Wave Prediction using Wave Rider Position Measurements and NARX Network In Wave Energy Conversion

Mohammed A. A. Desouky[1a], Ossama Abdelkhalik[2b]

[a]*Michigan Technological University, MI, USA*
[b]*Iowa State University, IA, USA*

Abstract

Several control methods of wave energy converters (WECs) need prediction in the future of wave surface elevation. Prediction of wave surface elevation can be performed using measurements of surface elevation at a location ahead of the controlled WEC in the upcoming wave. Artificial neural network (ANN) is a robust data-learning tool, and is proposed in this study to predict the surface elevation at the WEC location using measurements of wave elevation at ahead located sensor (a wave rider buoy). The Nonlinear Autoregressive with exogenous input network (NARX NN) is utilized in this study as the prediction method. Simulations show promising results for predicting the wave surface elevation. Challenges of using real measurements data are also discussed in this paper.

Keywords: Wave Energy Conversion, Wave Energy, Wave Prediction

1. Introduction

Wave forecasting is vital for marine and coastal activities. Several approaches are proposed in the literature for wave parameters prediction, including empirical, numerical and data learning methods.

[1]PhD student, Michigan Technological University, ME-EM department, 1400 Townsend Dr., 815 R. L. Smith ME-EM Building, Houghton, MI 49931, madesouk@mtu.edu
[2]Associate Professor, Iowa State University, Email: ossama@iastate.edu

Empirical models for wave prediction include the well-known JONSWAP empirical model [1]. Numerical approaches for wave parameters prediction usually use the following wave energy-balance equation:

$$\frac{\partial E\left(f,\theta,t\right)}{\partial t}+\mathcal{V}.\triangledown E=S=S_{in}+S_{nl}+S_{ds}, \tag{1}$$

where $E\left(f,\theta,t\right)$ is a two dimensional wave spectrum, f is the frequency, θ is the propagation direction, $\mathcal{V}=\mathcal{V}(f,\theta)$ is the (deep-water) group velocity, S is the net source function, S_{in} is the external wave making factors such as local wind and local current, S_{nl} is the non-linear energy transfer by wave-wave interactions, and S_{ds} is the dissipation[2]. Several developments were dedicated to improve the wave modeling through solving the energy-balance equation [3]. Statistical and data learning approaches include the soft-computing methods such as machine learning, genetic algorithms and fuzzy interface systems.

Wave surface elevation and the significant wave height predictions are of particular importance especially in the control problem of wave energy converters (WECs), in which it is desired to maximize the harvested energy. These predictions are usually based on time series analysis, deterministic wave propagation models, or on Kalman filtering. Reference [4], for instance, investigated the use of a deterministic propagation model, for WECs wave-by-wave control, assuming a single group velocity , and showed a considerable improvement in estimating the wave surface elevation. However, the results are limited to small-amplitude oscillations in small-amplitude waves due to the assumptions of Uni-directional/long-crested waves, heave-only motion, and small viscous effects. Reference [5] employed the genetic programming (GP) during real time wave forecasting. Reference [5] suggested that the better results obtained using GP might manage a large number of training pairs efficiently than the ANN.

However, ANNs have gained an enormous interest in recent years due to their capabilities. Prediction over horizon is one of the tasks that ANNs can execute with greater effectiveness over many other methods. Reference [6], for instant, employed the Classification and Regression Trees (CART) algorithm for building and evaluating regression trees for significant wave height prediction

using the wind and wave parameters as inputs. A comparison with artificial neural networks shows that the neural networks were marginally more accurate than regression trees from the error statistics point of view [6]. Reference [7] utilized different combinations of inputs for the neural network to predict the significant wave height, and showed that combining the wave direction and wind direction in one parameter improved the prediction accuracy. In addition, it was also shown that for long lead prediction of the significant wave height, the wind parameters have the most impact on prediction [7]. Reference [8] built an ANN with only the wave parameters as inputs. The results demonstrated a good accuracy in significant wave height prediction, especially in the near lead time for one or two hours [8] which is in agreement with the findings in reference [7]. Reference [9] estimated the monthly mean significant wave height using the ANN and regression methods with the wind speed, sea level pressure, air temperature and significant wave height as inputs. With different combinations of theses inputs, it was shown that the best results can be obtained only when taking all the above factors into consideration [9]. Batch-learning algorithms with back propagation were also employed for wave parameters prediction. Batch-learning algorithms involve complex computations, in which network parameters need to be fixed a priori [3]. Reference [3] used each of the sequential learning algorithms, the Minimal Resource Allocation Network (MRAN) and the Growing and Pruning Radial Basis Function (GAP-RBF) network, for wave prediction. The results show that the prediction obtained MRAN and GAP-RBF outperforms that obtained using Support Vector Regression (SVR) and Extreme Learning Machine (ELM) methods [3]. However, no statistical method exists in the literature that is dedicated to predict the wave surface elevation as far as the authors' knowledge goes.

In summary, several studies, in the literature, show that the ANN prediction outperforms several other methods such as regression methods, SVR, GP, deterministic models, and empirical models. Therefore, in this study, ANN will be utilized as a prediction tool. It is noted, however, that most of the work in the literature did not solve the proposed problem, in which it is attempted to

predict the wave surface elevation at a WEC location using measurements of the position of a wave rider ahead of the WEC, using statistical methods. Two cases are presented in this paper. In the first case, the significant wave height at the WEC location is predicted using the measurements from a nearby buoy. This case study is considered here as performance measure for the prediction tool and covered in the literature as discussed before. In the Second case, the wave surface elevation at the WEC location is predicted using the measurements from a nearby buoy. These two cases are of particular importance in WEC control problems in which the motion of a wave rider buoy at one point can be used to predict wave information at another point where a WEC is controlled for energy harvesting.

The paper is organized as follows: the second section gives a brief introduction about the ANN with special focus on the NARX NN. The results of numerical simulations of wave height prediction are presented in section 3. The real data that will be used in this study are reported in section 4.1. Further discussion on working with real data is in section 5.

2. Artificial Neural Networks

Artificial neural networks are computing systems that perform tasks by learning without task-specific programming. An ANN is a collection of connected neurons, where the neurons are organized in layers as shown in figure 1. Neuron is the basic computation unit in ANN. Each neuron receives input from some other neurons, from an external source (X_1,X_2,X_3,..) and bias b and computes an output Y. Each input has an associated weight W, which is assigned on the basis of its relative importance to other inputs as shown in Figure 1. The Neuron applies a function f (Activation Function) to the weighted sum of its inputs to compute the output Y, as shown in Eq. (2) and Figure 2.

$$Y = f\left(X_1.W_1 + X_2.W_2 + X_3.W_3 + ... + b\right) \tag{2}$$

The purpose of the activation function is to introduce non-linearity into the output of a neuron in order to represent nonlinear data in real world. Several

activation functions have been introduced in the literature. However, the most common ones are as follows:

1. The Sigmoid function which takes a real-valued input and squashes it to the range [0, 1]:
$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{3}$$
2. The tanh function which takes a real-valued input and squashes it to the range [-1, 1]:
$$tanh(x) = 2\sigma(2x) - 1 \tag{4}$$
3. The Rectified Linear Unit function (ReLU) which takes a real-valued input and thresholds it at zero (replaces negative values with zero)
$$f(x) = max(0, x) \tag{5}$$

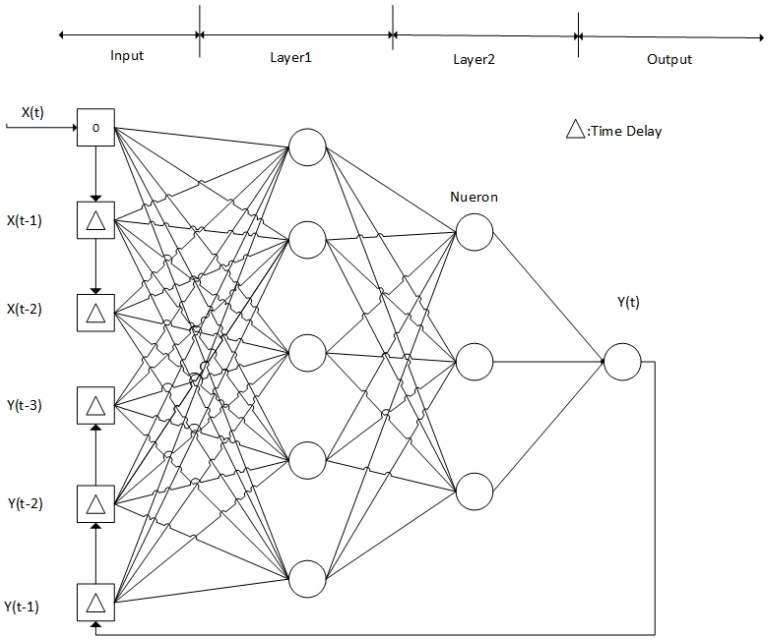


Figure 1: An example of NARX Network with three layers. Usually the input layer is not counted. Two-delay time line with two steps delay in input and three steps delay in output. Five neurons in the first layer, three neurons in the second layer and one neuron in the output layer.

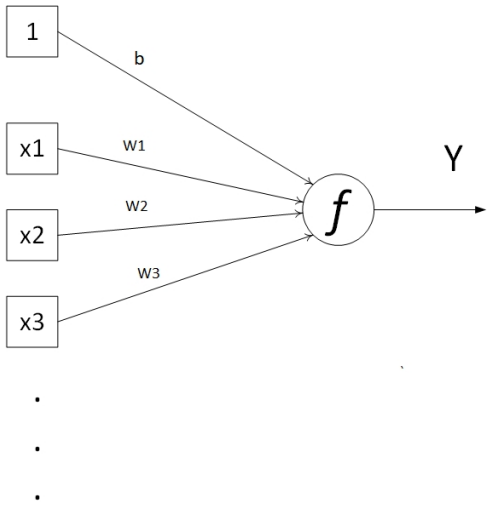


Figure 2: Single Neuron structure with several weighted inputs and bias. The activation function f takes the summation of the weighted inputs and bias to compute non-linear output.

The above activation functions are usually used in the middle layers of the ANN. However, in the last layer (output layer), a linear activation function is usually used to prepare the outputs unless a special task that require one of the non-linear activation functions is needed.

In general, a signal can be transmitted from one neuron to another, from the first input layer to the last output layer. Every neuron and the connections between neurons have a weight that gets updated during the learning process. Therefore, the neuron's output is a non-linear function of its inputs [10].[3].

Artificial neural networks are usually used in time series prediction applications as mentioned in the introduction section. The inherent nonlinearity of ANN models and their higher robustness to noise are the reasons that make the ANN models frequently outperform standard linear techniques in time series prediction applications. Time series prediction can be made for one-step-ahead or for a longer prediction horizons. In longer prediction horizons, the ANN

[3] `https://www.ngdc.noaa.gov/geomag/WMM/DoDWMM.shtm`

output is usually fed back to the input, unlike the one-step-ahead prediction. The recurrent neural networks (RNN), usually used in longer horizons prediction, are usually trained by means of temporal gradient-based variants of the back propagation algorithm. However, learning to perform tasks in which the temporal dependencies (present in the input/output signals) span long time intervals can be quite difficult using gradient-based learning algorithms. It has been reported that the nonlinear autoregressive with exogenous input (NARX) is more effective and doesn't suffer from this problem [10]. This advantage of the NARX is due to its input vector that is built through two tapped-delay lines: one sliding over the input signal and another sliding over the networks output [11].

2.1. NARX

NARX is an ANN that is well suited for modeling nonlinear systems [12, 13]. The reasons for that are its faster convergence in reaching the optimal weights of the connections between neurons and/or inputs, it is more effective compared to other ANNs [12, 13] and it is much better at discovering long time dependences than conventional RNN [12, 13]. The neutral development of NARX was after linear Autoregressive with exogenous inputs (ARX) model. The NARX can be considered a dynamic RNN with feedback connections and is based on the linear (ARX) model. An illustrative example of the NARX is shown in figure 1. Figure 1 shows an NARX Network with two step delay in input and three step delay in output. The NARX model equation is:

$$y\left(t\right)=f\left(y\left(t-1\right),y\left(t-2\right),...,y\left(t-n_{y}\right),u\left(t-1\right),u\left(t-2\right),...,u\left(t-n_{u}\right)\right), \tag{6}$$

Where $[1\cdots n_y]$ is the time delay line (TDL) span of the output and $[1\cdots n_u]$ is the TDL span of the input. The next value of the dependent output signal $y(t)$ is regressed on previous values of the output signal and previous values of an independent (exogenous) input signal u. A NARX network can be trained in two modes [14]:

1. Series-Parallel (open loop) (SP Mode): in which the true output is used instead of feeding back the estimated output. Therefore, the output's regressor is formed only by actual values of the system's output, as shown in figure 3a.
2. Parallel (closed loop) (P Mode): the estimated outputs are fed back and included in the output's regressor, as shown in figure 3b.

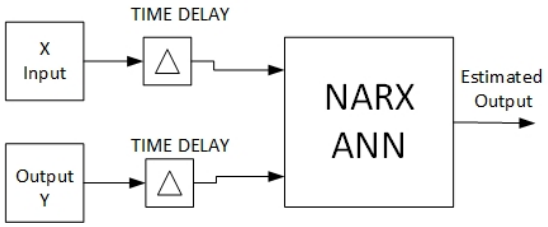
It is worth to note that, these two modes can be utilized during the prediction phase also. The key difference between the two modes is whether the time delay line of the output will use the actual output during the training and/or prediction (in SP mode) or will use the estimated output from the output layer (in P mode). Therefore , it is expected that for long prediction horizons, the best performance could be achieved by training the NARX in SP mode, as shown in figure 3a, to get better accuracy and performance and less training time; hence using the P mode, as shown in figure 3b, during the prediction phase. However, for one-step-ahead prediction, both the training and prediction phases will be executed in the SP mode as the real history measurements of the required output parameters are used during the prediction phases to estimate only one-step-ahead for the parameters of interest.

3. Numerical Case Study

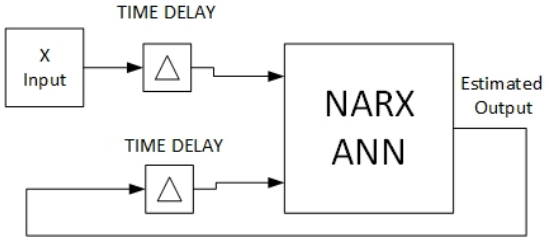
This section presents a case study with two tests in which there are two buoys, the first is a wave rider and the second one is a WEC. The WEC distance from the wave rider buoy is not constant, due to the surge motion. For the purpose of controlling the WEC to harvest energy, wave surface elevation prediction is needed at the WEC location. The wave rider buoy position measurements are used to predict the wave surface elevation at the WEC location in this case study, as illustrated in Figure 4.

A Bretschneider spectrum is used as an example to simulate the ocean wave. The formula for the Bretschneider (one-sided) ocean wave spectrum is:

$$S(\omega) = \frac{5}{16}\frac{\omega_m^4}{\omega^5}\ H_{1/3}^2\ e^{-5\omega_m^4/4\omega^4}, \tag{7}$$



(a) Series-Parallel Architecture of NARX ANN



(b) Parallel Architecture of NARX ANN

Figure 3: The structure of the NARX ANN in the two different modes

where ω is frequency in radians per second, ω_m is the modal (most likely) frequency given wave, and is $H_{1/3}$ the significant wave height. Figure 5 represents samples of the wave surface elevation that are generated at the wave rider and WEC positions using the Bretschneider spectrum. To simulate the measurements of the wave surface elevation at the wave rider and WEC locations, errors are generated in the simulated wave surface elevation.

A NARX network model is used for prediction. The measurements of the surface elevation of the wave rider are the input to the NARX model. While, the output of the NARX model is the estimated wave surface elevation of the WEC. The wave surface elevation at the WEC is predicted given also a lag feedback from the WEC surface elevation itself which is considered as a second input to the NARX.

The NARX model for this case study has three hidden layers with five neurons in each layer with Sigmoid function as activation function. First, a open loop mode is used during the training process. During the training phase, the inputs are the wave surface elevation time series of the wave rider, X, and the

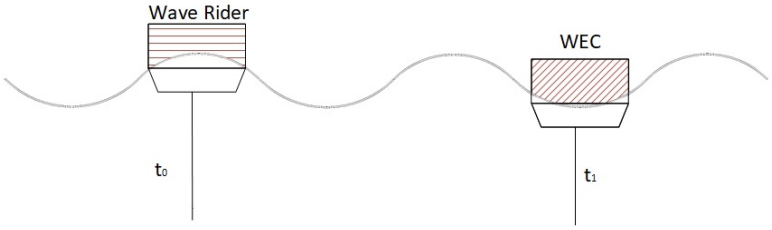


Figure 4: Wave surface elevation measured at the wave rider is used to predict the wave surface elevation at the WEC in the future.

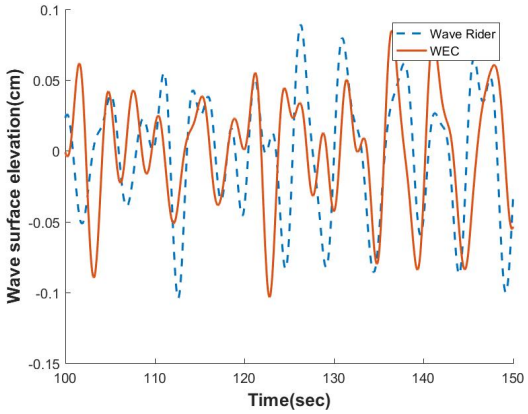


Figure 5: Samples of the simulated data of the wave surface elevation of the two buoys

actual wave surface elevation time series of the WEC, Y. The output is the estimated wave surface elevation at the WEC, Y'. This estimated wave surface elevation at the WEC is compared to the actual wave surface elevation at the WEC, Y, during the training phase. The loop in Figure 3a is used during the training phase. After training, the prediction phase starts, and the loop in Figure 3b is used, in which the trained NARX uses the output feedback to get the estimated wave surface elevation at the WEC, in addition to the measured wave surface elevation at the wave rider, X, as the input.

The sampling rate for the simulated data is 100 Hz. The waves at the rider arrive at different times at the WEC. In this study, we assumed that the waves arrive at the WEC after $10 + v(t)$ seconds, where $v(t)$ is a random variable with

normal distribution of zero mean and a variance of 0.2 seconds.

Morover, a random noise is added to the wave surface elevation at the WEC location to simulate the actual difference in the measurements of wave surface elevation between the two buoys. This random noise is assumed to have normal distribution, with zero mean and a variance of 15% of the true wave surface elevation at the WEC location. In this test, the input TDL is assumed to start at 9 sec and ends at 11 sec, while the output TDL starts at 0.01 sec and ends at 1 sec. These parameters of the first test are summarized in table 1.

Table 1: Simulation parameters of the first numerical test

Sampling rate	WEC Wave height random noises	
0.01 sec	$\pm 15\%$	
Time shift	Input TDL	Output TDL
$10 + \nu(0, 0.2)$ sec	[9-11] sec	[0.01-1] sec

The training phase in this test was conducted for a period of 1000 sec, with the same rate of 100 Hz.

Using the trained NARX, the wave surface elevation at the WEC location is then predicted for a different period of time, using the parallel NARX model presented in Figure 3b. Figure 6 shows the prediction results. As can be seen in figure 6, there is a good match between the prediction data (NARX output) and the target data which is the simulated measurements of the wave surface elevation at the WEC location.The performance of prediction can be further assessed by computing several validation factors: the root mean square error (RMSE), the Bias, the mean absolute error (MAE), the normalized root mean square error (NRMSE), the scatter index (SI) and the correlation coefficient (CC) [15]. Figure 7 shows the regression plot of the target and the predicted wave surface elevation, where the vertical axis represents the predicted wave surface elevation. Table 2 represents the values of the validation factors of this case which indicate a good prediction accuracy.

It is worth to note that the prediction is of type long prediction horizon.

Table 2: Validation factors for the first numerical test

RMSE (cm)	Bias (cm)	MAE (cm)	NRMSE (cm)	SI	CC
0.018	-1.7e-05	1.7e-05	0.360	0.36	0.93761

This prediction started by letting the NARX to get enough training and to demonstrate an acceptable performance in the open loop mode (series-parallel mode). Then the NARX is transformed to the closed loop mode (parallel mode), in which the inputs are the measurements of the wave surface elevation at the wave rider from 9 to 11 seconds ago (9-11 TDL), which is 201 measurements of the wave surface elevation using 100 HZ sampling rate, starting at time (-11 sec) to time (-9 sec). The estimated output from the NARX is fed back as input with (0.01-1 TDL), which is 100 values of the estimated wave surface elevation for 100 HZ sampling rate at the WEC location. The estimation process, however, cannot continue for longer periods of time, as the estimation error would accumulate over time.

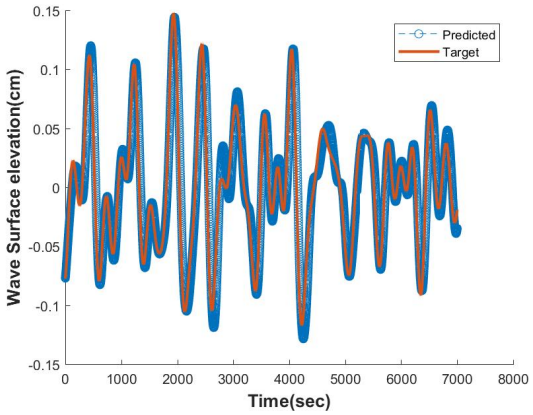


Figure 6: Prediction sample of the wave surface elevation at the WEC compared with actual measurements of the wave surface elevation

In the second test, the random noise of the time shift between the two buoys is assumed to have normal distribution with zero mean and a variance of 0.4 sec. The random noise of the wave surface elevation at the WEC location

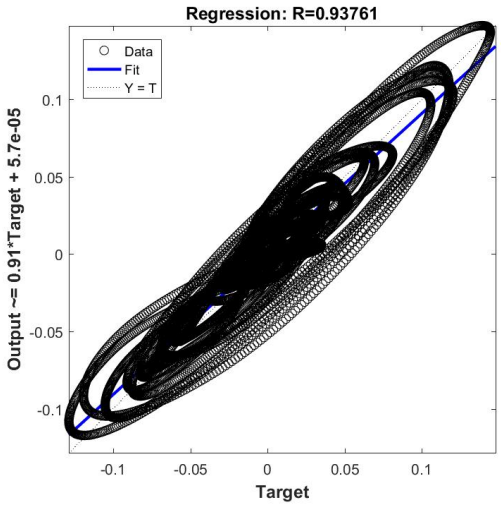


Figure 7: Regression relation between the target (actual wave surface elevation at the WEC) on the horizontal-axis and the predicted wave surface elevation (NARX output) on the vertical-axis . The equation of the vertical-axis represents the equation of the regression line and R is the correlation factor between the actual and predicted wave surface elevation. This figure is plotted using the plotregression function in MATLAB

is assumed to have normal distribution, with zero mean and a variance of 30% of the true wave surface elevation at the WEC location. In this test, the input TDL is assumed to start at 8 sec and ends at 12 sec, while the output TDL starts at 0.01 sec and ends at 2 sec. These parameters of the first test are summarized in table 3. In this test case, the trained NARX from the previous test is retrained using the new data. This leads to reducing the training time and may lead to better results than the previous test.

Table 3: Simulation parameters of the second numerical test

Sampling rate	WEC Wave height random noises	
0.01 sec	$\pm 30\%$	
Time shift	Input TDL	Output TDL
$10 + \nu(0, 0.4)$ sec	[8-12] sec	[0.01-2] sec

Table 4: Validation factors for the second numerical test

RMSE (cm)	Bias (cm)	MAE (cm)	NRMSE	SI	CC
0.004	-2.22e-04	2.22e-04	0.09	0.093	0.99575

The training phase in this test was conducted for a period of 1000 sec, with the same rate of 100 Hz.

Using the trained NARX, the wave surface elevation at the WEC location is then predicted for a different period of time, using the parallel NARX model presented in Figure 3b. Figure 8 shows the prediction results. As can be seen in Figure 8, there is a good match between the prediction data and the target data. Figure 9 shows the regression plot of the target and the predicted wave height, where the vertical axis represents the predicted wave surface elevation. In this case, the validation factors; RMSE, Bias, MAE, NRMSE, SI and CC are presented in table 4. The results demonstrate that the retrained NARX could be used to get better results.

This prediction is of type long prediction horizon. The inputs in this case are the measurements of the wave surface elevation at the wave rider from 8 to 12 seconds ago (8-12 TDL). This is equivalent to 401 measurements of the wave surface elevation at 100 HZ sampling rate started at (time -12 sec) to (time -8 sec). The estimated output from the NARX is feedback as inputs with (0.01-2 TDL). Which is 200 values of the estimated of wave surface elevation for 100 HZ sampling rate at the WEC location. This prediction type is a long horizon prediction where the estimated output from the NARX is used in fed back loop to continue the prediction. However, this process cannot continue for longer periods, as the estimation error would accumulate over time.

As can be seen in the previous two tests, the NARX is able to capture the underlying nonlinear relationship between the target data (wave surface elevation at WEC) and the input data (wave surface elevation at wave rider), despite the random noise added to the time shift and wave surface elevation in the target data. The correlation coefficient using NARX is close to 1 in both

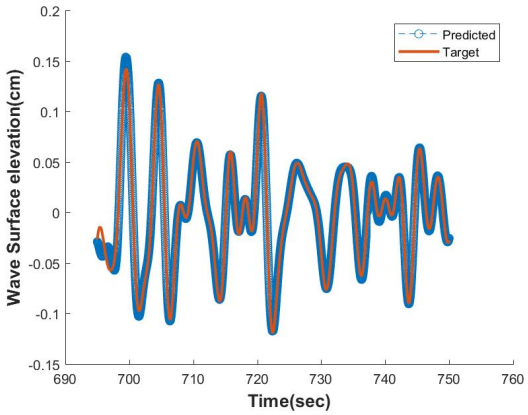


Figure 8: Prediction sample of the wave surface elevation at the WEC And compared with actual measurements of the wave surface elevation

tests which indicate high performance.

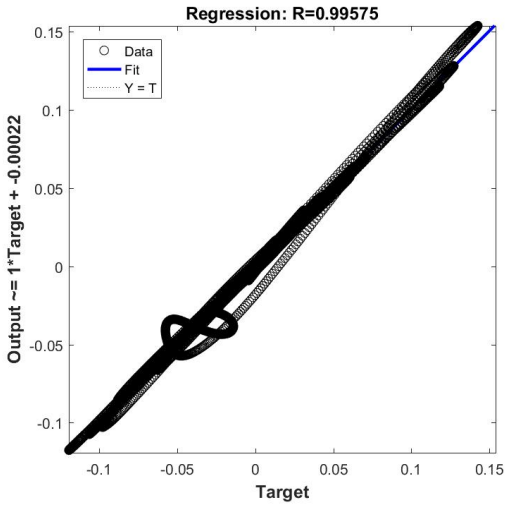


Figure 9: Regression relation between the target (actual wave surface elevation at the WEC) on the horizontal-axis and the predicted wave surface elevation (NARX output) on the vertical-axis . The equation of the vertical-axis represents the equation of the regression line and R is the correlation factor between the actual and predicted wave surface elevation. This figure is plotted using the plotregression function in MATLAB

4. Prediction of Significant Wave Height Using Real Ocean Data

In this section the NARX ANN model is tested using real measurements from buoys on the east coast of USA. This section, however, provides prediction only for the wave significant height for long prediction horizon . Section 5 discusses the challenges in using the NARX ANN method to predict the wave surface elevation using the real data. The following section describes the data used in this study.

4.1. Measured Data and Measuring Stations

The Coastal Data Information Program (CDIP) is an extensive network for monitoring waves and beaches along the coastlines of the United States. CDIP measures, analyzes, archives and disseminates these coastal environment data

on their website[4]. CDIP first began to use directional buoys, which measure sea surface temperature and wave direction in addition to wave energy. Later in 2007 CDIP started using the Datawell Buoys which are equipped with Iridium satellite communication capabilities, eliminating the need for shore stations and allowing for buoy deployments further from the coastline.

From the CDIP network, the following two buoys (stations) are selected for our case study (also shown in Figure 10):

1. Station-198 (NDBC 51207), in Kaneohe bay, HI
   (a) Most recent location: 21 28' 38" N, 157 45' 09" W
   (b) Instrument description: Datawell directional buoy
   (c) Most recent water depth: 81 m.
   (d) Measured parameters: wave energy, wave direction, sea temperature
2. Station-225 (NDBC 51210), in Kaneohe bay, WETS, HI
   (a) Most recent location: 21 28' 38" N, 157 45' 20" W
   (b) Instrument description: Datawell directional buoy
   (c) Most recent water depth: 80 m
   (d) Measured parameters: wave energy, wave direction, sea temperature

These two buoys are moored. The advantages of working with moored buoys are [3]:

1. Sensors locations on the buoys are selected such as to avoid atmospheric exposure problems that causes measurement errors
2. Sampling and averaging periods for the measurements are determined after accounting for buoy motion

In this section, it is assumed that the significant wave height is measured; which is the case for the data provided by CDIP. Each of the two stations provides measurements for the significant wave height every half hour. For the significant wave height, the measurements at one buoy location (e.g. Station

[4] `http://cdip.ucsd.edu`

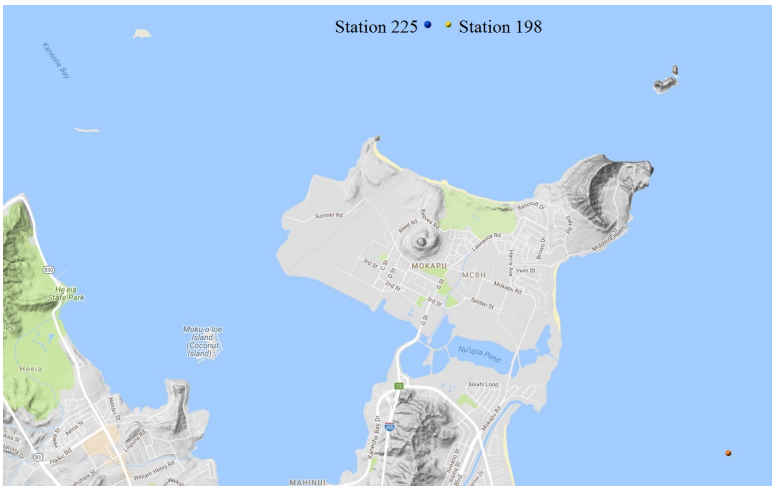


Figure 10: Location of stations 225 and 198 in Hawaii

198) can be used as it is at the other buoy location (station 225) at the current time, since the two buoys are close to each other. Prediction of the significant wave height in the future, however, needs a prediction algorithm. This section presents results of using NARX for prediction of significant wave height of long horizon prediction type.

4.2. Numerical Results

As mentioned above, the significant wave height is predicted at station 225 using the measurements of the significant wave height at station 198 as input for the prediction model, NARX. The data used in this study is collected in the period May 1-31, 2017. It is worth to note that it is better to collect data for several months or yeas to get better results. however, here in this case study one month is used for proof of concept.

The NARX for this case has three hidden layers and each layer has three neurons; the activation function is the Sigmoid function. The Serial-Parallel NARX ANN model, presented in Figure 3a, is used for training. In this test, the input TDL is assumed to start at 0.78 sec and ends at 2.34 sec, while the output TDL starts at 0.78 sec and ends at 1.56 sec. Figure 11 shows a good

matching between the predicted significant wave height and the real one as measured by station 225, during the training process.

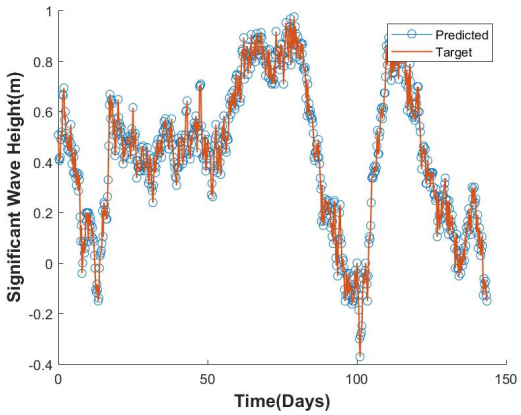


Figure 11: The prediction of significant wave height at buoy 225 matches the real data during the training phase

Figure 12 represents the predicted significant wave height at station 225 obtained using the trained NARX, as well as the actual (target) data as reported by station 225, for comparison purpose. Note that this prediction is obtained using only data from 198 as input and the estimated significant wave height at station 225 in the feedback loop. As shown in figure 12, The NARX captures the underlying nonlinear dynamics of the significant wave height as it changes over time. For further assessment for the prediction performance, figure 13 renders the correlation between the predicted and target values. The correlation coefficient R is above 0.987. It is here noted that if the significant wave height at buoy 225 is needed at current time of measurement (no prediction in the future), then the measurement at buoy 198 could be used. The validation factors: RMSE, Bias, MAE, NRMSE, SI and CC are computed for this test and are presented in table 5. To assess the accuracy of the prediction, a comparison in terms of the validation factors is conducted. Zheng et al.[16] provided the RMSE, MAE, Bias and CC for prediction of the power density function in table (1) in [16] using WW3-T639 (WaveWatch-III-T639) and WW3- T213 at the

Table 5: Validation factors for the significant wave height prediction case

	RMSE	Bias	MAE	NRMSE	SI	CC
NARX NN	0.0753	-0.0502	0.0502	0.1312	0.0977	0.9819
WW3-T639 at Korean Peninsula station(22101)[16]	1.91	-0.14	0.98	–	–	0.97

waters surrounding the Korean Peninsula, Japan and Taiwan Island. As shown in table 5, the NARX NN produces a comparable results compared to WW3 results in [16].

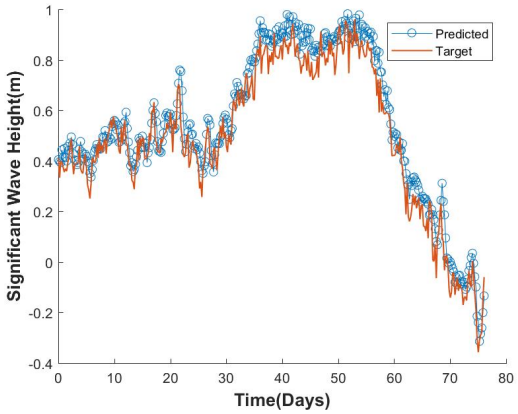


Figure 12: The prediction of significant wave height at buoy 225 matches the real data during the prediction phase

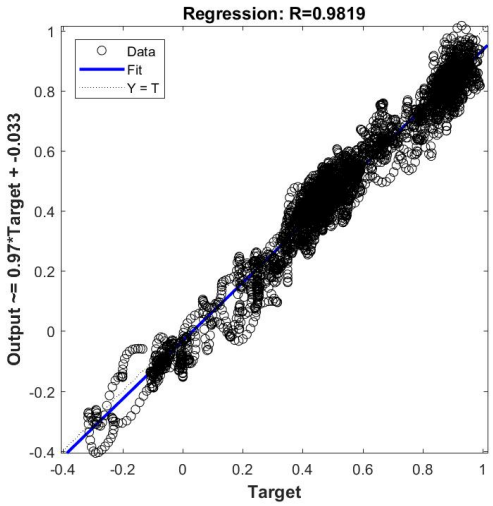


Figure 13: NARX output regression relation between the target on the horizontal-axis and the predicted significant wave height at station 225 (NARX output) on the vertical-axis. The equation of the vertical-axis represents the equation of the regression line and R is the correlation factor between the actual and predicted wave surface elevation. This figure is plotted using the plotregression function in MATLAB

5. Discussion

This section discusses some challenges in predicting the wave surface elevation using real data. A case study is conducted for that purpose in which the results of a long prediction horizon will be presented, as well as the results for a one-step-ahead prediction. Wave surface elevation measurements are collected at a rate of 1.28 Hz. In this case study, the measurements of the wave surface elevation from the two stations (198 and 225) are used. The goal is to compute the wave surface elevation at station 225, at current time, using measurements collected up to current time at station 198. This section shows that the ANN methods are not suitable for solving the problem of long prediction horizon; this canbe attributed to the fact that the real data are non-stationary [10, 17–19]. To highlight this issue, consider the case study presented in Section 3. The data used in Section 3 is generated via simulations. The time range was divided into

intervals. The mean value for the wave surface elevation is computed over each interval, and its value is shown in Figure 14 for the different intervals. Similarly, the variance values of the same set of data for the various intervals are shown in Figure 15. As it be seen from both figures, the mean value is almost

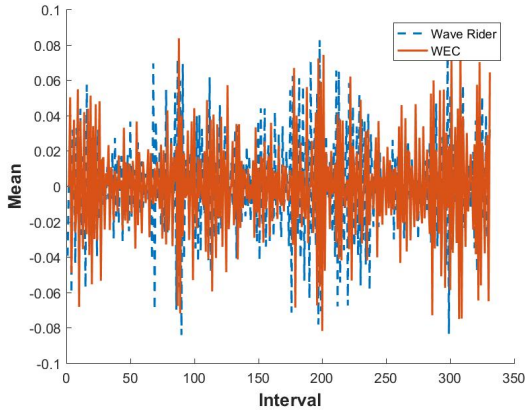


Figure 14: Mean values of the simulated data over time intervals (interval=300 sec.)

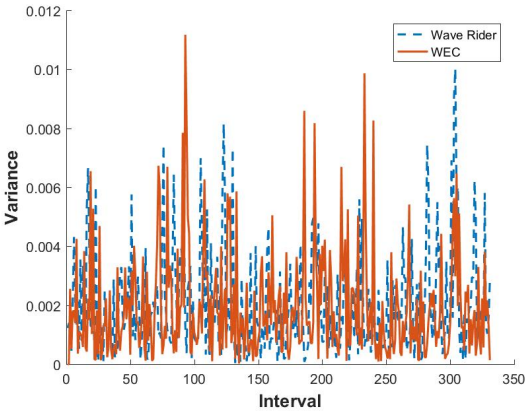


Figure 15: Variance values of the simulated data over time intervals (interval=300 sec.)

constant across time intervals (changes are in the order of 10^{-3}), and the variance value is also very small (in the order of 10^{-3}). Hence, the simulation data is stationary[17, 18] and can be used in the proposed NARX method directly without the need for preprocessing[10, 19].

For the real data obtained from CDIP , the mean and covariance values, over different time intervals, are plotted in Figures 16 and 17, respectively. Each time interval is 5 minutes. As can be seen in both figures, there is a significant change in the mean and variance values among different intervals. Therefore, the data statistical distribution of the time history is highly dispersed, which indicates the data is non-stationary and is not suitable to use with NARX directly.

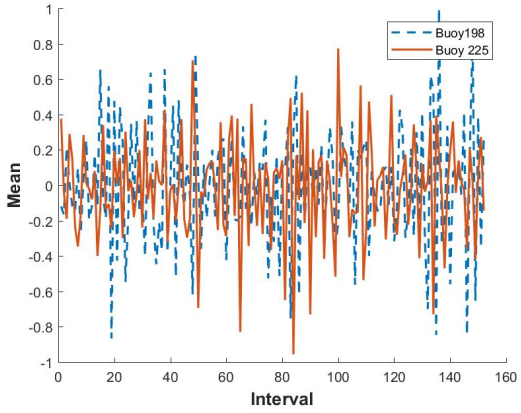


Figure 16: Mean values of data from stations 198 and 225 (interval=300 sec.)

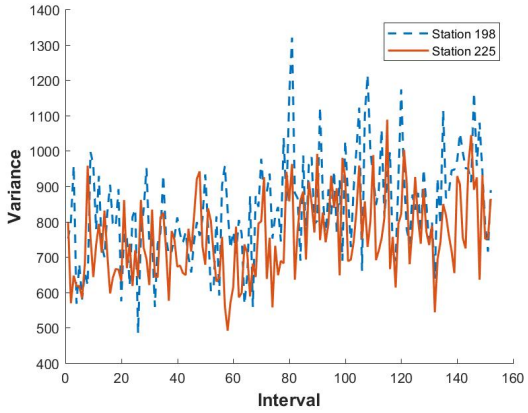


Figure 17: Variance values of data from stations 198 and 225 (interval=300 sec.)

In an attempt to improve the performance, couple of preprocessing steps are carried out. First, a coordinate transformation is performed to make all the

surface elevation values as positive numbers. Second, we take the square root (or Logarithm) of the wave surface elevation to enhance the data variance. Figure 18 represents the mean values of wave surface elevation during intervals of (300 sec.) after carrying out coordinate transformation and taking the Logarithm of the wave surface elevation. As expected, the range of the mean does not change, yet, the global average value is changed due to coordinate transformation. Figure 19 represents the variance values of wave surface elevation during intervals of (300 sec.) after carrying out coordinate transformation and taking the Logarithm of the wave surface elevation. The data variance is improved, however, further enhancement is still needed. Further Preprocessing steps can be carried out as follows:

1. Filtering: a fast Fourier transform is used to detect the sinusoids of small amplitudes and remove them as the very small amplitudes waves are assumed to be generated by external sources such as boats.
2. Data smoothing: the CDIP website provides the wave measurements at the buoys with a low sampling rate (1.28 Hz). Therefore, these show a lot of step changes which lead to losing valuable information about the real trend of the wave parameters. Consequently, a smoothing step is carried out. The power of fast Fourier series is employed to give intermediate points by interpolation which in return improve the training and prediction processes.

Three different training functions for NARX are attempted; these are:

1. Levenberg-Marquardt function: fast training process but not the best
2. BFGS Quasi-Newton function: not as fast as Levenberg-Marquardt function but give best performance
3. Bayesian Regularization function: the slowest between them however has the best performance results.

The NARX for this case has three hidden layers and each layer has three neurons; the activation function is the Sigmoid function. The Serial-Parallel

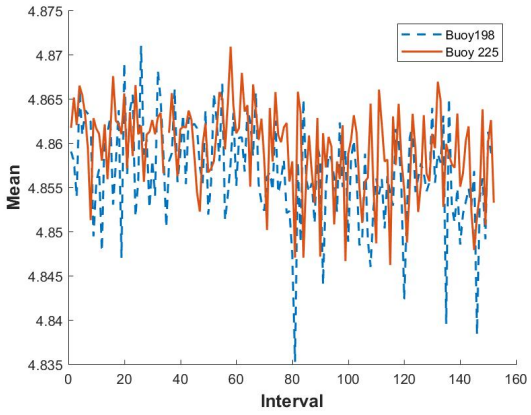


Figure 18: Mean values of data from stations 198 and 225 after coordinate transformation and taking the Logarithm of these data (interval=300 sec.)

NARX ANN model, presented in Figure 3a, is used for training. In this test, the input TDL is assumed to start at 0.78 sec and ends at 15.6 sec. The choice of the TDL period here is estimated according to the distance between the stations, and the minimum and maximum group velocities, which are given on the CDIP website. The output TDL starts at 0.78 sec and ends at 1.56 sec.

The results of using NARX for each of the training functions, and the pre-processing steps are summarized in table 6. Table 6 shows the performance of the NARX in terms of the mean square error (MSE) between the predicted wave surface elevation and the real data. The results in table 6 are obtained using training in SP mode and prediction in P mode. In other words, these results are of a long prediction horizon.

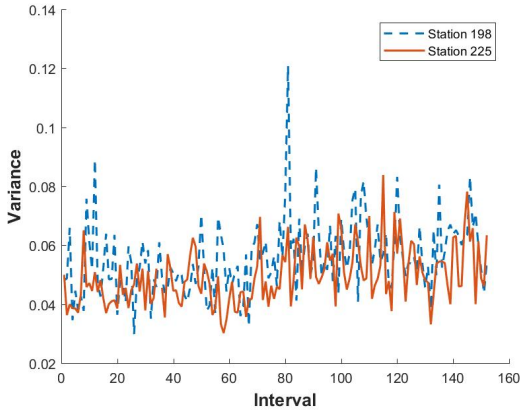


Figure 19: Variance values of data from stations 198 and 225 after coordinate transformation and taking the Logarithm of these data (interval=300 sec.)

Table 6: The MSE of the predicted wave surface elevation obtained using NARX with different preprocessing steps and training methods

Training Function	FFT Filter	FFT Smooth	MSE (cm) Training	MSE (cm) Prediction
Bayesian Regularization	x	x	3.1	2232.9
Bayesian Regularization	x		183.4	1599.3
Bayesian Regularization		x	20.5	1854.9
Bayesian Regularization			626.7	1825.6
BFGS Quasi-Newton	x	x	56.2	4906.7
BFGS Quasi-Newton	x		235.2	2014.9
BFGS Quasi-Newton		x	171.5	14114.5
BFGS Quasi-Newton			720.3	1952
Levenberg-Marquardt	x	x	3.1	1591.6
Levenberg-Marquardt	x		184.0	1605.8
Levenberg-Marquardt		x	20.6	1839.0
Levenberg-Marquardt			628.4	1843.3

Table 6 shows an improvement in the NARX performance, as evident from the decreased value of the MSE, when using the two proposed preprocessing steps for every training functions. However, this improvement is not sufficient. As shown in Table 6, the NARX is able to achieve acceptable MSE only during the training phase. During the prediction phase, however, the obtained MSE is still not acceptable. Figure 20 shows a comparison between the predicted wave surface elevation and the actual real data, during the training phase. It is evident that the two quantities match to an acceptable level. Figure 21 represents the regression plot during the training phase. The correlation coefficient R is close to 1 during training phase.

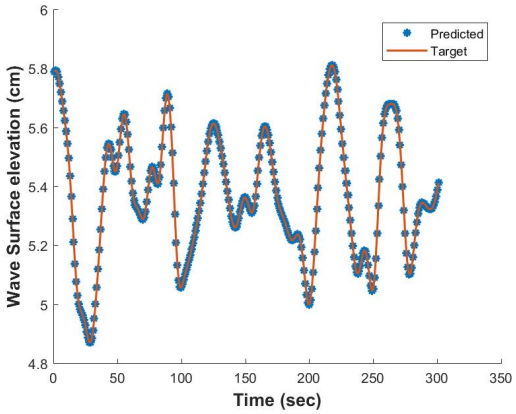


Figure 20: NARX output for station 255 compared to real data during training mode

While the obtained accuracy during the training phase can be accepted, the trained NARX produces poor accuracy during the prediction phase, as evident from the MSE values in Table 6; the predicted wave surface elevation is also plotted in Figure 22 versus the real data for comparison. Figure 23 represents the regression plot during the prediction phase of long prediction horizon. The correlation coefficient R is less than 0.8 during long horizon prediction phase. In general non-stationary data are the most difficult to predict [10, 19]. The real data used in this case study seem to be non-stationary as discussed above. It is worth mention that the CDIP provides the number of wave counts passing each

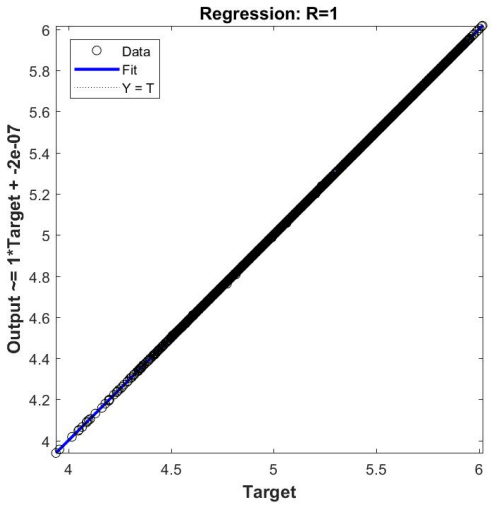


Figure 21: Regression relation between the target data and the predicted wave surface elevation using NARX for station 255 during the training phase

station in time intervals of half hour. Based on the CDIP data for stations 225 and 198, this count is different for the two stations despite the distance between them is about 350 m.

Figure 22 shows, however, an interesting match in the first few seconds. It is clear that the prediction gets worse over time due to the dependency on the estimated output history. Figure 24 shows the relation between the horizon time and the mean square error between the true wave surface elevation at station 255 and the estimated wave surface elevation from the NARX. As shown in fig. 24, when the horizon time is less than $4sec$ the MSE is good, while it significantly increases when horizon time increases.

In the case of only one-step-ahead prediction using SP mode, the history of the output data (wave surface elevation at the WEC, in this case station 225) should be available. Using that along with the current measurements of the wave surface elevation of the wave rider (in this case station 198), a one-step-ahead prediction is performed. A very good performance is achieved as shown in figs. 25 and . 26. Figure 25 shows a comparison between the predicted wave

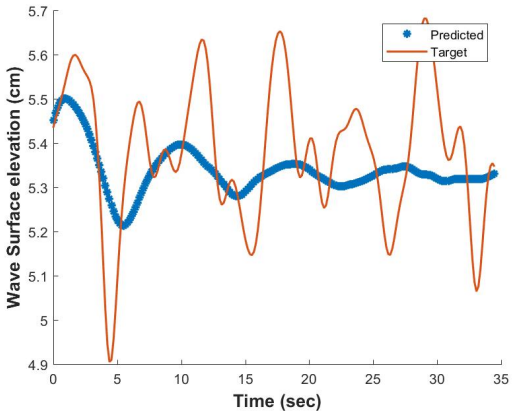


Figure 22: NARX output for station 255 compared to real (target) data during long prediction phase in P mode

surface elevation and the actual real data, during the one-step- ahead prediction phase. It is evident that the two quantities match with high performance level. Figure 26 represents the regression plot during the one-step- ahead prediction phase. The validation factors; RMSE, Bias, MAE, NRMSE, SI and CC for this case are provided in table 7.

In summary, even with the non-stationary real data from CDIP website, the NARX demonstrated its capability in the one-step- ahead prediction with SP mode to predict the wave surface elevation. These real data should be tested first to define the data type (non-stationary or stationary). Preprocessing steps may be required with some processing on the data to improve the data variance. For long prediction horizon, the NARX is not able to perform a good prediction performance. One of the main reasons of this problem is that the distance between the two station is long, 350 m, in which unexpected external disturbances may occur at station 255. For using this prediction strategy with wave-by-wave control, one should expect the real distance between the wave rider and WEC to be about $30m$. Therefore, the one-step- ahead prediction with SP will be enough for online control or long prediction horizon with small time horizon (horizon time $<$ 4 sec.). It is worth noting that the advantage

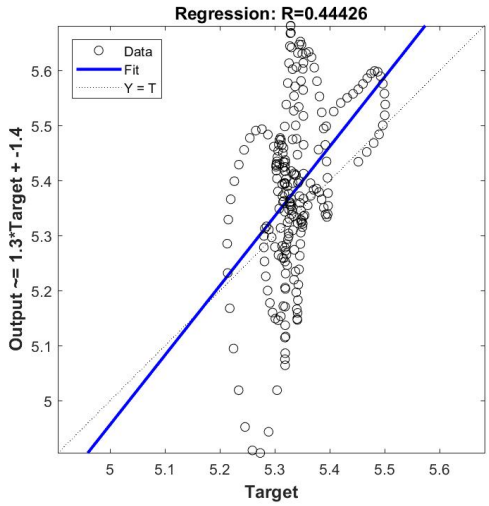


Figure 23: NARX output regression relation between the target on the x-axis and the predicted wave surface elevation at station 225 (NARX output) on the y-axis during the long prediction horizon phase in P mode. The equation of the y-axis represents the equation of the regression line and R is the correlation factor between the actual and predicted wave surface elevation. This figure is plotted using the plotregression function in MATLAB

Table 7: Validation factors for one-step-ahead prediction using SP mode

RMSE (cm)	Bias (cm)	MAE (cm)	NRMSE (cm)	SI	CC
8.74e-05	-1.73e-06	1.736e-06	1.625e-05	1.625e-05	1

of the statistical time series analysis methods (such as the NN) is the training part of the process. The purpose of this training part is capture the underlying dynamics from long enough time history of measurements such as the swell waves that can propagate for thousands of kilometers along great circle paths with little attenuation [20–23].

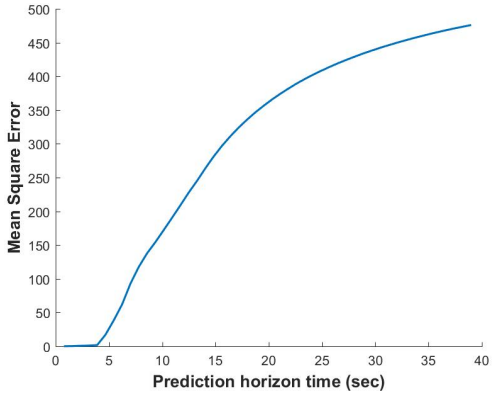


Figure 24: Mean square error, between the target data and the predicted wave surface elevation using NARX for station 255 during the long prediction phase with p mode, versus horizon time

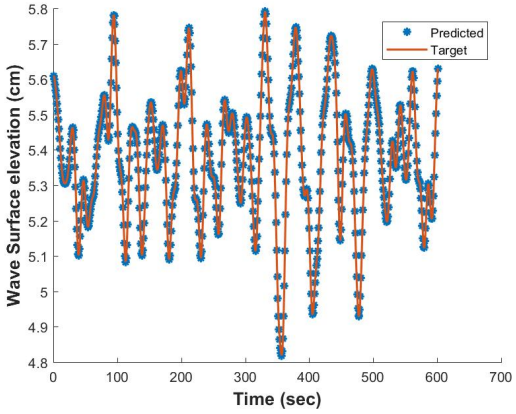


Figure 25: NARX output for station 255 compared to real (target) data during one-step-ahead prediction phase in SP mode

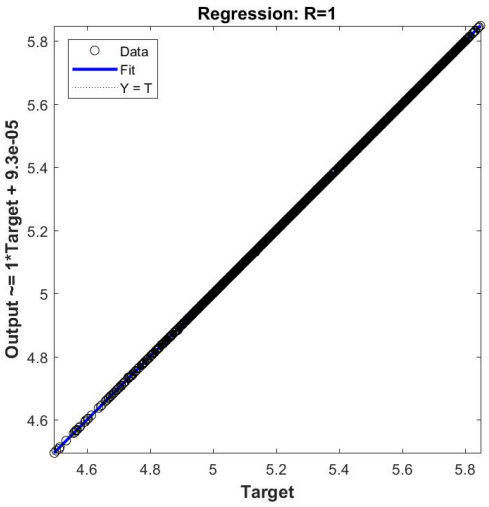


Figure 26: NARX output regression relation between the target on the x-axis and the predicted wave surface elevation at station 225 (NARX output) on the y-axis during the one-step-ahead prediction phase in SP mode. The equation of the y-axis represents the equation of the regression line and R is the correlation factor between the actual and predicted wave surface elevation. This figure is plotted using the plotregression function in MATLAB

6. Conclusion

In this study, prediction of wave surface elevation at a specific location is attempted using NARX and input data from ahead sensor, such as the wave surface elevation at ahead wave rider. This type of prediction is critical for ocean wave energy converters. The study is conducted on two different sets of data. The first set is a simulation data with added noise and the second is real data available through CDIP. The results confirm that the NARX can predict wave surface elevation to a great accuracy for long prediction horizon when using simulation data. When using real data the prediction is poor for long prediction horizon. However, high accuracy is fulfilled in case one-step-ahead prediction with SP mode. Using real data, however, it is possible to have a long prediction horizon with P mode for the significant wave height, at a future time, at one location using measurements, at current time, at ahead location.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant Number 1635362.

References

[1] K. Hasselmann, T. Barnett, E. Bouws, H. Cartwright, K. Enke, J. Ewing, H. Gienapp, D. Hasselmann, P. Kruseman, A. Meerburg, P. Mqller, D. Olbers, K. Richter, W. Sell, W. H., Measurements of wind-wave growth and swell decay during the joint north sea wave project (jonswap).

[2] K. Hasselmann, T. Barnett, E. Bouws, H. Cartwright, K. Enke, J. Ewing, H. Gienapp, D. Hasselmann, P. Kruseman, A. Meerburg, P. Mqller, D. Olbers, K. Richter, W. Sell, W. H., Ocean Wave Modeling, Springer US, Boston, MA, 1985.

[3] N. Krishna kumar, R. Savitha, A. A. Mamun, Regional ocean wave height prediction using sequential learning neural networks, Ocean Engineering 129 (2017) 605–612. doi:10.1016/j.oceaneng.2016.10.033.

[4] U. A.Korde, Near-optimal control of a wave energy device in irregular waves with deterministic-model driven incident wave prediction, Applied Ocean Research 53 (2015) 31–45. doi:doi.org/10.1016/j.apor.2015.07.007.

[5] S. Gaur, M. Deo, Real-time wave forecasting using genetic programming, Ocean Engineering 35 (11-12) (2008) 1166–1172. doi:doi.org/10.1016/j.oceaneng.2008.04.007.

[6] A. Etemad Shahidi, J. Mahjoobi, Prediction of significant wave height based on regression trees, in: Environmental Modelling and Software, 4th International Congress, Barcelona, Catalonia, Spain, 2008.

[7] B. Kamranzad, A. Etemad Shahidi, M. H. Kazeminezhad, Wave height forecasting in dayyer, the persian gulf, Ocean Engineering 38 (1) (2011) 248–255. doi:10.1016/j.oceaneng.2010.10.004.

[8] F. Hosseinibalam, Y. Dehghan, Application of artificial neural networks in wave height prediction in the persian gulf, Science and Engineering Publishing Company 3.

[9] K. Gunaydn, The estimation of monthly mean significant wave heights by using artificial neural network and regression methods, Ocean Engineering 35 (14-15) (2008) 1406–1415. doi:doi.org/10.1016/j.oceaneng.2008.07.008.

[10] D. P. Mandic, J. A. Chambers, Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability, John Wiley and Sons, Ltd, 2002. doi:10.1002/047084535X.

[11] P. Jos Maria, J. Menezes, G. A.Barreto, Long-term time series prediction with the narx network: An empirical evaluation, Neurocomputing 71 (16-18) (2008) 3335–3343. doi:doi.org/10.1016/j.neucom.2008.01.030.

[12] E. Diaconescu, The use of narx neural networks to predict chaotic time series, WSEAS Trans. Comp. Res. 3 (3) (2008) 182–191.
URL http://dl.acm.org/citation.cfm?id=1466884.1466892

[13] A. Tatli, S. Kahveciolu, Narx neural networks based time series prediction for amount of airworthiness time, in: Electrical, Electronics and Biomedical Engineering (ELECO), 2016 National Conference, IEEE, 2017.

[14] H. T. Hang Xie, Y.-H. Liao, Time series prediction based on narx neural networks: An advanced approach, in: Machine Learning and Cybernetics, International Conference, IEEE, 2009. doi:10.1109/ICMLC.2009.5212326.

[15] C. W. Zheng, C. Y. Li, Variation of the wave energy and significant wave height in the china sea and adjacent waters, Renewable and Sustainable Energy Reviews 43 (2015) 381–387. doi:10.1016/j.rser.2014.11.001.

[16] C. W. Zheng, C. Y. Li, X. Chen, J. Pan, Numerical forecasting experiment of the wave energy resource in the china sea, Advances in Meteorologydoi:10.1155/2016/5692431.

[17] S. Julius, Bendat, A. G. Piersol, Random Data: Analysis and Measurement Procedures, Fourth Edition, John Wiley and Sons, Ltd, 2010.

[18] M. B. Robert, R. Russell Rhinehart, Applied Engineering Statistics, Marcel Dekker, Inc., New York, NY, 1991.

[19] C. Chatfield, Time-Series Forecasting, Chapman and Hall/CRC, 2016.

[20] W. Munk, G. Miller, F. Snodgrass, N. Barber, Directional recording of swell from distant storms, Mathematical, Physical and Engineering Sciences 225 (1963) 505–584. doi:10.1098/rsta.1963.0011.

[21] A. White, J. McTigue, C. Markides, Wave propagation and thermodynamic losses in packed-bed thermal reservoirs for energy storage, Applied Energy 130 (2014) 648–657. doi:10.1016/j.apenergy.2014.02.071.

[22] C. W. Zheng, C. Y. Li, Propagation characteristic and intraseasonal oscillation of the swell energy of the indian ocean, Applied Energy 197 (2017) 342–353. doi:10.1016/j.apenergy.2017.04.052.

[23] C. W. Zheng, C. Y. Li, Analysis of temporal and spatial characteristics of waves in the indian ocean based on era-40 wave reanalysis, Applied Ocean Research 63 (2017) 217–228. doi:10.1016/j.apor.2017.01.014.