22nd Australasian Fluid Mechanics Conference AFMC2020
Brisbane, Australia, 7-10 December 2020
https://doi.org/10.14264/d97a9ed

Time-Variations of Wave-Energy and Forecasting Power Availability using Different Techniques

Avikesh Kumar[1], M.R. Ahmed[2] and M.G.M. Khan[1]

[1]School of Computing, Information and Mathematical Sciences

The University of the South Pacific, Suva, Fiji

[2]School of Engineering and Physics

The University of the South Pacific, Suva, Fiji

Abstract

Wave energy has captured the attention of island nations that have a significantly larger sea area compared to the land area. An accurate forecasting model is required to correctly predict the availability of wave energy with all the required variables clearly defined. In this work, wave energy forecasting models: artificial neural network, regression and time series techniques were developed using wave height and wave period data. The models were tested using the wave data measured near Muani, Kadavu in Fiji for 14 months. The performance of each model is assessed using mean squared error, root mean square error, mean absolute error and coefficient of determination. The empirical results reveal that the artificial neural network model is more efficient and accurate in forecasting wave energy in comparison to the regression and time series models.

Keywords

Wave energy; Forecasting model; Artificial neural network; Multiple linear regression; Time series models.

Introduction

The dynamics of sources of energy have changed over the decades, as there are interventions of more modern and powerful technologies in this fast changing digital world. The transition on the different forms of energy sources has shifted from natural fossil fuel energies to renewable energy sources with the prime focus to provide clean and eco-friendly energy. While wind, solar and hydro energy are proven and reliable sources for Fiji Islands, energy generated from ocean surface waves can bring more diversification in the production of energy.

Energy generated from ocean surface waves is not something new to the energy turf because many developed countries have accepted this form of energy with greater prospect, as energy field continues to expand through excellent researches. Ocean waves have a good potential to generate electricity (power) along the coastlines [9]. As cited by Ram et al. [8], USA has a wave power potential of 240 GW, Republic of China has a wave power potential of 125 GW while Coastline of India has a wave energy potential of 6 GW/year respectively. Fiji has a sea to land area ratio of 70 with a coastline of 1129 km having wave power potential of 29 GW and 0.5% of these wave energy is enough to meet the daily demand of the Fiji Islands. Based on its geographical location, wave energy may be the best substitute to the form of energy been used in Fiji Islands. However, before wave energy can directly compete on the same platform with different forms of energy available in the Island, it would be imperative to have a thorough assessment done on its sustainability and practicality. An excellent forecasting model will definitely provide insights on how feasible it would be to accept wave energy as a substitute to the existing forms of energy [1]. Statistical models can predict ocean wave energy more accurately over short spans compared to Physics models [9]. He further claims that more research was done on the flux and its components than on wave power. Thus in this paper, an attempt is made to develop artificial neural network (ANN), multiple linear regression (MLR) and different time series techniques to forecast the amount of power available at a site in Fiji Islands

Forecasting technique is a tool that dilutes the concentration of risks and uncertainties by providing with realistic pecuniary measures in operating and sustaining a system. Generally, an accurate forecasting model will ensure for fluent stability and create a mathematical relation that will minimize the uncertainties and maximize the reliability on chosen energy source. Many excellent statistical forecasting models that exist perform best based on the area of problem and its individual unique properties.

Time series models look for trends and cyclic seasonality in the previous data while identifying the relationships between the variables. Exponential smoothing models add weight to the past data, which decreases exponentially with time while forecasting. Thus, the forecast value for this model as explained by Winston [11] is

$$A_t = A_{t-1} - \alpha e_t, \tag{1}$$

where A_t is the forecast value of the initial value A_{t-1}, e_t is the error and α is the smoothing constant having a range of 0 to 1. A bigger α value would mean the model is becoming more sensitive to the present value while a smaller α value would mean a model with less variance.

Furthermore, Holt-Winter (HW) additive model is an extension of exponential smoothing model as it has the capacity to learn about some trends and constant seasonal variations within the data [7]. HW additive model has a form:

$$x_t = a + bt + \varepsilon_t, \tag{2}$$

where a is the base level at the beginning of period t, b is the per-period trend and ε_t is the error for period t [10]. HW multiplicative model is said to be a further extension to additive model. It has the power to understand the trends, seasonal variations in the data as well as the changing proportionality to the levels of the data as explained by Pongdatu & Putra [7]. HW multiplicative model has a form:

$$x_t = ab^t\varepsilon_t \tag{3}$$

where b is the percentage growth in the base level and ε_t is the random error factor with a mean of 1.

While time series models capture trends, seasonal variations and levels within the data, MLR on the other hand, is able to identify the correlation between the variables and considers the influence of each when forecasting [10]. The multiple linear regression model has a form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \varepsilon \qquad (4)$$

where y is the output variable, $x_1, ..., x_k$ are the input variables, β_0, ... β_k are the regression coefficients and ε is the error. The robustness of MLR model is based on its linearity, normality, independence and homoscedasticity test [10].

Data collected in a natural setting where there is lack of control over some variables, usually exhibits nonlinear patterns. In situations like these, where relationships between variables are unknown, though very few priori assumptions are presented but have enough data, ANN with its computational structure have the power to predict the outputs with greater accuracy [12]. Reikard from his work concluded that time series models perform better than physics models when predicting wave energy [9]. ANN is a mathematical model for predicting systems output inspired by the structure and the function of a human biological neural network. It has the ability to explain and solve complex, highly nonlinear functions and synthesis, which is hard to express mathematically. It can also perform with high speed of evaluation, robustness and is adaptive to any changes in the data sets. According to Hadadpour et al. [2], ANN has the ability to predict the wave parameters correctly. Although wave energy assessment has been done at sites in Fiji Islands, forecasting of availability using a statistical model or an ANN is not yet investigated.

There is no fixed architecture of an ANN model. The strength of an ANN model lies within its parameters, which are varied while investigating for an optimal model with the one that has the least generalization error [4]. Each layer compromises of multiple units connected completely with the next layer by a weights and a bias function. The input layer carries the total information, which is passed to the hidden layer by the nodes, and then it attaches with a small weight and a bias function that is passed through the activation function to get the desired output. A learning algorithm is then deployed to minimize the error by re-adjusting the weights of each nodes. It is the hidden nodes in the hidden layers, which is able to capture a pattern and then perform mapping between the variables. Figure 1 below displays a structure of an ANN model and summarizes the process involved in developing an optimum model.

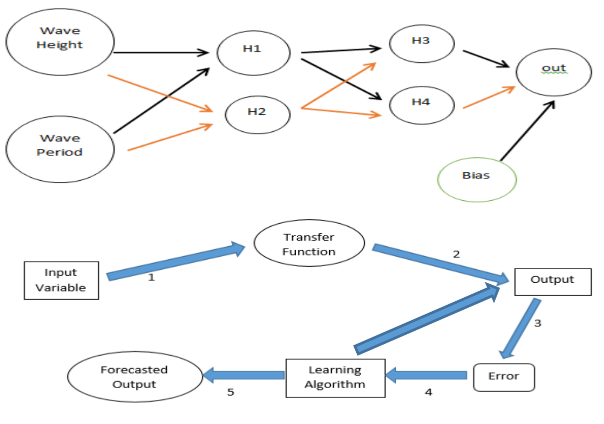

Figure 1. ANN model with its layers and processes involved.

Methodology

This section elaborates on the variables used while developing different forecasting models to predict the wave energy generated at a nearshore location in the Fiji Islands. Data were collected at Mauni Bay, a site near the Island of Kadavu at a latitude of $19^{0}9'52.80^{0}S$ with a longitude of $178^{0}8'39.34''E$ at a depth of $18m$, approximately $900m$ from the shore [8]. Wave height and wave period were the independent variables while the wave power was the dependent variable. The data was then divided into two groups: the training group had 80% of the data, which had 2192 data values while the testing group had remaining 20% of data, which had 549 data values.

Time Series Models

The training data were used to develop four time series models using the statistical EView software. Coefficients like $\alpha, \beta, and\ \gamma$ were obtained using the training data and were kept fixed while running the model using the test data. The equations used for each time series model are given below.

Exponential Smoothing Model

$$A_t = \alpha x_t + (1-\alpha)A_{t-1} \qquad (5)$$

Holt Winters Additive model

$$L_t = \alpha x_t + (1-\alpha)(L_{t-1} + T_{t-1}) \qquad (6)$$

$$T_t = \beta(L_t - L_{t=1})x_t + (1-\beta)T_{t-1} \qquad (7)$$

$$y_t = L_t + T_t + S + \varepsilon_t \qquad (8)$$

Holt winters Multiplicative Model

$$f_{t,k} = (L_t + kT_t)S_{t+k-c} \qquad (9)$$

$$L_t = \alpha \frac{x_t}{S_{t-c}} + (1-\alpha)(L_{t-1} + T_{t-1}) \qquad (10)$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \qquad (11)$$

$$S_t = \gamma \frac{x_t}{L_t} + (1-\gamma)S_{t-c} \qquad (12)$$

where L_t is the estimate of base level, T_t is the estimate of per-period trend, x_t is the estimate of the period t base level, S is the additive seasonality factor, ε_t is the error term, k is the number of forecast ahead, α, β and γ are smoothing constants with a range of 0 to 1 [7].

Multiple Regression Models

Three linear and two nonlinear regression models were developed using the training data in the R statistical software.

The linear regression models are

Mod.1 : $Power = \beta_0 + \beta_1.waveheight + \varepsilon$

Mod.2: $Power = \beta_0 + \beta_1.waveperoid + \varepsilon$

Mod.3: $Power = \beta_0 + \beta_1.waveheight + \beta_2.waveperiod + \varepsilon$

The two nonlinear models: a quadratic model and a Box-Cox transformation model developed are

Mod.4: $Power = \beta_0 + \beta_1.waveheight + \beta_2.(waveheight)^2 + \varepsilon$

Mod.5: $Power = (waveheight + waveperiod)^2 + \varepsilon$

The optimum regression model identified was based on the selection criteria, which were Akaike information criterion (AIC), Bayesian information criterion (BIC), sum squared error (SSE), residual error and coefficient of determination (R^2), and the four assumptions of an MLR model as mentioned in the previous section.

Artificial Neural Network

The strength of any ANN model for forecasting lies in its formation and its structure that allows multiple mapping between the variables. The functioning and performance of any ANN model solely lies in the selection of its parameters, which is identifying the right number of input variables, hidden nodes, hidden layers and the transfer function used.

An R software was used to develop thirty-four ANN models. Model parameters were continuously varied to achieve the optimum ANN model. The steps and functioning of each parameter in the model are described below:

1. The input variables identified were in the form of wave height and wave period while the output variable was wave power generated.
2. Using $x_n = (x_o - x_{\min})/(x_{\max} - x_{\min})$, the three variables were normalized, to ensure all computational problems are avoided by allowing same weight to each hidden layer and hidden nodes.
3. Inputs of $\{i_1, i_2, i_3, \ldots\ldots\ldots, i_n\}$ are passed to the hidden layer. In the hidden layer, the input node is attached with some weights. The model is trained to get closer to the targeted vectors $\{t_1, t_2, t_3, \ldots\ldots\ldots, t_Q\}$. The output from neuron i, O_i is connected to the input of neuron j through interconnection weight w_{ij}. Then finally the output will be
$$O_k = f\left(\sum_i^Q w_{ik} O_i + b\right)$$,where $f(x)$ is the transfer function.
4. The forecasted power is de-normalized using $\text{Predicted} = x_{normalized}(x_{\max} - x_{\min}) + x_{\min}$.
5. The attached weights are re-adjusted using a learning algorithm.

Once all the optimum models were identified, comparison was made based on the following model evaluators using the testing data. The measures of accuracy used in this paper are:

1. Mean squared Error (MSE) : $MSE = = \frac{1}{N}\Sigma(Y_t - A_t)^2$
2. Mean Absolute Error (MAE): $MAE = \frac{\Sigma|(Y_t - A_t)|}{N}$
3. Mean Absolute percentage Error $MAPE = \frac{100}{n}\left(\Sigma\left(\frac{|A_t - Y_t|}{A_t}\right)\right)$
4. Root Mean Squared Error : $RMSE = \sqrt{\frac{\Sigma(Y_t - A_t)^2}{n}}$
5. The goodness-of-fit measures considered are the coefficient of correlation. It is given by $R^2 = 1 - \frac{\Sigma(Y_t - A_t)^2}{\Sigma(Y_t - \bar{A})^2}$

The optimum model is the one with the lowest MSE, MAE, MAPE, RMSE and with the highest R^2 value.

Results

Using the training data, different time-series forecasting models are fitted. Based on the model evaluators presented in Table 1, the optimum times series model is found to be HW's Additive model. The model recorded highest R^2 value, lowest MSE and RMSE values and the third lowest MAPE. However, it recorded the highest MAE error value.

Model	MSE	RMSE	MAPE	MAE	R^2
S. Exp.	28.501	28.5011	56.80	**2.884**	0.7691
D. Exp	32.662	5.71503	56.67	3.042	0.7354
HW add.	**28.362**	**5.32559**	57.48	2.896	**0.7707**
HW mult	28.423	5.33130	**55.52**	2.886	0.7698

Table 1: Error values for time series model using training data.

Figure 2 presents the predicted and actual values of the power and the residual of HW Additive model using test data.

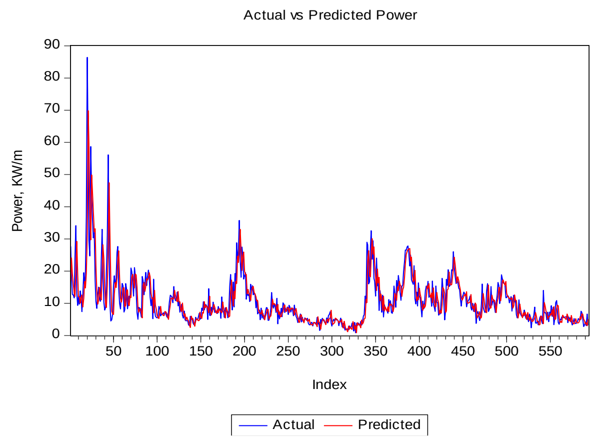

Figure 2. Graph of Predicted and Actual power for optimum time series model-HW additive model.

Similarly, based on the different model selection criteria, the optimum MLR model among the five models fitted in the training data was the Box-Cox nonlinear transformation model (Mod.5). The error measures of the five MLR models are presented in Table 2.

Model	AIC	BIC	SSE	Resid	R^2
Mod.1	11467.3	11484.4	25391.5	3.43	0.8570
Mod.2	15385.3	15402.4	155497	8.49	0.1245
Mod.3	10160.8	10183.5	13862.4	2.53	0.9219
Mod.4	10537.6	10560.3	16501.8	2.77	0.9071
Mod.5	**8679.3**	**8707.7**	**6979.7**	**0.12**	**0.9607**

Table 2: Error values of different MLR models using training data.

Model 5 is the best regression model since it has the highest R^2 value, lowest AIC, BIC, SSE and residual standard error compared to other MLR models that were developed. The second best MLR model was Model 3 which has the second highest R^2 value, second lowest AIC, BIC, SSE and residual error. The predicted and actual values of the power and the residual of Model 5 are shown in Figure 3.

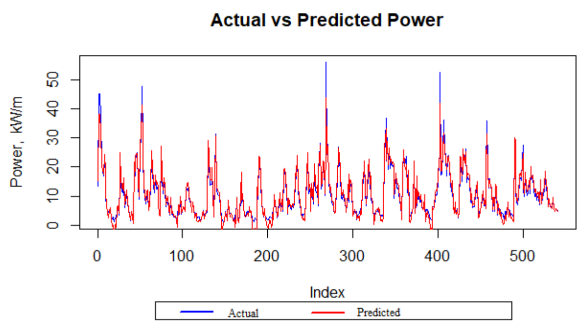

Figure 3. Actual vs Predicted power from best MLR model - Model 5.

The performance of the three best ANN models are presented in Table 3. ANN(7,2,1) is proposed to be the best ANN model using the training data set. It has the highest R^2 value, lowest MSE and RMSE values and the second lowest MAE and MAPE values. It is a two hidden layer network, which has 7 hidden nodes in the first layer with 2 nodes in the second hidden layer.

It uses the tangent hyperbolic transfer function over the logistics transfer function. The second best ANN model was ANN (4,3,1) which recorded the second highest R^2 value, lowest MAE and MSE values and second lowest MSE and RMSE values. These three best ANN models were achieved using the tangent hyperbolic transfer function.

Mod.	MSE	RMSE	MAPE	MAE	R^2
ANN(3,2,1)	0.0121	0.1100	1.3424	0.0746	0.9998
ANN(7,2,1)	**0.0076**	**0.0873**	1.2664	0.0654	**0.9999**
ANN(4,3,1)	0.0089	0.0944	**0.8300**	**0.056**	0.9998

Table 3: Error values for best three ANN model using training data

Figure 4 shows the predicted and actual values of the power and the residual of best ANN (7,2,1) model.

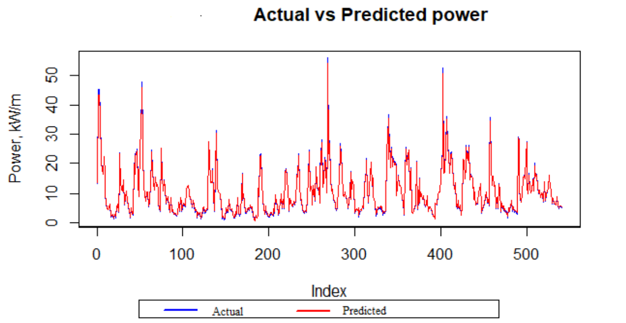

Figure 4: Comparison of actual and predicted power from ANN (7,2,1) model using test set data.

Discussion

In this section, a comparative study on the efficiency of the best three forecasting models HW additive model in time series, MLR Mod.5 and ANN (7,2,1) is conducted using testing data.

Table 4 displays the results of the error measures of the comparative analysis of the optimum models developed while forecasting the wave power generated.

Model	MSE	MAE	RMSE	MAPE	R^2
HW Add.	7.281	0.8024	2.6984	7.7758	0.5674
MLR mod. 5	2.701	**0.1707**	1.6434	20.1357	0.9591
ANN(7,2,1)	**0.165**	0.3106	**0.4061**	**4.9059**	**0.9975**

Table 4: Comparative analysis for optimum models developed using the testing data set.

From Table 4, it is evident that the proposed ANN (7,2,1) model outperforms the different models formed using the testing data. ANN (7,2,1) model when compared with rest of the models displayed the highest R^2 value, lowest MSE, RMSE, MAPE and second lowest MAE values. Similar results were reported by Hatalis et al. [3] and Mandal & Prabaharan [6] while using ANN models to forecast the wave power and its parameter. The second best model is the multiple regression model 5, which has the second highest R^2 value, second lowest MSE, RMSE and the lowset MAE value. However, this MLR model performed worst in terms of MAPE, recording the highest percentage error.

Conclusion

All the forecasting models developed are good at forecasting, but the robustness of a model totally depends on its individual unique properties. Wave data usually have some trends, seasonality and non-linear characteristics due to the nature of the waves. The comparative analysis within the optimum models formed illustrates that the artificial neural network model is a useful forecasting model, which can predict the power more accurately. The empirical results also revealed that the artificial neural network model is more efficient in forecasting wave energy in comparison to the regression and time series models.

As explained by Kohzadi et al. [5], ANN models perform best if the data exhibits non-linear or chaotic behaviour, which cannot be easily captured by other models. Thus, it can conclusively be noted that ANN models perform best provided alterations to its parameters are made based on the different scenarios presented.

References

[1] Fleix, A., Hernandez-Fontes, J. V., Lithgow, D., Mendoza, E., Posada, G., Ring, M., and Silva, R. (2019). Wave Energy in Tropical Regions: Deployment challenges, Environmental and Social Perspectives. *Journal of Marine Science and Engineering*, 7(7), 219 (DOI: 10.3390/jmse7070219).

[2] Hadadpour, S., Etemad-Shahidi, A., and Kamranzad, B. (2014). Wave Energy Forecasting using Artifical Neural Networks in the Caspian Sea. *Proceedings of the Institution of Civil Engineers - Maritime Engineering*, 167(1), 42-52 (DOI: 10.1680/maen.13.00004).

[3] Hatalis, K., Pradhan, P., Kishore, S., & Blum, R. (2014). Multi-step Forecasting of Wave Power Using a Nonlinear Recurrent Neural Network. . *IEEE.*

[4] Khashei, M., and Bijari, M. (2010). An Artifical Neural Network (p, d, q) Model for Timeseries Forecasting. *Expert System with Application,* 37(1), 479-489 (DOI: 10.1016/j.eswa.2009.05.044).

[5] Kohzadi, N., Boyd, M., Kermanshahi, B., and Kaastra, I. (1993). A Comparsion of Artifical Neural Network and Time Series Models for Forcasting Commoditiy Prices. *Neurocomputing*, 10(2), 169-181 (DOI: 10.1016/0925-2312(95)00020-8).

[6] Mandal, S., & Prabaharan, N. (2005). Ocean wave forecasting using recurrent neural networks. *Ocean Engineering 33 (2006)*, 1404-1410

[7] Pongdatu, G., and Putra, Y. (2018). Seasonal Time Series Forecasting using SARIMA and Holt-Winter's Exponential Smoothing. *IOP Conference Series. Materials Science and Engineering.* 407, 012153 (DOI: 10.1088/1757-899X/407/1/012153.

[8] Ram, K., Narayan, S., Ahmed, M.R., Nakavulevu, P., and Lee, Y.-H. (2014). In situ near-shore wave resource assessment in the Fiji Islands. *Energy for Sustainable Development, 23,* 170-178 (DOI: 10.1016/j.esd.2014.09.002).

[9] Reikard, G. (2013). Integrating wave energy into the power grid: Simulation and forecasting . *Ocean Engineering (73)*, 168-178.

[10] Uyank, G. K., and Guler, N. (2013). A Study of Multiple Linear Regression Analysis . *Procedia - Social and Behavioral Science*, 106, 234-240 (DOI: 10.1016/j.sbspro.2013.12.027).

[11] Winston, W. L. (2003). *Operation Research: Applications and Algorithms*. *Cengage Learning,*

[12] Zhang, G., Patuwo, B. E., and Hu, M. Y. (1998). Forecasting with Artifical Neural Networks: The state of the art. *International Journal of Forecasting*, 14(1), 35-62 (DOI: 10.1016/50169-2070(97)00044-7).