Simulation and Analysis of Sea-Level Change from Tide Gauge Station by using Artificial Neural Network Models
(Simulasi dan Analisis Perubahan Aras Laut dari Stesen Tolok Air Pasang Surut dengan menggunakan Model Rangkaian Neural Buatan)

MILAD BAGHERI1, ZELINA Z IBRAHIM2, LATIFAH ABD MANAF2, MOHD FADZIL AKHIR1 & WAN IZATUL ASMA WAN TALAAT1,*

1Institute of Oceanography and Environment, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu Darul Iman, Malaysia
2Department of Environment, Faculty of Environmental and Forestry, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor Darul Ehsan, Malaysia

Received: 9 October 2021/Accepted: 23 December 2021

ABSTRACT

Sea level change is one of the most certain results of global warming. Sea level change would increase erosion in coastal areas, result in intrusion into water supplies, inundate coastal marshes and other important habitats, and make the coastal property more vulnerable to erosion and flooding. This situation coincides with the massive socio-economic development of the coastal city areas. The coastal areas of the East Coast of Peninsular Malaysia are vulnerable to sea-level change, flooding, and extreme erosion events. The monthly Mean Sea Level (MSL) change was simulated by using two Artificial Neural Network (ANN) models, Feed Forward- Neural Network (FF-NN) and Nonlinear Autoregressive Exogenous- Neural Network (NARX-NN) models. Both models did well in recreating sea levels and their fluctuating patterns, according to the data. The NARX-NN model with architecture (5-6-1) and four lag options, on the other hand, got the greatest results. The findings of the model’s mean sea level rise simulation show that Kuala Terengganu would have a growing and upward trend of roughly 25.34 mm/year. This paper shows that the eastern coast of Malaysia is highly vulnerable to sea-level rise and therefore, requires sustainable adaptation policies and plans to manage the potential impacts. It recommends that various policies, which enable areas to be occupied for longer before the eventual retreat, could be adapted to accommodate vulnerable settlements on the eastern coast of Malaysia.

Keywords: Climate change; coastal city; FF-NN; NARX-NN; tide gauge; time series analysis

ABSTRAK

Perubahan paras laut adalah salah satu hasil pemanasan global yang paling pasti. Perubahan paras laut akan meningkatkan hakisan di kawasan pantai, mengakibatkan pencerobohan ke dalam bekalan air, memberi jari paya pantai dan habitat penting lain dan menjadikan harta pantai lebih terdedah kepada hakisan dan banjir. Keadaan ini bertepatan dengan pembangunan sosio-ekonomi yang besar di kawasan bandar pantai. Kawasan pantai di Pantai Timur Semenanjung Malaysia terdedah kepada perubahan paras laut, banjir dan kejadian hakisan yang melampau. Perubahan Purata Aras Laut (MSL) bulanan telah disimulasikan dengan menggunakan dua model Rangkaian Neural Buatan (ANN), Rangkaian Neural Feed Hadapan (FF-NN) dan Model Rangkaian Neural Eksogen Autoregresif Tak Linear (NARX-NN). Kedua-dua model itu berjaya mencipta semula paras laut dan corak turun naiknya, menurut data. Model NARX-NN dengan seni bina (5-6-1) dan empat pilihan ketinggalan, sebaliknya, mendapat hasil terbaik. Penemuan simulasi kenaikan paras laut purata model menunjukkan bahawa Kuala Terengganu akan mempunyai aliran meningkat dan meningkat kira-kira 25.34 mm/tahun. Kertas ini mendedahkan bahawa pantai timur Malaysia sangat terdedah kepada kenaikan paras laut dan oleh itu, memerlukan dasar dan rancangan penyesuaian yang mampu untuk mengurus kesan yang berpotensi. Ia mengesyorkan bahawa pelbagai dasar, yang membolehkan kawasan diduduki lebih lama sebelum berundur akhirnya, boleh disesuaikan untuk menampung penempatan yang terdedah di pantai timur Malaysia.

Kata kunci: Analisis siri masa; bandar pantai; FF-NN; NARX-NN; perubahan iklim; tolok air pasang surut
INTRODUCTION
Climate change and global warming are seen as a severe danger to both the environment and mankind in the twenty-first century, and as a result, they have attracted global interest (Bagheri et al. 2021c). Nelson and Serafin (1996) predicted that global temperatures will climb between 1.5 and 4.5 °C in the next century, whereas Wei (2015) predicted a significant rise in the mean annual temperature of 0.99-3.44 °C each century. A rise in sea level is one of the most serious consequences of climate change, and it poses a significant risk to coastal cities (Bagheri et al. 2021a). As a result of rising global temperatures, warming and expanding seas, and melting ice caps and glaciers, sea levels are rising. Furthermore, these alterations have an impact on marine life (Lan et al. 2013).

A rise in sea level is one of the major negative impacts of climate change, and it poses a significant risk to coastal regions (Bagheri et al. 2021d). The global average sea level rose by 0.19 m between 1901 and 2010, and the pace of change since the mid-nineteenth century has been faster than the average rate during the previous two millennia (Mimura 2013). During the 1901-2010 era, global sea level increase was probably 1.7 and 3.2 mm/year during the 1993-2010 period (IPCC 2014).

The trend of global warming in Malaysia has risen in the past 3 to 5 decades, and the mean temperature is projected to rise between 0.6-4.5 °C in 2060. Although Malaysia is classed as a nation with a moderate risk of climate change, extreme events such as floods, storms, droughts, and shifts in climatic patterns have occurred in Pahang, Johor, Kelantan, Kedah, and Terengganu (Bagheri et al. 2021b). These occurrences had demonstrated that global warming could have an adverse effect and threaten Malaysia (Tanggang 2007). Low-lying locations with strong socioeconomic activity and people at risk of flooding (Bagheri et al. 2019) because Malaysia has such a long coastline, many of its cities are located near it. The National Hydraulic Research Institute of Malaysia (NAHRIM) has conducted several studies on imminent global warming issues, particularly in the areas of sea-level change projections in Malaysia, coastal vulnerability assessments for high-risk areas, and the production of prospective sea-level change inundation maps (Awang & Hamid 2013). NAHRIM (2010a) performed research on how climate change affects sea-level change in Malaysia, as well as the anticipated pace of sea-level rise in Malaysia’s coastal region by 2100.

According to the study (NAHRIM 2010b), around 3.3 percent of the 1,963 km of coastline is classified as severely susceptible, with the northern reaches of the Kedah shoreline and the southern portions of the Terengganu beachfront being the most vulnerable. Ibrahim and Wibowo (2013) observed that flooding has always occurred in Terengganu (about 30% of the area), which is located on the East Coast of Peninsular Malaysia. For example, floods occur often every year in Terengganu’s Dungan area, particularly when the river’s water level rises a few meters beyond the danger threshold. Heavy rainfall is one of the features of the monsoon season on the East Coast of Peninsular Malaysia, which runs from October to March each year (Ariffin et al. 2016). The article’s precise goals are also stated. 1) Analysis of actual and projected data of sea level from 1991 to 2020, 2) Determine the ANN models for simulating sea level rise in the Eastern Coast of Peninsular Malaysia as a case study.

MATERIALS AND METHODS
DATA REQUIREMENT
Kuala Terengganu has 320 km of sandy coastline beaches that stretch from Besut in the north to Kemaman in the south along the South China Sea (SCS) (Bagheri et al. 2013) (Figure 1). The requisite data for this section of the study are reanalysis data and secondary data which were obtained from diverse websites and departments. The input data regarding simulation comprise Kuala Terengganu’s observed monthly tide gauge (from Cendering station), Sea Surface Temperature (SST), rainfall, wind, and Sea Level Pressure (SLP). The observed monthly tide data were gathered from the Department of Survey and Mapping Malaysia (JUPEM), geodetic survey division, vertical reference section infrastructure. The Malaysian Meteorological Department (MMD) provided data for the monthly rainfall and monthly wind speed. Also, the National Centre for Environmental Prediction (NCEP) provided reanalysis data for SST and SLP as depicted. Regarding the forecasting of changes in sea level tide analysis system software in JUPEM was utilized for the simulation of monthly tide records. While the monthly anticipated wind was obtained from MMD, the research center for the tropical climate change system in Universiti Kebangsaan Malaysia (UKM) was the source of the monthly anticipated rainfall.

TIDE GAUGE TIME SERIES ANALYSIS
Data series refer to a set of observations that are arranged in a specific way, and time-series data are data series that are ordered chronologically (Amerian & Voosoghi 2011). The use of MSL daily observation
enables the analysis of historical records at tide stations which offers evidence of nonlinear change in sea level.

The study needs to generate the sea level residual for sea-level rise simulation, but before that, it conducts some statistical investigation on the observed tide data to detect the outliers and missing data. Thus, box plot, histogram, and quantile-quantile plot are some of the graphical tests used to test for normality (whether the data are normally distributed) and detect outliers in time series data. The box plot could be suitable for symmetry testing which in most cases is an appropriate substitute for normality.

FIGURE 1. Cendering tide gauge station in Kuala Terengganu

FF-NN MODEL STRUCTURE

The input data were standardized or normalized to increase the efficiency of training (Zime 2014). In essence, it is essential to conduct the normalization of the output and input data to ensure that they are in an identical range of applied transfer functions. The essence of this was to restrict their range within the interval of 0-1 (Polo et al. 2015) because the middle layer’s Processing Elements (PEs) were allotted a sigmoidal activation function. Hence, this function’s shape plays a vital responsibility in ANNs learning (Ghamarnia & Jalili 2015). Thus, the following formula was used to normalize the data:

\[ X_{\text{normal}} = 0.5 + 0.5 \frac{x - x_{\text{mean}}}{x_{\text{max}} - x_{\text{min}}} \]

where \( x_{\text{normal}} \) is the normal of data; \( x_{\text{mean}} \) is the sea level residual observation means; \( x_{\text{min}} \) is the minimum; and \( x_{\text{max}} \) is the maximum sea level residual observation. In this study, for ANN structure, five input and one output data from 1991-2020 for simulation were used.

FF-NN model structure typically encompasses three distinguished layers namely the input layer, the hidden layer which has a sigmoid activation function (that processes the data), and an output layer (that produces the ANN results) (Tezel & Buyukyildiz 2016). Hence, the information in a distinctive FF-NN model moves only in a forward direction, from the input layer, via the hidden layers, and thereafter to the output layer and the network has no loops. Each layer consists of nodes known
as neurons, and each neuron in the proceeding layer is linked to another neuron (Polo et al. 2015).

Mathematically, the following expressions could be used to define a neuron:

\[
v = \sum_{i=1}^{n} w_i x_i - w_{n+1} = W x^T - w_{n+1} \tag{2}\]

\[
y = \Psi(v) \tag{3}\]

where \(v\) is the sum of all relevant products of weights and outputs from the previous layer \(i\); \(y\) is the activation of the node at hand; \(w = [w_1, w_2, ... , w_n]\) vector of weights; \(x[x_1, x_2, ... , x_n]\) vector of input signals; \(w_i\) bias; \(\Psi(v)\) activation function; and \(n\) is the neuron number.

NARX-NN MODEL STRUCTURE

Xie et al. (2009) posited that the nonlinear autoregressive network which has exogenous inputs NARX-NN model is frequently utilized in the identification area system. As a neural time, series device, the NARX-NN is a recurrent dynamic network that has feedback connections that encompass some network layers. Thus, the NARX-NN model is centered on the linear ARX model which is generally used in time-series modeling. The NARX-NN model’s defining equation is given as follows:

\[
y(t) = f(y(t-1), y(t-2), ... , y(t-n_y), u(t-1), u(t-2), ... , u(t-n_u)) \tag{4}\]

This indicates that the next value of the dependent output signal \(y(t)\) is regressed on the previous values of the output signal and previous values of an independent input signal. Hence, the NARX-NN model can be executed by using a feed-forward neural network for the estimation of the function \(f\).

Both the input and output of the NARX-NN model is built on a unique multilayer perceptron, which consists of neurons with adjustable synaptic bias and weights. As a result, the present and past values of the inputs, which denote independent inputs generated outside the network, and the delayed output values that the model output is regressed, denote the signal vector’s data window used by the input layer. The learning device multilayer perceptron networks require a resolution heuristic algorithm that ensures the greatest solution. Polo et al. (2015) noted that the Levenberg-Marquardt Algorithm (LMA) trained the network in MATLAB software. Diverse applications of the NARX-NN network exist, and it can be used as a predictor to forecast the subsequent input signal’s value. It can also be used for nonlinear filtering where the target output is the input signal’s noise-free version.

EVALUATION AND PERFORMANCE ASSESSMENT

The performance assessment is discussed about the simulation error, and it explains the validation between the observed and simulated data. The evaluation of model accuracy is essential to determine the superlative neural network architecture that produces the most precise and reliable simulated data (Khamis & Abdullah 2014).

Several methods of performance assessment are applied to measure the accuracy. This present study evaluates the performance with regards to Mean Square Error (MSE) and the coefficient of determination (\(R\)). The \(R\) measures a regression’s goodness of fit, and it is utilized in this study as a measure to appraise the degree of correlation between the trained network estimation and the experimental data (Mashaly et al. 2015). The correlation of determination (\(R\)) is given as follows:

\[
R = \frac{\sum_{i=1}^{n}(O_i - \overline{O})(P_i - \overline{P})}{\sqrt{\sum_{i=1}^{n}(O_i - \overline{O})^2} \sqrt{\sum_{i=1}^{n}(P_i - \overline{P})^2}} \tag{5}\]

where \(n\) represents the number of samples; \(O_i\) represents observed sea level residual; \(P_i\) represents simulated sea level; \(\overline{O}\) represents mean of observed values; and \(\overline{P}\) represents mean of simulated values. The \(R^2\) values lie between 0 and 1, with the value of 0 indicating the absence of correlation between observation and simulation values, whereas the value of 1 signifies maximum correlation between the values. The MSE is an alternative procedure of measuring performance assessment between the real values and an estimator (Hamzehie et al. 2014; Ranković et al. 2014). It is given as follows:

\[
\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (O_i - P_i)^2 \tag{6}\]

An MSE which is close to 0 indicates more accurate model responses.

RESULTS AND DISCUSSION

FF-NN ARCHITECTURE

This section presents the simulated results, and it uses five input data and one target (monthly data from 1991
to 2020) for simulation. Similarly, for the simulation of sea-level change from 1991 to 2020, it applied five inputs (we have some limitation of data so we choose input data from 1991 to 2020) as input data and one target (output) which is monthly sea level residual after selecting the best neural network architecture. It used the neural network toolbox in MATLAB (version R2013A) software. For the FF-NN Model, it applied three layers (interconnected neural network) consisting of an input layer, a hidden layer, and an output layer, and the best FF-NN architecture was selected after 100 runs. Thus, there are 5 input nodes (one each for the tide gauge, rainfall, wind, SST, and SLP) and the output layer will have one node (sea level residual). In hidden and output layers, each input would be multiplied by a corresponding weight, sum the product, and then process the sum using a nonlinear transfer function to generate a result. After training and running the model, the best simulation performance, ANN architecture (5-10-1), was selected with MSE and R for training, validation, and testing, respectively.

Thus, 5 inputs (tide gauge, rainfall, SLP, SST, and wind), 10 hidden layers, and one output layer were applied for the simulation of sea-level rise by the FF-NNModel. In this model, 60, 20, and 20% of data were used for training, validation, and testing sets, respectively. To select the best FF-NN model for sea-level rise simulation (1991-2020), the model was run 100 times and the best performance was ANN architecture (5-10-1) with MSE and R (Table 1). Based on Figure 2, the best performance and validation with the minimum error was found at epoch 3.

<table>
<thead>
<tr>
<th>Run</th>
<th>ANN Architecture</th>
<th>Neuron Number</th>
<th>Target Value</th>
<th>MSE</th>
<th>R (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Training 60% data</td>
<td>Validation 20% data</td>
<td>Testing 20% data</td>
</tr>
<tr>
<td>100</td>
<td>5-2-1</td>
<td>2</td>
<td>158</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>100</td>
<td>5-4-1</td>
<td>4</td>
<td>158</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>100</td>
<td>5-6-1</td>
<td>6</td>
<td>158</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>100</td>
<td>5-8-1</td>
<td>8</td>
<td>158</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>100</td>
<td>5-10-1</td>
<td>10</td>
<td>158</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>100</td>
<td>5-12-1</td>
<td>12</td>
<td>158</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>100</td>
<td>5-14-1</td>
<td>14</td>
<td>158</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>100</td>
<td>5-16-1</td>
<td>16</td>
<td>158</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>100</td>
<td>5-18-1</td>
<td>18</td>
<td>158</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>100</td>
<td>5-20-1</td>
<td>20</td>
<td>158</td>
<td>53</td>
<td>53</td>
</tr>
</tbody>
</table>

**FIGURE 2.** a) FF-NN model performance, training, validation, and testing result (1991-2020), b) The best validation performance by using MSE
NARX-NN MODEL ARCHITECTURE

In the NARX-NN model, five input data and one target (monthly data from 1991 to 2020) were also used for simulation to train the neural network. The sea level rates between the layers were acquired by training the neural network through a reverse calculation process in which the contribution or importance of each effect criteria on sea level was computed. Therefore, the contribution or importance of each criterion, the future sea-level rates, was determined.

In this model, 60% of data was used for training, 20% for validation, and 20% for the testing phase. For the training, validation, and testing of the NARX-NN model, 60, 20 and 20% of observed data, respectively, were used. The model was run 100 times to get the best performance. The best NARX-NN model was selected with architecture (5-6-1) and four lags.

This model was selected as the best model with MSE of 0.000292, 0.000269, and 0.000261 for the training set, for the validation set, and for the testing set, respectively, and their respective R of 93, 92, and 94%. For sea-level change simulation by NARX-NN Model, 5 inputs (simulated time series data same as FF-NN model), 6 hidden layers, and one output layer were applied. In this model, 60, 20, and 20% of data were used for training, validation, and testing, respectively. To select the best NARX-NN model for sea-level change simulation (1991-2020), the model was run 100 times and the best performance was ANN architecture (5-6-1) with MSE and R (Table 2) and (Figure 3(a), 3(b)).

The result illustrates that the best validation performance is at epoch 4 where the error was up to 1.66192 e-4. This indicates that the model is well trained and well fitted. This indicates a good fit and therefore the NARX-NN model is well trained and could be used to simulate sea levels from 1991-2012. Based on this result of learning, the NARX-NN model with architecture (5-6-1) and with four Lag is suitable for the simulation of sea-level rise from 1991-2020. Figure 3(c) indicates the simulated sea-level rise from 1991-2020 which has an upward trend.

The result of the FF-NN model and NARX-NN model have indicated in Figure 4. Since the performance of the NARX-NN model was better than the FF-NN model (based on MSE and R), the study applied the simulation result of the NARX-NN model with architecture (5-6-1) and with four Lag to estimate the sea-level changes for Kuala Terengganu. The observed and simulated sea-level changes from 1991-2020 by the NARX-NN model are shown in Figure 5.

From 1991 to 2020, the result shows an increase trend in sea level in the Kuala Terengganu coastal area, with a minimum rate of 1.10 mm/year and a maximum rate of 79.26 mm/year. Changes in sea level trends indicate that Kuala Terengganu coastal city area is expected to experience a sea-level rise in a long time which is an indication of hazard in this area that could cause inundation, flood, and erosion in the future in Kuala Terengganu coastal area. The result of this study is consistent with the findings of Awang and Abd Hamid (2013) and NAHRIM (2010a, 2010b).
The NARX-NN model, being one of the better ANN models, predicts an increased trend in sea level along the Kuala Terengganu coast in the future. Because it mimicked the pattern of a sea-level well, researchers may use this simulation model to forecast future sea-level shifts. Sea level fluctuations, particularly SST and SLP, are connected to tide gauge shifts and climate variations. The connection between SST and sea level residual is positive, whereas the association between SLP and sea level residual is negative.

Despite the use of NARX-NN model and limited data, experience shows that such exercises are effective in simulating coastal city area long-term expansion and highlighting significant concerns among coastal policymakers. The output of the NARX-NN model was utilized to generate research on coastal city sustainability shoreline modification. It is important to investigate and estimate the risks causing parameters for the study region to develop the coastal city danger model.
FIGURE 4. The FF-NN and NARX-NN model were used to compare simulated sea-level rise (1991 to 2020) for the Kuala Terengganu area.

FIGURE 5. The NARX-NN model simulation trend of sea-level rise (1991 to 2020) for the Kuala Terengganu region.
CONCLUSION

Sea level simulation is an important issue in coastal areas that are vulnerable to sea-level changes. This study attempts to predict the future sea-level changes in the coastal area of Kuala Terengganu until 2020 through the application of ANN models. The ANN models can support time-series data, and inputs can be changed by the researcher for different study areas. The backpropagation training algorithm has difficulties when trying to follow the internal processes of the procedure.

The results of both models were compared, and both performed well in replicating sea levels and their shifting patterns. However, based on least MSE and maximum R, the NARX-NN model with architecture (5-6-1) and four lag selections had the best performance. This model simulated the pattern of the sea level well and was therefore used for the simulation of sea-level changes at Kuala Terengganu. According to the modelling results, the Kuala Terengganu coast would see an increase trend in sea level of roughly 25.34 mm/year in the future. Low-lying areas with high populations and socio-economic activities are at risk of being inundated. These potential sea-level rise rates can be applied as a guide for planning and implementation agencies, and local authorities in their development planning to avoid huge development in critical areas.

The result can also be used to estimate some hazard events such as erosion and flood in this area in the future. This result would be important for managers and policymakers in future risk management. Furthermore, the study may enhance the ability of those involved in the planning, policy, and decision-making in evaluating the effects of climate changes on the shoreline and coastal area as well as assist them in the development of long-term and short-term projections of land-use suitability in the coastal area.

ACKNOWLEDGEMENTS

Funding for this project has been provided by the Universiti Putra Malaysia (UPM) RUGS 4 with Project Number (03-04-11-1477RU) and RUGS 6 with Project Number (03-01-12-1664RU) programs. Acknowledgement also goes to the INOS Higher Institution of Centre of Excellence (Vot 66928) for partially supporting the extension of the study.

REFERENCES


*Corresponding author; email: wia@umt.edu.my