mathematics

MDPI

Article

Prediction of Sea Level Using Double Data Decomposition and Hybrid Deep Learning Model for Northern Territory, Australia

Nawin Raj [1,*], Jaishukh Murali [1], Lila Singh-Peterson [2] and Nathan Downs [1]

[1] School of Mathematics, Physics and Computing, Springfield Campus, University of Southern Queensland, Springfield, QLD 4300, Australia; u1147091@umail.usq.edu.au (J.M.); nathan.downs@usq.edu.au (N.D.)
[2] School of Agriculture and Environmental Science, Springfield Campus, University of Southern Queensland, Springfield, QLD 4300, Australia; lila.singh-peterson@usq.edu.au
* Correspondence: nawin.raj@usq.edu.au

Abstract: Sea level rise (SLR) attributed to the melting of ice caps and thermal expansion of seawater is of great global significance to vast populations of people residing along the world's coastlines. The extent of SLR's impact on physical coastal areas is determined by multiple factors such as geographical location, coastal structure, wetland vegetation and related oceanic changes. For coastal communities at risk of inundation and coastal erosion due to SLR, the modelling and projection of future sea levels can provide the information necessary to prepare and adapt to gradual sea level rise over several years. In the following study, a new model for predicting future sea levels is presented, which focusses on two tide gauge locations (Darwin and Milner Bay) in the Northern Territory (NT), Australia. Historical data from the Australian Bureau of Meteorology (BOM) from 1990 to 2022 are used for data training and prediction using artificial intelligence models and computation of mean sea level (MSL) linear projection. The study employs a new double data decomposition approach using Multivariate Variational Mode Decomposition (MVMD) and Successive Variational Mode Decomposition (SVMD) with dimensionality reduction techniques of Principal Component Analysis (PCA) for data modelling using four artificial intelligence models (Support Vector Regression (SVR), Adaptive Boosting Regressor (AdaBoost), Multilayer Perceptron (MLP), and Convolutional Neural Network–Bidirectional Gated Recurrent Unit (CNN-BiGRU). It proposes a deep learning hybrid CNN-BiGRU model for sea level prediction, which is benchmarked by SVR, AdaBoost, and MLP. MVMD-SVMD-CNN-BiGRU hybrid models achieved the highest performance values of 0.9979 (d), 0.996 (NS), 0.9409 (L); and 0.998 (d), 0.9959 (NS), 0.9413 (L) for Milner Bay and Darwin, respectively. It also attained the lowest error values of 0.1016 (RMSE), 0.0782 (MABE), 2.3699 (RRMSE), and 2.4123 (MAPE) for Darwin and 0.0248 (RMSE), 0.0189 (MABE), 1.9901 (RRMSE), and 1.7486 (MAPE) for Milner Bay. The mean sea level (MSL) trend analysis showed a rise of 6.1 $\pm$ 1.1 mm and 5.6 $\pm$ 1.5 mm for Darwin and Milner Bay, respectively, from 1990 to 2022.

Keywords: deep learning (DL); sea level (SL); convolutional neural network (CNN); bidirectional gated recurrent unit; support vector regression (SVR); adaptive boosting regressor (AdaBoost); multivariate variational mode decomposition (MVMD); successive variational mode decomposition (SVMD); principal component analysis (PCA)

MSC: 68T07; 68T09; 68W01

Citation: Raj, N.; Murali, J.; Singh-Peterson, L.; Downs, N. Prediction of Sea Level Using Double Data Decomposition and Hybrid Deep Learning Model for Northern Territory, Australia. *Mathematics* **2024**, *12*, 2376. https://doi.org/10.3390/math12152376

Academic Editors: Laura Cornejo-Bueno and Jorge Pérez-Aracil

Received: 14 May 2024
Revised: 18 July 2024
Accepted: 26 July 2024
Published: 30 July 2024

1. Introduction

Global sea level rise is a complicated environmental issue exacerbated by human activities and climate change [1]. The phenomenon of sea level rise (SLR) is currently a matter of global concern due to its significant implications for coastal regions and ecosystems. According to the IPCC (2021), by 2050, the global mean sea level (GMSL) is projected to rise by at least 0.15 up to 0.26 m [2]. Climate change, which is mostly caused by human

activities, including the burning of fossil fuels, deforestation, and industrial processes, is largely responsible for the rise in sea levels worldwide [3]. These impacts, together with the thermal expansion of the ocean [4,5], are the primary causes of sea level rise [5]. This occurs when oceans absorb heat as the Earth's atmospheric temperature continues to increase due to continued greenhouse gas production. The absorption of heat causes sea water to expand, leading to an increase in the volume of the oceans [6]. In addition, the melting of glaciers and polar ice caps adds to the volume of oceans. These combined impacts are attributed to a gradual increase in sea levels, which pose a serious threat to coastal areas [7].

The effects of sea level rise are wide-ranging and broad. As sea levels rise, coastal towns and cities are increasingly vulnerable to flooding [8,9], which can harm infrastructure and public facilities, causing property loss and disrupting beach amenity and daily life. Furthermore, drinking water supplies, aquifers that contain freshwater, and agricultural land are also at risk of saltwater interference. Seawater intrusion is a threat to coastal habitats, including wetlands and estuaries, and can have a negative impact on biodiversity [10].

In recognition of these far-reaching impacts, the need to adapt to changing coastlines is a reality for all coastal communities. Data pertaining to what future sea levels may look like provide a foundation for planning and adaptation strategies that ensure that the impacts associated with SLR are mitigated or at least minimalized.

This study focusses on the Northern Territory region of Australia and aims to present a method developed to assess and predict future sea level rise. The potential for coastal inundation is further compounded by tropical cyclones and storm surges which frequently occur in this area. In fact, on average, the Northern Territory experiences at least two cyclone events in a year [11]. The effects of tropical cyclones are often disastrous, particularly when storm surges coincide with spring tides [11]. This would be an important investigation and future study for the region to anticipate such occurrences with sea level rise. In the same region, a ground-based study [12] found that the majority of local people were likely to relocate from their houses due to sea water intrusion in coming years. Northern Territory has coastal wetlands that lie between 0 and 5 m above the mean sea level and are therefore more vulnerable to SLR and sea water inundation [13]. In such an environment, accurate and reliable information regarding future sea level changes is extremely important for the development of timely decisions regarding adaptation and mitigation options.

A sea level dataset analysis can be challenging with the generation of large volumes of time series data and its dependence on related oceanic parameters. For processing such complicated, high-dimensional data, the advancement of deep learning architecture has provided a suitable platform to process multi-inputs for data modelling and forecasting [14,15]. The deep learning models provide the advantages of neural layers which are capable of processing and providing higher accuracy in predictions [16]. Extraction of spatial information from the input data is highly suited for hybrid models such as Convolutional Neural Network–Bidirectional Gated Recurrent Unit (CNN-BiGRU) [17]. It is able to handle large datasets and effectively extract connections in the dataset [17,18].

Feature extraction techniques have also become very important in data modelling as they help to extract valuable information from raw signals [19]. A study in [20] used multi-signal wavelet decomposition to extract temporal multiscale variations from satellite altimetry data and successfully tracked the spatial–temporal distribution. Sea level signal was decomposed into its intrinsic mode functions (IMFs) in [21] using Successive Variational Mode Decomposition (SVMD), which were used as feature inputs for sea level rise forecasting. The Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) technique was effectively used in [22] to extract IMFs for sea level rise prediction.

Most studies have either used a single standalone machine learning model [23–25] or employed single data decomposition with machine learning models [21,22,26] for SLR predictions. To date, no SLR study has employed PCA analysis with data decomposition to optimize the data inputs. Hence, this study utilizes CNN-BiGRU as its objective model and implements double data decomposition for feature extraction. To compare and analyse how double data decomposition enhances prediction accuracy, the study also uses Multilayer Perceptron (MLP), Support Vector Regression (SVR), and Adaptive Boosting Regressor (AdaBoost) as benchmark models. While many studies have used single data decomposition techniques, this study is novel as it is the first study, to the best of authors' knowledge, to explore a double data decomposition method for sea level prediction in Australia. These exploratory approaches culminate in a new double data decomposition method of Multivariate Variational Mode Decomposition (MVMD) and Successive Variational Mode Decomposition (SVMD), followed by the Principal Component Analysis (PCA) for optimal dimensionality reduction. This method presents a new approach for sea level rise study in Australia. Specifically, the objectives of this research are the following:

1. To employ the double data decomposition method (MVMD-SVMD) for sea level signal to extract intrinsic mode functions (IMFs).
2. To reduce input data dimensionality by applying PCA analysis.
3. To predict sea level using a hybrid deep learning CNN-BiGRU model.
4. To benchmark the CNN-BiGRU model with SVR, AdaBoost, and MLP models.
5. To provide an estimate of future sea levels for both study sites.

The remainder of this paper is organized as follows. Section 2 provides the details on oceanic and sea level data sources, locations, and characteristics. It also includes data input correlations, decomposition process and dimensionality reduction, evaluation metrics formulae, and the model background. Section 3 presents the results of data modelling and predictions.

2. Materials and Methods

2.1. Study Area and Dataset

The historical dataset for Darwin and Milner Bay in the Northern Territory of Australia from the "Australian Baseline Sea Level Monitoring Project Hourly Sea Level and Meteorological Data" (http://www.bom.gov.au/oceanography/projects/abslmp/data/index.shtml, accessed on 15 July 2023) provided by the Bureau of Meteorology (BOM), Australia, was used for this study. This dataset includes hourly tide gauge data for the study locations from 1990 to 2022. Tide gauge observations provide continuous data that are useful for modelling future sea level predictions [27]. The dataset includes oceanic data such as sea level height (in m), water temperature (in °C), air temperature (in °C), barometric pressure (in hPa), wind direction (in degrees True), wind gust (in m/s), and wind speed (in m/s) at 60 min intervals. With the goal of tracking sea levels along Australia's coastline, BOM developed the Australian Baseline Sea Level Monitoring Project. The tide gauge is a part of a water observation station that gathers data using installed sensors. The study locations are Darwin and Milner Bay with latitudes and longitudes of 12°27′40″ S, 130°50′30″ E and 12°23′11″ S, 130°51′43″ E, respectively, in the state of Northern Territory, Australia. The research area map of the two coastal cities is displayed in Figure 1 below.

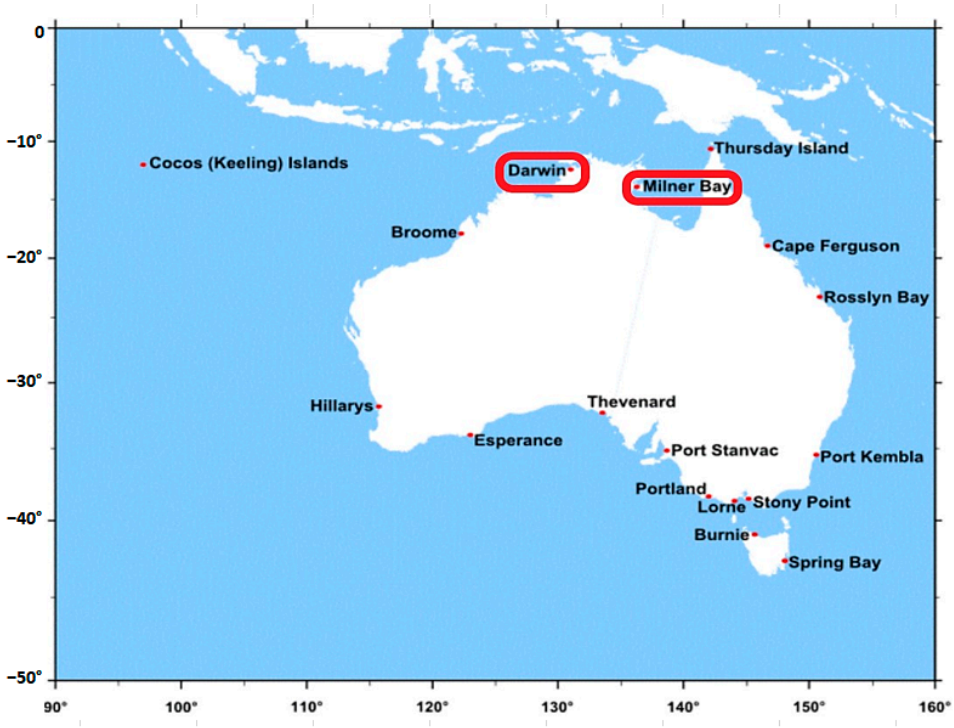

Figure 1. Australia map showing the sea level study sites. The red outline is around the study locations.

2.2. Data Preprocessing and Input Selection

A critical component of every data analysis study is ensuring the quality of the data and providing scientific solutions for missing values. One of the most important aspects of the provision of accurate SLR modelling relates to the quality of available data and methods of analysis for sea level rise forecasting. This involves data preprocessing and feature extraction for Artificial Intelligence (AI) modelling. Data preprocessing is an essential foundation for SLR forecasting that determines the quality of sea level predictions [28]. Data preparation entails a number of stages, including the removal of inaccurate or missing data points, the use of means from a specific year or interpolation to fill in gaps, and the standardization of the data to ensure consistency in units and scales [29]. This study used the linear interpolation method to address the missing values in the dataset. Lagged values were used as model inputs in many types of research, which is known to improve prediction accuracy. Figure 2 below shows the antecedent behaviour of lags in Darwin using the Auto Correlation Function (ACF) and Partial Autocorrelation Function (PACF) graphs for the sea level series values in terms of variations in their separation from one another. In addition, PACF is also used because of its capacity to yield the series' partial correlation with its very own lagged readings. This helps to determine how many historical lags are useful to include in the prediction model. Four lags of the sea level time series were chosen as input for the AI model after the graphical analysis for each series in Darwin and Milner Bay dataset. Figure 2 shows the ACF-PACF analysis for the Darwin site. Table 1 lists the oceanic parameters and lagged values used in the Artificial Intelligence models. The lags and other oceanic metrics were evaluated using a correlation test as shown in Tables 2 and 3 when compared with sea level as a target variable for prediction.

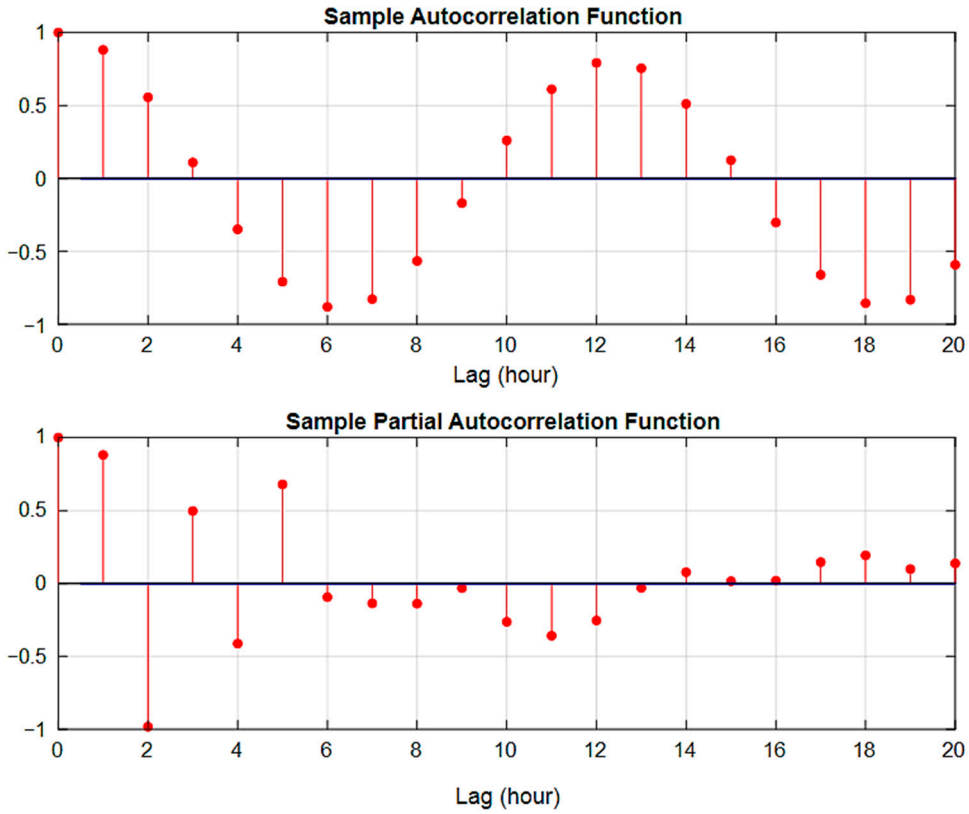

Figure 2. ACF-PACF analysis for the Darwin study site. The red dots shows the correlation of lags for the Darwin sea level dataset.

Table 1. Input features and description.

Input Oceanic Features
Air Temperature (°C)
Water Temperature (°C)
Wind Direction (in degrees True)
Wind Gust (m/s)
Wind Speed (m/s)
Barometric Pressure (hPa)
Sea Level Lags (t − 1, t − 2, t − 3, t − 4) (m)
Decomposed Intrinsic Mode Functions (IMFs) (m)

Table 2. Sea level correlation with lagged values and decomposed IMFs.

	L1	**L2**	**L3**	**L4**	**IMF1**	**IMF2**	**IMF3**	**Sea Level**
L1	1.0000	0.8809	0.5559	0.1093	0.0869	0.8263	0.0362	0.8809
L2	0.8809	1.0000	0.8809	0.5559	0.0853	0.4960	−0.0116	0.5559
L3	0.5559	0.8809	1.0000	0.8809	0.0835	0.0399	−0.0443	0.1093
L4	0.1093	0.5559	0.8809	1.0000	0.0820	−0.4261	−0.0385	−0.3489
Sea Level	0.8809	0.5559	0.1093	−0.3489	0.0875	0.9466	0.0588	1.0000

Table 3. Sea level correlation with observed oceanic parameters.

	Water Temp.	Air Temp.	Pressure	Wind Direction	Wind Gust	Wind Speed	Sea Level
Water Temp.	1.0000	0.5471	−0.5596	0.1819	0.0483	−0.0053	0.0410
Air Temp.	0.5471	1.0000	−0.4722	0.2535	0.0599	0.0187	−0.0362
Pressure	−0.5596	−0.4722	1.0000	−0.2906	−0.1957	−0.1244	−0.0451
Wind Direction	0.1819	0.2535	−0.2906	1.0000	0.0533	0.0730	−0.0501
Wind Gust	0.0483	0.0599	−0.1957	0.0533	1.0000	0.7858	0.0176
Wind Speed	−0.0053	0.0187	−0.1244	0.0730	0.7858	1.0000	0.0094
Sea Level	0.0410	−0.0362	−0.0451	−0.0501	0.0176	0.0094	1.0000

2.3. Data Normalization

Every one of the model's input and target datum were normalized in order to make modelling easier. Equation (1) was used to scale the values to the [0,1] range:

$$x_n = \frac{x_{actual} - x_{min}}{x_{max} - x_{min}} \tag{1}$$

Equation (2) was used to return the values to their original scale following predictions produced using the trained model:

$$x_{actual} = x_n(x_{max} - x_{min}) + x_{min} \tag{2}$$

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

2.4. Double Data Decomposition by Multi-Variational and Successive Variational Mode Decomposition (MVMD-SVMD)

This study uses two data decomposition techniques of MVMD and SVMD for sea level signal decomposition. The steps involved in the MVMD are as follows [30]:

(i) The data matrix **X** with dimensions **N** × **T** is defined. **N** and **T** are the number of variables and time points, respectively.
(ii) The data matrix is normalized.
(iii) The objective function is defined as follows and is minimized to decompose the data into **U** and **V** modes:

$$\min_{\mathbf{U},\mathbf{V}} \sum_{n=1}^{N} \|\mathbf{X_n} - \mathbf{U_n}\mathbf{V}\|_2^2 + \lambda \sum_{i=1}^{N} \|\mathbf{U}_n\|_1 \tag{3}$$

where $\mathbf{X_n}$ represents the n-th channel of **X**, $\mathbf{U_n}$ is the mode of the n-th channel, **V** is the common mode across all channels, and **λ** is a regularization parameter that balances the sparsity of **U**.
(iv) The optimization problem is solved iteratively using the alternating direction method of the multiplier (ADMM).
(v) The modes are updated iteratively until the convergence criteria are met.
(vi) The $\mathbf{X_n}$ modes are reconstructed ($\mathbf{X_n} \approx \mathbf{U_n}\mathbf{V}$) for each channel n.

This also utilizes the SVMD method [31] with the decomposed modes of MVMD. The steps of SVMD are as follows:

(i) Original signal **X** and residual $\mathbf{R}_0 = \mathbf{X}$ are set up.

(ii) The residual $\mathbf{R_{k-1}}$ is decomposed into $\mathbf{U_k V}_k$ modes and updated in the next step. The decomposed modes are computed by the following function:

$$\mathbf{U_k V_k} = \underset{\mathbf{U,V}}{\mathbf{argmin}} \|\mathbf{R_{k-1}} - \mathbf{UV}\|_2^2 + \lambda_\mathbf{k} \|\mathbf{U_n}\|_1 \quad (4)$$

where $\lambda_\mathbf{k}$ is the regularization parameter for sparsity.

(iii) The residual $\mathbf{R_{k-1}}$ is updated for the next iteration, as follows:

$$\mathbf{R_k} = \mathbf{R_{k-1}} - \mathbf{U_k V_k} \quad (5)$$

(iv) The process of decomposition is iterated until the residual $\mathbf{R_k}$ attains convergence.

(v) After $\mathbf{K}$ iterations are completed, the original signal is reconstructed using the extracted modes, as follows:

$$\mathbf{X_n} \approx \sum_{\mathbf{k=1}}^{\mathbf{K}} \mathbf{U_k V_k} \quad (6)$$

Figures 3–5 show the graphs of decomposed IMFs after the decompositions were applied using the above algorithms in MATLAB software (version R2019b by The MathWorks, Inc located in Natick, MA, USA) [30,31].

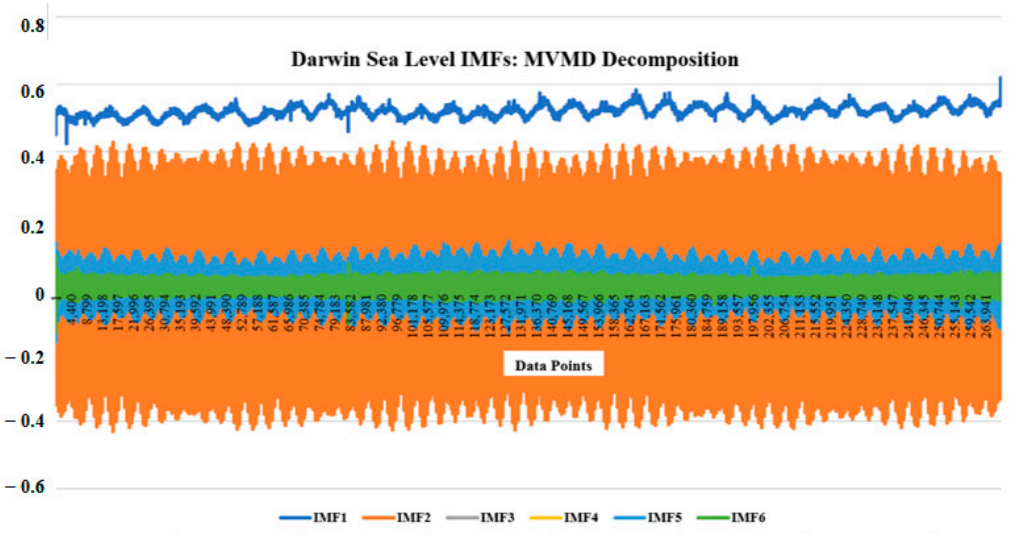

Figure 3. Decomposed sea level IMFs obtained from MVMD decomposition for the Darwin study site. The IMFs represent the local characteristics of the sea level signal and represent intrinsic modes of oscillation that reflect the sea level signal variability.

Each IMF was further decomposed by Successive Variational Mode Decomposition (SVMD), which was first presented in [31]. This allows for the sequential extraction of modes for more feature information from the MVMD decomposed IMFs. Thus, SVMD offers less computing overhead and is based on the same concepts as Variational Mode Decomposition (VMD). The input signal is subjected to consecutive applications of Variational Model Extraction (VME), which adds restrictions to stop convergence on formerly obtained modes. In the process to extract the intrinsic modes in the decomposition algorithm, the convergence on previous extracted modes is restricted to ensure that the algorithm continues to extract new modes until all relevant modes are identified [32]. To locate the signal containing the most compact spectrum, the optimization problem is handled while guaranteeing the reconstruction error is minimized via the Lth mode extracted sum. With 100 data points, the intrinsic mode decomposition (IMF) of the sea level (lag) in Darwin is shown in Figure 4 below. It should be noted that IMFs are derived from the original sea level signal via the decomposition process which focusses on the properties of the signal such as the zero mean, the number of extrema, and symmetric envelopes; hence, they are unitless [33].

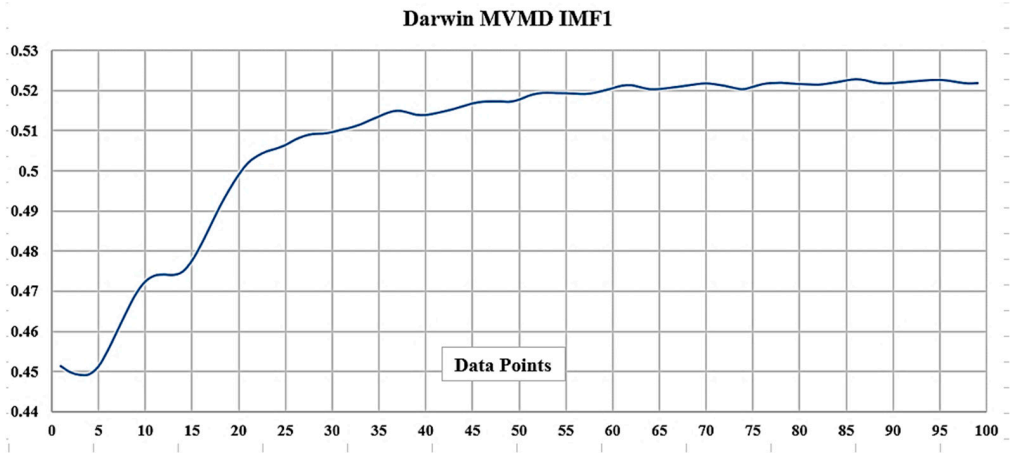

Figure 4. Darwin sea level IMF1, shown for 100 data points obtained from MVMD decomposition.

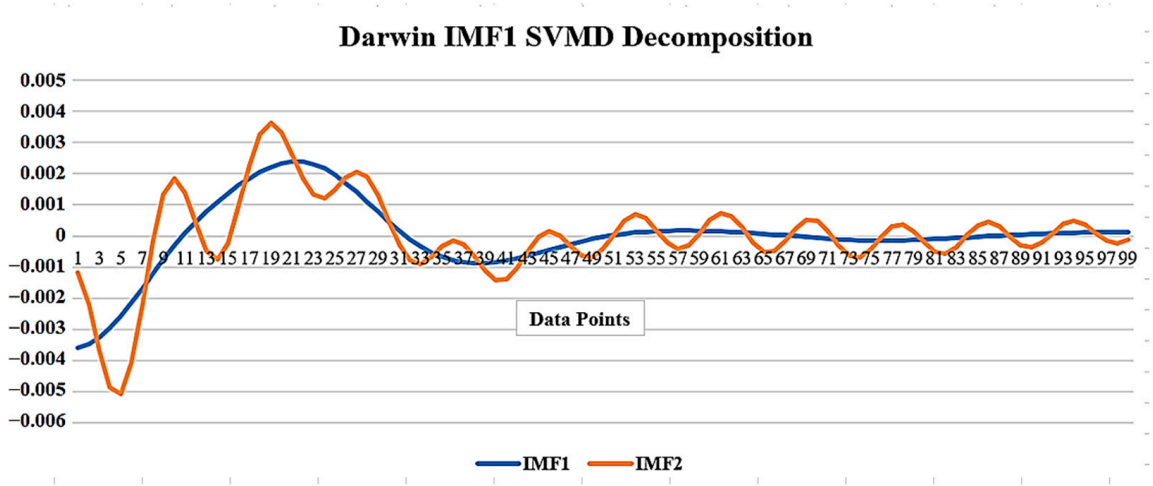

Figure 5. Darwin sea level IMF1, 100 data points shown for decomposition using SVMD. This is a snapshot of SVMD decomposed IMF 1, and 2 modes derived from MVMD IMF1.

Both MVMD and SVMD are an extension of Variational Mode Decomposition (VMD) [34], where MVMD decomposes the multivariate signals to extract both common and individual components simultaneously, and SVMD decomposes the signals into their constituent modes or components iteratively. Using double data decomposition provides many advantages. Firstly, it enhances the interpretability of the sea level signal, which contains different levels of components such as trends, seasonality, and residuals with irregular fluctuations, whereby single decomposition may not capture all levels of complexity in it. Secondly, each decomposition technique helps to reduce the noise and unwanted variations in the signal. Thirdly, two successive decomposition steps help to separate and isolate different features in the sea level signal more efficiently.

2.5. Input Feature Selection Using Principal Component Analysis (PCA)

PCA was used for data reduction in this study. The sea level signal was decomposed into six IMFs by the MVMD and then each IMF was further decomposed into two IMFs. This added 12 columns of input data, which could have added more computational load into the model. Hence, PCA was utilized to reduce the data into two columns of PCA1 and PCA2. This is critical to ensure optimization of the computational process. PCA is an unsupervised statistical-based method for dimensionality reduction [35]. This is performed by transforming correlated variables into a smaller set (principal components) that retain

the original information and are linear combinations of the original variable [36]. The summary of steps involved in the PCA analysis is as follows [37,38]:

(i) Each IMF of sea level data was standardized by subtracting the mean and divided by standard deviation.
(ii) Matrix A is formulated where each row corresponds to the IMF and each column to a sample point.
(iii) The covariance is performed on matrix A to obtain covariance matrix B.
(iv) Eigenvalue decomposition of matrix B is computed to obtain eigenvectors and eigenvalues of matrix B.
(v) The eigenvectors are sorted in descending order.
(vi) Matrix W is obtained from the selection of the top two eigenvectors.
(vii) The original matrix A of IMF data is projected on the selected components to obtain the dimensionality reduction in the IMFs, as follows:

$$\mathbf{Y} = \mathbf{AW} \tag{7}$$

where **Y** is the transformed data matrix, **A** is the standardized data matrix, and **W** is the matrix of selected principal components.

Figure 6 shows the PCA reduction in Darwin sea level IMFs. PCA identifies patterns and correlation in the sea level IMFs and compresses the data but preserves the important information for further analysis [39].

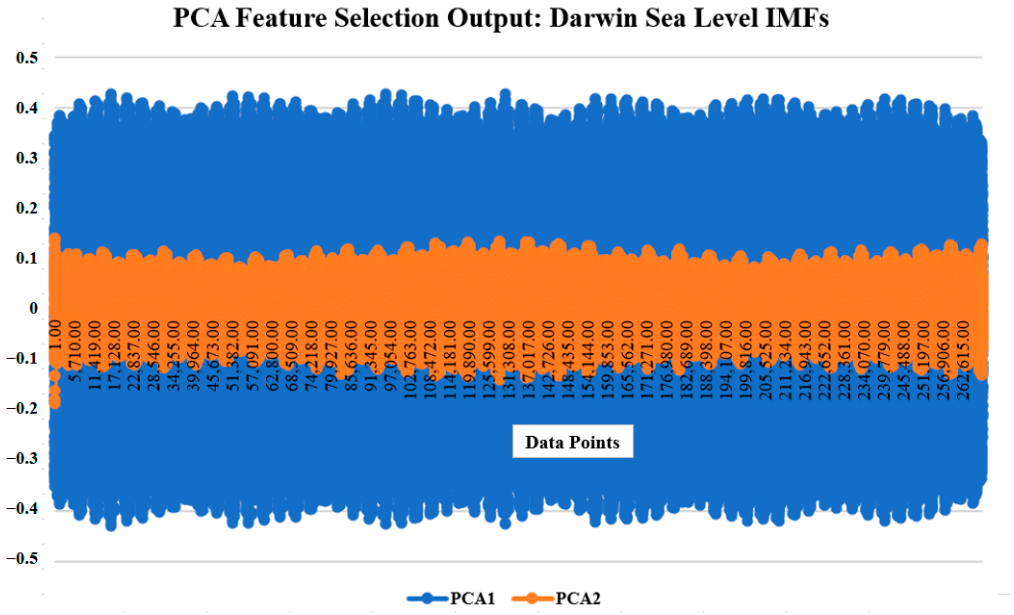

Figure 6. Darwin sea level IMFs feature selection using Principal Component Analysis (PCA).

2.6. Data Partition

Although there is no rigid guideline for data partitioning, Table 4 provides a detailed breakdown of how the sea level dataset was split which was based on the split ratio used in previous research [21,22,40,41].

Table 4. Data partition for the study sites broken down with the dates.

Partition	Training Set (60%)	Validation Set (20%)	Testing Set (20%)
Sea Level Dataset	January 1990–February 2011	March 2011–February 2017	March 2017–December 2022

2.7. Model Theoretical Background

This section provides the theoretical background for all models in this study, plus their optimized parameters in modelling and predictive model evaluation metrics. It outlines the statistical and mathematical functions used for model algorithms and formulae.

2.7.1. Convolutional Bidirectional Gated Recurrent Unit (CNN-BiGRU)

The CNN-BiGRU hybrid architecture in the study combines the layers of 1D CNN and two layers of bidirectional GRU for regression modelling. It allows the input dataset to pass through CNN and BiGRU, consequently utilizes the benefits of both platforms in processing the data for training the hybrid model. Hence, it leverages the strengths of both architectures to process sequential datasets. It obtains the superior ability of CNN for feature extraction and then allows the BiGRU to capture sequential dependencies. The parameter tuning is performed using grid search in Python (Python version 3.8.5 by Python Software Foundation located in Fremont, CA USA). 'GridSearchCV' with scikit-learn. The grid search includes hyperparameters for the hybrid model through the optimization of parameters that provide the best values such as weight regularization, batch, and dropout. Hence, a hybrid CNN-BiGRU model has the advantage of utilizing the capabilities of CNN and BiGRU's deep learning architecture.

CNN is a feedforward and backward neural network that can be considered to work in two parts: a feature extraction and a classification part [42]. CNN is different from a traditional neural network by the inclusion of multiple layers, pooling, local connections, and sharing of weights [43,44]. BiGRU is a deep learning architecture with two Gated Recurrent Units (GRUs). Thus, the CNN-BiGRU model can successfully capture temporal dependencies in the data and uses these two gating mechanisms to selectively update the hidden state of the network on each step, making it an effective tool for forecasting [45,46]. The GRU shown in Figure 7 includes a reset and update gate which have the following equations:

$$\text{Update gate}: \; z = \delta(W_z h_{t-1} + U_z x_t) \tag{8}$$

$$\text{Reset gate}: \; r = \delta(W_r h_{t-1} + U_r x_t) \tag{9}$$

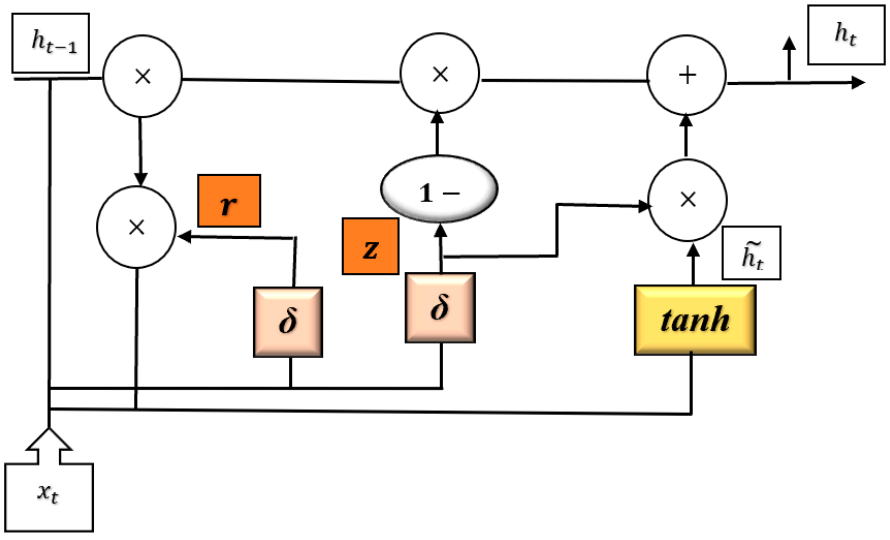

Figure 7. An overview of the structure of the GRU cell. The dark brown cells represent the gates and light brown cells are the sigmoid function.

The GRU takes in two inputs, a previous cell state (h_{t-1}) and training data input (x_t), resulting in a cell output (h_t). These are passed through the following components within the GRU unit [47]:

×—Elementwise multiplication;
+—Vector addition;
δ—Sigmoid function;
tanh—Hyper tangent function.

This study used BiGRU, which takes input in both forward and backward processes. This helps to increase its efficiency for capturing important information from the input dataset [48]. Figure 8 below shows the Python modelling using the CNN-BiGRU model after the MVMD-SVMD IMF feature extraction. Figure 9 shows the associated validation and training loss during the process.

```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv1d_5 (Conv1D)           (None, None, 64)          256

 max_pooling1d_5 (MaxPooling  (None, None, 64)         0
 1D)

 bidirectional_10 (Bidirecti  (None, None, 64)         18816
 onal)

 bidirectional_11 (Bidirecti  (None, 64)               18816
 onal)

 dense_5 (Dense)             (None, 1)                 65

=================================================================
Total params: 37,953
```

Figure 8. CNN-BiGRU modelling in Python showing the arrangement of layers, output shape, and parameters.

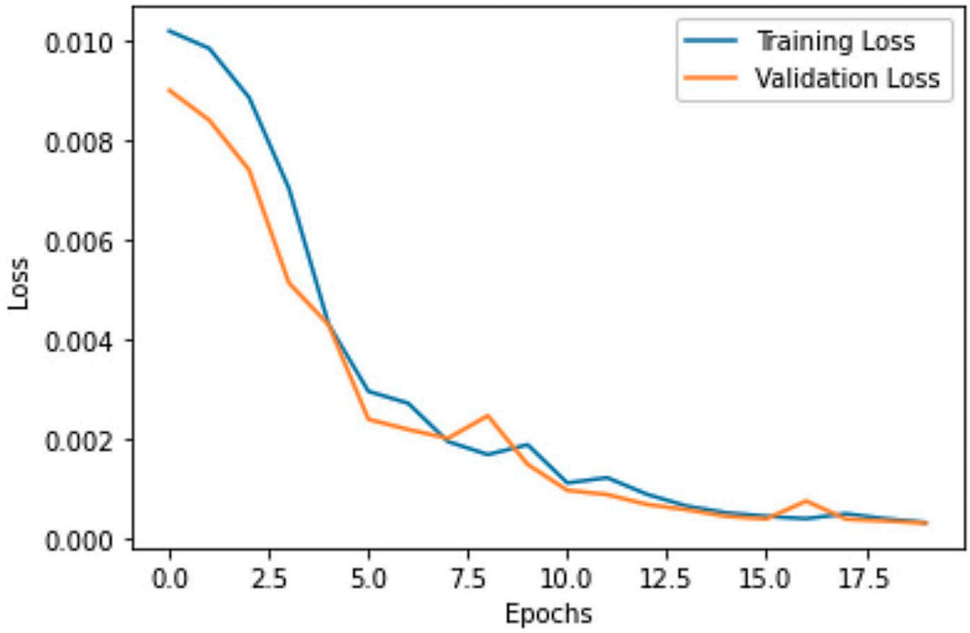

Figure 9. CNN-BiGRU training and validation loss.

2.7.2. Multilayer Perceptron (MLP)

Multilayer Perceptron is a feedforward and backward artificial neural network of fully connected neurons which consists of nonlinear activation functions [49]. Generally, it contains an input layer, one or more hidden layers, and an output layer [50], as shown below in Figure 10.

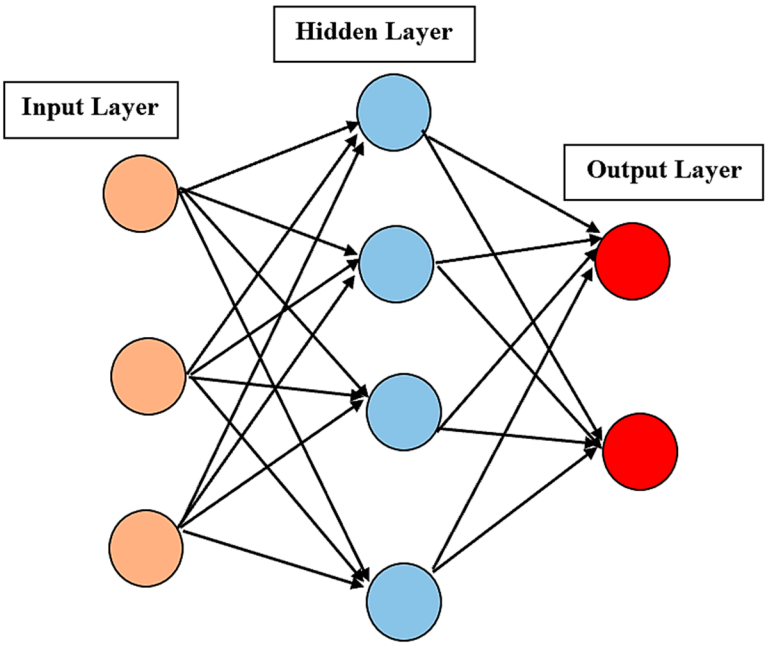

Figure 10. General structure of MLP network architecture. The brown circles show input layer, light blue circles are hidden layer and red circle show the output layer.

Each layer in MLP has nodes that connects to every other node in the subsequent layer. The input layer distributes the input data and the nodes have linear activation functions with no thresholds [51]. The hidden layer has hidden unit nodes which have thresholds with weights and nonlinear activation functions. The MLP is fed forward and backward in its information processing. Finally, the information is passed to the output layer nodes, which have linear activation functions to provide the data output. The MLP model utilized in this study has two hidden layers with 20 and 10 neurons, respectively, with an initial rate optimizer of 0.001. The hidden layer uses Rectified Linear Unit (ReLu) as the activation function and Adaptive Moment Estimation (ADAM) algorithm to minimize the loss function.

2.7.3. Adaptive Boosting Regressor (AdaBoost)

Boosting is based on the general problem of producing an accurate prediction rule by combining moderately inaccurate rules-of-thumb [52]. First introduced by Freund and Schapie [52], the adaptive boosting algorithm is used to serve as a meta-estimator in the structure of the AdaBoost Regressor. By modifying the sample weights (w_i) of items in accordance with predictor errors, this approach incrementally improves the efficiency of a regressor that has been fitted to a dataset (X, y). Figure 11 shows an ensemble ($i = 1 \ldots I$) of four weak learners (f_i) that are trained sequentially on the data (X_i, y_i), which are sampled from (X, y) with replacement [53]. The confidence measures β_i are computed where a low value implies high confidence in the prediction.

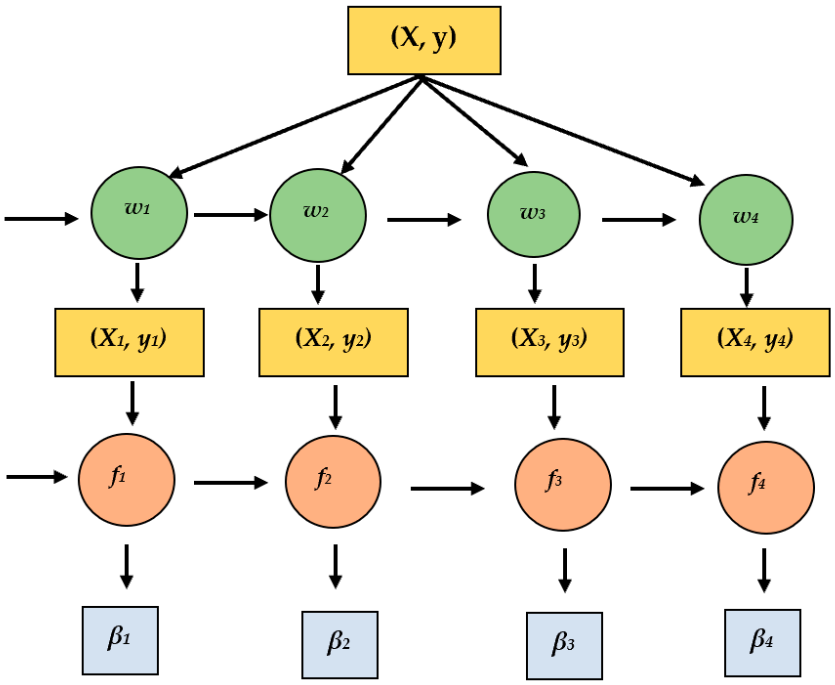

Figure 11. An example of AdaBoost network architecture with four weak learners.

2.7.4. Support Vector Regression (SVR)

The Support Vector Regression (SVR) is based on the Support Vector (SV) algorithm developed by Vapnik and associates [54]. SVR, a further variation of support vector machines (SVMs), was effectively used for time series in [55]. The SVR was presented in 1997 by Vapnik, Steven Golowich, and Alex Smola [56] and uses the same principles of SVM. This technique efficiently learns a variable's importance for characterizing the input and output relationship [57]. SVR is a supervised learning algorithm and trains data using a symmetrical loss function whereby high and low estimates are equally penalized [58]. The algorithm attempts to find the optimum solution by minimizing the prediction error. Formally, this can be presented with the equations below [59] and Figure 12 shows the arrangement of error and hyperplane.

$$\text{Subject to}: \underset{w,b,\xi,\xi^*}{min} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi + \xi^*)$$

$$\left\{ \begin{array}{l} w\phi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \\ y_i - w\phi(x_i) - b \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, i = 1, ...n. \end{array} \right\} \quad (10)$$

where C = penalty parameter, ε = an insensitive loss function, ξ_i, ξ_i^* = slack variables, and $\|w\|^2$ = Euclidean norm.

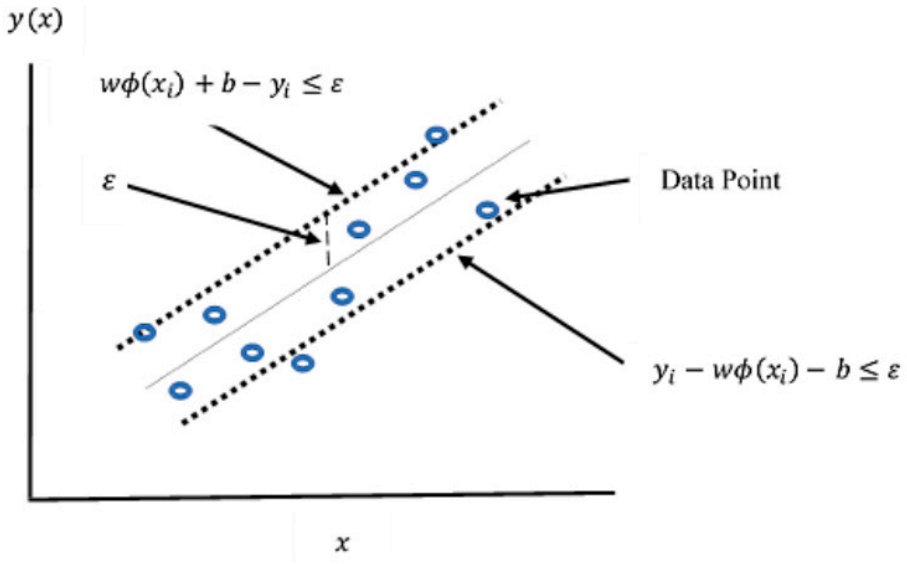

Figure 12. The SVR principle showing the arrangement of error around the hyperplane for the optimization algorithm. The blue circles denote data points, dotted line represents the band around the regression line for the margin of deviation.

2.8. Summary of Selected Model Parameters in Python

This section provides the values of optimized parameters used in Python. The values were obtained using GridSearchCV from the scikit-learn library. This approach has been effectively used in many past studies to obtain the best parameters for predictive models [60–62]. It helps to provide the best combinations of hyperparameters for machine learning models [63,64].

(i) CNN-BiGRU Modelling Parameters

CNN layer: filters = 64, kernel size = 3, activation function = 'relu' (Rectified linear Unit), loss = 'mean_squared_error', optimizer = 'adam', dropout = 0.1, early stopping: monitor = 'val_loss', patience = 3, restore_best_weights = True, kernel regularizer = l2(0.01).

BiGRU layer: GRU units = 32, activation function = 'relu' (Rectified linear Unit), loss = 'mean_squared_error', optimizer = 'adam', dropout = 0.1, early stopping: mode = minimum, patience = 20, kernel regularizer = l2(0.01).

(ii) MLP Modelling Parameters

hidden_layer_sizes = (20, 10), learning_rate_init = 0.001, max_iter = 5000, activation = 'relu'(Rectified linear Unit), solver = 'adam', dropout = 0.1, loss = 'mean_squared_error', monitor = 'val_loss', patience = 5, restore_best_weights = True.

(iii) AdaBoost Modelling Parameters

DecisionTreeRegressor(max_depth = 3), n_estimators = 50, random_state = 42.

(iv) SVR Model Parameters

n_samples = 1000, n_features = 20, noise = 0.1, random_state = 42), kernel = 'rbf', C = 1.0, epsilon = 0.1, gamma = 'scale'.

2.9. Performance Evaluation Metrics for AI Models

Eight statistical measures that are obtained from Equations (10)–(17) on the predicted outcomes are used in evaluating the models' performance, and tests are run on the test dataset. Table 5 shows the data quantities used in the computations.

Table 5. Symbols used in Equations (11)–(18).

Symbol	Quantity
DO_i	Observed Data
MDO	Mean Observed Data
DS_i	Simulated Data
MDS	Mean Simulated Data

1. Pearson's correlation coefficient (r):

$$r = \left[\frac{\sum_{i=1}^{n} (DO_i - MDO)(DS_i - MDS)}{\sqrt{\sum_{i=1}^{n} (DO_i - MDO)^2 \sum_{i=1}^{n} (DS_i - MDS)^2}} \right]^2 \quad (11)$$

2. Willmott's index of agreement (d):

$$d = 1 - \left[\frac{\sum_{i=1}^{n} (DO_i - DS_i)^2}{\sum_{i=1}^{n} (|DS_i - MDO| + |DO_i - MDS|)^2} \right] \quad (12)$$

3. Nash–Sutcliffe coefficient (NS):

$$NS = 1 - \left[\frac{\sum_{i=1}^{n} (DO_i - DS_i)^2}{\sum_{i=1}^{n} (DO_i - MDO)^2} \right], \ -\infty \leq NS \leq 1 \quad (13)$$

4. Legates and McCabe's index (LM):

$$LM = 1 - \left[\frac{\sum_{i=1}^{n} |DS_i - DO_i|}{\sum_{i=1}^{n} |DO_i - MDS|} \right], \ 0 \leq L \leq 1 \quad (14)$$

5. Root-mean-squared error (RMSE):

$$RMSE = \sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^{n} (DS_i - DO_i)^2} \quad (15)$$

6. Mean absolute error (MABE):

$$MABE = \frac{1}{n} \sum_{i=1}^{n} |(DS_i - DO_i)| \quad (16)$$

7. Relative root-mean-squared error (RRMSE):

$$RRMSE = \frac{\sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^{n} (DS_i - DO_i)^2}}{\frac{1}{n} \sum_{i=1}^{n} DO_i} \times 100 \quad (17)$$

8. Mean absolute percentage error (MAPE):

$$MAPE = \frac{1}{\mathrm{N}} \left(\sum_{N}^{\mathrm{i}=1} \left| \frac{(\mathrm{SWL}_{\mathrm{for}} - SWL_{\mathrm{obs}})}{SWL_{obs}} \right| \right) \times 100 \quad (18)$$

3. Results

3.1. Efficiency Performance Metrics

The results of this study include the computation of evaluation metrics of the predicted SLR with observed data for the models (SVR, AdaBoost, MLP, CNN-BiGRU). The

benchmark models applied in this study help to provide a comparative analysis for developing a robust and objective model. Such metrics also form an important aspect of machine learning that utilize datasets for prediction [65]. Here, our analysis helps to show whether combining the benefits of two different ML platforms can improve performance in target variable prediction. This study uses four performance evaluation metrics, including Pearson's correlation coefficient, Willmott's index of agreement, Nash–Sutcliffe coefficient and Legates and McCabe's index which have been successfully used in many past studies [21,22,41,65].

Pearson's correlation coefficient is used in [66] as a measure of the predictive accuracy of models in genome-wide prediction. It measures the strength of association between observed and predicted values [67,68]. Willmott's index of agreement [69] is a statistical measure which measures the degree of association between observed and predicted values of a machine learning model. The index reflects the bias and variability between the observed and predicted values [70]. It has been used in the evaluation of sea level prediction previously [41], ground water salinization risk prediction [71], and significant wave height prediction [40]. The Nash–Sutcliffe coefficient [72] is a widely used evaluation metric in hydrology [73,74]. It determines the relative magnitude of the residual variance compared to measured data variance [75]. Legates and McCabe's index [76] is an evaluation metric which also assesses the goodness of fit between model predictions and observed values. It overcomes some of the limitations of other evaluation metrics that measure the correlation by taking into account both the additive and proportional differences between the observed and predicted values of a machine learning model [77]. Hence, these metrics were computed for all models in this study for both the Milner Bay and Darwin study sites. Tables 6 and 7 below show the model performance results. Of the four machine learning models used to predict SLR, MVMD-SVMD-CNN-BiGRU (shown in the bottom row of each table) performed the best.

Table 6. Milner Bay efficiency performance metrics.

Model	Correlation Coefficient (r)	Willmott's Index of Agreement (d)	Nash–Sutcliffe' Coefficient (NS)	Legates and McCabes' Index (L)
MVMD-SVMD-SVR	0.9406	0.9362	0.8800	0.6509
MVMD-SVMD-AdaBoost	0.9571	0.9439	0.9074	0.6891
MVMD-SVMD-MLP	0.9974	0.9970	0.9943	0.9280
MVMD-SVMD-CNN-BiGRU	0.9980	0.9979	0.9960	0.9409

Table 7. Darwin efficiency performance metrics.

Model	Correlation Coefficient (r)	Willmott's Index of Agreement (d)	Nash–Sutcliffe' Coefficient (NS)	Legates and McCabes' Index (L)
MVMD-SVMD-SVR	0.9878	0.9856	0.9687	0.8263
MVMD-SVMD-AdaBoost	0.9881	0.9869	0.9730	0.8453
MVMD-SVMD-MLP	0.9963	0.9952	0.9900	0.9046
MVMD-SVMD-CNN-BiGRU	0.9984	0.9980	0.9959	0.9413

3.2. Error Metrics

Four error metrics (RMSE, MABE, RRMSE, MAPE) were utilized in the evaluation of the machine learning model variance in this study. RMSE and MABE metrics are widely used for predictive model evaluation [77]. RMSE is a measure of the average variance between the predicted and observed values [78]. MAE measures the average magnitude of errors between predicted and observed values [79]. RRMSE provides a standard deviation of the residuals and normalizes the RMSE by the mean [80], while MAPE provides a measure of the mean absolute percentage deviation [78]. Tables 8 and 9 show the error metrics for all models. Again, the variance between the predicted and modelled SLR was the lowest when using the MVMD-SVMD-CNN-BiGRU model. This was shown to be the case for both Darwin and Milner Bay.

Table 8. Milner Bay error metrics.

Model	RMSE	MABE	RRMSE	MAPE
MVMD-SVMD-SVR	0.1362	0.1120	10.9161	10.7201
MVMD-SVMD-AdaBoost	0.1197	0.0997	9.5904	9.5417
MVMD-SVMD-MLP	0.0297	0.0231	2.3814	2.1551
MVMD-SVMD-CNN-BiGRU	0.0248	0.0189	1.9901	1.7486

Table 9. Darwin error metrics.

Model	RMSE	MABE	RRMSE	MAPE
MVMD-SVMD-SVR	0.2813	0.2312	6.5577	6.8155
MVMD-SVMD-AdaBoost	0.2613	0.2058	6.0926	5.9953
MVMD-SVMD-MLP	0.1589	0.1269	3.7050	3.7037
MVMD-SVMD-CNN-BiGRU	0.1016	0.0782	2.3699	2.4123

4. Discussion

4.1. Result Evaluation

Tables 6 and 7 show the efficiency performance metrics of Darwin and Milner Bay. Pearson's correlation coefficient (r) shows the degree of association of predicted and observed variables [81] of sea level at both study sites. The MVMD-SVMD-CNN-BiGRU hybrid deep learning model shows the highest r values of 0.9980 and 0.9984 for Milner Bay and Darwin, respectively. MVMD-SVMD-MLP also achieves high r values of 0.9974 and 0.9963 for Milner Bay and Darwin, respectively. MVMD-SVMD-AdaBoost and MVMD-SVMD-SVR recorded values greater than 0.9 and performed well with data decomposition inputs. To further support the r value performance metric, this study also utilized Willmott's index of agreement (d), Nash–Sutcliffe' coefficient (NS), and Legates and McCabes' index (L). These have been widely used in many past studies [21,22,65] alongside correlation coefficient to effectively compare AI model performance. All results show consistent metrics, for both study sites. The MVMD-SVMD-CNN-BiGRU hybrid model introduced in this study shows the best performance with values of 0.9979 (d), 0.996 (NS), and 0.9409 (L); and 0.998 (d), 0.9959 (NS), and 0.9413 (L) for Milner Bay and Darwin, respectively. Legates and McCabes' index shows the most difference between MVMD-SVMD-SVR and MVMD-SVMD-CNN-BiGRU for both study sites. This index also confirms the superior performance of MVMD-SVMD-CNN-BiGRU for SLR prediction. Such error metric computations are an essential aspect of AI model performance evaluation [82] and each indicates the overall robust performance of MVMD-SVMD-CNN-BiGRU.

Root-mean-squared error (RMSE), mean absolute error (MABE), mean absolute percentage error (MAPE), and relative root-mean-squared error (RRMSE) were also calculated for each model in this study. MVMD-SVMD-CNN-BiGRU attained the lowest error values of 0.1016 (RMSE), 0.0782 (MABE), 2.3699 (RRMSE), and 2.4123 (MAPE) for Darwin and 0.0248 (RMSE), 0.0189 (MABE), 1.9901 (RRMSE), and 1.7486 (MAPE) for Milner Bay. Figures 13 and 14 present scatterplots of all models between the observed and predicted datasets in the testing phase. MVMD-SVMD-MLP and MVMD-SVMD-CNN-BiGRU show a close scatter of points, indicating a higher accuracy of prediction. This is also confirmed by the histogram of absolute prediction error where the lower bracket bins in both models have higher values than MVMD-SVMD-SVR and MVMD-SVMD-AdaBoost (Figure 15). Figure 16 shows the time series comparison of 100 data points with similar patterns for the predicted and observed SLR dataset.

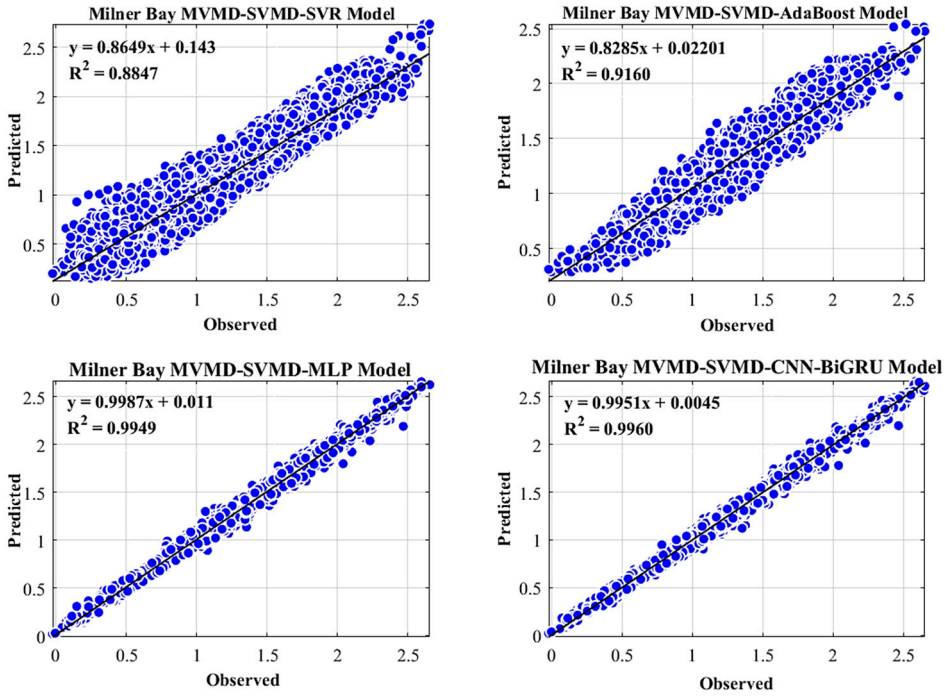

Figure 13. Scatterplot of Milner Bay for all AI models with correlation.

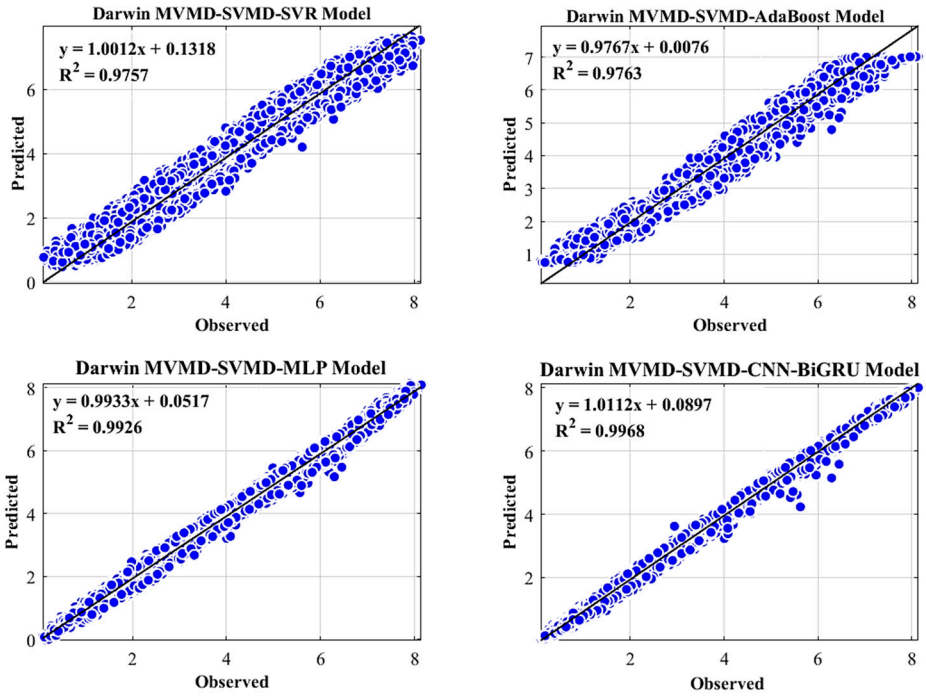

Figure 14. Scatterplot of Darwin for all AI models with correlation.

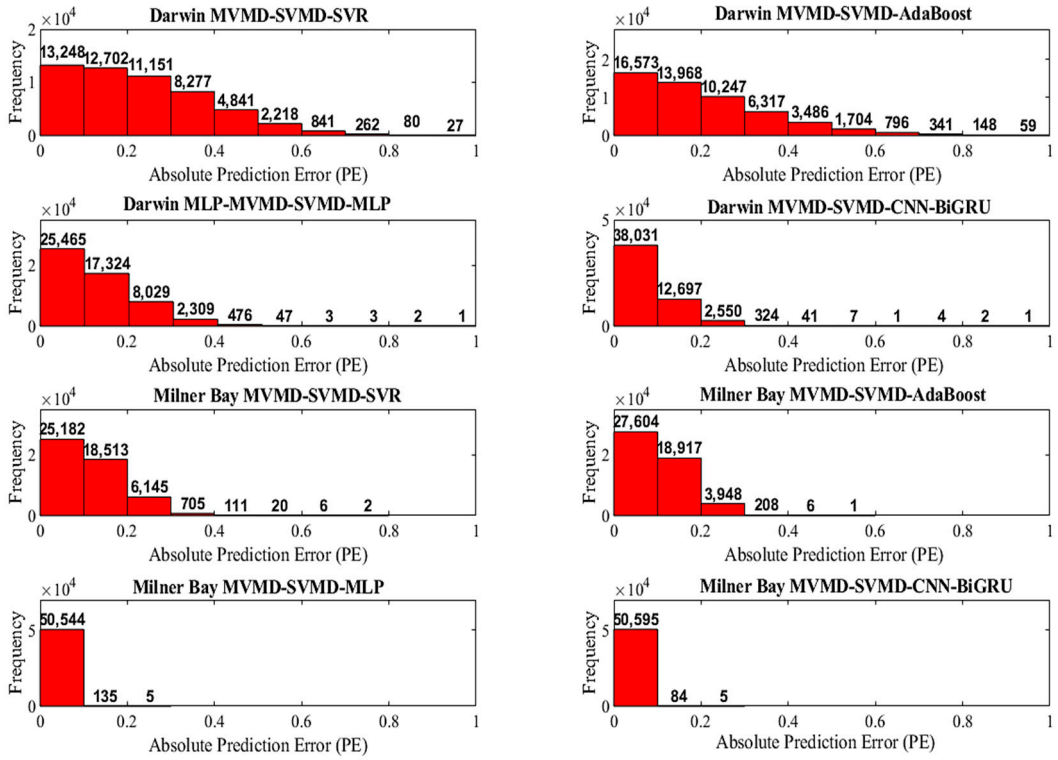

Figure 15. Absolute prediction error histogram of all models for both study sites.

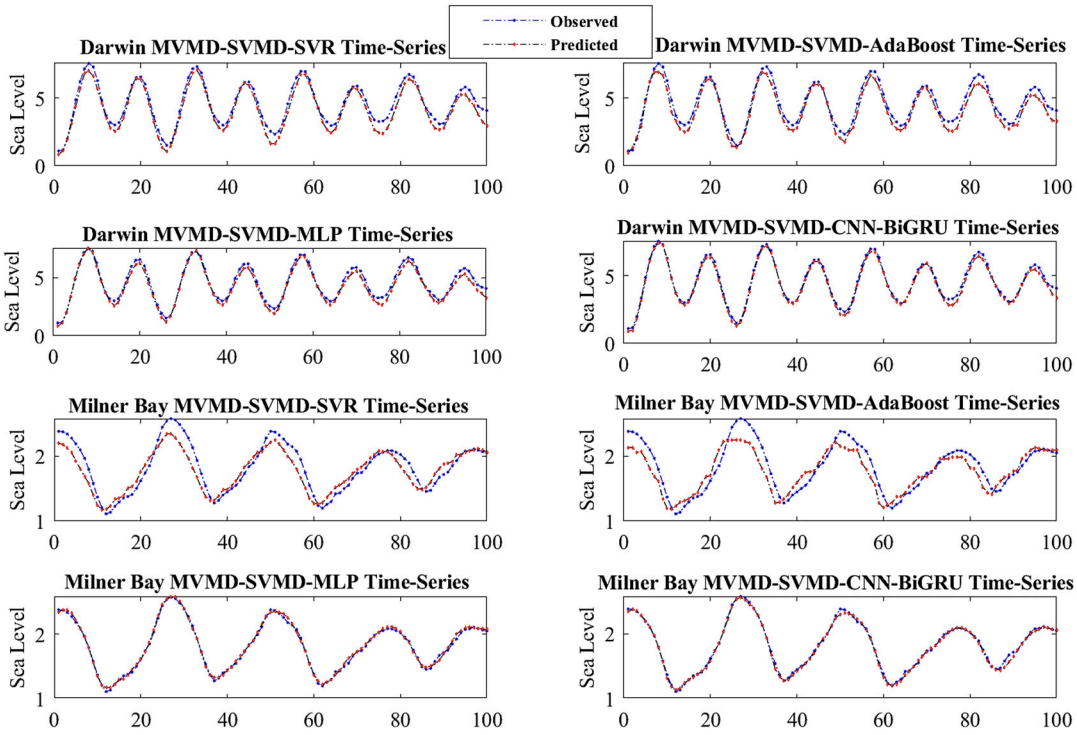

Figure 16. Time series comparison for 100 data points from the testing phase of all models for both study sites.

4.2. Mean Sea Level Trend Evaluation with MVMD-SVMD-CNN-BiGRU Model Mean Predictions

Figures 17 and 18 show mean sea level trend for Darwin and Milner Bay as 6.1 ± 1.1 mm and 5.6 ± 1.5 mm, respectively. These values, including uncertainties, coincide with the

given range of 4–6 mm/yr in [83], which is notably higher than the global average as reported. With the linear projection of the trend, the mean sea level of Darwin and Milner Bay are predicted to be about 4.37 m and 1.33 m, respectively, in the year 2032.

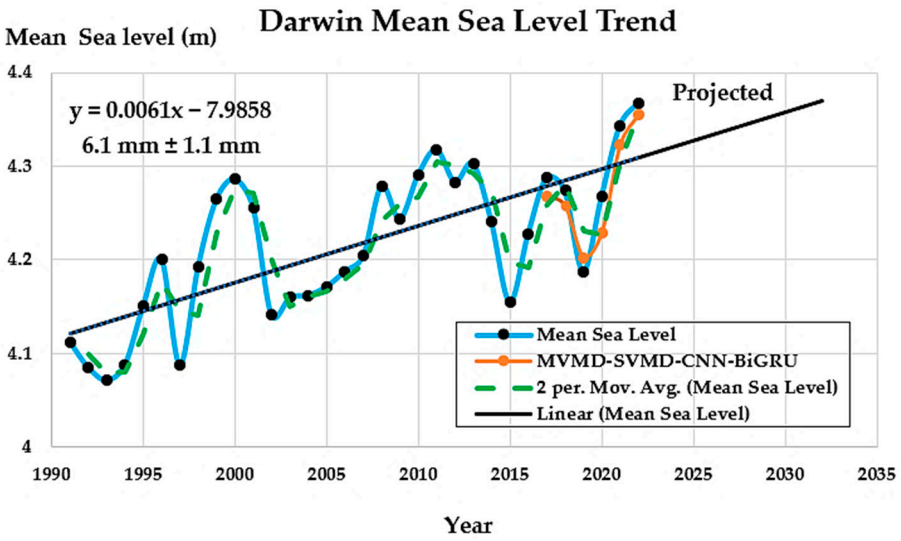

Figure 17. Mean sea level trend for Darwin with projection and MVMD-SVMD-CNN-BiGRU mean sea level prediction comparison.

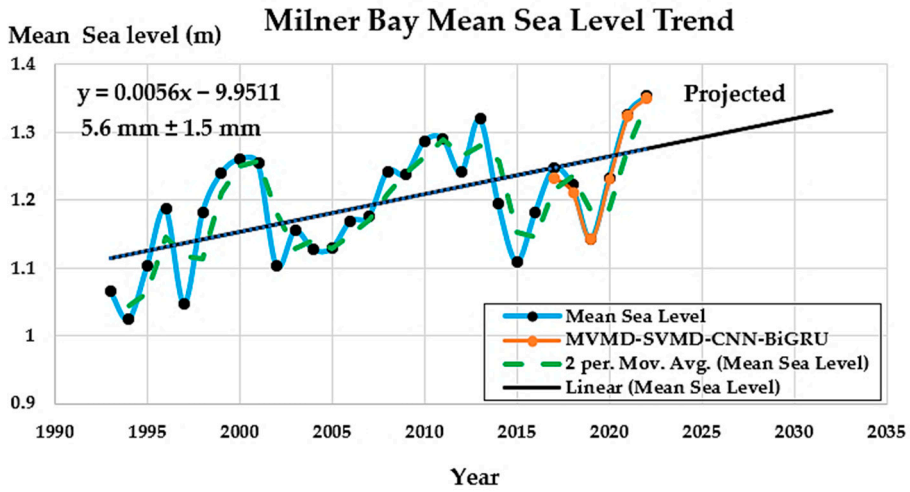

Figure 18. Mean sea level trend for Milner Bay with projection and MVMD-SVMD-CNN-BiGRU mean sea level prediction comparison.

4.3. Limitations, Challenges, and Model Interpretability

While machine learning models have evolved and expanded to efficiently train data for prediction, there are associated limitations and challenges that need to be considered [84–86]. Firstly, an important aspect of sea level prediction is the availability of sufficient data for AI modelling. The site should also be reliable and state the accuracy of data with minimum missing values. Secondly, the prediction of sea level also needs to include information on other associated oceanic and climate parameters for the same timeline. AI models rely on input data information to extract underlying features and relationships with the target variable for prediction. AI model interpretability in the early years focused on approaches to visualize the internal representations of predictive models [87]. Ribeiro et al. [88] proposed a model agnostic method known as the Local Interpretable Model-agnostic Explanation (LIME) by fitting interpretable local surrogate models around individual instances of the feature dataset [87]. Such approaches that highlight and explain

the importance of the parameters used in data-driven AI models improve the overall interpretation of the results and model performance, as shown below.

Figure 19 shows how positive and negative feature weights are assigned for a single instance in the CNN-BiGRU prediction model. In this case, PCA1, Lag 5, and wind direction display strong feature weightings. While LIME does provide information on feature weights in the model, it uses linear regression and may not capture the nonlinearities in the model [89]. Furthermore, the LIME explanation is on individual prediction rather than the global behaviour [89]. This area of AI interpretability is an emerging area of research that promises to provide more techniques that can give a more granular insight into the use of predictive AI models [87,90].

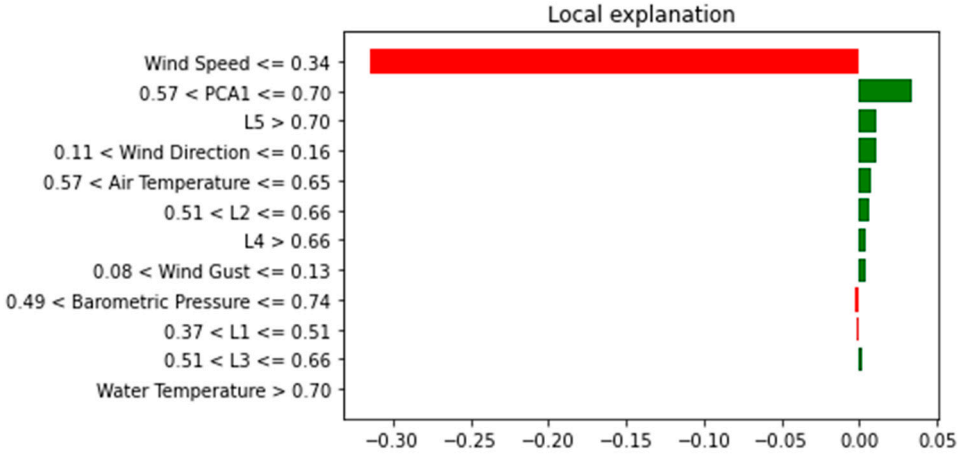

Figure 19. LIME evaluation of CNN-BiGRU prediction model showing feature weights for an instance of the Darwin prediction test dataset.

5. Conclusions

This study has successfully shown the importance of sea level data decomposition and feature extraction approach for artificial intelligence modelling. Since the new deep learning models are highly capable of handling more data inputs, the extraction of intrinsic features through IMFs provides more information on data inputs for prediction. Furthermore, the use of PCA for IMF input selection helps with dimensionality reduction for data modelling. This study shows that double data decomposition of sea level signal can achieve high prediction accuracy of AI models. Both study sites of Darwin and Milner Bay show an increase in mean sea level. While there are some studies for the Northern Territory, very few have quantified the impacts of sea level rise and trend. Given the rate of SLR and the cascading impacts associated with coastal erosion and inundation, there is an urgent need for a better understanding and accurate predictive ability of coastal and wetland changes. In this study, a method for achieving this is presented that aims to assist relevant stakeholders to make informed adaptation decisions for the region. The estimated SLR values of 6.1 $\pm$ 1.1 mm and 5.6 $\pm$ 1.5 mm for Darwin and Milner Bay are significant, indicating the need for appropriate adaptation strategies, particularly in consideration of the prevalence of tropical cyclones and storm surges in this area. The AI models were evaluated using LIME to obtain an insight into the model black box. Challenges and limitations to sea level study such as data availability and missing values are highlighted.

Author Contributions: Conceptualization, N.R.; methodology, N.R.; software, N.R.; validation, N.R.; formal analysis, N.R.; investigation, N.R., J.M.; resources, N.R., J.M.; data curation, J.M.; writing—original draft preparation, N.R., J.M.; writing—review and editing, N.R., L.S.-P., N.D.; visualization, N.R.; supervision, N.R.; project administration, N.R.; funding acquisition, N.R., J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received funding from University of Southern Queensland (UniSQ) 2023 Undergraduate Research Project Grant.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: This paper acknowledges the Australian Bureau of Meteorology Data Portal site from where the historical wave dataset for Northern Territory was obtained.

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

References

1. Church, J.A.; Clark, P.U.; Cazenave, A.; Gregory, J.M.; Jevrejeva, S.; Levermann, A.; Merrifield, M.A.; Milne, G.A.; Nerem, R.S.; Nunn, P.D. *Sea Level Change*; Cambridge University Press: Cambridge, UK, 2013.
2. Allan, R.P.; Hawkins, E.; Bellouin, N.; Collins, B. *IPCC, 2021: Summary for Policymakers*; IPCC: Geneva, Switzerland, 2021.
3. Masson-Delmotte, V.; Zhai, P.; Pörtner, H.-O.; Roberts, D.; Skea, J.; Shukla, P.R. *Global Warming of 1.5 C: IPCC Special Report on Impacts of Global Warming of 1.5 °C above Pre-Industrial Levels in Context of Strengthening Response to Climate Change, Sustainable Development, and Efforts to Eradicate Poverty*; Cambridge University Press: Cambridge, UK, 2022.
4. Pörtner, H.-O.; Roberts, D.C.; Masson-Delmotte, V.; Zhai, P.; Tignor, M.; Poloczanska, E.; Weyer, N. The ocean and cryosphere in a changing climate. In *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*; Cambridge University Press: Cambridge, UK, 2019; Volume 1155.
5. Schneider, D. The rising seas. *Sci. Am.* **1997**, *276*, 112–117. [CrossRef]
6. Laffoley, D.; Baxter, J.M. *Ocean Deoxygenation: Everyone's Problem: Causes, Impacts, Consequences and Solutions: Summary for Policy Makers*; International Union for Conservation of Nature (IUCN): Gland, Switzerland, 2019.
7. Sweet, W.V.; Kopp, R.E.; Weaver, C.P.; Obeysekera, J.; Horton, R.M.; Thieler, E.R.; Zervas, C. *Global and Regional Sea Level Rise Scenarios for the United States*; NASA: Washington, DC, USA, 2017.
8. Williams, S.J. Sea-level rise implications for coastal regions. *J. Coast. Res.* **2013**, *63*, 184–196. [CrossRef]
9. Neumann, J.E.; Price, J.; Chinowsky, P.; Wright, L.; Ludwig, L.; Streeter, R.; Jones, R.; Smith, J.B.; Perkins, W.; Jantarasami, L. Climate change risks to US infrastructure: Impacts on roads, bridges, coastal development, and urban drainage. *Clim. Chang.* **2015**, *131*, 97–109. [CrossRef]
10. Werner, A.D.; Bakker, M.; Post, V.E.; Vandenbohede, A.; Lu, C.; Ataie-Ashtiani, B.; Simmons, C.T.; Barry, D.A. Seawater intrusion processes, investigation and management: Recent advances and future challenges. *Adv. Water Resour.* **2013**, *51*, 3–26. [CrossRef]
11. Hennessy, K.; Page, C.; McInnes, K.; Walsh, K.; Pittock, B.; Bathols, J.; Suppiah, R. Climate change in the Northern Territory. *Consult. Rep. North. Territ. Dep. Infrastruct. Plan. Environ.* **2004**, *64*, 1–64.
12. Zander, K.K.; Petheram, L.; Garnett, S.T. Stay or leave? Potential climate change adaptation strategies among Aboriginal people in coastal communities in northern Australia. *Nat. Hazards* **2013**, *67*, 591–609. [CrossRef]
13. Miloshis, M.; Valentine, E. Sea Level rise and potential mitigation of saline intrusion in Northern Australia. In *2013 IAHR Congress*; Tsinghua University Press: Beijing, China, 2013; pp. 9158–9167.
14. Tsietso, D.; Yahya, A.; Samikannu, R.; Tariq, M.U.; Babar, M.; Qureshi, B.; Koubaa, A. Multi-Input deep learning approach for breast cancer screening using thermal infrared imaging and clinical data. *IEEE Access* **2023**, *11*, 52101–52116. [CrossRef]
15. Nosratabadi, S.; Mosavi, A.; Duan, P.; Ghamisi, P.; Filip, F.; Band, S.S.; Reuter, U.; Gama, J.; Gandomi, A.H. Data science in economics: Comprehensive review of advanced machine learning and deep learning methods. *Mathematics* **2020**, *8*, 1799. [CrossRef]
16. Gavrishchaka, V.; Yang, Z.; Miao, R.; Senyukova, O. Advantages of hybrid deep learning frameworks in applications with limited data. *Int. J. Mach. Learn. Comput.* **2018**, *8*, 549–558.
17. Braakmann-Folgmann, A.; Roscher, R.; Wenzel, S.; Uebbing, B.; Kusche, J. Sea level anomaly prediction using recurrent neural networks. *arXiv* **2017**, arXiv:1710.07099.
18. Gudelek, M.U.; Boluk, S.A.; Ozbayoglu, A.M. A deep learning based stock trading model with 2-D CNN trend detection. In Proceedings of the 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, HI, USA, 27 November–1 December 2017.
19. Salau, A.O.; Jain, S. Feature extraction: A survey of the types, techniques, applications. In Proceedings of the 2019 International Conference on Signal Processing and Communication (ICSC), Noida, India, 7–9 March 2019.
20. Luo, W.; Yu, Z.-Y.; Xiao, S.-J.; Zhu, A.-X.; Yuan, L.-W. Exploratory method for spatio-temporal feature extraction and clustering: An integrated multi-scale framework. *ISPRS Int. J. Geo-Inf.* **2015**, *4*, 1870–1893. [CrossRef]
21. Raj, N. Prediction of sea level with vertical land movement correction using deep learning. *Mathematics* **2022**, *10*, 4533. [CrossRef]
22. Raj, N.; Gharineiat, Z.; Ahmed, A.A.M.; Stepanyants, Y. Assessment and prediction of sea level trend in the South Pacific Region. *Remote Sens.* **2022**, *14*, 986. [CrossRef]
23. Balogun, A.-L.; Adebisi, N. Sea level prediction using ARIMA, SVR and LSTM neural network: Assessing the impact of ensemble Ocean-Atmospheric processes on models' accuracy. *Geomat. Nat. Hazards Risk* **2021**, *12*, 653–674. [CrossRef]

24. Nieves, V.; Radin, C.; Camps-Valls, G. Predicting regional coastal sea level changes with machine learning. *Sci. Rep.* **2021**, *11*, 7650. [CrossRef] [PubMed]
25. Sithara, S.; Pramada, S.; Thampi, S.G. Sea level prediction using climatic variables: A comparative study of SVM and hybrid wavelet SVM approaches. *Acta Geophys.* **2020**, *68*, 1779–1790. [CrossRef]
26. Hsieh, C.-M.; Chou, D.; Hsu, T.-W. Using modified harmonic analysis to estimate the trend of sea-level rise around Taiwan. *Sustainability* **2022**, *14*, 7291. [CrossRef]
27. Antunes, C.; Taborda, R. Sea level at Cascais tide gauge: Data, analysis and results. *J. Coast. Res.* **2009**, *I*, 218–222.
28. García, S.; Luengo, J.; Herrera, F. *Data Preprocessing in Data Mining*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 72.
29. Alasadi, S.A.; Bhaya, W.S. Review of data preprocessing techniques in data mining. *J. Eng. Appl. Sci.* **2017**, *12*, 4102–4107.
30. ur Rehman, N.; Aftab, H. Multivariate variational mode decomposition. *IEEE Trans. Signal Process.* **2019**, *67*, 6039–6052. [CrossRef]
31. Nazari, M.; Sakhaei, S.M. Successive variational mode decomposition. *Signal Process.* **2020**, *174*, 107610. [CrossRef]
32. Pang, B.; Nazari, M.; Tang, G. Recursive variational mode extraction and its application in rolling bearing fault diagnosis. *Mech. Syst. Signal Process.* **2022**, *165*, 108321. [CrossRef]
33. Chen, X.; Zhang, X.; Church, J.A.; Watson, C.S.; King, M.A.; Monselesan, D.; Legresy, B.; Harig, C. The increasing rate of global mean sea-level rise during 1993–2014. *Nat. Clim. Chang.* **2017**, *7*, 492–495. [CrossRef]
34. Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. *IEEE Trans. Signal Process.* **2013**, *62*, 531–544. [CrossRef]
35. Greenacre, M.; Groenen, P.J.; Hastie, T.; d'Enza, A.I.; Markos, A.; Tuzhilina, E. Principal component analysis. *Nat. Rev. Methods Primers* **2022**, *2*, 100. [CrossRef]
36. Zhang, Y.X. Artificial neural networks based on principal component analysis input selection for clinical pattern recognition analysis. *Talanta* **2007**, *73*, 68–75. [CrossRef] [PubMed]
37. Maćkiewicz, A.; Ratajczak, W. Principal components analysis (PCA). *Comput. Geosci.* **1993**, *19*, 303–342. [CrossRef]
38. Kurita, T. Principal component analysis (PCA). In *Computer Vision: A Reference Guide*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 1–4.
39. Goyal, D.; Mongia, C.; Sehgal, S. Applications of digital signal processing in monitoring machining processes and rotary components: A review. *IEEE Sens. J.* **2021**, *21*, 8780–8804. [CrossRef]
40. Raj, N.; Prakash, R. Assessment and prediction of significant wave height using hybrid CNN-BiLSTM deep learning model for sustainable wave energy in Australia. *Sustain. Horiz.* **2024**, *11*, 100098. [CrossRef]
41. Raj, N.; Pasfield-Neofitou, S. Assessment and Prediction of Sea Level and Coastal Wetland Changes in Small Islands Using Remote Sensing and Artificial Intelligence. *Remote Sens.* **2024**, *16*, 551. [CrossRef]
42. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 53. [CrossRef] [PubMed]
43. Panahi, M.; Sadhasivam, N.; Pourghasemi, H.R.; Rezaie, F.; Lee, S. Spatial prediction of groundwater potential mapping based on convolutional neural network (CNN) and support vector regression (SVR). *J. Hydrol.* **2020**, *588*, 125033. [CrossRef]
44. Bhatt, D.; Patel, C.; Talsania, H.; Patel, J.; Vaghela, R.; Pandya, S.; Modi, K.; Ghayvat, H. CNN variants for computer vision: History, architecture, application, challenges and future scope. *Electronics* **2021**, *10*, 2470. [CrossRef]
45. Mahajan, R.; Mansotra, V. Predicting geolocation of tweets: Using combination of CNN and BiLSTM. *Data Sci. Eng.* **2021**, *6*, 402–410. [CrossRef] [PubMed]
46. Raj, N.; Brown, J. Prediction of Mean Sea Level with GNSS-VLM Correction Using a Hybrid Deep Learning Model in Australia. *Remote Sens.* **2023**, *15*, 2881. [CrossRef]
47. Dey, R.; Salem, F.M. Gate-variants of gated recurrent unit (GRU) neural networks. In Proceedings of the 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), Boston, MA, USA, 6–9 August 2017.
48. Zulqarnain, M.; Ghazali, R.; Aamir, M.; Hassim, Y.M.M. An efficient two-state GRU based on feature attention mechanism for sentiment analysis. *Multimed. Tools Appl.* **2024**, *83*, 3085–3110. [CrossRef]
49. Taud, H.; Mas, J.-F. Multilayer perceptron (MLP). In *Geomatic Approaches for Modeling Land Change Scenarios*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 451–455.
50. Desai, M.; Shah, M. An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN). *Clin. eHealth* **2021**, *4*, 1–11. [CrossRef]
51. Delashmit, W.H.; Manry, M.T. Recent developments in multilayer perceptron neural networks. In Proceedings of the Seventh Annual Memphis Area Engineering and Science Conference, MAESC, Memphis, TN, USA, 11 May 2005.
52. Freund, Y.; Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [CrossRef]
53. Feng, D.-C.; Wang, W.-J.; Mangalathu, S.; Hu, G.; Wu, T. Implementing ensemble learning methods to predict the shear strength of RC deep beams with/without web reinforcements. *Eng. Struct.* **2021**, *235*, 111979. [CrossRef]
54. Montesinos López, O.A.; Montesinos López, A.; Crossa, J. Support vector machines and support vector regression. In *Multivariate Statistical Machine Learning Methods for Genomic Prediction*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 337–378.
55. Müller, K.-R.; Smola, A.J.; Rätsch, G.; Schölkopf, B.; Kohlmorgen, J.; Vapnik, V. Predicting time series with support vector machines. In *International Conference on Artificial Neural Networks*; Springer: Berlin/Heidelberg, Germany, 1997; pp. 999–1004.

56. Vapnik, V.; Golowich, S.E.; Smola, A.J. Support vector method for function approximation, regression estimation and signal processing. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 1997; pp. 281–287.
57. Zhang, F.; O'Donnell, L.J. Support vector regression. In *Machine Learning*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 123–140.
58. Awad, M.; Khanna, R.; Awad, M.; Khanna, R. Support vector regression. In *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*; Springer Nature: Berlin/Heidelberg, Germany, 2015; pp. 67–80.
59. Basak, D.; Pal, S.; Patranabis, D.C. Support vector regression. *Neural Inf. Process.-Lett. Rev.* **2007**, *11*, 203–224.
60. Kartini, D.; Nugrahadi, D.T.; Farmadi, A. Hyperparameter tuning using GridsearchCV on the comparison of the activation function of the ELM method to the classification of pneumonia in toddlers. In Proceedings of the 2021 4th International Conference of Computer and Informatics Engineering (IC2IE), Depok, Indonesia, 14–15 September 2021.
61. Ravindiran, G.; Rajamanickam, S.; Kanagarathinam, K.; Hayder, G.; Janardhan, G.; Arunkumar, P.; Arunachalam, S.; AlObaid, A.A.; Warad, I.; Muniasamy, S.K. Impact of air pollutants on climate change and prediction of air quality index using machine learning models. *Environ. Res.* **2023**, *239*, 117354. [CrossRef] [PubMed]
62. Kim, M.; Jang, J.; Jeon, S.; Youm, S. A Study on Customized Prediction of Daily Illness Risk Using Medical and Meteorological Data. *Appl. Sci.* **2022**, *12*, 6060. [CrossRef]
63. Alhakeem, Z.M.; Jebur, Y.M.; Henedy, S.N.; Imran, H.; Bernardo, L.F.; Hussein, H.M. Prediction of ecofriendly concrete compressive strength using gradient boosting regression tree combined with GridSearchCV hyperparameter-optimization techniques. *Materials* **2022**, *15*, 7432. [CrossRef] [PubMed]
64. Herdian, C.; Widianto, S.; Ginting, J.A.; Geasela, Y.M.; Sutrisno, J. The Use of Feature Engineering and Hyperparameter Tuning for Machine Learning Accuracy Optimization: A Case Study on Heart Disease Prediction. In *Engineering Applications of Artificial Intelligence*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 193–218.
65. Raj, N.; Brown, J. An EEMD-BiLSTM algorithm integrated with Boruta random forest optimiser for significant wave height forecasting along coastal areas of Queensland, Australia. *Remote Sens.* **2021**, *13*, 1456. [CrossRef]
66. Waldmann, P. On the use of the Pearson correlation coefficient for model evaluation in genome-wide prediction. *Front. Genet.* **2019**, *10*, 899. [CrossRef] [PubMed]
67. Schober, P.; Boer, C.; Schwarte, L.A. Correlation coefficients: Appropriate use and interpretation. *Anesth. Analg.* **2018**, *126*, 1763–1768. [CrossRef] [PubMed]
68. Sedgwick, P. Pearson's correlation coefficient. *BMJ* **2012**, *345*, e4483. [CrossRef]
69. Willmott, C.J.; Robeson, S.M.; Matsuura, K. A refined index of model performance. *Int. J. Climatol.* **2012**, *32*, 2088–2094. [CrossRef]
70. Deo, R.C.; Şahin, M. Application of the artificial neural network model for prediction of monthly standardized precipitation and evapotranspiration index using hydrometeorological parameters and climate indices in eastern Australia. *Atmos. Res.* **2015**, *161*, 65–81. [CrossRef]
71. Dhaoui, O.; Antunes, I.M.; Benhenda, I.; Agoubi, B.; Kharroubi, A. Groundwater salinization risk assessment using combined artificial intelligence models. *Environ. Sci. Pollut. Res.* **2024**, *31*, 33398–33413. [CrossRef]
72. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [CrossRef]
73. Duc, L.; Sawada, Y. A signal-processing-based interpretation of the Nash–Sutcliffe efficiency. *Hydrol. Earth Syst. Sci.* **2023**, *27*, 1827–1839. [CrossRef]
74. Lin, F.; Chen, X.; Yao, H. Evaluating the use of Nash-Sutcliffe efficiency coefficient in goodness-of-fit measures for daily runoff simulation with SWAT. *J. Hydrol. Eng.* **2017**, *22*, 05017023. [CrossRef]
75. Coffey, M.E.; Workman, S.R.; Taraba, J.L.; Fogle, A.W. Statistical procedures for evaluating daily and monthly hydrologic model predictions. *Trans. ASAE* **2004**, *47*, 59–68. [CrossRef]
76. Legates, D.R.; McCabe, G.J. A refined index of model performance: A rejoinder. *Int. J. Climatol.* **2013**, *33*, 1053–1056. [CrossRef]
77. Legates, D.R.; McCabe, G.J., Jr. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**, *35*, 233–241. [CrossRef]
78. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Peerj Comput. Sci.* **2021**, *7*, e623. [CrossRef] [PubMed]
79. Willmott, C.J.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82. [CrossRef]
80. Jayasinghe, W.L.P.; Deo, R.C.; Ghahramani, A.; Ghimire, S.; Raj, N. Deep multi-stage reference evapotranspiration forecasting model: Multivariate empirical mode decomposition integrated with the boruta-random forest algorithm. *IEEE Access* **2021**, *9*, 166695–166708. [CrossRef]
81. Asuero, A.G.; Sayago, A.; González, A. The correlation coefficient: An overview. *Crit. Rev. Anal. Chem.* **2006**, *36*, 41–59. [CrossRef]
82. Bennett, N.D.; Croke, B.F.; Guariso, G.; Guillaume, J.H.; Hamilton, S.H.; Jakeman, A.J.; Marsili-Libelli, S.; Newham, L.T.; Norton, J.P.; Perrin, C. Characterising performance of environmental models. *Environ. Model. Softw.* **2013**, *40*, 1–20. [CrossRef]
83. Clark, G.F.; Hunter, C. Coasta: Climate change. In *Australian State of Environment*; Australian Government Department of Agriculture, Water and the Environment: Canberra, ACT, Australia, 2021.
84. Zhou, L.; Pan, S.; Wang, J.; Vasilakos, A.V. Machine learning on big data: Opportunities and challenges. *Neurocomputing* **2017**, *237*, 350–361. [CrossRef]

85. L'heureux, A.; Grolinger, K.; Elyamany, H.F.; Capretz, M.A. Machine learning with big data: Challenges and approaches. *IEEE Access* **2017**, *5*, 7776–7797. [CrossRef]
86. Dargan, S.; Kumar, M.; Ayyagari, M.R.; Kumar, G. A survey of deep learning and its applications: A new paradigm to machine learning. *Arch. Comput. Methods Eng.* **2020**, *27*, 1071–1092. [CrossRef]
87. Shah, V.; Konda, S.R. Neural Networks and Explainable AI: Bridging the Gap between Models and Interpretability. *Int. J. Comput. Sci. Technol.* **2021**, *5*, 163–176.
88. Ribeiro, M.T.; Singh, S.; Guestrin, C. Model-agnostic interpretability of machine learning. *arXiv* **2016**, arXiv:1606.05386.
89. Zafar, M.R.; Khan, N. Deterministic local interpretable model-agnostic explanations for stable explainability. *Mach. Learn. Knowl. Extr.* **2021**, *3*, 525–541. [CrossRef]
90. Linardatos, P.; Papastefanopoulos, V.; Kotsiantis, S. Explainable ai: A review of machine learning interpretability methods. *Entropy* **2020**, *23*, 18. [CrossRef]