International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys.13, No.1 (Apr-23)

http://dx.doi.org/10.12785/ijcds/130166

Predicting Future Global Sea Level Rise From Climate Change Variables Using Deep Learning

Kazi Md. Abir Hassan

Department of Electrical and Electronic Engineering, Islamic University of Technology (IUT)
Gazipur, Bangladesh

E-mail address: abirhassan@iut-dhaka.edu

Received 3 Apr. 2022, Revised 24 Dec. 2022, Accepted 6 Feb. 2023, Published 16 Apr. 2023

Abstract: Rapid climate change accelerates global temperature rise, causing thermal expansion of seawater and melting of ice-based lands, such as ice sheets and glaciers; these anomalies eventually result in global sea level rise. Since the beginning of satellite records, the sea level has risen significantly faster in recent decades than in prior decades, affecting people living in coastal areas directly as well as indirectly causing many environmental abnormalities. It is now possible to continuously monitor the level of seawater using current technology, but to battle this problem, it is necessary to understand the current scenario as well as predict the future scenario of sea level so that people may prepare and researchers can develop a viable solution, which is the main objective of this study. Here, 29 years of data on variables that are closely related to climate change, such as global temperature anomaly, ocean heat content change, carbon dioxide level in the atmosphere, and mass variation in Antarctica and Greenland, was gathered to build a multivariant prediction model using advance deep learning algorithms such as Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU), WaveNet (a type of Deep Convolutional Neural Network), and Deep Hybrid Network to predict the future scenario of global sea level rise. The results indicate that each method performs up to a certain level, but the deep hybrid model performed best in terms of accurately detecting the pattern of the dataset where MAE is 5.77 and RMSE is 7.67. Deep learning algorithms are admirable at identifying patterns in time series datasets, and with the necessary optimization, they can also assist in uncovering future data.

Keywords: Sea Level Rise, Climate Change, Deep learning, RNN, LSTM, GRU, WaveNet, Hybrid Network

1. INTRODUCTION

Climate change has become a catchphrase in modern life, as evidenced by the fact that the globe has already begun to experience the devastating effects of climate change. Climate change is defined as changes in weather patterns, which are influenced by a variety of natural and man-made variables. A fallacy is continued that climate change means only the temperature rise; the temperature rise is merely the beginning of the chain of many unwanted events that already started to become a threat to many livings including humans in many regions around the world [1]. To encounter daily needs in the current world, fossil fuels are burned, resulting in the production of greenhouse gas. This greenhouse gas acts as a cover over the atmosphere of earth, that causes trapping heat from the sun; normally, the heat from the sun is released at night, but because of the cover created by greenhouse gas, the heat cannot be released, thus temperature continues to rise [2]. One of the greenhouse gases that contribute to global temperature rising is carbon dioxide. Forest zones are important to the ecology because they absorb poisonous gases like carbon and are often thought of as carbon sinks. Meanwhile, as global deforestation increases, carbon sinks are dwindling, resulting in temperature rises [3]. The temperature is rising not just on land, but also on the surface of the ocean. As the temperature of the ocean rises, thermal expansion occurs in seawater; additionally, as the temperature rises, ice sheets and glaciers melt, resulting in the sea level increases [4]. Sea level rise is a concern not just in coastal areas, but it may also create various natural abnormalities in the long run, such as erosion, atypical floods, salt contamination of agricultural land, and habitat loss for many living things. So, all these anomalies are directly linked to climate change. The way the climate is changing in recent years, the pace of rising sea level has been particularly rapid in recent decades, as seen in Fig. 1, resulting directly in coastal flooding [5]; currently, it is estimated that more than 110 million people living on land that is below the current high tide and it is not stopping there rather in the future the scenario becomes more acute that it is estimated that the land of 150 million people will go below the high tide line permanently by 2050 [6]. As this sea-level rise originated from other variables namely ocean temperature rising, glacier melting, and loss of ice sheets; ocean temperature rising significantly depends on the emission of excess greenhouse gas and one of the greenhouse gases that play a vital role in temperature rising is carbon dioxide as mentioned earlier; analyzing these variables can assist in determining how the current pattern of sea-level rise is behaving, as well as how it will act in the future.

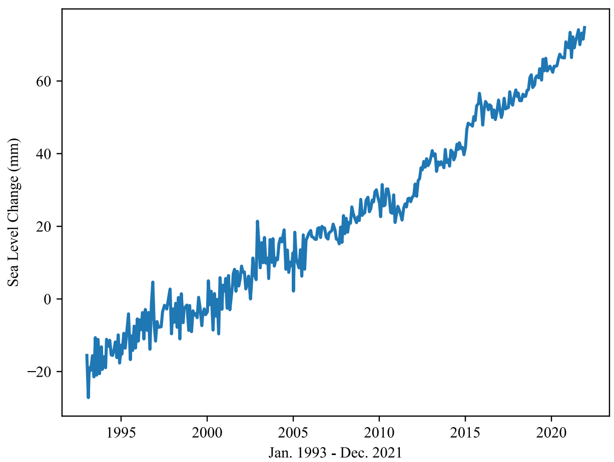

Figure 1. Global sea level rising, source: NASA

To study sea-level rise patterns and to generate future scenarios deep learning models can be a helpful tool, as in this modern-day the heavy computational power allows us to consider multiple variables altogether and can produce future data with high accuracy. Various approaches were taken previously to build the prediction model, single variable prediction model is one of them [7]; different machine learning algorithms such as regression, decision tree, random forest, or k-nearest neighbor were also employed to build the prediction model. Considering only one variable to build a prediction model can provide a better outcome [8] but as here it can be seen that sea level rise depends on multiple parameters, it might fail to capture the whole scenario from the aspect of another variable of the problem to predict future data. Thus, considering all those parameters that are varying over time is also necessary to view the whole scenario; that is why a prediction model is required that can analyze multiple variables altogether and produce future data. Although this multivariant prediction model is new and requires much attention because of its complexity, it might not be able to perform up to the mark [9]. But as this type of complex problem requires analyzing multiple variables, thus it is better to use a multivariant prediction model. Even though a multivariant prediction model is complex, a multivariant prediction model that is built with proper adjustment through trial and error can predict future data with very high accuracy [10]. Previously sea level rise was predicted in terms of Antarctica-related variables [11], global warming [7], greenhouse gas emission [12], as Antarctica's mass declines, global temperature rising, and increasing greenhouse gas emissions are directly connected with sea level rising. Also, it was taken an attempt to build a prediction model using global temperature, precipitations, and carbon emission altogether [13]. In addition, the sea level rising impact was studied in detail in specific regions like for Australian coast [14], Asia [15]-[16], Europe [17], North America [18], and South America [19]; very few times globally sea level impact is tried to predict using advance deep learning algorithms or to relate with multiple climate change variables. In this research, important climate change indicators such as global temperature anomaly, ocean heat content change, carbon dioxide level in the atmosphere, and Antarctica and Greenland mass variation all are used together to observe and anticipate a future global sea-level rise.

The remaining of this paper is presented in the following sequence. Section 2 delves into the methodology including a brief overview of major deep learning models that are used here. Section 3 explains the results of the models through error analysis, and Section 4 concludes the paper by presenting future predicted data.

2. METHODOLOGY

The entire approach can be broken down into three parts: gathering relevant data, preparing the data, and feeding the data to deep learning models for future data prediction.

A. *Collecting Datasets*

To build a multivariant prediction model multiple datasets are required. Here five datasets were collected, which are directly connected to climate change, named global average temperature anomaly, ocean heat content change, carbon dioxide level in the atmosphere, Antarctica and Greenland mass variation, and global sea-level rise shown in Fig. 1 and 2, are collected from [20].

B. *Preparing Datasets*

As it is known that to get better performance from the deep learning model, the raw data collected from any source cannot be used directly rather it is required to preprocess the dataset. To do so, the data collected in A are required to be prepared to be used in the deep learning models.

• Indexing: In time series analysis, it is critical to keep track of the numbers in order since chaotic data might generate anachronism and have a significant impact on the model. To avoid this problem, the full dataset was indexed according to the appropriate order.

• Reduction: While preparing the dataset it was found that certain datasets contain over 100 years of data for example global average temperature anomaly, however, some datasets have only 30 years of data, for example, carbon dioxide level in the atmosphere. But it is essential to have

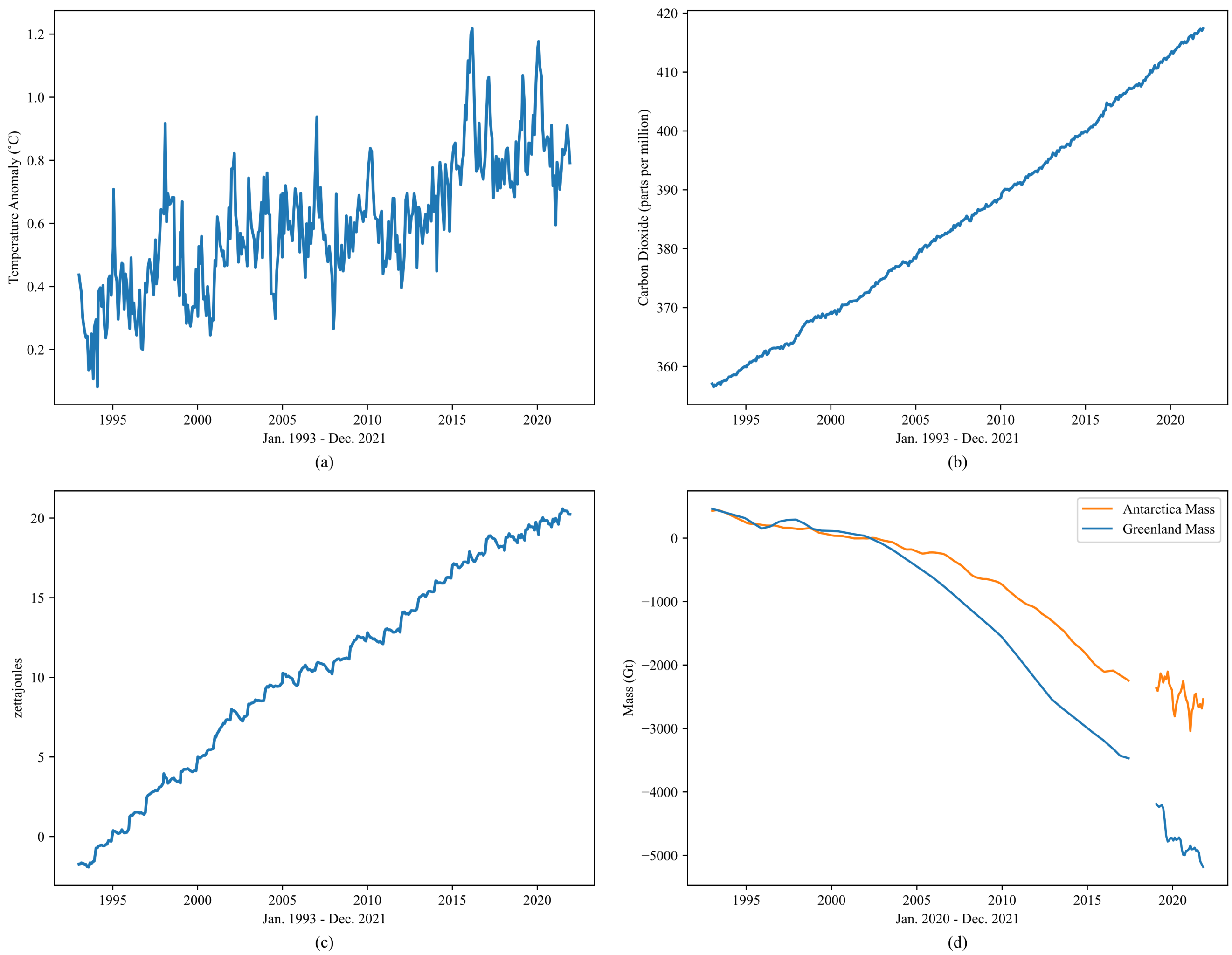

Figure 2. Variables that are collected to build multivariant prediction model (a) global average temperature anomaly, source: NASA (b) carbon dioxide level in atmosphere, source: NASA (c) ocean heat content changes, source: NASA (d) Antarctica and Greenland mass variation, source: IMBIE and NASA

an equal quantity of data for each variable to construct a multivariant prediction model to gain higher performance. To adjust, all data was collected from January 1993 to December 2021, a total of 348 months and the remaining data was discarded.

• Cleaning: In Fig. 2, Antarctica and Greenland mass variation, it is noticed that there is a gap in missing data. If there is missing data in a time series analysis then it is difficult for the deep earning model to capture the scenario of that specific period of the time where the data is missing. Conventionally this missing data is required to be filled manually or use the mean value when the data is normally distributed or use the regression method. For single variable forecasting, WaveNet performed with very high accuracy as mentioned here [4], thus, in this case, WaveNet was used to fill up the missing data.

• Transformation: In a multivariant prediction model, multiple variables are measured on a different scale and should play a distinct role in a prediction model. That is why it is required to normalize the whole data set so that all variables put an equal influence on the model as well as do not get biased on a particular variable. Here MinMax scaling is used to normalize the dataset so that all the values are in the range between 0 and 1. The mathematical formulation MinMax scaling can be expressed as

$$x_{scaled} = \frac{x - min(x)}{max(x) - min(x)} \tag{1}$$

C. Deep Learning Models

Five different networks were employed to create the prediction model for predicting future data. All of these algorithms have unique characteristics that help to achieve accurate prediction of time series data. Below is a quick description of each of them.

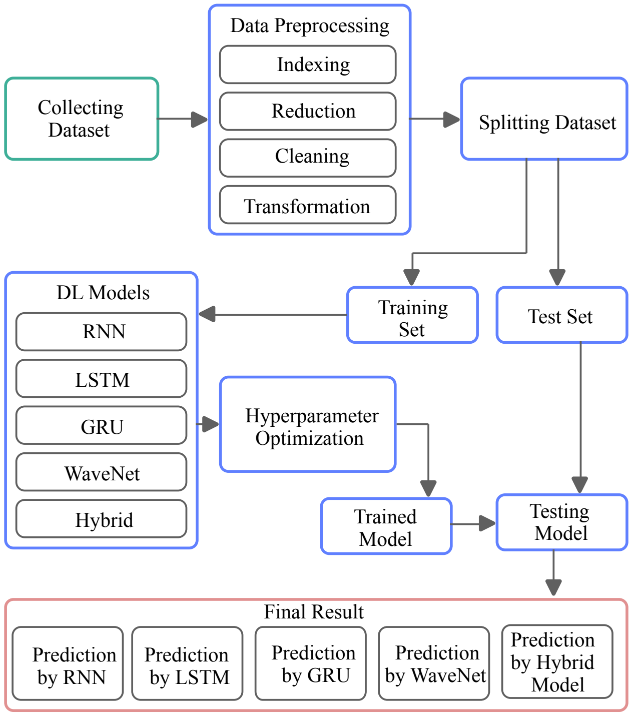

Figure 3. Summarization of the methodology

1. Recurrent Neural Network (RNN)
Because of having the ability to read sequential data, the recurrent neural network is one of the most often employed sequence models in neural networks. It can retain significant information from previous nodes and apply it to the current node. Thus, the current node is not only affected by current input information but also the previous node's information. Because time series are always influenced by past time step data, this sort of sequential model may be extremely useful in predicting future data. According to the information passed between nodes, RNN can be both unidirectional and bidirectional. In a unidirectional model, the model's current node is only impacted by previous nodes, not subsequent nodes. On the other hand, the bidirectional model is impacted by both previous and subsequent nodes.

Fig. 4 illustrates a sample of recurrent neural network topology. Where $x_1, x_2, x_3, \ldots, x_n$ is the input of the model and $y_1, y_2, y_3, \ldots, y_n$ is the output of the corresponding node. Also, $a_0, a_1, a_2, a_3, \ldots, a_n$ is the activation of the corresponding nodes. From the diagram, it can be seen that y_3 is influenced not just by the input of x_3, but also by x_1 and x_2. A deep recurrent neural network may also be built by stacking many layers of this model. Even though it is computationally costly, it can provide significantly better results than a single-layer model.

2. Long Short Term Memory (LSTM)
Long Short-Term Memory (LSTM) is an RNN architecture

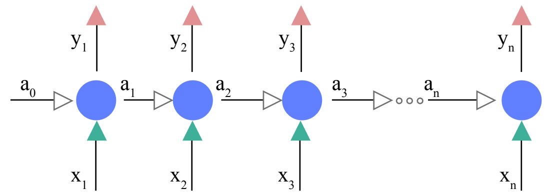

Figure 4. Sample of single-layer RNN

that was created to solve the vanishing gradient problem. For employing backpropagation to train RNNs, the long-term gradients that are backpropagated might reach zero or infinity, which is a significant concern when training a deep model. Because LSTM enables gradients to pass unaltered, it partially solves this problem. Gradients will not fall to zero, but they may still approach infinity as a result of this. In addition, LSTM allows for greater control over the mixing and flow of the inputs based on the learned weight. As a result, the model produces a lot of flexible output.

A cell, an input gate, an output gate, and a forget gate are the components of a typical LSTM unit. The three gates control the flow of information in and out of the cell, and the cell remembers values across arbitrary periods. The input gate determines what information from the current phase can be added. The next hidden state is determined by the output gate, while the previous inputs are stored in the hidden state. The forget gate, also known as the remember vector, instructs the cell state on which data to discard. Cell state is another word for long–term memory which enables the storage of data from earlier intervals in the LSTM cell.

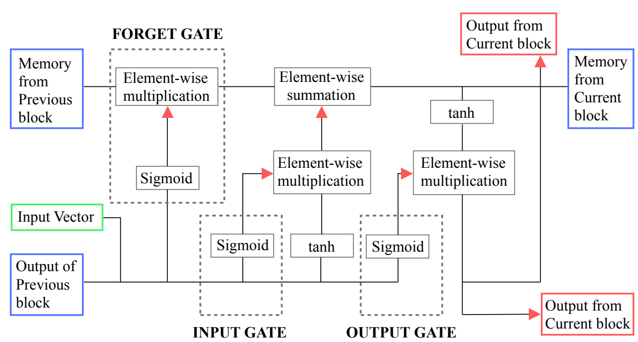

Figure 5. Single cell of LSTM [21]

3. Gated Recurrent Unit (GRU)
The gated recurrent unit is a redesigned version of the long short term memory. It manages information to flow without the need for a memory unit, whereas LSTM does. Also, it has two main gates one is reset gate and another is update gate, which makes one less gate than LSTM. The reset gate, which is more like an LSTM forget gate, is used for short-term memory or handling concealed states. As the sigmoid function is used in this gate so the values will be between 0 and 1. The update gate specifies how much information must be kept from the previous state as well as how much information will propagate from the preceding tier.

This type of network operates similarly to LSTM in most scenarios, although GRU is faster and less expressive. GRU is considerably quicker than LSTM for training the same size dataset since it has fewer parameters, making it computationally cheaper, but when the dataset is substantially bigger, LSTM outperforms GRU [22].

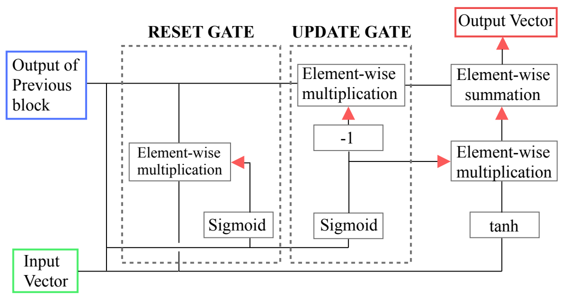

Figure 6. Single cell of GRU [23]

4. WaveNet

WaveNet is a modification of a deep convolutional neural network. A usual convolutional neural network is based on convolution operation more specifically causal convolution. This causal convolution performs convolution operations in a given sequence without relying on future time steps, and this order is maintained in the following time step. However, in causal dilated convolution, a particular variable called dilation rate is present, based on which a specific value will be excluded. This allows it to cover a much larger dataset in a shorter amount of time, as well as lower the complexity of the neural network, allowing it to train the network pretty quickly. For example, if the dilation rate is 1 then all the values will be taken, there will be no values that will be eliminated; a dilation rate of 2 means that every second element of the series will be taken which will make one value eliminated as it is depicted in Fig. 7. In this way, waveNet grows exponentially by reducing data and complexity.

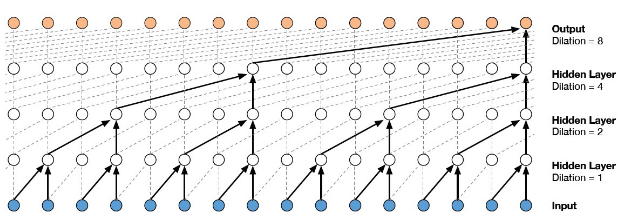

Figure 7. Dilated causal convolution operation. [24]

A multivariant prediction model incorporates numerous variables, the network becomes quite complicated; however, because of the dilation rate, a waveNet can decrease this complexity. By employing several dilation rates, such as 1, 2, 4, 8, and so on, the data will be omitted in a certain order, allowing the model to not only perceive the situation by skipping 1 or 2 or 4 or 8 values one at a time but also to predict the next scene more accurately.

5. Hybrid Deep Learning Model

Deep hybrid models are built by combining different deep learning approaches to create a new model that can offer achieving a better outcome. When numerous distinct models are joined, their unique features are also integrated, resulting in some fascinating results, and in some cases, it can even assist to achieve extremely high accuracy. However, each model has both distinct particular characteristics and disadvantages. These disadvantages must be addressed by trial and error; this is time-intensive and can make a model considerably more complex. If the hybrid model is well-constructed, it has the potential to produce the greatest results.

By far a total of four algorithms are discussed including their special features. Because of having recursive properties, the recurrent neural network can remember all information through time steps. In addition, intelligent recursive networks like LSTM and GRU can determine which information is required to pass on to the next tier as well as which information is less signifying to pass on to the next tier. By doing so they can perform significantly better to predict future data. Then waveNet has a special feature named dilated convolution and it helps to achieve a significantly better result to predict future data as can be seen in [8]. The parameters of waveNet that are used to build a hybrid model are shown in Table I.

TABLE I. Parameters of WaveNet

Filters	*Padding*	*Karnel Size*	*Stride*	*Dilation Rate*
32	Causal	2	1	1, 2, 4, 8, 16, 32

Lastly, a dense layer is a neural network with several hidden layers in which nodes of each layer are completely connected. This hybrid model is constructed by stacking one layer on top of the other, with each layer having its activation function; two activation functions were utilized here: hyperbolic tangent function and rectified linear unit (ReLU). The mathematical expression of ReLU is

$$f(x) = x^{+} = max(0, x) \tag{2}$$

Where x is the input to the function. The architecture of the hybrid network is shown in Table II.

The number of layers and variables in a deep hybrid network must be controlled, as they might increase the complexity of the network, slow down iteration, affect convergence, decrease efficiency and increase model training time. This particular structure yields the best results after testing numerous ones and optimizing hyperparameters. Because at first LSTM has more control over the mixing and flow of the inputs based on the learned weight as well as generates flexible output. Then, because of dilated convolution, some values will be dropped after a predetermined

TABLE II. PARAMETERS OF HYBRID MODEL

Layer Type	*Parameters*	*Activation Function*
LSTM	22464	tanh
LSTM	35072	tanh
Conv1D	4128	tanh
Conv1D	2080	tanh
Conv1D	2080	tanh
Conv1D	2080	tanh
Conv1D	2080	tanh
Conv1D	2080	tanh
GRU	7872	tanh
GRU	2080	tanh
Dense	792	ReLU
Total Parameters		92536
Total Trainable Parameters		92536
Total Non-Trainable Parameters		0

interval and passed to the GRU layer as necessary. As mentioned earlier GRU can generate a better result with fewer data points, thus, the GRU layer is used after the Conv1D layer. Final results will be obtained from the dense layer of the completely linked dense neural network, which will be activated by the ReLU function.

3. RESULT ANALYSIS

When analyzing a time series prediction model, it is important to look at how well the model captured the pattern of the data, such as if it was able to capture ups and downs, trends, or seasonality. The models utilized in this scenario have diverse characteristics, therefore each model will not only predict different values but also reveal different parts of the dataset. In comparison to other algorithms, RNN is the simplest method utilized here. As a result, it is projected that RNN will not be able to capture complicated data situations; instead, it'll be able to capture a simplified perspective of the dataset and predict future data based on that. RNN can hardly predict any data accurately or near to any correct data, as seen in Fig. 8, and so will do badly in the evaluation matrix. Furthermore, two separate evaluation matrices are employed to track the performance of these algorithms in the test set: one is Mean Absolute Error (MAE) and another is Root Mean Square Error (RMSE). To provide a fair comparison, all models are trained with the same amount of data, epoch number, and learning rate.

Then LSTM is used to forecast future data; typically, LSTM performs well in time series when the dataset is large because LSTM contains a memory block that can store the relevant features of the dataset and pass them on to the next tier. Although the dataset is not very large, LSTM outperformed RNN in this case, as seen in Fig. 9. Furthermore, it can be seen in the figure that some projected data was quite near to the real data, and the trend is almost captured, which is also reflected in the evaluation matrix.

As it is mentioned earlier, GRU is a slightly updated and modified version of LSTM, it has the extra advantage that it is computationally less expensive and it performs well in the smaller dataset. So, with a smaller dataset like this case, GRU performs better than LSTM to predict future data as can be seen in Fig. 10 and it took much less time to train the model. Furthermore, GRU could predict the trend of the dataset more accurately than other algorithms, as it can be seen that the predicted data is gradually going upwards.

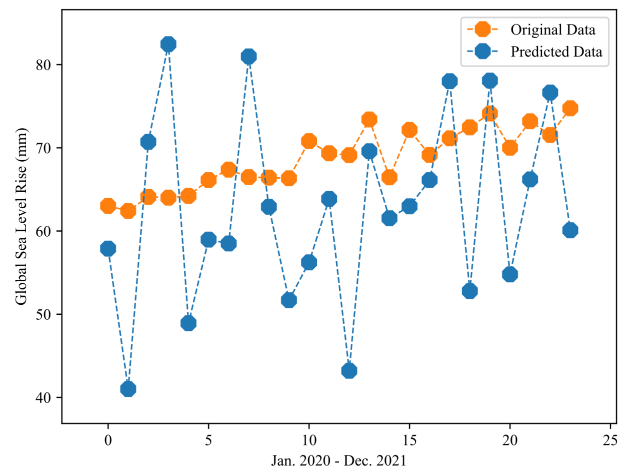

Figure 8. Comparison of original data and predicted data using RNN

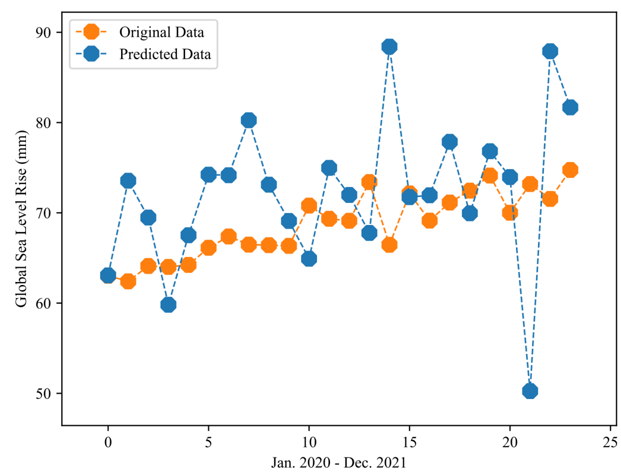

Figure 9. Comparison of original data and predicted data using LSTM

Because of dilated convolution, WaveNet is regarded as an excellent time series analysis prediction model. It is more efficient in capturing the situation of a large dataset and produces excellent results as found here [8], but in multivariant cases, it produces results that are nearly identical to LSTM or GRU. In the prediction model, there was no substantial improvement in performance. However, it was able to determine the trend of the dataset despite many fluctuations, shown in Fig. 11.

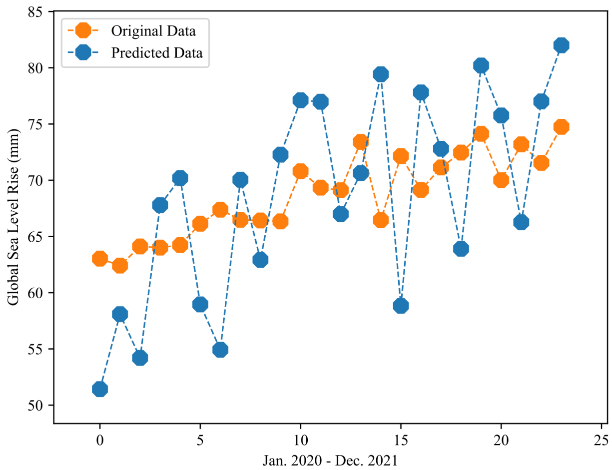

Figure 10. Comparison of original data and predicted data using GRU

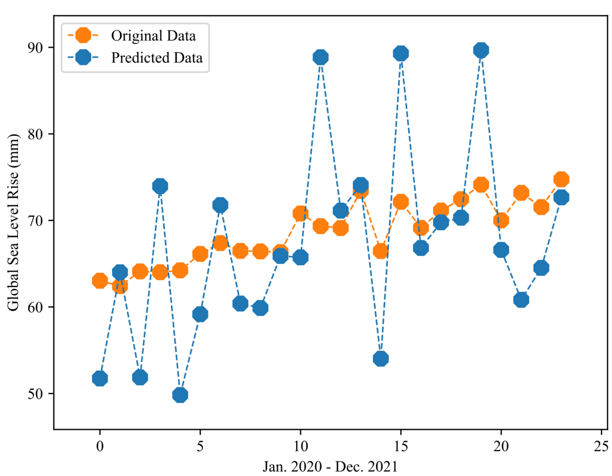

Figure 11. Comparison of original data and predicted data using WaveNet

The hybrid model was proposed to evaluate how well these deep learning models together could predict future data. As previously said, each algorithm utilized here has unique characteristics, which is why examining this hybrid model is necessary. If the specific characteristics are properly employed, this model should outperform the other algorithms tested in this study. Fig. 12 illustrates that the hybrid model can predict future data extremely well because it was able to encapsulate the scenario of the dataset. Some projected data was quite similar to the actual data, and the trend of future data was nearly identical to the original data.

Finally, all of these prediction models are compared using two evaluation matrices, revealing that the hybrid model outperformed the other models in terms of predicting future values. As a result, this model may be deemed to be dependable for predicting unknown future data, as evidenced by its strong performance in Fig. 12 and the evaluation matrix.

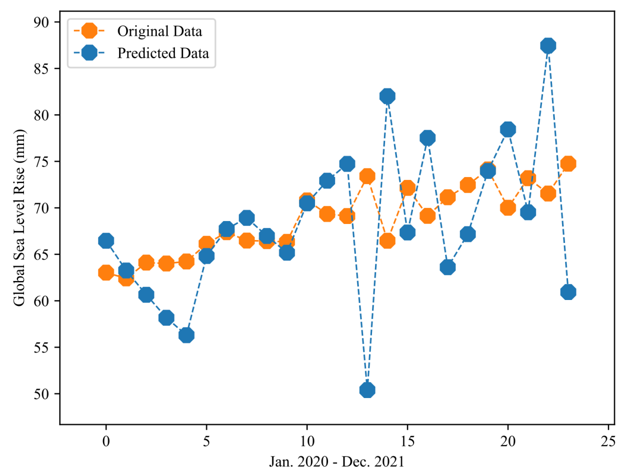

Figure 12. Comparison of Original data and predicted data using Hybrid Network

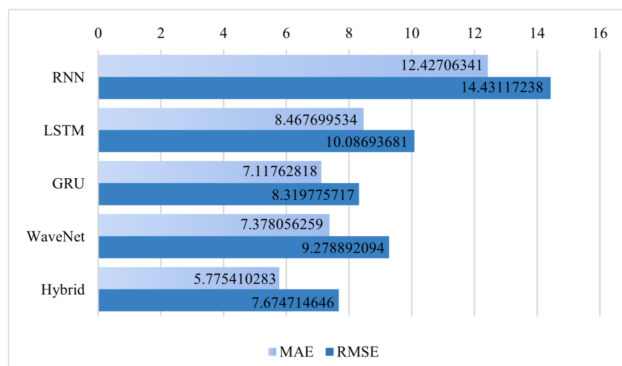

Figure 13. Comparison of evaluation matrix

4. CONCLUSION

Sea level rise is one of the consequences of climate change, and it is becoming a more solemn concern as time passes, as people of the coastal region of many countries already started to lose their homes to the ocean. Thus, research is required to better understand the pattern; this will permit us to predict the pattern in advance, allowing us to take the required actions based on the data. Predicting the future data is challenging as here multiple parameters are related to each other and one of them can be changed swiftly which will change the other parameters as well, then, the whole prediction model might become unreliable. Assuming that all of the variables used to construct the prediction model will not change overnight, an attempt was made to predict future scenarios of sea level rise using as many climate change variables as possible, and the paper is finally concluded by displaying the future scenario of sea level rise in the next fifty years using the best performing prediction model from the previous section where it can be seen that within next fifty years sea level will rise about 250 mm.

References

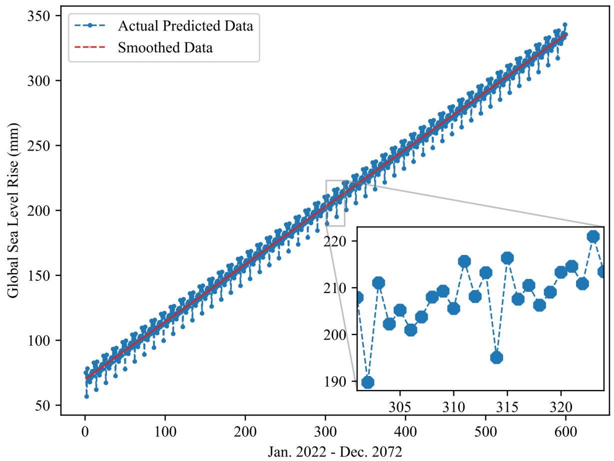

Figure 14. Predicted future data

Kazi Md. Abir Hassan received his BS degree in electrical and electronic engineering from the Islamic University of Technology (IUT), Bangladesh in 2021. His research interests include 4G, 5G, and NB-IoT, as well as machine learning and deep learning.