) model estimated on the January data set when only the frequencies up to $\omega = 1.2\,rad/s$ are considered.

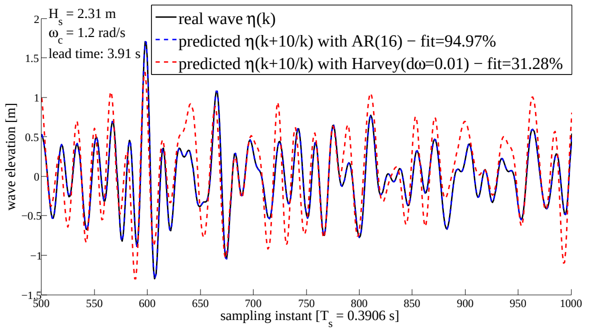

Figure 9: For 10 samples ahead the prediction with an AR(16) model is almost perfect, whereas a Harvey cyclical model shows a good prediction of the phases but an overestimation of the amplitudes.

$H_s = 2.31m$		$\omega_c = 2rad/s$	$\omega_c = 1.2rad/s$	
inputs	*structure*	$fit(5)$	$fit(5)$	$fit(20)$
10	$3-5-1$	76.29%	93.90%	4.40%
	$3-7-1$	90.71%	92.90%	0.88%
	$4-6-1$	87.55%	86.05%	−70.52%
15	$3-5-1$	83.11%	83.75%	−151.79%
	$3-7-1$	80.49%	96.99%	40.81%
	$4-6-1$	78.29%	93.23%	11.45%

Table 5: Neural networks results on January high energy data set

a pure black box where the choices for the regression order, the structure and the estimation algorithm can only be guessed and improved with experimental simulations.

More effort can be put into finding a proper structure for a neural network which may be comparable or even more accurate than an AR model, but it is the authors opinion that before undertaking such a task, the needs and the requirements for the forecasts must be specified so that it is possible to evaluate what is, in real applications of wave energy (e.g. control and optimisation of wave energy converters), the required accuracy.

A clearer comparison between the performance of some of the different models is shown in Fig. 11 and in Fig. 12, where the goodness-of-fit index is shown for all the forecasting horizons from 1 to 50 samples (approximately 0.39 to 19.53 seconds).

5 Conclusion

This study was focused on the problem of short term wave prediction, which is a central topic in the wave energy field, in order to allow a better effectiveness and economic viability of any WEC. It was treated as a pure

$H_s = 0.31m$		$\omega_c = 3.5rad/s$	$\omega_c = 1.2rad/s$	
inputs	*structure*	$fit(5)$	$fit(5)$	$fit(20)$
10	$3-5-1$	47.85%	93.22%	−11.42%
	$3-7-1$	48.63%	96.91%	46.01%
	$4-6-1$	49.99%	95.53%	16.65%
15	$3-5-1$	55.73%	98.49%	−57.89%
	$3-7-1$	54.31%	94.61%	31.55%
	$4-6-1$	49.09%	94.25%	12.80%

Table 6: Neural networks on May low energy data set

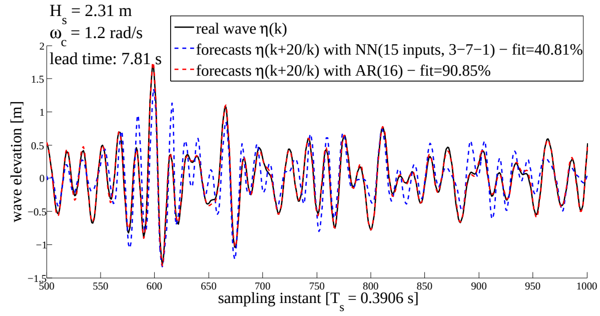

Figure 10: The 20-sample-ahead prediction of an high energy wave system, with a $\omega_c = 1.2\,rad/s$, for an AR(16) model is almost perfect, and outperforms a neural networks with 15 inputs and a $3-5-1$ structure.

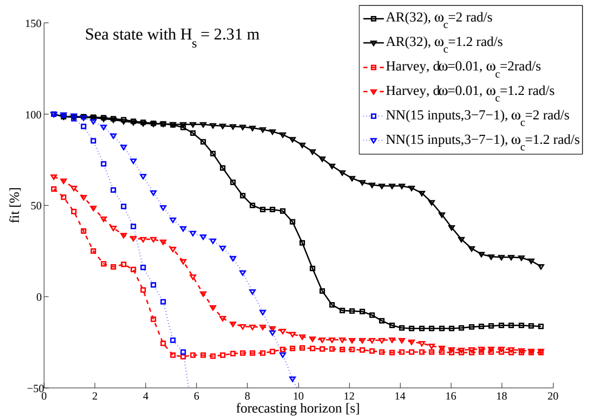

Figure 11: Performance of some of the models, over different forecasting horizons, for the January high energy data set.

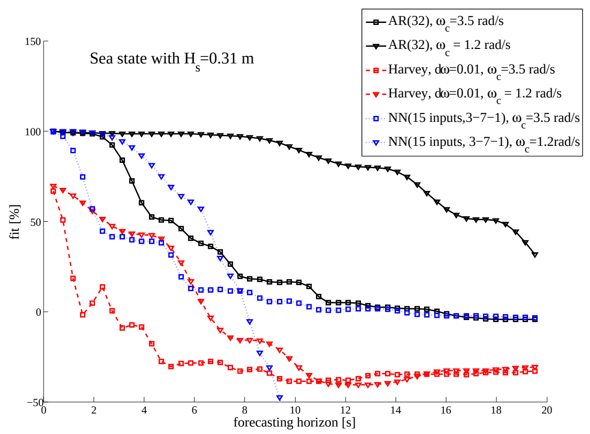

Figure 12: Performance of some of the models, over different forecasting horizons, for the May low energy data set.

univariate time series forecasting problem and several possible solutions were proposed. Real data from the Galway bay were available for testing the proposed solutions, and some interesting analysis was provided in section 2. In particular, a very valuable tool for the predictability analysis, independently from any particular solution, was proposed in section 2.3, whose application showed how lower frequency waves are easier to predict and, from a wave energy point of view, high frequency components, which carry lower energy (as revealed by the Fourier analysis provided in section 2.1), should be filtered out before the prediction.

The most straightforward models outlined were harmonic models where the wave elevation is explicitly represented as a sum of sines and cosines, on the basis of linear wave theory. It was underlined how many issues (particularly the high complexity of the resultant models) arise due to the problem of the choice of frequencies when they are kept constant, so that reasonable predictions are only achieved for 5-10 samples (2-4 seconds) in the future, if only low frequencies are predicted. Cyclical models with adaptive frequencies could have been considered, but then they become non-linear and the complexity will be even higher.

Then, an analysis of AR models, in section 3.2, highlighted how they implicitly represent cyclical models where the frequencies are easily estimated with linear least squares (as they are related to the regression coefficients). The amplitudes and phases of each harmonic component is, moreover, implicitly adaptive to the recent observations due to the regression terms of the model, so that only a batch estimate of the model offered very good accuracy up to 30 samples-ahead (almost 12 seconds) predictions for the low frequency components of the waves. It was shown also how the frequencies are automatically estimated in the significant range of the sample spectrum of the training data set.

A comparison with neural networks, finally, showed how it would not be very appealing, in the authors opinion, to further undertake this more complicated direction, even because they do not offer any possibility of analysis and extraction of the characteristics of the real process from the model, which would instead be very straightforward with AR models.

It is fundamental, of course, that further work should be made in order to provide some indications and constraints about the required accuracy of the forecasts (and required prediction horizon), so that the capability of the proposed models can be properly judged. Such a work will involve a study of the inter-connections between wave absorbers, wave excitation and control architecture, and will be fundamental before any further attempt to improve the results of this work is eventually undertaken.

Acknowledgements

The authors are grateful to the Irish Marine Institute for providing real sea observations. In addition, the funding provided under the IRCSET scheme is acknowledged.

References

[1] J. Falnes. *Ocean Waves and Oscillating Systems*. Cambridge University Press, 2002.

[2] J. Tedd and P. Frigaard. Short term wave forecasting, using digital filters, for improved control of wave energy converters. *The Proceedings of International Offshoree and Polar Engineering (ISOPE)*, pages 388, 394, 2007.

[3] J. Falnes. On non-causal impulse response functions related to propagating water waves. *Applied Ocean Research*, 17:379–389, 1996.

[4] K. Budal, J. Falnes, et al. The norwegian wave-power buoy project. *The Second International Symposium on Wave Energy Utillization*, Trondheim, June 1982.

[5] M. R. Belmont et al. Filters for linear sea-wave prediction. *Ocean Enginnering*, 33:2332–2351, 2006.

[6] B. Holmes V.Voronovich and G. Thomas. A preliminary numerical and experimental study of wave prediction. *Proc. 6th European Wave and Tidal Energy Conference*, 2005.

[7] F. Fusco and J. Ringwood. Linear models for short term wave forecasting. *Proc. World Renewable Energy Congress (WRECX)*, pages 1030–1035, 2008.

[8] A.A.E. Price and A.R. Wallace. Non-linear methods for next wave estimation. *Proceedings of the 7th European Wave and Tidal Energy Conference, Porto*, 2007.

[9] G. Nolan, J.V. Ringwood, and B. Holmes. Short term wave energy variability off the west coast of Ireland. *Proceedings of the 7th European Wave and Tidal Energy Conference, Porto, Portugal (EWTEC)*, 2007.

[10] M. K. Ochi. *Ocean waves. The stochastic approach*. Cambridge University Press, 1998.

[11] G.G. Stokes. On the theory of oscillatory waves. *Transactions of the Cambridge Philosophical Society*, 8:441–455, 1847.

[12] M. K. Ochi. *Applied probability and stochastic processes*. Whiley Inter-science, 1990.

[13] C.E. Shannon. The mathematical theory of communication. *Bell System Technical Journal*, (27):379–423, 1948.

[14] X. Hong and S. A. Billings. Time series multistep-ahead predictability estimation and ranking. *Journal of Forecasting*, 18:139–149, 1999.

[15] A. C. Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, 1989.

[16] P. C. Young, Diego J. Pedregal, and Wlodek Tych. Dynamic harmonic regression. *Journal of Forecasting*, 18:369–394, 1999.

[17] G. M. Jenkins G. E. P. Box and G. C. Reinsel. *Time Series Analysis: Forecasting and Control*. Prentice-Hall, 1994.

[18] M. Norgaard, O. Ravn, N.K. Poulsen, and L.K. Hansen. *Neural Networks for Modelling and Control of Dynamic Systems*. Springer, 2000.