

## Modelling of Sea Water Level During High Tide Using Statistical Method and Neural Network

(Permodelan Terhadap Aras Air Laut Semasa Air Pasang dengan Menggunakan Kaedah Berstatistik dan Rangkaian Neural)

Firdaus Mohamad Hamzah<sup>a\*,b</sup>, Wong Khai Seen<sup>c</sup>, Izamarlina Asshaari<sup>a,b</sup>, Mohd Saifullah Rusiman<sup>d</sup>, Mohd Khairul Amri Kamarudin<sup>e</sup> & Shamsul Rijal Muhammad Sabri<sup>f</sup>

<sup>a</sup>Department of Engineering Education, Faculty of Engineering & Built Environment, Universiti Kebangsaan, Malaysia

<sup>b</sup>Centre for Engineering Education Research, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia

<sup>c</sup>Department of Civil Engineering Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, Malaysia

<sup>d</sup>Faculty of Applied Sciences and Technology, Universiti Tun Hussien Onn Malaysia, Malaysia

<sup>e</sup>Faculty of Applied Social Sciences, Universiti Sultan Zainal Abidin, Malaysia

<sup>f</sup>School of Mathematical Sciences, Universiti Sains Malaysia, Malaysia

\*Corresponding author: fir@ukm.edu.my

Received 15<sup>th</sup> June 2022, Received in revised form 1<sup>st</sup> August 2022

Accepted 1<sup>st</sup> September 2022, Available online 15<sup>th</sup> November 2022

### ABSTRACT

Recently, the rise of sea level has caused an increase in rising tides that affected about three million locations around the world. The tide rising phenomenon has been occurring in Peninsular Malaysia since the 20th century. The purpose of this study is to determine the most critical station and forecast three stations located on the West Coast of Peninsular Malaysia. The Box Plot analysis method has been used in this study; the results shown that Bagan Datuk Station is the most critical station. This is due to the maximum tide's value of Bagan Datuk Station experienced the highest increment of 0.45 m, compared to Port Klang station and Permatang Sedepa Station with only 0.2 m increment in 10 years. However, these three stations are also experiencing rising tides. Thus, the focus of managing coastal structures should be given to all these three stations as well. In addition, for forecasting, the Artificial Neural Network (ANN) forecasting model provides better forecasting results compared to the Autoregressive Integrated Moving Average (ARIMA) model for long-term forecast. In this study, the artificial Neural Network (ANN) forecasting model obtained value of RMSE 0.05642 at Bagan Datuk Station compared to the RMSE value of 0.0928 obtained from the ARIMA model at the same station. Besides, MAE value of ANN method, 0.04387 compared to the MAE value of ARIMA which is worth 0.06391 at Bagan Datuk Station. This study can conclude that the Artificial Neural Network (ANN) forecasting model is better in high tide forecasting.

**Keywords:** Tidal rising; Coastal flooding; Forecasting; Artificial Neural Network (ANN); ARIMA

### ABSTRAK

Kebelakangan ini, peningkatan air pasang yang disebabkan oleh faktor peningkatan aras air laut telah mempengaruhi hampir tiga juta lokasi di seluruh dunia. Fenomena peningkatan air pasang telah berlaku di Semenanjung Malaysia bermula dari abad ke-20. Tujuan kajian ini dijalankan adalah untuk menentukan stesen paling kritikal dan meramalkan ketiga-tiga stesen yang terletak di Pantai Barat Semenanjung Malaysia. Kaedah analisis Plot Kotak telah digunakan dalam kajian ini, keputusan menunjukkan Stesen Bagan Datuk adalah stesen yang paling kritikal. Hal ini disebabkan oleh nilai maksimum Stesen Bagan Datuk mengalami peningkatan yang paling tinggi iaitu dengan nilai 0.45 m, berbanding dengan stesen Pelabuhan Klang dan Stesen Permatang Sedepa yang mengalami peningkatan 0.2 m sahaja dalam julat 10 tahun. Walau bagaimanapun, ketiga-tiga stesen ini juga mengalami peningkatan air pasang. Oleh itu, tumpuan pengawasan dalam pengurusan struktur pesisiran pantai perlu diberi kepada ketiga-tiga stesen ini. Di samping itu, untuk peramalan, model peramalan Rangkaian Neural Buatan (ANN) memberikan hasil peramalan yang lebih baik berbanding dengan model Auto Regresi Purata Bergerak (ARIMA) untuk peramalan jangka panjang. Dalam kajian ini, model peramalan Rangkaian Neural Buatan (ANN) memperoleh nilai RMSE 0.05642 di Stesen Bagan Datuk, berbanding dengan nilai RMSE 0.0928 yang

diperoleh daripada Model ARIMA. Selain itu, nilai MAE bagi kaedah ANN, 0.04387 berbanding dengan nilai MAE bagi ARIMA yang bernilai 0.06391 di Stesen Bagan Datuk. Melalui kajian ini, dapat disimpulkan bahawa model peramalan Rangkaian Neural Buatan (ANN) lebih baik dalam peramalan air pasang.

*Kata kunci:* Kenaikan air pasang; Banjir air pasang; Peramalan; Rangkaian Neural Buatan (ANN); ARIMA

## INTRODUCTION

Sea level rising has been considered as one of the major threats to especially coastal areas around the world since global warming appeared in the 20th century (Feng et al. 2018). Sea level rise is mainly due to the melting of glaciers and ice during global warming (Li et al. 2019; Lindsey & Lumpkin 2020; Yin et al. 2015). The factor that causes global warming is because of the increase in the concentration of carbon dioxide (CO<sub>2</sub>) in the atmosphere. Carbon dioxide is a global warming and acidification gas. In addition, production such as gases from vehicles, logging of flora, use of air conditioners and others that will produce global warming gases (Boyd et al. 2015; Johnson et al. 2015; Zhang et al. 2018). The rising of sea level at the same time will cause rising of tidal, thus, forecast of tidal are known to be important issues in the design of coastal structures, tidal energy use, prevention of marine disasters and military affairs (Olbert et al. 2017). Previous research, at Baltic Sea in the eastern and south-western regions of the sea, the rate of increase is the fastest and water level extremes are influenced by an increase in water volume maxima of 3-4 mm/year while the Bangal Delta's, Bangladesh water level are rising at a rate of 4.0 between 4.8 mm per year. (Feist et al. 2021; Pindsoo et al. 2020). This is because of the weather and climate uncertainty (Yusofa et al. 2022).

Knowledge of tidal levels and the ability to predict the tidal levels has a direct implication to engineering profession, especially for coastal engineering (Pramanik et al. 2016). Forecasting is classified as a qualitative method, where the historical data on variables are predicted either inappropriate or unavailable and forecasts can be used based on expert judgment (Islam et al. 2015; Srivastava et al. 2016). Forecasting cannot be performed with harmonic models alone because this method requires long historical data and has high errors (Liu et al. 2019). This will bring challenges to the accurate of tide forecasts (Consoli et al. 2014). There are various kinds of forecasting techniques and they are generally competitive (Espinoza 2015; Graff et al. 2017). Therefore, the use of forecasting models such as ARIMA and Artificial Neural Network (ANN) to correct forecasting errors. Throughout the literature, the Auto Regressive Integrated Moving Average (ARIMA) and Artificial Neural Network (ANN) models have been applied for forecasting sea level from critical station.

Boxplots are also known as box and whisker plot method. It is useful in identifying its properties and comparing distributions. Boxplot provides basic information about distribution. It graphically depicts numerical data sets according to their quartiles. Box plots only work well when there is sufficient data to provide statistics. In a boxplot, it

has a minimum value, lower quartile, median, upper quartile, maximum value (Praveen et al. 2018; Vignesh et al. 2018).

The ARIMA model, also known as the Box-Jenkins model and is commonly used in analysis and forecasting. ARIMA model is widely used for forecasting technique in social science and extensively for time series. The use of ARIMA in time series forecasting is essential with uncertainty as it does not assume knowledge of any underlying model. Model ARIMA relies on the historical values of the series as well as previous error terms for the prediction (Fattah et al. 2018; Manoj 2014; Shu et al. 2014). ARIMA models are now widely used for many applications, such as nature, medicine, and engineering. It is also more robust and efficient than other more complex structural models, especially in short-term prediction (Adebiyi et al. 2014; Tealab et al. 2017; Tylkowski & Hojan 2019). Akaike Information Criterion (AIC) and The Bayesian Information Criterion (BIC) are both penalized likelihood criteria. They are sometimes used for choosing best predictor subsets in regression and often used for comparing non-nested models, which ordinary statistical tests cannot do (Reddy et al. 2017). In developing ARIMA model, analysis of autocorrelation function (ACF) need to be performed for chosen ARIMA model. In validating the model, diagnostics checking need to be developed (Alsharif et al. 2019; Bakar & Rosbi 2017).

The ANN method has been widely used in forecasting selected natural phenomena over the years (Abiodun et al. 2018). In coastal engineering, the ANN model is used to evaluate the stability of shield units and rock breakers. As a data-driven technique, the neural network can represent non-horizontal and non-periodic complex relationships. It is also suitable for high and low tide forecasts (Meena & Agrawal 2015). Multi-layer Perceptron (MLP) is a popular learning technique in ANN. It is basically a mathematical model inspired by behaviour of the human brain and nervous system. MLP is basically made up of three layers; input layer, hidden layer, and output layer (Mohapatra et al. 2019; Olatunji et al. 2019).

This paper aims to determine the most critical station. Next, generate tidal forecasting model using ARIMA and ANN method. Finally, the most suitable model for tidal forecasting is determined in this study. Artificial Neural Network (ANN) model is better than model ARIMA model in time series dataset (Garg et al. 2016; LOLA et al. 2018). The results were obtained by comparing the root-mean-square-error (RMSE) and mean-absolute error (MAE) for ARIMA and ANN models respectively. Generally, the RMSE and MAE are not strong explanatory by themselves, but they can be used to provide a more accurate comparison between the forecasting models. In general, the smaller the value, the

better the model (Dikshita & Singh 2019; Jahanshahi et al. 2019; Prahutama & Mustafid 2016; Srivastava et al. 2016).

## STUDY AREA

### CASE STUDY AND DATA PREPARATION



FIGURE 1. Location of Case Study

The study area was conducted at the states of Perak and Selangor. The study areas selected were station Bagan Datuk, Port Klang and Permatang Sedepa. Bagan Datuk is located south of Perak and is the most south western region of Perak. The area borders the Perak River which separates Manjung and Central Perak. Port Klang is a major port in Malaysia. Originally known as Swettenham Harbor, it is the largest and busiest port in Malaysia. Permatang Sedepa is an offshore lighthouse in the Straits of Malacca.

TABLE 1. Data on Study Area

Station	Data duration	Coordinate	Period
Bagan Datuk	2008-2017	4.00° N 100.75° E	10 years
Permatang Sedepa	2008-2017	2.88° N 100.98° E	10 years
Port Klang	2008-2017	3.00° N 101.38° E	10 years

Data is the most important element in this study. It is the basis for the analysis and discussion of this study. For this study, tidal daily history data were required. Therefore, the daily data used in this study are from the Pusat Cerapan Bumi, UKM and the National Institute of Hydraulic Research (NAHRIM). Table 1 above shows the details of the summarized station.

## METHODOLOGY

### BOXPLOTS

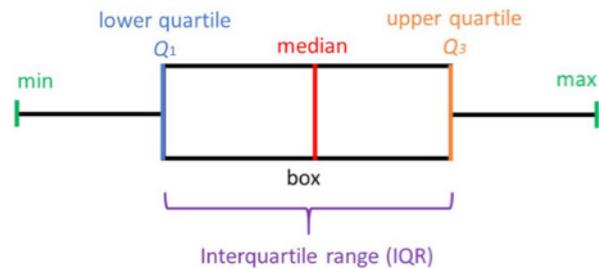


FIGURE 2. Example of Boxplots

Boxplots able to identify the values: minimum value, lower quartile, median, upper quartile, maximum value. Lower quartile ( $Q_1$ ) is the value such that 25% of the data lie at. Median, which known as second quartile ( $Q_2$ ) is indicates as midpoint. If the data set has an odd number of observations, then the median is the middle observation (ranked data). When number of data is even, median is the mean of the two middle observations. For upper quartile ( $Q_3$ ) is the value such that 75% of the data lie at. The interquartile range (IQR) is the length of the box and envelops all of the data between  $Q_3$  and  $Q_1$ . Interquartile Range (IQR) is little affected by the presence of outliers and is a measure that is very useful for comparing two data sets. Fences are the limits above and below the box (generally not visualized) that are used to flag possible fence is the lower limit computed as  $Q_1 - 1.5 \times IQR$  (Kendrick et al. 1989; Praveen et al. 2018; Vignesh et al. 2018).

### Auto Regressive Integrated Moving Average (ARIMA)

The model ARIMA  $(p, d, q) \times (P, D, Q)_s$  can be expressed as:

$$\phi_p(B^s)(1 - B^s)^d \phi_P(B)(1 - B^s)^D y_t = \delta + \theta_Q(B^s)\theta_Q(B)\epsilon_t \quad (1)$$

Where:

$\delta$  is constant value.

$y_t$  is the original values of time series.

$\epsilon_t$  is random noise.

$B$  is backshift operator,  $BY_t = Y_{t-1}$

$p$  is autoregressive model AR order

$d$  is the differentiation order

$q$  is moving average model MA order

$P$  is autoregressive seasonal model SAR order

$Q$  moving average seasonal model SMA order

There are several steps to ARIMA modeling. The first step is the introduction section. Typically, the ARIMA model has only three parameters  $(p, d, q)$  where  $p$  represents the seasonal AR sequence,  $d$  for the seasonal differential sequence, and  $q$  for the seasonal MA sequence. The number

of terms used in the AR and MA models sections is not arbitrarily selected. Box-Jenkins has proposed structural methods to determine the best model to use. There are 5 stages to obtain ARIMA model (Fernandez et al. 2017):

1. Time series exploratory analysis
2. Model identification
3. Model Estimation
4. Model diagnostic examination
5. Forecasting

For the efficiency measurement of all approaches, three performance statistics were used in this study to estimate performance between observed and predicted tidal water values, agreement index (d), Root-mean-square-error (RMSE) and mean absolute error (MAE).

$$d = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (|y_i - x_i| + |x_i - y_i|)^2} \quad (2)$$

Root-mean-square-error (RMSE) can be expressed by:

$$RMSE = \sqrt{\left(\frac{1}{n} \sum_{i=1}^n [y_i - x_i]^2\right)} \quad (3)$$

Mean absolute error (MAE) can be defined by:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - x_i| \quad (4)$$

$n$  is the number of observations,  $x$  is the measured and  $y$  is the predicted value (Collantes-duarte & Rivas-echeverria 2015; Srivastava et al. 2016).

Penalized-likelihood criteria:

Akaike information criteria (AIC) and Bayesian Information Criteria (BIC) are penalized-likelihood criteria. They are used to select the best subset predictors in regression.

Equations for ARIMA non-seasonal models:

$$AIC = -2 \log \log (L) + 2(p + q + k)$$

Where  $L$  is the maximum value of the probability function,  $p$  is the arrangement of the autoregressive part and  $q$  is the arrangement of the moving average parts.  $k$  represents the ARIMA model shortcut.

Bayesian Information Criteria (BIC) is a model selection criterion between a limited set of models. It is based on probability function, and is closely related to Akaike Information Criteria (AIC).

The BIC equation is:

$$BIC = AIC + (\log \log (T) - 2)(p + q + k)$$

The lower these criteria (AIC or BIC) for the model, the better the model fits the data (Ongbali et al. 2018; Ray & Bhattacharyya 2015).

Artificial Neural Network (ANN)

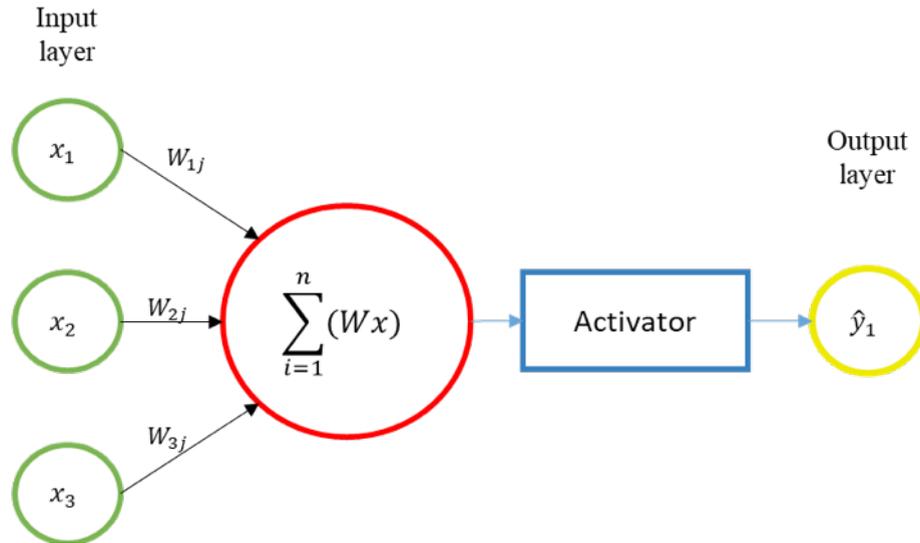


FIGURE 3. Simplified Figure of ANN

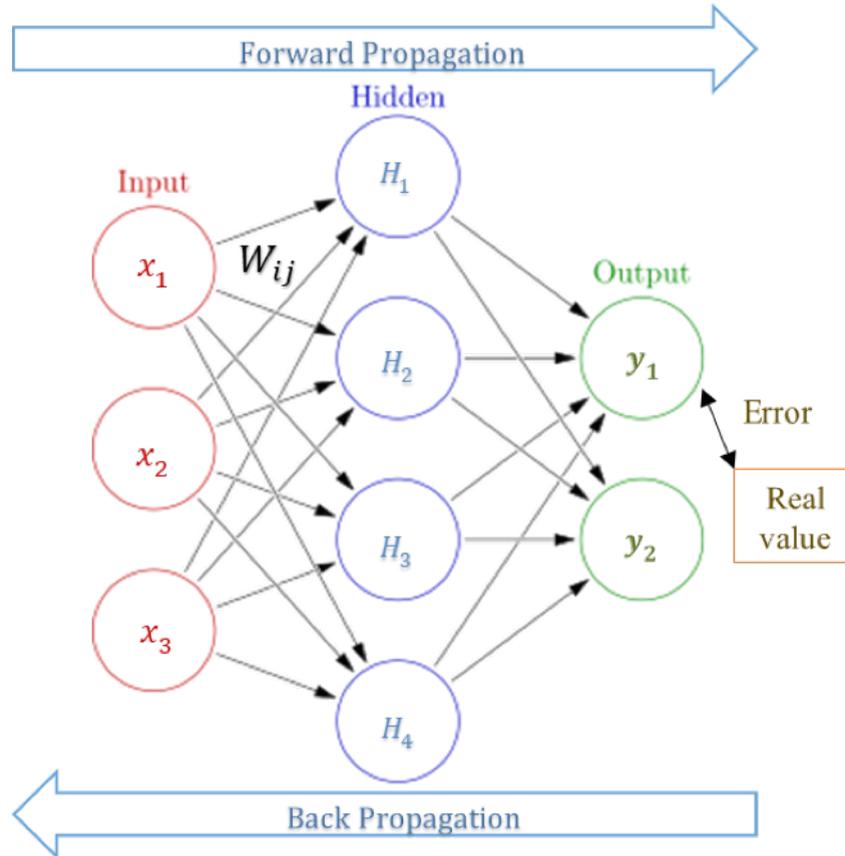


FIGURE 4. Structural Forecasting Methods of Artificial Neural Networks (ANN)

A typical three-layered ANN with the input layer I, hidden layer H, and output layer O. The most commonly used transfer function in geoscience is the sigmoid function  $f(x)$  described by equation (Yin et al. 2015):

$$f(x) = (1 + e^{-x})^{-1} \quad (5)$$

Whereas, the differentiation can be expressed by:

$$\frac{df}{dx} = f(x)[1 - f(x)] \quad (6)$$

The error function in output neuron can be defined as:

$$E = \frac{1}{2} \sum_n (T_j - A_j)^2 \quad (7)$$

Where,  $T_j$  dan  $A_j$  are the target and output neuron respectively. The gradient descent algorithm adjusts the weight according to the gradient error:

$$\Delta W_{ij} = -\eta \frac{\partial E}{\partial W_{ij}} \quad (8)$$

Where  $\eta$  is the learning rate and general form of the  $\frac{\partial E}{\partial W_{ij}}$  is expressed as

$$\frac{\partial E}{\partial W_{ij}} = \delta_j^n \cdot A_i^{n-1} \quad (9)$$

Substituting equation (9) to equation (8), the gradient error becomes:

$$\Delta W_{ij} = -\eta \delta_j^n \cdot A_i^{n-1} \quad (10)$$

Where  $A_i^{n-1}$  is the output value of the sub-layer associated with the weighting ( $W_{ij}$ ). If neuron th is one of the output neurons, then neuron is a hidden layer neuron with the equation:

$$\delta_j = (T_j - Y_j) Y_j (1 - Y_j) \quad (11)$$

Finally, the weighting value can be expressed in equation:

$$\tilde{W}_{ij}^m = W_{ij}^{m-1} + \Delta W_{ij}^m = W_{ij}^{m-1} + \eta \delta_j^n A_i^{n-1} \quad (12)$$

The steps of Artificial Neural Network model are (Gao et al. 2017; Mohanty & Mohanty 2015; Shamsuddin et al. 2008; Tealab et al. 2017):

1. All weight, network threshold levels are randomly distributed in small range.
2. The first dataset observation inserted in input layer, each feature in one node.
3. Forward propagation: From left to right, neurons are activated by each activation of neurons restricted by the weight.
4. Compare output with real value to obtain error gradient.

- 5. Back propagation: From right to left, error propagates again. Update the weight according to how much of the error.
- 6. f. Increase of iteration repeated steps a-e until the selected error criterion is satisfied

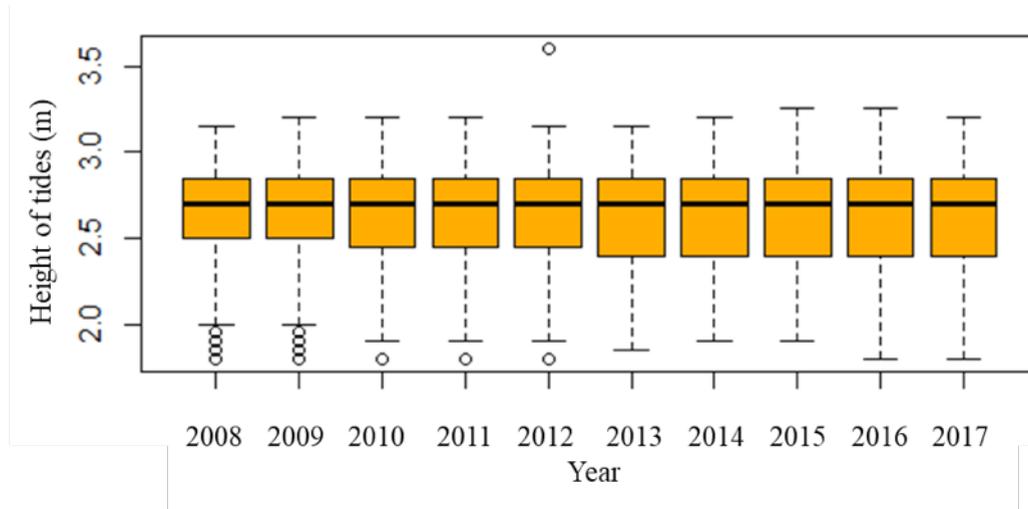


FIGURE 5. Boxplots for Bagan Datuk Station from 2008-2017

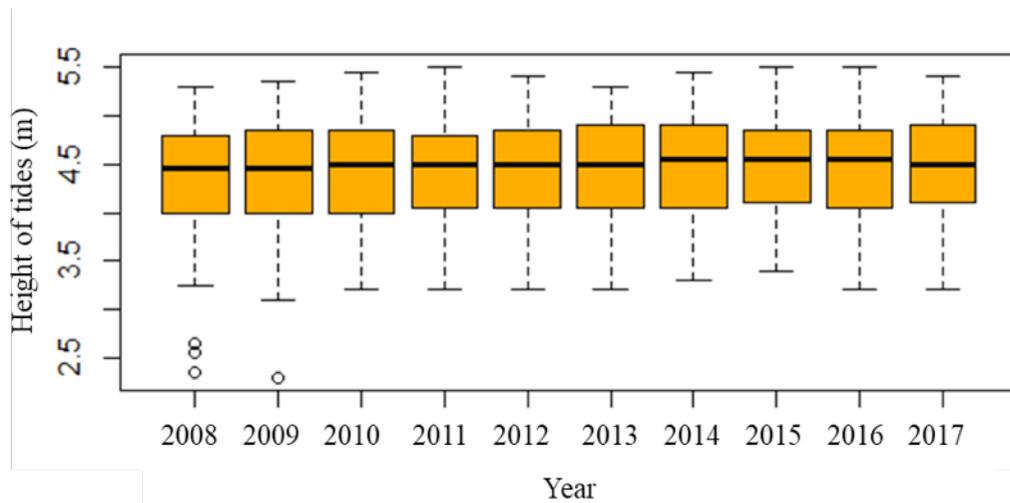


FIGURE 6. Boxplots for Port Klang Station from 2008-2017

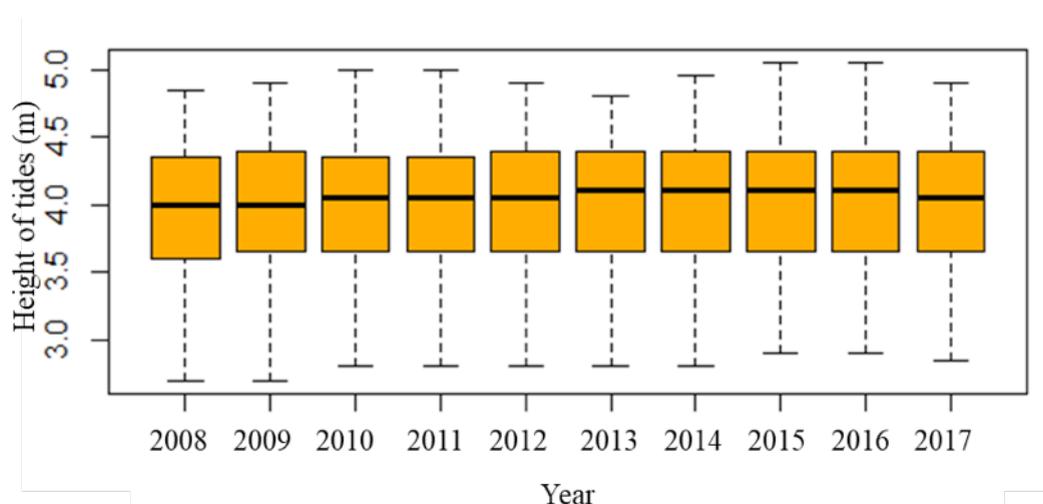


FIGURE 7. Boxplots for Permatang Sedepa Station from 2008-2017

Figure 5, 6 and 7 shown the historical tidal data (2008-2017) obtained from the Department of Hydrography Malaysia for Bagan Datuk Station, Port Klang station and Permatang Sedepa station respectively with box plot analysis. In Figure 5, Bagan Datuk station has an abnormal maximum value in 2012, which has a maximum value exceeding 3.5 m. In Figure 6, Port Klang station have abnormal maximum value in 2008 and 2009, which is the minimum value below 2.5m. This is because the data collection using an incorrect technique or the value is affected by external issues. However, for Figure 7, Permatang Sedepa station do not has any abnormal value. This shown that the historical data collected using the correct technique so that able to perform higher accuracy. Overall, Figure 5, 6 and 7 shows that the tide has experienced a consistent increase in the last 10 years.

From the box plot analysis in experiment, table below shows a comparison of the box plot analysis for all three stations from year 2008 to 2017. Through box plot, the most critical stations are determined. In this analysis, the critical stations will be given high attention so that management and adaption of the coastline can be control well before tide flooding or sea level rising.

TABLE 2. Comparison of Boxplots Analysis between All Three Stations

Station	Bagan Datuk		Pelabuhan Klang		Permatang Sedepa	
	2008	2017	2008	2017	2008	2017
Year	2008	2017	2008	2017	2008	2017
Min	1.80	1.80	2.35	2.30	2.70	2.70
First Quartile	2.50	2.45	4.00	4.05	3.60	3.65
Median	2.70	2.70	4.45	4.50	4.00	4.05
Mean	2.64	2.64	4.39	4.44	3.95	4.00
Third Quartile	2.85	2.85	4.80	4.85	4.35	4.40
Max	3.15	3.60	5.30	5.50	4.85	5.05

Based on Table 2 above, the maximum value of Bagan Datuk Station has experienced the highest increase from 3.15 m to 3.6 m in the last 10 years. For Port Klang and Permatang Sedepa stations, they also experienced an increase in the tide with a value of 5.3 m increased to 5.5 m for Port Klang Station and 4.85 m to 5.05 m for Permatang Sedepa Station. The increasing difference for Station Bagan Datuk indicates that the tide at this station is experiencing a higher increment. Therefore, in this study, Bagan Datuk

Station was identified as the most critical station because it experienced an increase of 0.45 m compared to Port Klang station and Permatang Sedepa Station which experienced an increase of 0.2 m only in 2008 to 2017. However, all three the station is also experiencing an increase in tides. Therefore, the focus of supervision in managing coastal structures should be given to these three stations. However,

Bagan Datuk Station experienced the most increase and is known as a critical station. More focus should be given to this station.

FORECASTING MODEL ANALYSIS

ARIMA MODEL

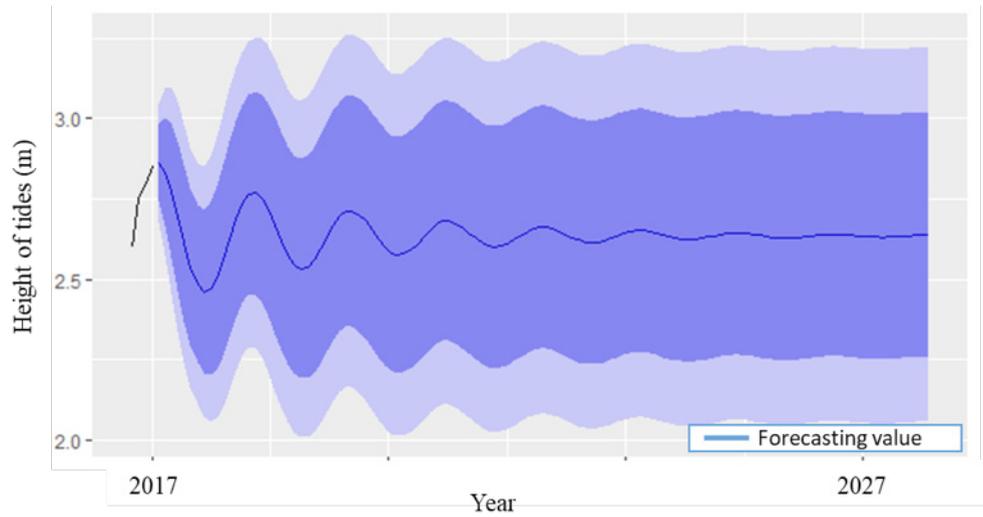


FIGURE 8. ARIMA (3, 0, 1) model for Bagan Datuk station in 10 years

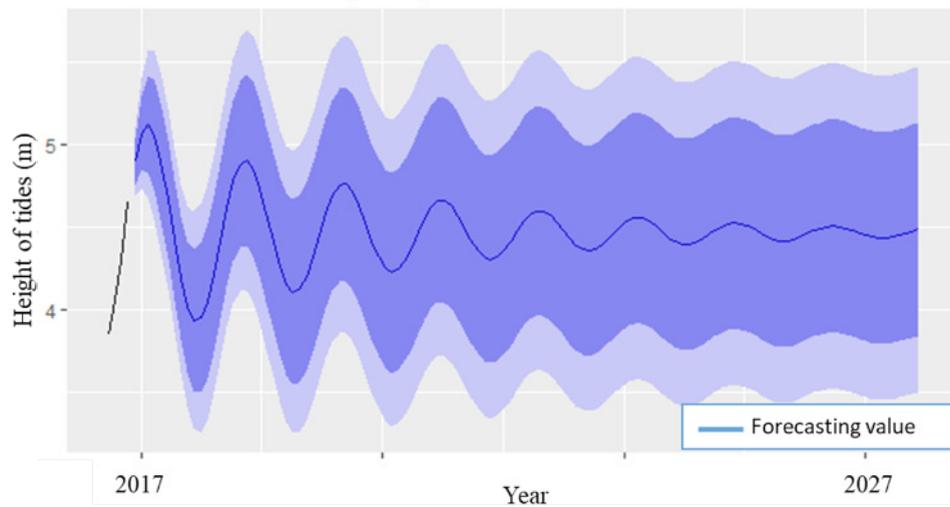


FIGURE 9. ARIMA (2, 0, 3) model for Port Klang station in 10 years

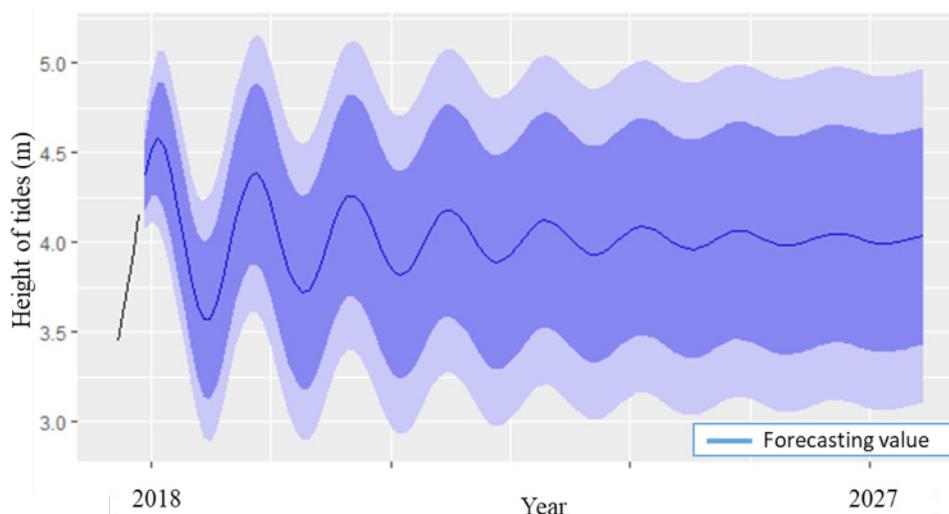


FIGURE 10. ARIMA (2, 0, 2) model for Permatang Sedepa station in 10 years

Figure 8, 9 and 10 show the ARIMA forecasting model for Bagan Datuk, Port Klang and Sedepa Station in 10 years. The square-rooted error (RMSE) and mean absolute error (MAE) are shown in Table 3 below. In general, the value of RMSE should always greater or equal to the MAE. The higher value of RMSE and MAE will indicate more error. The forecasting value (blue stripe) shown in Figures 8, 9 and 10 found to be straightening and approaching the median of the model. This indicates that the forecasting method, model ARIMA does not have good potential in this tidal forecasting.

In choosing the best forecasting model, the information criteria of each model need to be compared. This is because the lower the criteria for the model being investigated, the better the model fits the data. Thus, value of AIC and BIC is very important in ARIMA order selection. From figure 8, 9 and 10, ARIMA (3,0,1) model for Bagan Datuk station, ARIMA (2,0,3) model for Port Klang station and ARIMA (2,0,2) model for Permatang Sedepa station is selected. This is because those order of ARIMA provided smallest AIC and BIC values. The smallest AIC and BIC values will be selected as the best forecasting model. The value of AIC and BIC of those stations is compared at the table below:

TABLE 3. Value of AIC and BIC in three station

Stations	Order	AIC value	BIC value
Bagan Datuk	(3,0,1)	341.83	361.22
	(2,0,0)	346.27	359.19
	(3,0,2)	342.54	365.15
Port Klang	(2,0,3)	343.18	365.80
	(2,0,0)	346.27	359.19
Permatang Sedepa	(3,0,1)	343.83	361.22
	(2,0,2)	342.28	361.67
	(3,0,2)	342.54	365.15
	(2,0,0)	346.27	346.49

TABLE 4. Comparison of errors with ARIMA model

Station	RMSE	MAE
Bagan Datuk	0.0930	0.06399
Port Klang	0.10741	0.07759
Permatang Sedepa	0.15281	0.09283

Table 4 shows the comparison of RMSE and MAE values at the three stations with ARIMA forecasting model. In ARIMA forecasting the results show that Bagan Datuk Station has the smallest error, with RMSE and MAE values of 0.0930 and 0.06399 respectively only. Lower RMSE values indicate lesser error between measured values (historical data) and calculated (predictive values) for tidal

changes (Srivastava et al. 2016). Therefore, lower values for RMSE and MAE will indicate better forecasting. In the ARIMA forecasting model, Bagan Datuk Station has the lowest error value therefore it is the station with the best forecasting performance.

ARTIFICIAL NEURAL NETWORK (ANN) MODEL

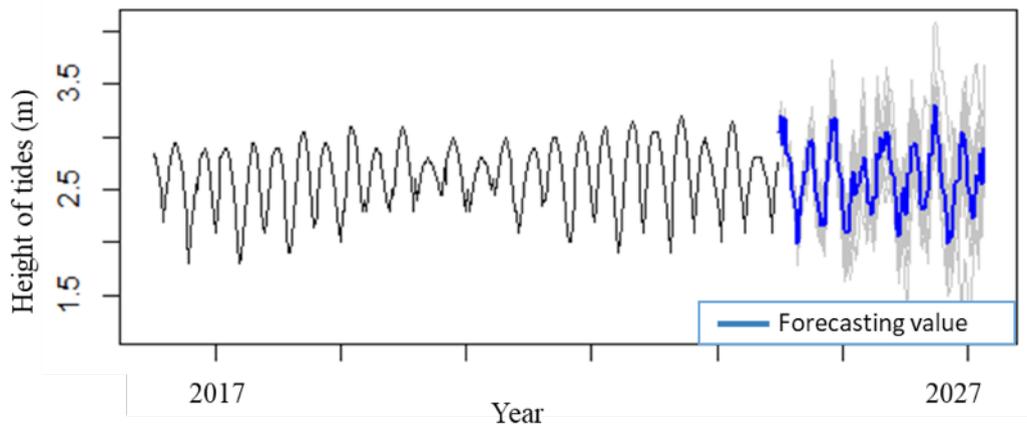


FIGURE 11. ANN model for Bagan Datuk Station in 10 years

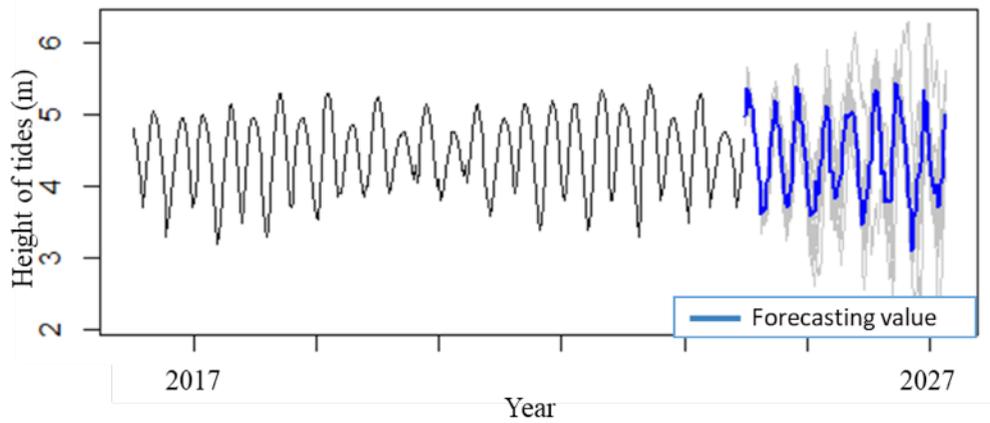


FIGURE 12. ANN model for Port Klang Station in 10 years

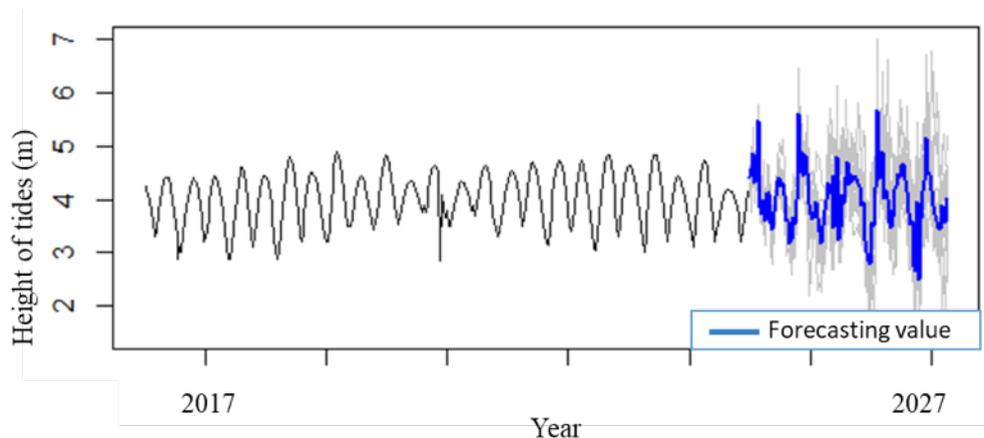


FIGURE 13. ANN model for Permatang Sedepa Station in 10 years

Figure 11, 12 and 13 show the ANN forecasting model for Bagan Datuk, Port Klang and Sedepa Station in 10 years. The square-rooted error (RMSE) and mean absolute error (MAE) are shown in Table 4 below. In general, the value of RMSE should always greater or equal to the MAE. The higher value of RMSE and MAE will indicate more error. The forecast value (blue stripe) in Figure 8, 9 and 10 shows a relatively normal trend compared to historical data. This indicates that the forecasting method, model ANN has good potential in this tidal forecasting.

TABLE 5. Comparison of errors with Artificial Neural Network (ANN) model

Station	RMSE	MAE
Bagan Datuk	0.05642	0.04387
Port Klang	0.08242	0.06495
Permatang Sedepa	0.13318	0.08023

Table 5 shows the comparison of RMSE and MAE values at the three stations with Artificial Neural Network (ANN) forecasting model. In ANN forecasting the results show that Bagan Datuk Station has the smallest error, with RMSE and MAE values of 0.05642 and 0.04387 respectively only. Lower RMSE values indicate lesser error between measured values (historical data) and calculated (predictive values) for tidal changes. Therefore, lower values for RMSE and MAE will indicate better forecasting. In the ANN forecasting model, Bagan Datuk Station has the lowest error value therefore it is the station with the best forecasting performance.

#### MOST SUITABLE FORECASTING MODEL

In comparison, Artificial Neural Network model has a lower RMSE value and MAE value than the ARIMA model. The value shown in table 6 indicate that model ANN have RMSE values of 0.05642, 0.08242 and 0.13318 as well as MAE values 0.04387, 0.06495 and 0.08023 at Bagan Datuk Station, Port Klang and Permatang Sedepa respectively.

ANN model has lower value compared to ARIMA models, with RMSE values 0.0928, 0.10741 and 0.15281 and MAE values 0.06391, 0.07759 and 0.09283 at all three stations.

TABLE 6. Comparison of RMSE and MAE at All Three Stations

Stations	RMSE		MAE	
	ANN	ARIMA	ANN	ARIMA
Bagan Datuk	0.05642	0.0928	0.04387	0.06391
Port Klang	0.08242	0.10741	0.06495	0.07759
Permatang Sedepa	0.13318	0.15281	0.08023	0.09283

The best forecasting stations is Bagan Datuk Station with the lowest RMSE and MAE values (0.05642 and 0.04387). In Agrawal's study, which is a study of rain prediction, his study showed that the model of Artificial Neural Network (ANN) forecasting has better performance compared to ARIMA (Agrawal 2006). The results achieved in this study are in corresponding with the Agrawal study. Therefore, using the ANN approach along with coastal structural strategies is very useful to save costs and time for a forecasting time series.

From the study, we can find that the level of accuracy of the ANN model prediction compared to the ARIMA model is not very significant (Frausto-Solis et al. 2008). It can be said that both models achieve good predictive performance, based on the relatively low prediction errors of both models (Yaseen et al. 2018). However, the performance of the ANN model is better than the ARIMA model in terms of forecasting accuracy by comparing errors between the two models. In the ARIMA forecasting Figures, such as Figures 8, 9 and 10, shown that the ARIMA model pattern is approach to the median. Meanwhile, for the ANN model, the forecasting model pattern in the Figure 11, 12 and 13 has shown a normal pattern with historical patterns. Therefore, this study is corresponding with other studies, which is the ANN model is better than the ARIMA model in time series forecasting (Adebiyi et al. 2014; Ibrahim et al. 2009).

## CONCLUSION

Tidal forecasting has been analysed. The present analysis uses 10 years observations as inputs. The ARIMA forecasting method and the ANN are both good methods of forecasting (Dai & Chen 2019). Through the best forecasting station, Bagan Datuk station, this study able to determine that the forecasting model performs better than ARIMA in tidal forecasting. This result can be ascertained by comparing error: Root mean square error and mean absolute error. In this study, the ANN model obtained a value of 0.05642, compared to the value obtained from the ARIMA Model, 0.0928 for error RMSE. As well as error MAE with a value of 0.04387 (ANN) compared to 0.06391 (ARIMA). The forecasting of the ARIMA model is actually not so suitable for long-term forecasting. This is because the amplitude of the ARIMA forecasting model will become weaker over the long term and finally forecast back to the median (Emamzadeh et al. 2016; Munandar 2019). This study reveals that ANN model can be used as an appropriate forecasting tool for tidal prediction. ARIMA model, is not as good as Artificial Neural Network (ANN) model in tidal prediction. In this study, the present of error of model Artificial Neural Network (ANN) and ARIMA indicate that better model should be discover for better accuracy. Therefore, future study should be focused on obtaining good forecasting values. More attention should be paid to predicting long-term value based on short historical data. Future studies could also explore the potential of linear and non-linear hybrid models in dynamic time series data prediction.

## ACKNOWLEDGEMENTS

The authors would like to thank Universiti Kebangsaan Malaysia for their financial support under the research grant, GUP-2020-013.

## REFERENCES

- Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A. E. & Arshad, H. 2018. State-of-the-art in artificial neural network applications: A survey. *Heliyon* 4(11): e00938. doi:10.1016/j.heliyon.2018.e00938
- Adebiyi, A. A., Adewumi, A. O. & Ayo, C. K. 2014. Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics* 2014: 9–11. doi:10.1155/2014/614342
- Agrawal, K. 2006. Modelling and prediction of rainfall using artificial neural network and ARIMA techniques. *J. Ind. Geophys. Union* 10(2): 141–151.
- Alsharif, M. H., Younes, M. K. & Kim, J. 2019. Time series ARIMA model for prediction of daily and monthly average global solar radiation: The case study of Seoul, South Korea. *Symmetry* 11(2): 1–17. doi:10.3390/sym11020240
- Bakar, N. A. & Rosbi, S. 2017. Autoregressive Integrated Moving Average (ARIMA) Model for Forecasting Cryptocurrency Exchange Rate in High Volatility Environment: A New Insight of Bitcoin Transaction. *International Journal of Advanced Engineering Research and Science* 4(11): 130–137. doi:10.22161/ijaers.4.11.20
- Boyd, P. W., Lennartz, S. T., Glover, D. M. & Doney, S. C. 2015. Biological ramifications of climate-change-mediated oceanic multi-stressors. *Nature Climate Change* 5(1): 71–79. doi:10.1038/nclimate2441
- Collantes-duarte, J. & Rivas-echeverría, F. 2015. Time Series Forecasting using ARIMA, Neural Networks and Neo Fuzzy Neurons 2 Time Series: ARIMA Model 2 Artificial Neural Networks. *Time* (JANUARY 2002). Retrieved from <http://www.wseas.us/e-library/conferences/switzerland2002/papers/464.pdf>
- Consoli, S., Recupero, D. R. & Zavarella, V. 2014. A survey on tidal analysis and forecasting methods for Tsunami detection. *Science of Tsunami Hazards* 33(1): 1–56.
- Dai, J. & Chen, S. 2019. The application of ARIMA model in forecasting population data. *Journal of Physics: Conference Series* 1324(1). doi:10.1088/1742-6596/1324/1/012100
- Dikshita & Singh, H. 2019. Estimating and forecasting volatility using arima model: A study on NSE, India. *Indian Journal of Finance* 13(5): 37–51. doi:10.17010/ijf/2019/v13i5/144184
- Emamzadeh, E. S. S., Soltani, J., Mashal, M., Kalanaki, M. & Asadolahzadeh, T. 2016. Performance Evaluation of MLP and RBF Neural Networks to Estimate the Soil Saturated Hydraulic Conductivity. *Modern Applied Science* 11(3): 1. doi:10.5539/mas.v11n3p1
- Espinoza, S. D. M. 2015. Modelos de regresión para el pronóstico de series temporales con estacionalidad creciente. *Computacion y Sistemas* 18(4): 821–831. doi:10.13053/CyS-18-4-1552
- Fattah, J., Ezzine, L., Aman, Z., El Moussami, H. & Lachhab, A. 2018. Forecasting of demand using ARIMA model. *International Journal of Engineering Business Management* 10: 1–9. doi:10.1177/1847979018808673
- Feist, S. E., Hoque, M. A., Islam, M. A., Ahmed, K. M. & Fowler, M. 2021. Recent Trends in Inland Water Level Change in Coastal Bangladesh—Implications of Sea Level Rise in Low-Lying Deltas. *Global and Planetary Change* 206(103639).
- Feng, J. L., Li, W. S., Wang, H., Zhang, J. L. & Dong, J. X. 2018. Evaluation of sea level rise and associated responses in Hangzhou Bay from 1978 to 2017. *Advances in Climate Change Research* 9(4): 227–233. doi:10.1016/j.accre.2019.01.002
- Fernandez, F. R., Po, R., Montero, N. & Addawe, R. 2017. Prediction of South China sea level using seasonal ARIMA models. *AIP Conference Proceedings* 1905. doi:10.1063/1.5012237

- Frausto-Solis, J., Pita, E. & Lagunas, J. 2008. Short-term streamflow forecasting: ARIMA vs neural networks. *Proceedings of the American Conference on Applied Mathematics* 402–407. Retrieved from <http://dl.acm.org/citation.cfm?id=1415583.1415653>
- Gao, G., Lo, K. & Fan, F. 2017. Comparison of ARIMA and ANN Models Used in Electricity Price Forecasting for Power Market. *Energy and Power Engineering* 09(04): 120–126. doi:10.4236/epe.2017.94b015
- Garg, N., Sharma, M. K., Parmar, K. S., Soni, K., Singh, R. K. & Maji, S. 2016. Comparison of ARIMA and ANN approaches in time-series predictions of traffic noise. *Noise Control Engineering Journal* 64(4): 522–531. doi:10.3397/1/376398
- Graff, M., Escalante, H. J., Ornelas-Tellez, F. & Tellez, E. S. 2017. Time series forecasting with genetic programming. *Natural Computing* 16(1): 165–174. doi:10.1007/s11047-015-9536-z
- Ibrahim, M. Z., Zailan, R., Ismail, M. & Lola, M. S. 2009. Forecasting and Time Series Analysis of Air Pollutants in Several Area of Malaysia. *American Journal of Environmental Sciences* 5(5): 625–632. doi:10.3844/ajessp.2009.625.632
- Islam, T., Srivastava, P. K., Dai, Q., Gupta, M. & Zhuo, L. 2015. An introduction to factor analysis for radio frequency interference detection on satellite observations. *Meteorological Applications* 22(3): 436–443. doi:10.1002/met.1473
- Jahanshahi, A., Jahanianfard, D., Mostafaie, A. & Kamali, M. 2019. An Auto Regressive Integrated Moving Average (ARIMA) Model for prediction of energy consumption by household sector in Euro area. *AIMS Energy* 7(2): 151–164. doi:10.3934/energy.2019.2.151
- Johnson, V. R., Brownlee, C., Milazzo, M. & Hall-Spencer, J. M. 2015. Marine microphytobenthic assemblage shift along a natural shallow-water CO<sub>2</sub> gradient subjected to multiple environmental stressors. *Journal of Marine Science and Engineering* 3(4): 1425–1447. doi:10.3390/jmse3041425
- Kendrick, J. S., Williamson, D. F., D, P., Parker, R. A. & Kendrick, J. S. 1989. The box plot: a simple visual method to interpret data. *Ann Intern Med* 110: 916
- The Box Plot: A Simple Visual Method to Interpret Data. *Academia and Clinic* 10(July 1989): 916–921. doi:10.1059/0003-4819-110-11-916
- Li, S., Liu, L., Cai, S. & Wang, G. 2019. Estuarine, Coastal and Shelf Science Tidal harmonic analysis and prediction with least-squares estimation and inaction method. *Estuarine, Coastal and Shelf Science* 220(February): 196–208. doi:10.1016/j.ecss.2019.02.047
- Lindsey, R. & Lumpkin, R. 2020. In some ocean basins, sea level rise has been as much as 6-8 inches (15-20 centimeters) since the start of the satellite record 1–9.
- Liu, J., Shi, G. & Zhu, K. 2019. High-precision combined tidal forecasting model. *Algorithms* 12(3). doi:10.3390/A12030065
- Lola, M. S., Zainuddin, N. H., Tajuddin, M., Abdullah & Ponniah. 2018. Improving the Performance of Ann-Arima Models for Predicting. *Improving The Performance Of Ann-Arima Models For Predicting Water Quality In The Offshore* 13(1): 27–37.
- Manoj, K. 2014. An Application Of Time Series Arima Forecasting Model For Predicting Sugarcane Production In India. *Studies in Business and Economics* 9(1): 81–94.
- Meena, B. L. & Agrawal, J. D. 2015. Tidal level forecasting using ANN. *Procedia Engineering* 116(1): 607–614. doi:10.1016/j.proeng.2015.08.332
- Mohanty, A. & Mohanty, S. 2015. Prevention of transient instability and reactive power mismatch in a Stand-alone wind-diesel-tidal hybrid system by an ANN based SVC. *Aquatic Procedia* 4(Icwrcoe): 1529–1536. doi:10.1016/j.appro.2015.02.198
- Mohapatra, S. K., Swain, J. K. & Mohanty, M. N. 2019. Detection of Diabetes Using Multilayer Perceptron. *Advances in Intelligent Systems and Computing*, hlm. Vol. 846. Springer Singapore. doi:10.1007/978-981-13-2182-5\_11
- Munandar, D. 2019. Multilayer perceptron (MLP) and autoregressive integrated moving average (ARIMA) models in multivariate input time series data: Solar irradiance forecasting. *International Journal on Advanced Science, Engineering and Information Technology* 9(1): 220–228. doi:10.18517/ijaseit.9.1.6426
- Olatunji, O. O., Akinlabi, S., Madushele, N., Adediji, P. A. & Felix, I. 2019. Multilayer perceptron artificial neural network for the prediction of heating value of municipal solid waste. *AIMS Energy* 7(6): 944–956. doi:10.3934/energy.2019.6.944
- Olbert, A. I., Comer, J., Nash, S. & Hartnett, M. 2017. High-resolution multi-scale modelling of coastal flooding due to tides, storm surges and rivers inflows. A Cork City example. *Coastal Engineering* 121(January): 278–296. doi:10.1016/j.coastaleng.2016.12.006
- Ongbali, S. O., Igboanugo, A. C., Afolalu, S. A., Udo, M. O. & Okokpujie, I. P. 2018. Model Selection Process in Time Series Analysis of Production System with Random Output. *IOP Conference Series: Materials Science and Engineering* 413(1). doi:10.1088/1757-899X/413/1/012057
- Pindsoo, K. & Soomere, T. 2020. Basin-Wide Variations in Trends in Water Level Maxima in the Baltic Sea. *Continental Shelf Research* 193(104029).
- Prahitama, A. & Mustafid. 2016. Modelling tide prediction using linear model and adaptive neuro fuzzy inference system (ANFIS) in Semarang, Indonesia. *ARPJ Journal of Engineering and Applied Sciences* 11(11): 6972–6976.
- Pramanik, M. K., Biswas, S. S., Mukherjee, T. & Kumar Roy, A. 2016. Sea Level Rise and Coastal Vulnerability along the Eastern Coast of India through Geo-spatial Technologies. *Journal of Remote Sensing & GIS* 4(2). doi:10.4172/2469-4134.1000145

- Praveen, V., Delhi Narendran, T., Pavithran, R. & Thirumalai, C. 2018. Data analysis using box plot and control chart for air quality. *Proceedings - International Conference on Trends in Electronics and Informatics, ICEI 2017* 2018-Janua(June): 1082–1085. doi:10.1109/ICOEI.2017.8300877
- Ray, S. & Bhattacharyya, B. 2015. Availability in Different Source of Irrigation in India : A Statistical Approach. *Availability in Different Source of Irrigation in India: A Statistical Approach* 5: 109–116. doi:10.5923/c.ije.201501.16
- Reddy, J. C. R., Ganesh, T., Venkateswaran, M. & Reddy, P. R. S. 2017. Forecasting of Monthly Mean Rainfall in Coastal Andhra. *International Journal of Statistics and Applications* 7(4): 197–204. doi:10.5923/j.statistics.20170704.01
- Shamsuddin, S. M., Sallehuddin, R. & Yusof, N. M. 2008. Artificial neural network time series modeling for Revenue forecasting. *Chiang Mai Journal of Science* 35(3): 411–426.
- Shu, M. H., Hung, W. J., Nguyen, T. L., Hsu, B. M. & Lu, C. 2014. Forecasting with fourier residual modified ARIMA model - An empirical case of inbound tourism demand in New Zealand. *WSEAS Transactions on Mathematics* 13: 12–21.
- Srivastava, P. K., Islam, T., Singh, S. K., Petropoulos, G. P., Gupta, M. & Dai, Q. 2016. Forecasting Arabian Sea level rise using exponential smoothing state space models and ARIMA from TOPEX and Jason satellite radar altimeter data. *Meteorological Applications* 23(4): 633–639. doi:10.1002/met.1585
- Tealab, A., Hefny, H. & Badr, A. 2017. Forecasting of nonlinear time series using ANN. *Future Computing and Informatics Journal* 2(1): 39–47. doi:10.1016/j.fcij.2017.05.001
- Tylkowski, J. & Hojan, M. 2019. Time decomposition and short-term forecasting of hydrometeorological conditions in the south Baltic Coastal zone of Poland. *Geosciences (Switzerland)* 9(2). doi:10.3390/geosciences9020068
- Vignesh, V., Pavithra, D., Dinakaran, K. & Thirumalai, C. 2018. Data analysis using box and whisker plot for stationary shop analysis. *Proceedings - International Conference on Trends in Electronics and Informatics, ICEI 2017* 2018-Janua(May): 1072–1076. doi:10.1109/ICOEI.2017.8300874
- Yaseen, Z. M., Allawi, M. F., Yousif, A. A., Jaafar, O., Hamzah, F. M. & El-Shafie, A. 2018. Non-tuned machine learning approach for hydrological time series forecasting. *Neural Computing and Applications* 30(5): 1479–1491. doi:10.1007/s00521-016-2763-0
- Yin, J., Wang, N. & Hu, J. 2015. A hybrid real-time tidal prediction mechanism based on harmonic method and variable structure neural network. *Engineering Applications of Artificial Intelligence* 41: 223–231. doi:10.1016/j.engappai.2015.03.002
- Yusofa, Z. & Saadb\*, N. A. 2022. Successful Approaches to Integrated Water Resources Management: A Mini Review. *Jurnal Kejuruteraan* 34(2)(01): 181-190.
- Zhang, Y., Wang, T., Li, H., Bao, N., Hall-Spencer, J. M. & Gao, K. 2018. Rising levels of temperature and CO<sub>2</sub> antagonistically affect phytoplankton primary productivity in the South China Sea. *Marine Environmental Research* 141(August): 159–166. doi:10.1016/j.marenvres.2018.08.011