

Article

Evaluation of Multivariate Adaptive Regression Splines and Artificial Neural Network for Prediction of Mean Sea Level Trend around Northern Australian Coastlines

Nawin Raj ¹ and Zahra Gharineiat ^{2,*}

¹ School of Sciences, Springfield Campus, University of Southern Queensland, Toowoomba, QLD 4300, Australia; Nawin.Raj@usq.edu.au

² School of Civil Engineering and Surveying, Springfield Campus, University of Southern Queensland, Toowoomba, QLD 4300, Australia

* Correspondence: Zahra.gharineiat@usq.edu.au

Abstract: Mean sea level rise is a significant emerging risk from climate change. This research paper is based on the use of artificial intelligence models to assess and predict the trend on mean sea level around northern Australian coastlines. The study uses sea-level times series from four sites (Broom, Darwin, Cape Ferguson, Rosslyn Bay) to make the prediction. Multivariate adaptive regression splines (MARS) and artificial neural network (ANN) algorithms have been implemented to build the prediction model. Both models show high accuracy ($R^2 > 0.98$) and low error values (RMSE < 27%) overall. The ANN model showed slightly better performance compared to MARS over the selected sites. The ANN performance was further assessed for modelling storm surges associated with cyclones. The model reproduced the surge profile with the maximum correlation coefficients ~ 0.99 and minimum RMS errors ~ 4 cm at selected validating sites. In addition, the ANN model predicted the maximum surge at Rosslyn Bay for cyclone Marcia to within 2 cm of the measured peak and the maximum surge at Broome for cyclone Narelle to within 7 cm of the measured peak. The results are comparable with a MARS model previously used in this region; however, the ANN shows better agreement with the measured peak and arrival time, although it suffers from slightly higher predictions than the observed sea level by tide gauge station.

Keywords: ANN; MARS; mean sea level; prediction; Australia; tide gauge

Citation: Raj, N.; Gharineiat, Z. Evaluation of Multivariate Adaptive Regression Splines and Artificial Neural Network for Prediction of Mean Sea Level trend around Northern Australian Coastlines. *Mathematics* **2021**, *9*, 2696. <https://doi.org/10.3390/math9212696>

Academic Editor: Akemi Galvez Tomida

Received: 23 September 2021

Accepted: 21 October 2021

Published: 24 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Estimating regional mean sea-level rise is critical to understanding and characterizing current and future causes of sea-level change. During the last two decades, the rate of sea-level rise around the Australian region has not been consistent, with the sea-level trend due the dynamic influence induced by internal climate modes about three times greater than the global mean sea-level around the north and north-west of Australia [1,2]. Previous studies of sea-level trends employing various approaches to accommodate different spatial and temporal resolutions of the data around Australia (e.g., [3,4]). Ref. [5] used the multivariate regression (MR) model to integrate satellite altimetry and tide gauge data, leading to a significant improvement of temporal resolution of sea-level data from satellite altimetry around northern Australian coastlines. However, the MR model slightly underestimates the observed sea-level variability. This is due to the fact that the linear MR model may not be able to closely fit the reality of sea-level features, which are naturally non-linear and vary greatly over time [6–8] also highlighted the non-linearity of mean sea-level rise based on analysis of the world's longest records.

To overcome this limitation, ref. [9] employed the multivariate adaptive regression splines (MARS) by Friedman. The MARS is a non-parametric and multi-stage regression

model, in which spline functions are used to fit data with an arbitrary regression function [10]. Ref. [9] showed that the MARS method outperformed the multivariate regression in modelling sea levels during extreme sea-level events, increasing the mean value of R^2 from the multivariate regression model (0.45) to the MARS model (0.62). It was also found that the mean RMSE value over the study region was decreased from 8.21 cm by the multivariate regression model to 6.73 cm by the MARS model. They highlighted that the MARS has been successful in modelling the non-linear relationship between number of variables with multi-collinearity, and/or a high-level interaction.

Lately, artificial neural networks (ANN) have become popular for forecasting of the sea due to its ability to handle data with a non-linear relationship, working based on uncertainties, and learning from experience levels [11]. Also, ANN does not depend on the assumptions of the functional model, the probability distribution, or the smoothness of the underlying data [12]. Several studies have used ANN models to predict sea level and associated tidal signals (e.g., [13–15]) and the modelling results have been considerably improved compared to other existing models. Since the ANN and MARS models can model out the non-linearity in a given set of data, this study aims to compare the performance of ANN model to the previously used model (MARS) in order to select the best method for predicting sea level at the northern Australian region. The paper is organised as follows. The data and study area are presented in Section 2. Section 3 explains the theoretical background of the MARS and ANN model, as well as the interpretation of the results. Section 4 demonstrates the modelling outcomes for four selected tide gauges located around the northern Australian coastal region using both the ANN model and MARS model. Finally, the last section presents the conclusion and recommendations for future studies.

2. Data and Study Area

The 26 years (1993–2019) of hourly sea surface heights (SSHs) from four tide gauges (Figure 1) were extracted from the Australian National Tidal Centre (NTC, <http://www.bom.gov.au/oceanography/>, accessed on 23 January 2020). The tide gauge datasets have been checked and corrected for apparent errors including datum and time shifts. We carefully chose tide gauges that had a complete time series for more than 20 years. The mean sea surface (MSS) was then removed from individual tide gauge time series. The inverse barometric was also corrected using the two-dimensional gravity waves model (MOG2D-IB) to decrease the noise in estimating the part of sea level variations that are not associated with atmospheric pressure [16].

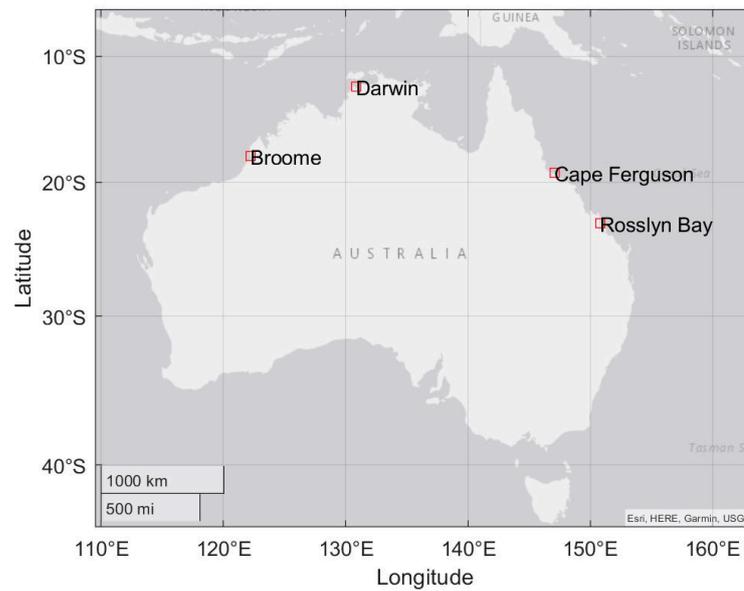


Figure 1. The location of tide gauges (red squares) used for this study.

3. Methodology

3.1. Theoretical Background

3.1.1. Multivariate Adaptive Regression Splines (MARS) Model

MARS is a non-linear and non-parametric regression approach which was initially developed by Jerome H. Friedman in 1991 [17]. It is a stepwise linear regression that is capable of taking higher-dimensional inputs. This is done over an equivalent interval to explore complex and non-linear relationships between the response and input variables [17]. MARS predictive modelling has been widely used in many machine-learning applications [18–21].

The MARS method generates forecasts by learning the relationships between the response and predictor variables. The training data sets are divided into splines which are separate piecewise linear segments of different gradients.

The model is a weighted sum of basis functions (BFs) defined in pairs which form based on a knot and subgroups to define an inflection region [17,19]. MARS characterises data either globally or using linear regression between any two knots.

The linear combination of the basis functions can be given as follows [22]:

$$f(x) = C_0 + \sum_{i=1}^k C_i BF(x) \tag{1}$$

where, $C_{i...k}$ are unknown coefficients estimated by the least squares method, k is the number of terms considered in the model, $BF(x)$ is the basis function based on knots from a piecewise linear basis function. During the development of the model, the basis functions are selected based on the following generalized cross validation (GCV).

$$GCV(N) = \left(\frac{1}{n} \frac{\sum_{i=1}^k (y_i - \hat{y}_i)^2}{\left(1 - \frac{C(M)}{n}\right)^2} \right) \tag{2}$$

where, n is the number of data points, y_i is the actual value of data point i , \hat{y}_i is the predicted value for data point i and $C(M)$ is the penalty factor defined as follows:

$$C(M) = M + dM \quad (3)$$

where, d is the cost penalty factor of each basis function in the optimisation. Over fitting can result when several basis functions are selected in the forward phase, therefore, deleting some basis functions in the backward phase is important to select the optimised model.

3.1.2. Artificial Neural Network (ANN) Model

ANNs are based on a series of parallel architectures that are connected by nodes called artificial neurons [23–25]. These networks use learning capabilities obtained from inputs, which can be effectively used for the prediction of mean sea level as available data are fairly large. One of the main strengths of the neural network architecture is that it improves its own problem-solving ability by continually learning from trial and error. Once this is done over time, the network is able to detect patterns and processes in the data. An artificial neuron contains five main components: inputs, weights, sum function, activation function and outputs [26]. In this network, units are placed as layers that are connected to allow the information to flow unidirectionally. It passes from the input units through the units located on the hidden layers and then to the units on the output layer [27]. A set of weighted inputs allows each artificial neuron in the system to give related outputs. The effect of the weights is calculated by the sum function which is calculated by Equation (4).

$$(net)_j = \sum_{i=1}^n w_{ij} x_i + b \quad (4)$$

where “(net) $_j$ is the weighted sum of the j th neuron for the input received from the preceding layer with n neurons, w_{ij} is the weight between the j th neuron in the preceding layer, x_i is the output of the i th neuron in the preceding layer, b is a fixed value as internal addition and Σ represents the sum function” [26,28,29]. The weights provide an important link for the ANN memory and significant information is fed through the network for optimization by backward propagation [30]. These weights are changed as the input values are read by the network to reduce the difference between the predicted and target values. The activation function processes the net input obtained through the sum function and provides the output values. The output is created using a sigmoid function as follows:

$$(out)_j = f(net)_j = \frac{1}{1 + e^{-\alpha(net)_j}} \quad (5)$$

where α is a constant used to control the slope of the semi-linear region [26].

The sigmoid function is used in the ANN algorithm to convert the linear inputs into non-linear signals. This is extremely important for the learning of higher order polynomials beyond one degree for deeper networks [31,32]. The differentiable nature of the activation function enables the much needed backpropagation process which otherwise will not be possible.

3.2. Data Preparation

The first step in data preparation is to identify and remove inconsistent and incorrect values from the dataset. These values have a value of -9999 in the set as recording may not have been available at the particular observation point. This study uses the visualisation method of the boxplot, a data filtering technique in Excel and regression comparison using Cook’s distance method [33] to remove the outliers. Cook’s distance D_i of observation k is given as:

$$D_i = \frac{\sum_{i=1}^n (\hat{y}_i - (\hat{y}_{i(k)}))^2}{pMSE} \quad (6)$$

where, \hat{y}_i is the i th fitted response value, $\hat{y}_{i(k)}$ is the i th fitted response value when the fit excludes observation k , MSE is the mean squared error, p is the number of coefficients in the regression model.

The next step is to determine if the dataset is stationary. In order to fit a stationary model, it is very important to determine that the data is a realisation of a stationary process [34]. This study uses the augmented Dickey Fuller (ADF) test to confirm the stationarity. The augmented Dickey Fuller test is for larger datasets and tests the null hypothesis that unit root is present in the data. A greater negative value of the ADF statistic than the critical value confirms that null hypothesis can be rejected, and no unit root is present. The following results in Table 1 confirm the Broome dataset's stationarity criteria, and the same procedure was used for all study area locations to confirm stationarity.

Table 1. Augmented Dickey Fuller (ADF) results to determine stationarity.

ADF Statistic: -16.195319	
Results of the Test:	
Test Statistic	-16.19532
<i>p</i>-value	4.130155×10^{-29}
#Lags Used	74.00000
Observations Used	1.373840×10^5
Critical Value (1%)	-3.430398
Critical Value (5%)	-2.861561
Critical Value (10%)	-2.566781

The next step is to consider the number of lags that can be used as inputs for the modelling process. The autocorrelation (ACF) and partial autocorrelation (PACF) function analysis is used to determine the input of the sea level prediction model. Autocorrelations measure the association between current and past values and indicate which past time series values are most useful in predicting future values [35–37]. The partial autocorrelation function (PACF) provides the partial correlation of a stationary time series with its own lagged values, which regresses the time series values at all shorter lags [38,39]. Cross-correlation is a measure of similarities between two time series which helps to identify important patterns [40,41].

Figure 2 shows all lags with ACF and PACF for all datasets, and three significant lags were selected for inputs in the MARS and ANN model.

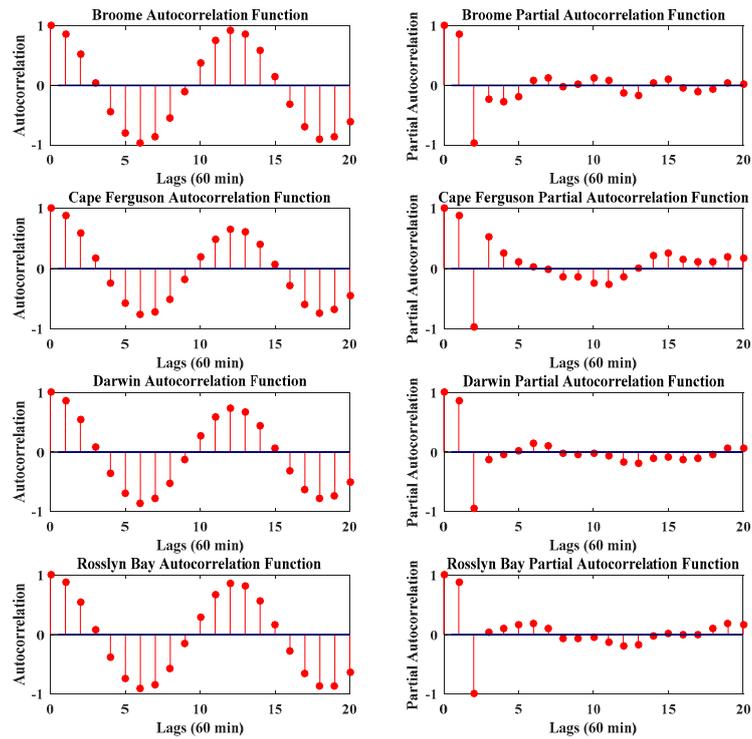


Figure 2. Autocorrelation (ACF) and partial autocorrelation (PACF) for all data locations.

Considering, sea level SL as the time series variable for the 60-min interval, three significant lags of (SL, SL_{t-2}, SL_{t-1}) are then used with the other input variables of water temperature WT , air temperature AT , barometric pressure BP , residuals R , adjusted residuals AR , wind direction WD , wind gust WG , and wind speed WS . Table 2 shows the cross-correlation results of the inputs and target variable SL for the Broome dataset.

Table 2. Cross-correlation results of the inputs and target variable sea level for Broome tide gauge station.

Broome	Water Temp	Air Temp	Barometric Pressure	Residuals	Adjusted Residuals	Wind Direction	Wind Gust	Wind Speed	Sea Level
Water Temp	1	0.8	−0.74	0.034	−0.29	0.29	0.13	0.078	0.025
Air Temp	0.8	1	−0.63	0.028	−0.24	0.33	0.002	−0.021	0.07
Barometric Pressure	−0.74	−0.63	1	−0.22	0.21	−0.32	−0.24	−0.19	−0.029
Residuals	0.034	0.028	−0.22	1	0.91	0.0023	−0.089	−0.098	0.066
Adjusted Residuals	−0.29	−0.24	0.21	0.91	1	−0.13	−0.2	−0.18	0.072
Wind Direction	0.29	0.33	−0.32	0.0023	−0.13	1	0.031	0.043	0.024
Wind Gust	0.13	0.002	−0.24	−0.089	−0.2	0.031	1	0.92	−0.03
Wind Speed	0.078	−0.021	−0.19	−0.098	−0.18	0.043	0.92	1	−0.024
Sea Level	0.025	0.07	−0.029	0.066	0.072	0.024	−0.03	−0.024	1

3.3. Data Normalisation

All of the model input data were normalized [42–44] to make the range of 0–1 for modelling by Equation (7):

$$x_n = \frac{x_{\text{actual}} - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \tag{7}$$

After the prediction using the then trained model, values are returned to the original values using Equation (8):

$$x_{\text{actual}} = x_n (x_{\text{max}} - x_{\text{min}}) + x_{\text{min}} \tag{8}$$

where x is the input data value, x_{min} is the overall minimum and x_{max} is the overall maximum value.

3.4. Model Development

All predictive models were developed in MATLAB, R and the Python programming environment under an Intel i7, 3.40 GHz system. The main objective of this study was to apply MARS and ANN for forecasting mean sea level around northern Australian coastlines. The process of developing the forecasting model was using the 60% training set and predicting the last 20% of data after validation (20%).

3.4.1. MARS Model Development

To develop the MARS model for the prediction of mean sea level, the following parameters are selected to obtain optimal results (please refer to Table 3).

Table 3. Parameters of multivariate adaptive regression splines (MARS) model.

maxFuncs	c	maxInteractions	Cubic
20	0	2	piecewise-cubic

The ‘maxFuncs’ parameter is the maximal number of basis functions in the model, which are included in the forward model building phase before pruning in the backward phase. The default value of the intercept term is 21 and the recommended value for this parameter is two times the number of basis functions in the final model [17]. The ‘c’ parameter is the generalized cross-validation (GCV) penalty per knot, this is taken to be in

the range of about 2–4. A value of 0 penalizes only terms, not knots; hence it can be useful with large data and low noise. According to [17], the default value for c is 3. The ‘max-Interactions’ value is 2 in the model that specifies the maximum degree of interactions between input variables. When set to 1 for additive modelling, there will be no interactions. Generally, a low degree interaction is taken with 1 as default value, however, higher degrees can be used when needed. The ‘cubic’ condition can be given as piecewise-cubic (true) or piecewise-linear (false) type of modelling [17]. The model in this study uses piecewise-cubic modelling to obtain higher predictive performance as data is ‘smoother’ and less noisy.

3.4.2. Artificial Neural Network (ANN) Model Development

To develop the ANN model for predicting mean sea level, Table 4 shows the parameters selected to obtain the optimal results.

Table 4. Parameters of artificial neural network (ANN) model.

Algorithm	Performance Criteria	Inputs	Hidden Layers	Output Layer
Levenberg-Marquardt (trainlm)	Mean Squared Error (mse)	11	10	1

Figure 3 shows the ANN network architecture, the hidden layer and output layer has the weights, fixed value of b and the sigmoid function. The Levenberg Marquardt algorithm is one of the commonly used training algorithms for a multi-layer neural network [45].

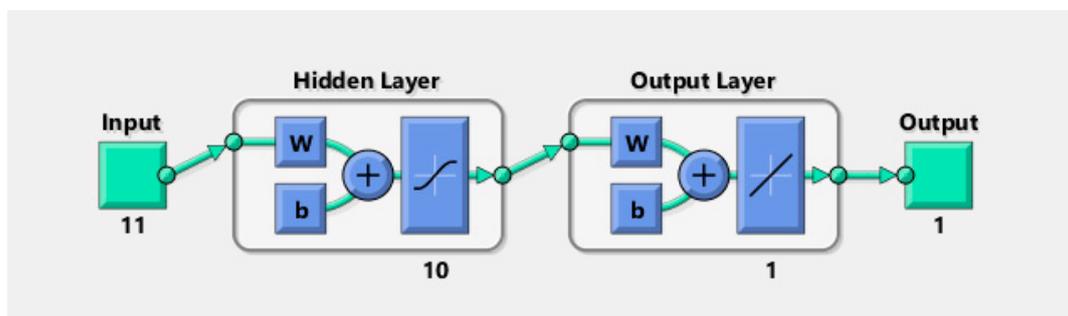


Figure 3. ANN network architecture for data modelling (11–10–1–1).

This is a gradient descent technique to minimize the error for the training phase in modelling. In this study, ANNs inputs are trained (Figure 4) and validated (Figure 5) with the Levenberg Marquardt algorithm. The ‘mse’ is the default performance function for feed forward neural network, it is the mean magnitude of the squares of the error. A well-trained ANN has a low ‘mse’ at the end of the training phase, which signifies a smaller distance between predicted and actual values.

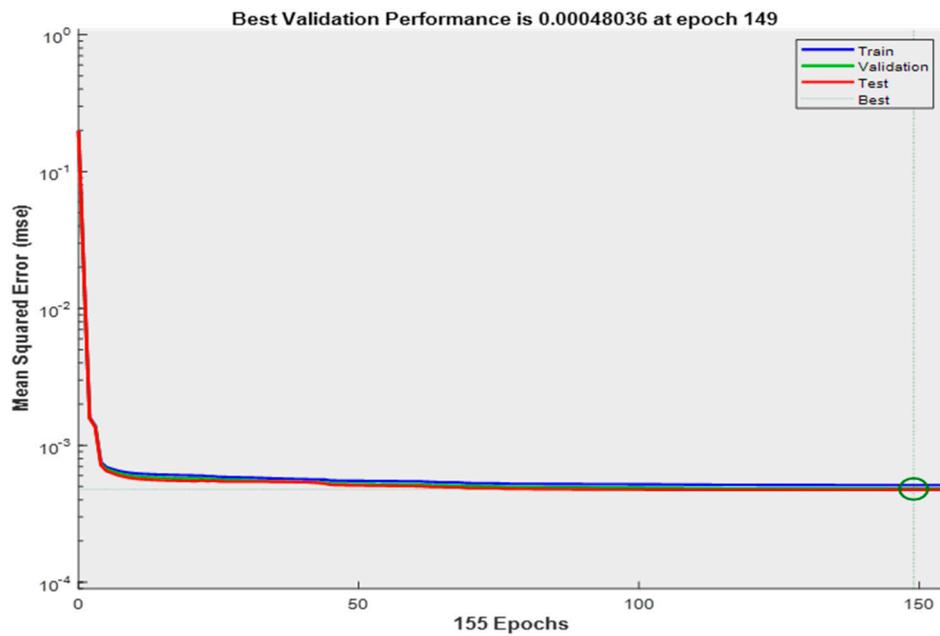


Figure 4. Training performance showing mean squared error (mse) with epochs.

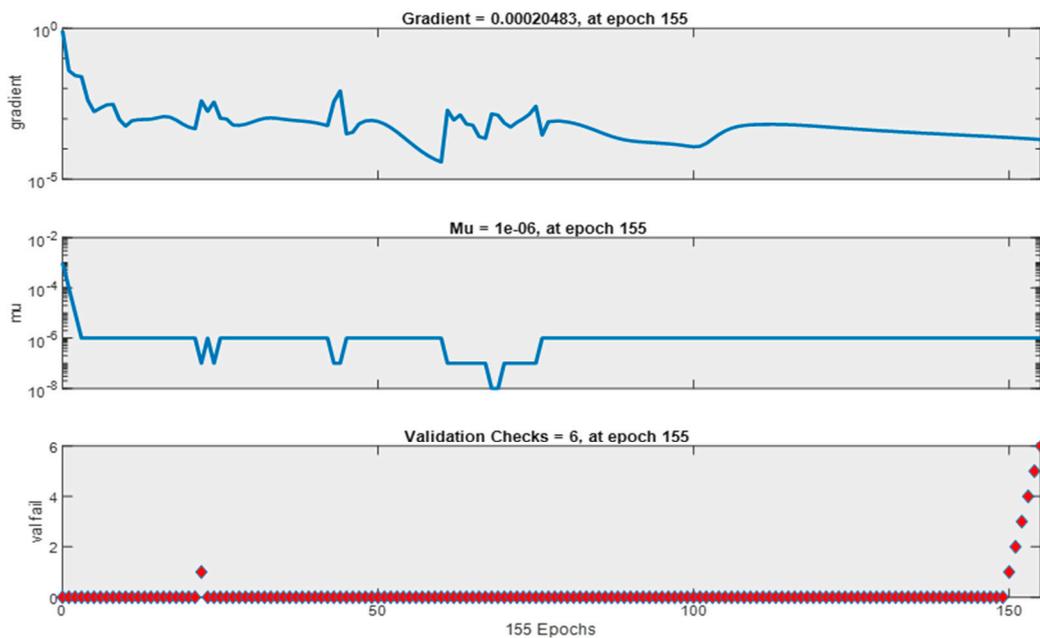


Figure 5. ANN training state.

3.5. Model Performance Criteria

Several statistical metrics were used in this study to evaluate the performance of the MARS and ANN model. The commonly used model score metrics such as Pearson’s correlation coefficient (R), Nash–Sutcliffe coefficient (ENS), Willmott’s index of agreement (d), root mean square error (RMSE; Wm-2), mean absolute error (MAE; Wm-2), including the relative root mean square error (RRMSE; %) and Legates and McCabes index (L) have been calculated to compare the performance. The mathematical forms are as follows,

1. Correlation coefficient (R)

$$R^2 = \left[\frac{\sum_{i=1}^n (DO_i - MDO)(DS_i - MDS)}{\sqrt{\sum_{i=1}^n (DO_i - MDO)^2} \sqrt{\sum_{i=1}^n (DS_i - MDS)^2}} \right]^2 \tag{9}$$

where:

DO_i = Data observed, DS_i = Data simulated,

MDO – Mean of observed data, MDS – Mean of simulated data

2. Nash–Sutcliffe Coefficient (ENS)

$$E_{NS} = 1 - \left[\frac{\sum_{i=1}^n (DO_i - DS_i)^2}{\sum_{i=1}^n (DO_i - MDO)^2} \right] \tag{10}$$

$-\infty \leq E_{NS} \leq 1$

3. Willmott’s Index of agreement (d)

$$d = 1 - \left[\frac{\sum_{i=1}^n (DO_i - DS_i)^2}{\sum_{i=1}^n (|DS_i - MDO| + |DO_i - MDS|)^2} \right] \tag{11}$$

4. Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^n (DS_i - DO_i)^2} \tag{12}$$

5. Mean Absolute Error (MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |DS_i - DO_i| \tag{13}$$

6. Relative Root Mean Square Error (RRMSE)

$$RRMSE = \frac{\sum_{i=1}^n |DS_i - DO_i|}{\sum_{i=1}^n DO_i} \times 100 \tag{14}$$

7. Legates and McCabes index (L)

$$L = 1 - \left[\frac{\sum_{i=1}^n |DS_i - DO_i|}{\sum_{i=1}^n |DO_i - MDS|} \right] \tag{15}$$

$0 \leq L \leq 1$

4. Results and Discussion

In this section, results obtained from the MARS and ANN model for forecasting mean sea level around northern Australian coastlines were assessed to validate the study. The forecasted values using all models in this study were analysed in terms of the predictive accuracy. The comparison was made based on the eight statistical performance criteria (Equations (9)–(15)). Both model metrics show highly accurate results in forecasting the mean sea level (SL) at all data locations in this study (Tables 5 and 6).

Table 5. Training metrics for both models at all locations.

Location	Model	R ²	RMSE	MAE	RRMSE	d	ENS	LEGATES
Broome	MARS	0.9944	0.2185	0.0931	4.0375	0.9944	0.9888	0.9455
Broome	ANN	0.9959	0.1873	0.0698	3.4607	0.9959	0.9918	0.9591
Cape Ferguson	MARS	0.9958	0.0621	0.0393	3.5402	0.9955	0.9915	0.9290
Cape Ferguson	ANN	0.9973	0.0489	0.0337	2.9032	0.9972	0.9947	0.9386

Darwin	MARS	0.9862	0.2663	0.1227	6.4002	0.9862	0.9726	0.9091
Darwin	ANN	0.9892	0.2360	0.1071	5.6718	0.9891	0.9785	0.9206
Rossllyn Bay	MARS	0.9952	0.1032	0.0564	4.2664	0.9951	0.9905	0.9378
Rossllyn Bay	ANN	0.9961	0.0931	0.0502	3.8511	0.9960	0.9923	0.9447

Table 6. Testing metrics for both models at all locations.

Location	Model	R2	RMSE	MAE	RRMSE	d	ENS	LEGATES
Broome	MARS	0.9918	0.2636	0.0990	4.7889	0.9919	0.9836	0.9417
Broome	ANN	0.9939	0.2278	0.0743	4.1397	0.9940	0.9877	0.9563
Cape Ferguson	MARS	0.9966	0.0558	0.0368	3.1877	0.9964	0.9931	0.9337
Cape Ferguson	ANN	0.9974	0.0488	0.0333	2.7919	0.9973	0.9947	0.9399
Darwin	MARS	0.9960	0.1478	0.0974	3.4788	0.9957	0.9914	0.9270
Darwin	ANN	0.9970	0.1262	0.07	2.9703	0.9968	0.9937	0.9409
Rossllyn Bay	MARS	0.9975	0.0777	0.0503	3.1130	0.9974	0.9948	0.9459
Rossllyn Bay	ANN	0.9980	0.0692	0.0448	2.7736	0.9979	0.9959	0.9519

The important aspect in evaluation of any prediction model is to find the error values such as RMSE, MAE, RRMSE. Figure 6 shows a three-dimensional bar graph of the percentage RMSE, RRMSE and MAE generated by MARS and ANN models for all locations analyzed using the hourly predicted data in the testing phase. The analysis of the predicted mean sea level revealed a greater efficacy of the ANN when compared with the MARS model. All RMSE metrics at all locations generated by the ANN model were slightly lower than those of the MARS model (RMSE = 26.36% vs. 22.78% for Broome, 5.58% vs. 4.88% for Cape Ferguson, 14.78% vs. 12.62% for Darwin and 7.77% vs. 6.92% for Rossllyn Bay). The percentage RRMSE and MAE show the same trend with lower error values in ANN model predictions.

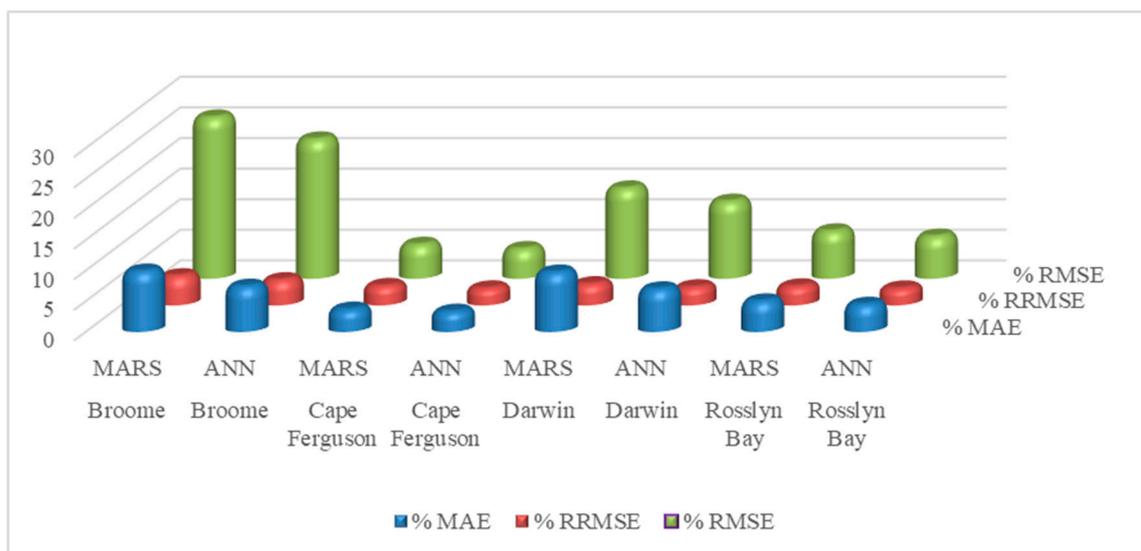


Figure 6. Three-dimensional bar graph of the percentage RMSE, RRMSE, MAE for all study sites generated by MARS and ANN in the testing phase.

Further to error evaluation of the models, dimensionless index such as Willmott’s index of agreement (d), Nash–Sutcliffe coefficient (ENS) and Legates and McCabes index (L) also provide significant measure of model performances. A radar plot is very useful in displaying multivariate data in the form of a two-dimensional chart. Figure 7 is a radar plot of the three indices (d, ENS, L) showing the accuracy of the models at each data study

location in the testing phase. Each star spans greater than a radius of 0.94 where values closer to 1 indicate higher accuracy in the prediction of the mean sea level.

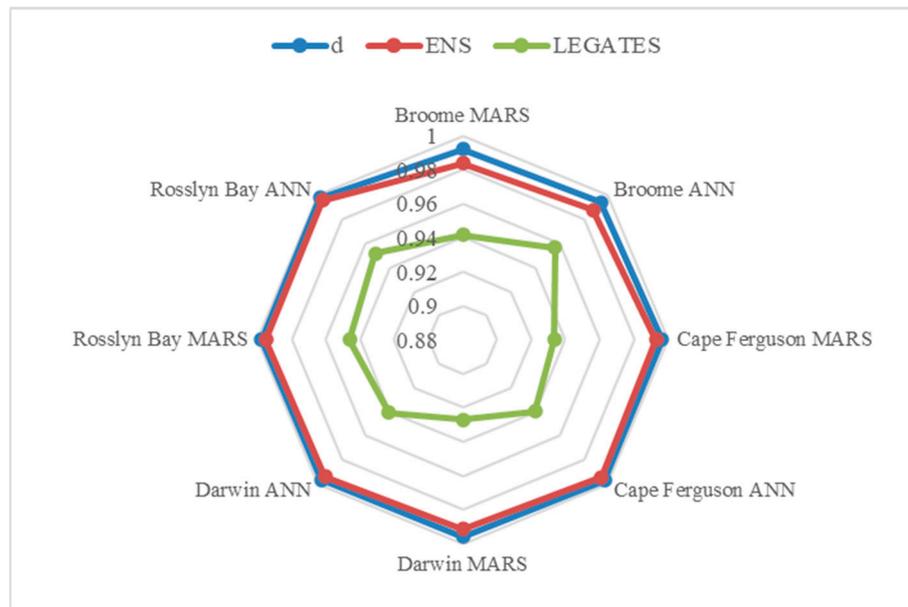


Figure 7. Radar the absolute prediction error distribution of MARS and ANN model.

The box plot is an exploratory data analysis technique which shows significant patterns in a data set [46]. Even though it is simple, it is not distorted by a few extreme values and effectively visualises the measure of dispersion [47]. Figure 8 shows the box plot of absolute prediction errors for each study site. Evidently, each ANN plot is narrower than the MARS plot indicating a much greater spread of errors in MARS when compared with the ANN model.

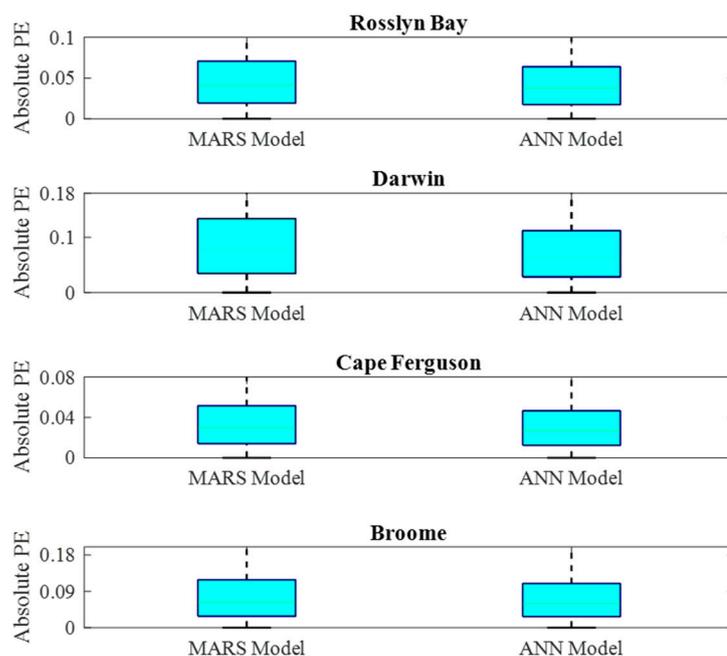


Figure 8. Boxplots show the absolute prediction error distribution of the MARS and ANN model.

To further analyse the result metrics, Figure 9 shows the absolute prediction error (PE) presented in the testing phase. This is obtained in error brackets of 0.1 increment; the frequency of absolute prediction error is shown in bars of the histogram. Consistent with the results presented earlier, the best prediction is shown by the ANN model as a high frequency of errors are on the lower bracket of the histogram (Figure 9).

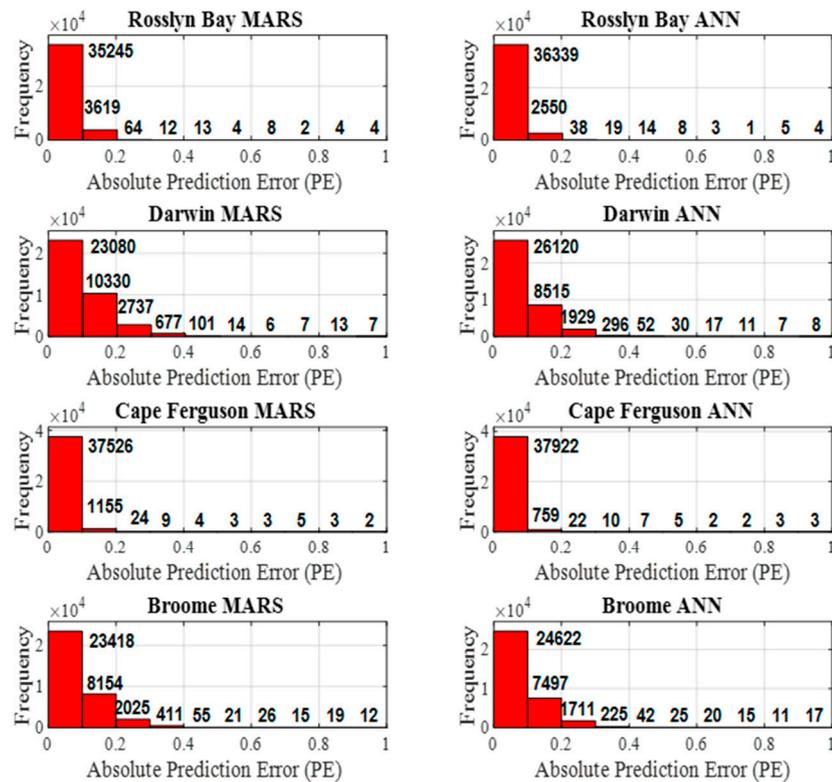


Figure 9. Histograms show the absolute prediction error distribution bars of the MARS and ANN model.

The scatter plot is a powerful analytical tool [48,49]. Figure 10 plots observed and predicted data with a correlation coefficient to visualise the accuracy of MARS and ANN model. As seen in the graphs, all data sites show a high correlation in both models (>0.99). However, ANN has shown better forecasting performance comparatively.

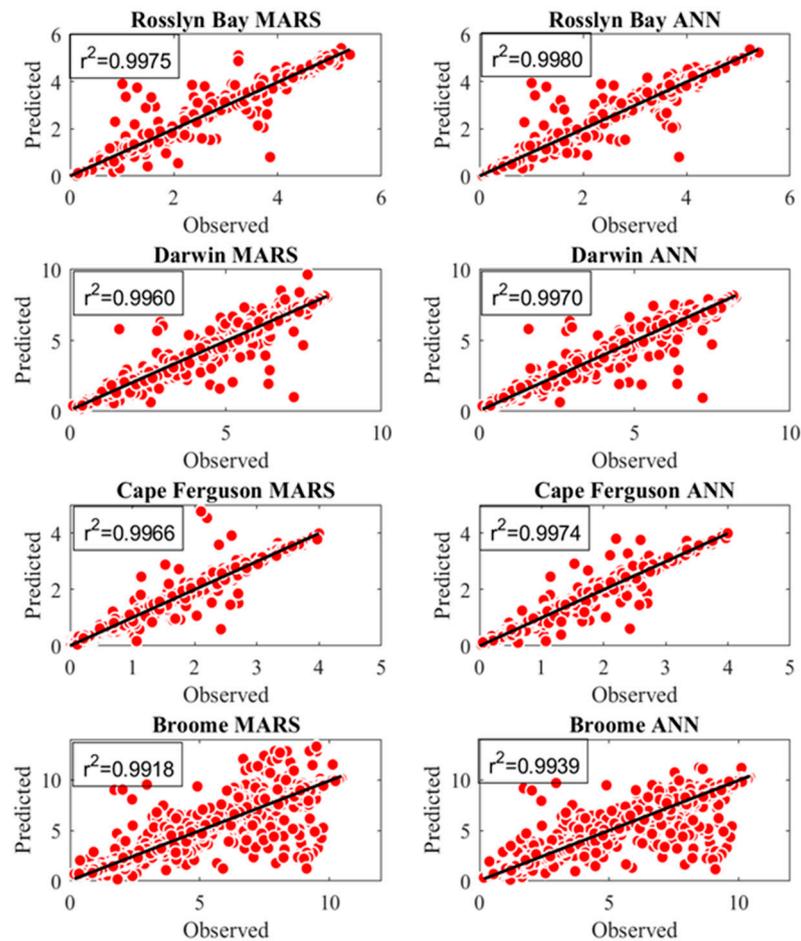


Figure 10. Scatterplot shows the line of best fit of the MARS and ANN model.

A Taylor diagram is used to visualise the degree of correspondence between predicted and observed data. This diagram shows the correlation coefficient and the standard deviation of the two models by a single point on a two-dimensional (2-D) plot. Furthermore, it provides the relative merits of MARS and ANN models on how they have performed with respect to the observed reference point. The close proximity of both models with high correlation value on the Taylor plot further confirms the reliable prediction performance. The presence of ANN (in blue) below the MARS (in red) on the plot implies greater accuracy in prediction results (Figure 11).

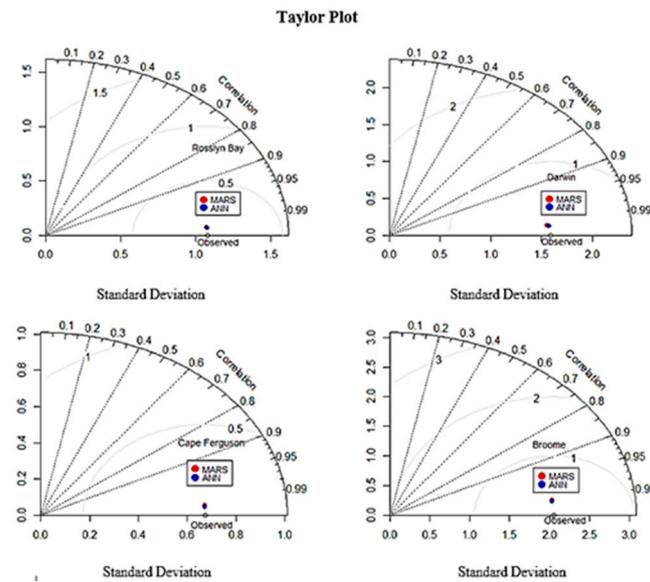


Figure 11. Taylor plot shows the observed, predicted MARS and ANN with standard deviation and correlation. The horizontal and vertical axes show standard deviation values with correlation values on the arc.

Figure 12 shows an hourly sea surface height time series from 20–31 December 2018, both model curves closely follow the observed data points. This confirms the ability of both the models to accurately forecast mean sea level at each data site (see Tables 5 and 6 for RSMSE and R2 values).

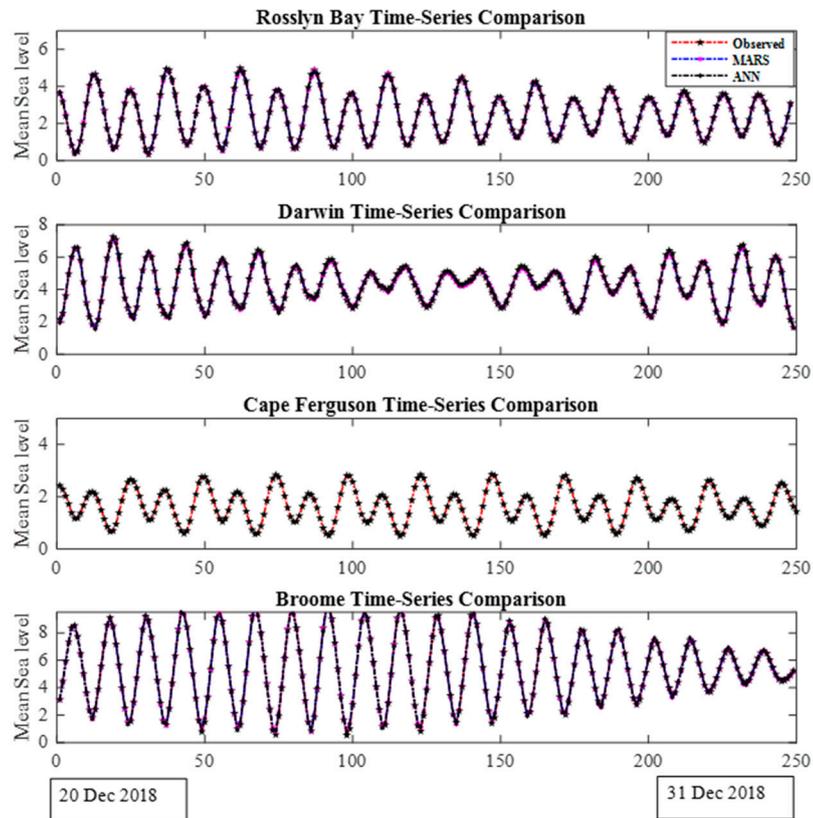


Figure 12. Time series plot shows the predicted MARS and ANN data points with observed data.

Since the ANN model outperforms the MARS model, we have further evaluated the ability of the ANN model to predict sea level variations during a storm surge. For this purpose, we have selected two tropical cyclones over the study region: severe tropical cyclone Marcia (15 February–1 March 2015) and severe tropical cyclone Narelle (5–15 January 2013) and analysed the modelled sea level at nearby tide gauge stations Rosslyn Bay and Broome, respectively. Figure 13 shows tropical cyclone tracks near tide gauge stations Broome and Cape Ferguson.

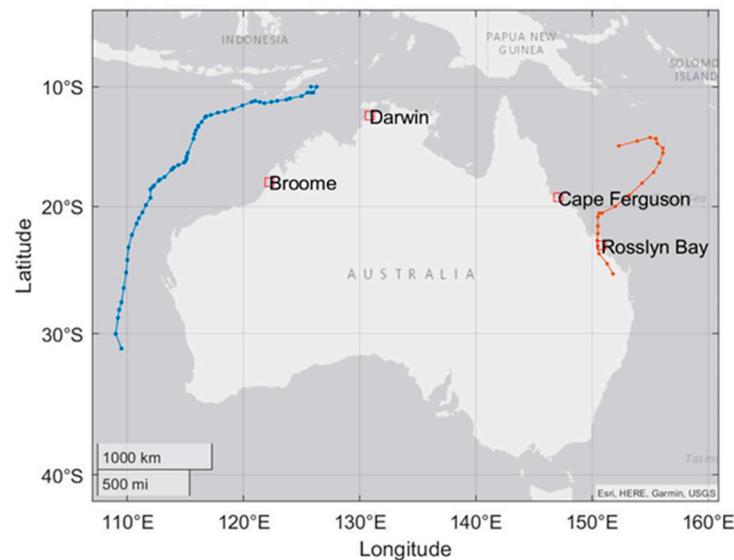


Figure 13. Cyclone track locations have been shown in red line for severe tropical cyclone Marcia (15 February–1 March 2015) and blue line for severe tropical cyclone Narelle (5–15 January 2013). The locations of tide gauges used for this study are shown in red squares.

The results indicate good agreement between the modelled sea-level anomalies (SLAs) by the ANN model and recorded SLAs by the chosen tide gauge stations during = cyclone periods, namely Marcia and Narelle. The maximum correlation coefficients ~ 0.9 and minimum RMS errors ~ 6 cm are calculated between high water levels associated with cyclones estimated using the ANN model and measured at chosen tide gauge stations. The ANN model can successfully model a storm surge that happened during the cyclones, but it has slightly overestimated the surge peak (Figure 14). The modelled peak surge by ANN is 2.16 m (no tide) as compared to the measured peak of 2.14 that occurred at Rosslyn Bay site on 7 February 2015 during cyclone Marcia (Figure 14, upper panel). The modelled peak storm surge by the ANN model is slightly higher at 5.86 m than the 5.95 observed at the site observed on 9 January 2013 (Figure 14, lower panel).

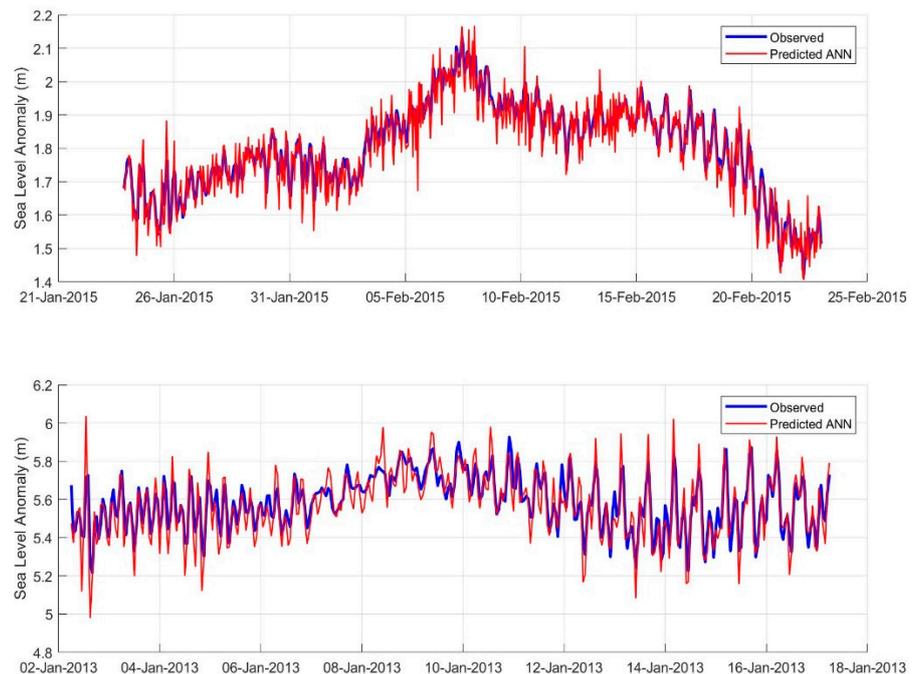


Figure 14. Comparisons between modelled and measured sea level variations during cyclones at selected sites: Upper Panel: Rosslyn Bay gauge, cyclone Marcia, Lower Panel: Broome gauge, cyclone Narelle. Sea levels measured at validation sites are shown in blue, modelled by ANN are shown in red.

5. Conclusions

In this study, the ANN and MARS models are used to predict mean sea level based on 26 years of data from four tide gauge stations (Broome, Cape Ferguson, Darwin and Rosslyn Bay) located around northern Australian coastlines. The performance of the ANN model has been compared with the MARS previously modelled in the study region (c.f. [9]). The analysis of the predicted mean sea level revealed a greater efficacy of the ANN when compared with the MARS model. All RMSE metrics at all locations generated by the ANN model were slightly lower than those of the MARS model (RMSE = 26.36% vs. 22.78% for Broome, 5.58% vs. 4.88% for Cape Ferguson, 14.78% vs. 12.62% for Darwin and 7.77% vs. 6.92% for Rosslyn Bay). The percentage RRMSE and MAE show the same trend with lower error values in ANN model predictions. Scatter plots of all data sites showed a high correlation in both models (>0.99). However, ANN has shown better forecasting performance comparatively.

The ANN model performance was further evaluated for modelling high water levels associated with cyclones. The model reproduced the surge profile with the maximum correlation coefficients ~ 0.99 and minimum RMS errors ~ 4 cm at the selected validating site. In addition, the ANN model predicted the maximum surge at Rosslyn Bay for cyclone Marcia to within 2 cm of measured peak and the maximum surge at Broome for cyclone Narelle to within 7 cm of the measured peak. The results are comparable with the MARS model previously used in this region. However, the ANN model shows better agreement with the measured peak and arrival time, although it suffers from slightly higher predictions than the observed sea level by tide gauge station.

In this study, we have removed the tidal signal from observed sea level, however, [50] showed that the inclusion of the tidal variations in the storm surge models, giving higher explained variances of the measured peak surge. It is proven that the inclusion of tides in the ANN model will probably not have a large effect on the RMS errors due to the

deterministic nature of the tidal sea level variability [51]. Therefore, it is possible to obtain an improvement in the hindcast skill of the ANN model by including the tides. This implies that the surge can be predicted independently of the tide in this region. Future studies can apply this technique to predict sea levels in other parts of the world.

Author Contributions: Conceptualization, Z.G.; data curation, Z.G.; formal analysis, N.R.; investigation, Z.G.; methodology, N.R.; software, N.R.; validation, N.R. and Z.G.; writing—original draft, N.R. and Z.G.; writing—review & editing, N.R. and Z.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The tide gauges data are taken from the Australian National Tidal Centre (NTC, <http://www.bom.gov.au/oceanography/>).

Acknowledgments: We would like to thank the anonymous reviewers and editors for their constructive comments on this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Church, J.A.; White, N.J. Sea-Level Rise from the Late 19th to the Early 21st Century. *Surv. Geophys.* **2011**, *32*, 585–602.
- White, N.J.; Haigh, I.D.; Church, J.A.; Koen, T.; Watson, C.; Pritchard, T.R.; Watson, P.J.; Burgette, R.J.; McInnes, K.L.; You, Z.-J.; et al. Australian sea levels—Trends, regional variability and influencing factors. *Earth-Sci. Rev.* **2014**, *136*, 155–174. <https://doi.org/10.1016/j.earscirev.2014.05.011>.
- Watson, P.J. An Assessment of the Utility of Satellite Altimetry and Tide Gauge Data (ALT-TG) as a Proxy for Estimating Vertical Land Motion. *J. Coast. Res.* **2019**, *35*, 1131–1144. <https://doi.org/10.2112/jcoastres-d-19-00031.1>.
- Gharineiat, Z.; Deng, X. Description and assessment of regional sea-level trends and variability from altimetry and tide gauges at the northern Australian coast. *Adv. Space Res.* **2018**, *61*, 2540–2554. <https://doi.org/10.1016/j.asr.2018.02.038>.
- Deng, X.; Gharineiat, Z.; Andersen, O.B.; Stewart, M.G. Observing and modelling the high water level from satellite radar altimetry during tropical cyclones. In *IAG 150 Years*; Springer: Cham, Switzerland, 2015; pp. 491–500. https://doi.org/10.1007/1345_2015_108.
- Jevrejeva, S.; Grinsted, A.; Moore, J.; Holgate, S. Nonlinear trends and multiyear cycles in sea level records. *J. Geophys. Res. Space Phys.* **2006**, *111*. <https://doi.org/10.1029/2005jc003229>.
- Watson, P.J. Acceleration in European Mean Sea Level? A New Insight Using Improved Tools. *J. Coast. Res.* **2017**, *331*, 23–38. <https://doi.org/10.2112/JCOASTRES-D-16-00134.1>.
- Watson, P.J. Acceleration in U.S. Mean Sea Level? A New Insight using Improved Tools. *J. Coast. Res.* **2016**, *322*, 1247–1261. <https://doi.org/10.2112/jcoastres-d-16-00086.1>.
- Gharineiat, Z.; Deng, X. Application of the Multi-Adaptive Regression Splines to Integrate Sea Level Data from Altimetry and Tide Gauges for Monitoring Extreme Sea Level Events. *Mar. Geodesy* **2014**, *38*, 261–276. <https://doi.org/10.1080/01490419.2015.1036183>.
- Quirós, E.; Felicísimo, M.; Cuartero, A. Testing Multivariate Adaptive Regression Splines (MARS) as a Method of Land Cover Classification of TERRA-ASTER Satellite Images. *Sensors* **2009**, *9*, 9011–9028. <https://doi.org/10.3390/s91109011>.
- Boye, C.B.; Amoah, V.I. Performance Evaluation for Mean Sea Level Prediction using Multivariate Adaptive Regression Spline and Artificial Neural Network. *Ghana Min. J.* **2018**, *18*, 1–8. <https://doi.org/10.4314/gm.v18i1.1>.
- Pashova, L.; Popova, S. Daily sea level forecast at tide gauge Burgas, Bulgaria using artificial neural networks. *J. Sea Res.* **2011**, *66*, 154–161. <https://doi.org/10.1016/j.seares.2011.05.012>.
- Conyers, Z.A.; Grant, R.; Roy, S.S. Sea Level Rise in Miami Beach: Vulnerability and Real Estate Exposure. *Prof. Geogr.* **2019**, *71*, 278–291. <https://doi.org/10.1080/00330124.2018.1531037>.
- Meena, B.; Agrawal, J. Tidal Level Forecasting Using ANN. *Procedia Eng.* **2015**, *116*, 607–614. <https://doi.org/10.1016/j.proeng.2015.08.332>.
- Lee, H.S. and S. Kim. Sea-Level Records Analysis with Improved Empirical Mode Decomposition (EMD) and Artificial Neural Networks (ANN). In Proceedings of the 27th International Ocean and Polar Engineering Conference, San Francisco, CA, USA, 25 June 2017.
- Gharineiat, Z.; Deng, X. Spectral Analysis of Satellite Altimeter and Tide Gauge Data around the Northern Australian Coast. *Remote. Sens.* **2020**, *12*, 161. <https://doi.org/10.3390/rs12010161>.
- Friedman, J.H., Multivariate adaptive regression splines. *Ann. Stat.* **1991**, *19*, 1–67.

18. Mansouri, I.; Ozbakkaloglu, T.; Kisi, O.; Xie, T. Predicting behavior of FRP-confined concrete using neuro fuzzy, neural network, multivariate adaptive regression splines and M5 model tree techniques. *Mater. Struct.* **2016**, *49*, 4319–4334.
19. Keshtegar, B.; Mert, C.; Kisi, O. Comparison of four heuristic regression techniques in solar radiation modeling: Kriging method vs RSM, MARS and M5 model tree. *Renew. Sustain. Energy Rev.* **2018**, *81*, 330–341. <https://doi.org/10.1016/j.rser.2017.07.054>.
20. Kisi, O. Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree. *J. Hydrol.* **2015**, *528*, 312–320. <https://doi.org/10.1016/j.jhydrol.2015.06.052>.
21. Yang, C.-C.; Prasher, S.O.; Lacroix, R.; Kim, S.H. Application of multivariate adaptive regression splines (mars) to simulate soil temperature. *Trans. ASAE* **2004**, *47*, 881–887. <https://doi.org/10.13031/2013.16085>.
22. Deo, R.C.; Kisi, O.; Singh, V.P. Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. *Atmospheric Res.* **2017**, *184*, 149–175. <https://doi.org/10.1016/j.atmosres.2016.10.004>.
23. Kourou, K.; Exarchos, T.P.; Exarchos, K.P.; Karamouzis, M.V.; Fotiadis, D.I. Machine learning applications in cancer prognosis and prediction. *Comput. Struct. Biotechnol. J.* **2014**, *13*, 8–17. <https://doi.org/10.1016/j.csbj.2014.11.005>.
24. Bhattacharya, B.; Solomatine, D. Neural networks and M5 model trees in modelling water level–discharge relationship. *Neurocomputing* **2005**, *63*, 381–396. <https://doi.org/10.1016/j.neucom.2004.04.016>.
25. Nazari, A.; Khalaj, G.; Riahi, S.; Bohlooli, H.; Kaykha, M.M. RETRACTED: Prediction total specific pore volume of geopolymers produced from waste ashes by ANFIS. *Ceram. Int.* **2012**, *38*, 3111–3120. <https://doi.org/10.1016/j.ceramint.2011.12.011>.
26. Topçu, I.B.; Sarıdemir, M. Prediction of compressive strength of concrete containing fly ash using artificial neural networks and fuzzy logic. *Comput. Mater. Sci.* **2008**, *41*, 305–311. <https://doi.org/10.1016/j.commatsci.2007.04.009>.
27. Rodriguez-Galiano, V.; Sanchez-Castillo, M.; Olmo, M.C.; Chica-Rivas, M. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geol. Rev.* **2015**, *71*, 804–818. <https://doi.org/10.1016/j.oregeorev.2015.01.001>.
28. Abdullah, S.N.; Zeng, X. Machine learning approach for crude oil price prediction with Artificial Neural Networks–Quantitative (ANN-Q) model. In Proceedings of the 2010 International Joint Conference on Neural Networks (IJCNN), Barcelona, Spain, 18–23 July 2010; pp. 1–8. <https://doi.org/10.1109/ijcnn.2010.5596602>.
29. Kankar, P.; Sharma, S.C.; Harsha, S. Fault diagnosis of ball bearings using machine learning methods. *Expert Syst. Appl.* **2011**, *38*, 1876–1886. <https://doi.org/10.1016/j.eswa.2010.07.119>.
30. Agatonovic-Kustrin, S.; Beresford, R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *J. Pharm. Biomed. Anal.* **2000**, *22*, 717–727. [https://doi.org/10.1016/S0731-7085\(99\)00272-1](https://doi.org/10.1016/S0731-7085(99)00272-1).
31. Nwankpa, C.; Ijomah, W.; Gachagan, A.; Marshall, S. Activation functions: Comparison of trends in practice and research for deep learning. *arXiv* **2018**, arXiv:1811.03378.
32. Ghimire, S.; Deo, R.C.; Downs, N.J.; Raj, N. Global solar radiation prediction by ANN integrated with European Centre for medium range weather forecast fields in solar rich cities of Queensland Australia. *J. Clean. Prod.* **2019**, *216*, 288–310. <https://doi.org/10.1016/j.jclepro.2019.01.158>.
33. Cook, R.D. Detection of Influential Observation in Linear Regression. *Technometrics* **1977**, *19*, 15–18.
34. Metcalfe, A.V.; Cowpertwait, P.S. *Introductory Time Series with R*; Springer: Berlin/Heidelberg, Germany, 2009; <https://doi.org/10.1007/978-0-387-88698-5>.
35. Jiang, X.; Adeli, H. Wavelet Packet-Autocorrelation Function Method for Traffic Flow Pattern Analysis. *Comput. Civ. Infrastruct. Eng.* **2004**, *19*, 324–337. <https://doi.org/10.1111/j.1467-8667.2004.00360.x>.
36. Shimamura, T.; Kobayashi, H. Weighted autocorrelation for pitch extraction of noisy speech. In *IEEE Transactions on Speech and Audio Processing*; IEEE: New York, NY, USA, 2001; Volume 9, pp. 727–730.
37. Muhtarov, P.; Kutiev, I. Autocorrelation method for temporal interpolation and short-term prediction of ionospheric data. *Radio Sci.* **1999**, *34*, 459–464. <https://doi.org/10.1029/1998rs900020>.
38. Morf, M.; Vieira, A.; Kailath, T. Covariance Characterization by Partial Autocorrelation Matrices. *Ann. Stat.* **1978**, *6*, 643–648. <https://doi.org/10.1214/aos/1176344208>.
39. Dégerine, S.; Lambert-Lacroix, S. Characterization of the partial autocorrelation function of nonstationary time series. *J. Multivar. Anal.* **2003**, *87*, 46–59. [https://doi.org/10.1016/S0047-259X\(03\)00025-3](https://doi.org/10.1016/S0047-259X(03)00025-3).
40. Dean, R.T.; Dunsmuir, W.T.M. Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. *Behav. Res. Methods* **2015**, *48*, 783–802. <https://doi.org/10.3758/s13428-015-0611-2>.
41. Mouatadid, S.; Raj, N.; Deo, R.C.; Adamowski, J.F. Input selection and data-driven model performance optimization to predict the Standardized Precipitation and Evaporation Index in a drought-prone region. *Atmospheric Res.* **2018**, *212*, 130–149. <https://doi.org/10.1016/j.atmosres.2018.05.012>.
42. Kotsiantis, S.B.; Kanellopoulos, D.; Pintelas, P.E. Data preprocessing for supervised learning. *Int. J. Comput. Sci.* **2006**, *1*, 111–117.
43. Deo, R.C.; Ghimire, S.; Downs, N.J.; Raj, N. Optimization of windspeed prediction using an artificial neural network compared with a genetic programming model. In *Handbook of Research on Predictive Modeling and Optimization Methods in Science and Engineering*; IGI Global, Hershey, United States: 2018; pp. 328–359.
44. Ghimire, S.; Deo, R.C.; Downs, N.J.; Raj, N. Self-adaptive differential evolutionary extreme learning machines for long-term solar radiation prediction with remotely-sensed MODIS satellite and Reanalysis atmospheric products in solar-rich cities. *Remote. Sens. Environ.* **2018**, *212*, 176–198. <https://doi.org/10.1016/j.rse.2018.05.003>.

45. Moré, J.J. The Levenberg-Marquardt algorithm: Implementation and theory. In *Numerical Analysis*; Springer: Berlin/Heidelberg, Germany, 1978; pp. 105–116. <https://doi.org/10.1007/bfb0067700>.
46. Williamson, D.F.; Parker, R.A.; Kendrick, J.S. The Box Plot: A Simple Visual Method to Interpret Data. *Ann. Intern. Med.* **1989**, *110*, 916–921. <https://doi.org/10.7326/0003-4819-110-11-916>.
47. Schwertman, N.C.; Owens, M.A.; Adnan, R. A simple more general boxplot method for identifying outliers. *Comput. Stat. Data Anal.* **2004**, *47*, 165–174. <https://doi.org/10.1016/j.csda.2003.10.012>.
48. Cleveland, W.S.; McGill, R. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *J. Am. Stat. Assoc.* **1984**, *79*, 531–554.
49. Keim, D.A.; Hao, M.C.; Dayal, U.; Janetzko, H.; Bak, P. Generalized Scatter Plots. *Inf. Vis.* **2009**, *9*, 301–311. <https://doi.org/10.1057/ivs.2009.34>.
50. Burston, J.; Symonds, A.; Scheel, F. Application of D-Flow FM for Storm Surge Modelling: Case Study of TC Yasi. Available online: https://www.griffith.edu.au/__data/assets/pdf_file/0014/107312/Burston_DFlow_OCEANSCIENCES14_poster.pdf, (accessed on 22 October 2021).
51. Høyer, J.L.; Andersen, O.B. Improved description of sea level in the North Sea. *J. Geophys. Res. Space Phys.* **2003**, *108*. <https://doi.org/10.1029/2002jc001601>. <https://doi.org/>