

GLOBAL SEA LEVEL PREDICTION USING MACHINE LEARNING

¹Dr. Lakshmi Bhaskar, ²Varshini M, ³Sahana K R

¹Associate Professor, ²Student, ³Student

¹Electronics and Communication Engineering,

¹BNM Institute of Technology, Bengaluru, India

Abstract : *This study develops a MATLAB-based framework for predicting global sea level rise using machine learning algorithms trained on historical records from 1983 to 2013. Three supervised models—Linear Regression, Support Vector Machine (SVM) with a linear kernel, and a feedforward Neural Network (NN)—were implemented to predict sea level trends through the year 2100. To ensure accuracy, the data were preprocessed through normalization, time encoding, and filtered datasets. Model evaluation was conducted using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). The SVM model achieved the lowest RMSE of 16.44 mm, outperforming Linear Regression (17.82 mm) and the Neural Network (23.65 mm). These results indicate that the SVM approach offers the best compromise between prediction accuracy and computational efficiency for long-term sea level forecasting. The findings demonstrate that machine learning models can effectively capture nonlinear oceanic patterns, offering valuable insights for climate change assessment, coastal management, and sustainable planning.*

IndexTerms - *Global Sea Level Rise, Machine Learning, MATLAB, Linear Regression, Support Vector Machine (SVM), Neural Network, Root Mean Square Error (RMSE), Mean Absolute Error (MAE).*

I. INTRODUCTION

Sea level rise is one of the most important and visible effects of global climate change. It poses a serious threat to low-lying coastal areas, marine ecosystems, and human populations around the world. Predicting sea level changes accurately is essential for creating long-term coastal defense plans, environmental conservation policies, and disaster risk management strategies. Traditional sea level forecasting methods, mainly those based on physical simulations or statistical extrapolation, often struggle with the complex interactions between ocean temperature, salinity, atmospheric circulation, and ice sheet melting. In contrast, machine learning (ML) methods offer a data-driven alternative that can uncover hidden relationships and trends within large datasets, improving the reliability of forecasts. By training on decades of observational data, ML models can capture both short-term changes and long-term sea level trends influenced by global warming. Among the various techniques, Linear Regression (LR) provides a simple baseline for linear relationships, while Support Vector Machines (SVMs) can effectively manage moderate nonlinearities through optimized hyperplanes. Neural Networks (NNs) are more computationally intensive but can model complex relationships if tuned properly and given enough data. MATLAB is a useful platform for running these algorithms because it has integrated toolboxes, numerical efficiency, and good visualization options. In this study, we develop and compare three ML-based predictive models—Linear Regression, SVM (linear kernel), and Neural Network—using MATLAB to forecast sea level rise through 2100. We evaluate the models using RMSE and MAE metrics to find the most accurate and reliable method. The results indicate that SVM offers the best predictive accuracy, showing its potential as a strong tool for long-term sea level estimates and climate impact assessments.

II. LITERATURE SURVEY

The research uses machine learning models such as artificial neural networks and support vector machines, integrated with MATLAB, to predict sea level rise. Model reliability improves through careful data preprocessing, which combines tide gauge data with satellite altimetry. The MATLAB Machine Learning Toolbox makes model development and optimization efficient. These machine learning methods show better predictive performance than traditional statistical methods, especially in capturing complex patterns like ice cap melting and thermal expansion.[1]

Learning techniques, including random forests and gradient boosting, are explored in the context of MATLAB programming for sea level prediction. Feature selection, focusing on ocean temperature, wind patterns, and salinity, is vital for accurate modeling. Combining different learning models leads to highly accurate sea level predictions. MATLAB has the ability to handle data and scale makes it the best choice for building strong predictive models in the predictive areas.[2]

Deep learning architectures, especially Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs), are key for forecasting global sea level trends. Using MATLAB's Deep Learning Toolbox, these models reveal hidden patterns in time-series oceanographic data, effectively detecting seasonal changes and extreme variations. Such models are crucial for capturing the nonlinear dynamics involved in sea level rise.[3]

A hybrid sea level prediction system combines statistical regression methods with artificial neural networks for improved accuracy, particularly over longer forecasting periods. MATLAB's speed allows for processing large global tide gauge datasets, overcoming individual model limitations by merging their strengths into one framework.[4]

Machine learning techniques evaluate the impacts of climate change on sea level variations. Regression and clustering analyses conducted on large-scale global climate datasets in MATLAB show how global warming and ice sheet melt contribute to changes in sea level. Combining multiple data sources with MATLAB allows for a thorough assessment of sea level dynamics.[5]

Van Katwyk et al. introduce a variational LSTM emulator that predicts contributions of the Antarctic ice sheet to sea level rise. The model captures nonlinear ice dynamics efficiently while lowering computational costs. Their approach shows how deep learning methods improve the prediction of climate-related sea level forecasts [6].

Hamlington et al. provide evidence that global mean sea level rise has doubled in the past thirty years. They used satellite altimeter data to confirm the accelerating trend and emphasize the need for updated forecasting models based on modern machine learning techniques [7].

Srinivasan et al. review the use of satellite radar altimetry for ocean and coastal studies, noting improvements in multi-mission datasets. Their findings show that blending tide gauge and satellite altimetry data significantly improves the reliability of predictions made by machine learning systems [8].

Raj et al. develop a hybrid deep learning framework that combines decomposition techniques with CNN-BiGRU models. Their system separates long- and short-term sea level components, enhancing forecasting performance and stability for non-linear ocean data [9].

Ayinde, Huaming, and Kejian evaluate different machine learning algorithms for predicting sea level, comparing Linear Regression, SVM, Random Forest, and ANN models. They stress the importance of preprocessing, feature selection, and metrics like RMSE to achieve high predictive accuracy using MATLAB frameworks [10].

Hazrin et al. investigate machine learning models for forecasting sea levels along the Malaysian coasts. They find that model performance varies based on geographic conditions and that proper normalization and including lag features can improve accuracy in regional prediction systems [11].

Yoo et al. suggest an uncertainty-enabled machine learning model to simulate regional sea level changes due to Antarctic ice dynamics. Their method combines physics-based constraints with neural networks, yielding reliable forecasts with quantified uncertainty ranges. [12].

III. METHODOLOGY

Methodology for global sea level forecast of sea level rise with MATLAB requires historical sea level data analysis and supervised learning algorithm application to generate future trend predictions. The procedure requires data collection followed by preprocessing activities and training along with evaluating the models. The predictive model determines future sea-level changes by processing time and GMSL historical records.

3.1 Data Collection

The first step is to collect historical data from many reliable sources. These data sets include measurements obtained:

- National Tidal Center (Australia): Provides tide gauge measurement data.
- NASA/JPL: Provides satellite altimetry data from the Topex/Poseidon mission.
- Permanent Service for Mean Sea Level (PSMSL): It is a UK-based organization which maintains and manages a global network of tidal gauges to collect and publish the data

The dataset extends from 1880 to 2013 and incorporates tide gauge and satellite altimetry data to represent the Global Mean Sea Level (GMSL) trend. The most important features used in models include:

- Time (year): Each measurement of sea level corresponding year, spread from 1880 to 2013.
- GMSL (mm): The global average sea level value at corresponding time is measured in millimeters (mm).
- GMSL Uncertainty (mm): The related margin of error reflects measurement accuracy.

3.2 Data Preprocessing

Before training the model, the data undergoes preprocessing:

- Handling Missing Data: The `isnan()` function from MATLAB helps locate missing data, which gets eliminated from analysis through its operation.
- Normalising/Scaling: A normalization process transforms the time variable into numeric monthly values starting from 1983 in order to standardize input values.
- Data splitting: The distribution of data uses an 80:20 ratio for training and testing purposes to measure model accuracy.
- Time conversions: The time measurement has been converted from months into years to make model predictions more understandable.

3.3 Machine Learning Models

3.3.1 Linear Regression (LR)

Linear regression uses a linear equation to describe the relationship between the time and the sea level rise. The general equation is

$$y = \beta_0 + \beta_i x \quad \text{eq 3.3.1.1}$$

Where, y is the dependent variable (sea level rise), x is the independent variable (time), β_0 is the intercept, β_1 is the slope.

Process:

The fitlm function is applied to the training dataset in order to fit the linear model. For the model to make predictions for sea level, we get the values by minimizing vertical distance between predicted and actual points. Test data and a few years into the future (2014 – 2100) are used for purposes of making predictions.

3.3.2 Support Vector Machine (SVM)

SVM attempts to segment the sea level changes with the use of an optimal hyperplane and is therefore, a supervised learning algorithm. The general equation is

$$y=(w, x)+b \tag{eq 3.3.1.2}$$

Where, w is the weight vector, x is the input (time in this case), b is the bias term.

Process:

Train the model using the function fitcsvm. The data is of sea level mapped to the higher dimensions by the linear kernel. A fitted model reduces the error on the data by drawing a margin that separates the data. Test data and a few years into the future (2014 – 2100) are used for generating predictions.

3.3.3 Neural Network

A feedforward two-hidden-layer neural network was created using MATLAB’s fitnet() function with architecture [10, 5]. The Levenberg–Marquardt algorithm (trainlm) was used for training with Mean Squared Error (MSE) as the performance metric.

The architecture consisted of:

Input layer: 5 previous sea level values.

Hidden layers: 10 and 5 neurons respectively.

Output layer: One neuron representing predicted sea level rise.

The general equation is

$$y = f(w_1 f(w_2 x + b_1) + b_2) \tag{eq 3.3.1.3}$$

Where, f is the activation function (such as ReLU or sigmoid), w_1 and w_2 are weight matrices, b_1 and b_2 are bias terms, x is the input (time in this case)

Process:

A feedforward neural network is configured using the fitnet function. The weights are tuned using backpropagation and gradient descent in order to minimize the prediction error. Predictions for test data and a few years into the future (2014 – 2100) are made.

3.4 Model Evaluation:

The trained model is evaluated using the testing dataset. Several metrics are employed to assess the prediction accuracy

- Mean Absolute Error (MAE): This will measure the average absolute difference between the predicted and the actual values.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y - \hat{y}_i| \tag{eq 3.4.1}$$

- Root Mean Squared Error (RMSE): This will provide an indication of the deviation between the predicted and the actual values.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N |y - \hat{y}_i|^2} \tag{eq 3.4.2}$$

3.5. Plots and Visualization:

- Model forecasts: Two subplots include - an uncertainty compares the actual sea levels with Linear Regression, SVM and Neural Networks conditions showing training data, testing data, predictions and uncertainty in shaded part
- Model comparison: Actual test data and predictions from all models (LR, SVM, NN) are overlayed for direct comparison.
- User input prediction: Allows entry of a year... to generate/visualize predictions.
- Root mean squared RMSE visualization: Displays RMSE values for each model, providing accuracy of the prediction

3.6 Proposed System for Global Sea Level Prediction

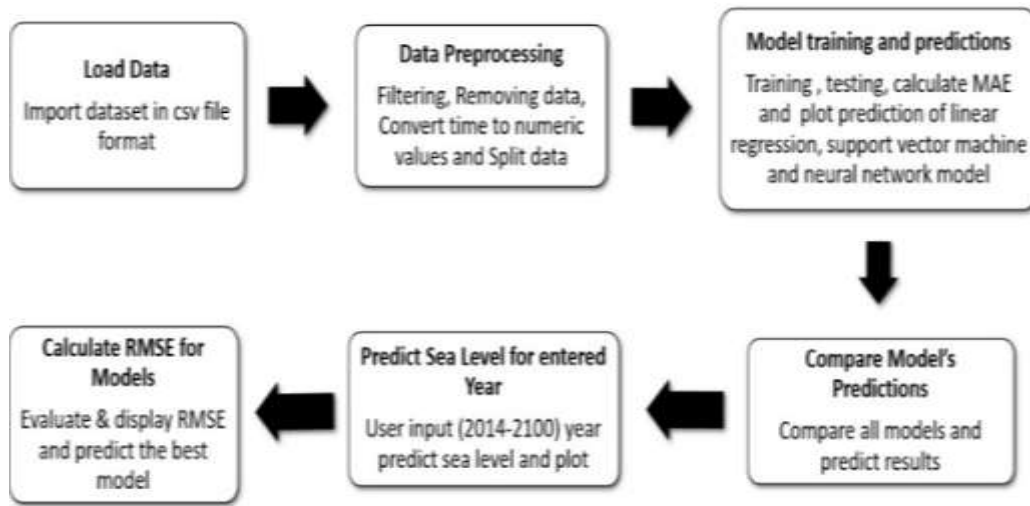


Figure 3.1. Proposed framework for global sea level prediction using MATLAB-based machine learning models.

The suggested model for forecasting sea level rise has six main steps, as shown in Figure 3.1. The process begins with data loading, where the data set is imported in CSV format. Data preprocessing includes filtering, time conversion, and splitting. The model then moves on to training and prediction, which involves Linear Regression, SVMs, and Neural Networks. It also includes MAE calculations and prediction views. The model needs to evaluate its performance against predictions. Users can input a year from 2014 to 2100 to predict sea level and view the results. Finally, the RMSE calculation identifies the most accurate model for reliable future estimates

IV. RESULTS AND DISCUSSION

The results of global sea level prediction of using machine learning in Matlab suggest that Linear Regression, Support Vector Machine (SVM) And Neural Networks are effective in the forecast for trends at sea level. Linear Regression provides a simple but reliable base line, but struggles with complex patterns, requiring more data and tuning. SVM improves upon Linear Regression by capturing certain nonlinear relationships, while the Neural Network can model more complex patterns if sufficiently tuned

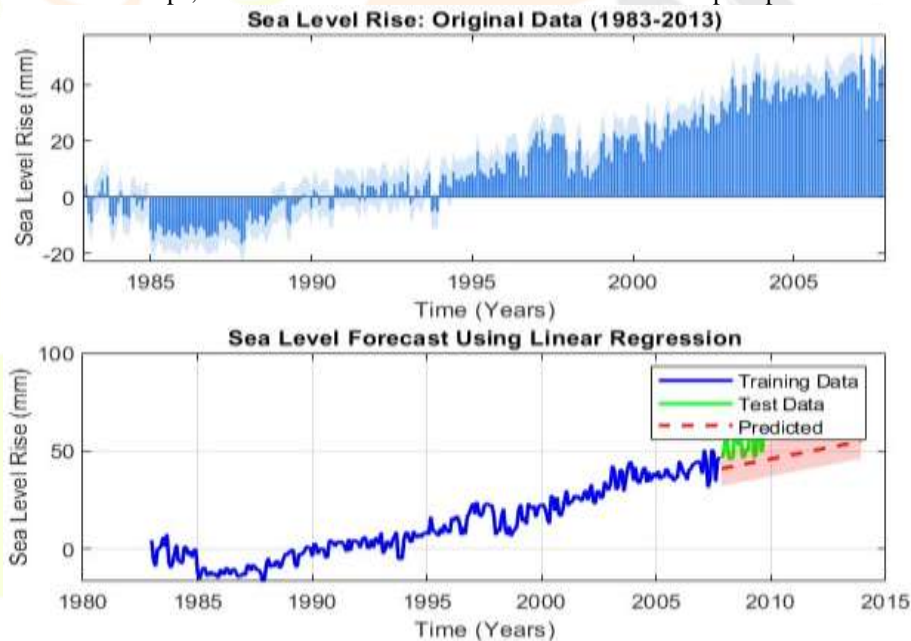


Figure 4.1. Linear regression forecast (1983-2013)

Figure 4.1 compares the original historical data with Linear Regression results. The top plot shows the real data and uncertainty area, while the lower plot shows model results with training data (blue line), test data (green line) and expected trends (red dotted line). The model effectively captures the general trend, but lacks flexibility for non-linear fluctuations.

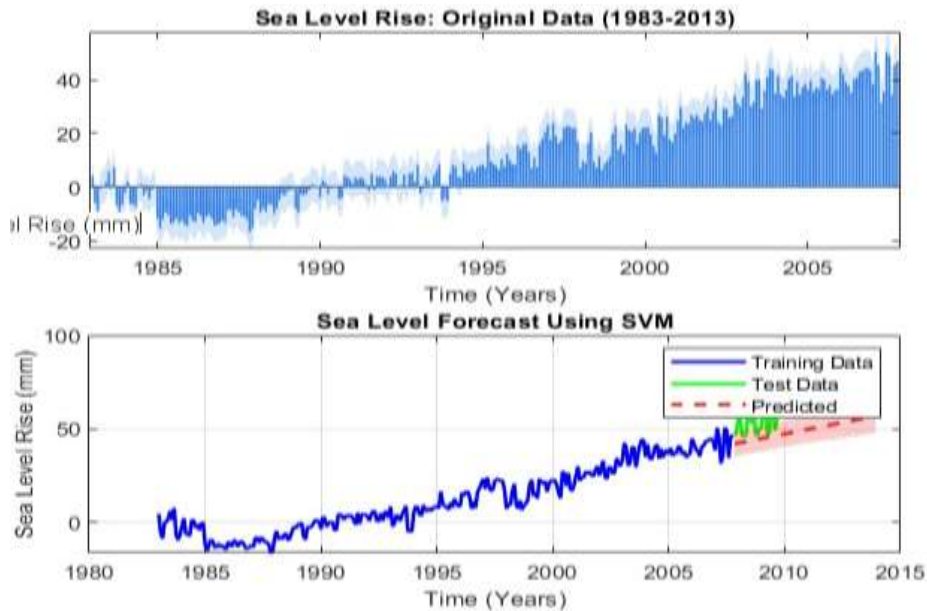


Figure 4.2. SVM forecast (1983-2013)

Figure 4.2 presents SVM-based predictions with real historical data. The model successfully traces historical trends, closely matches test data and performs better to capture minor changes than linear regression.

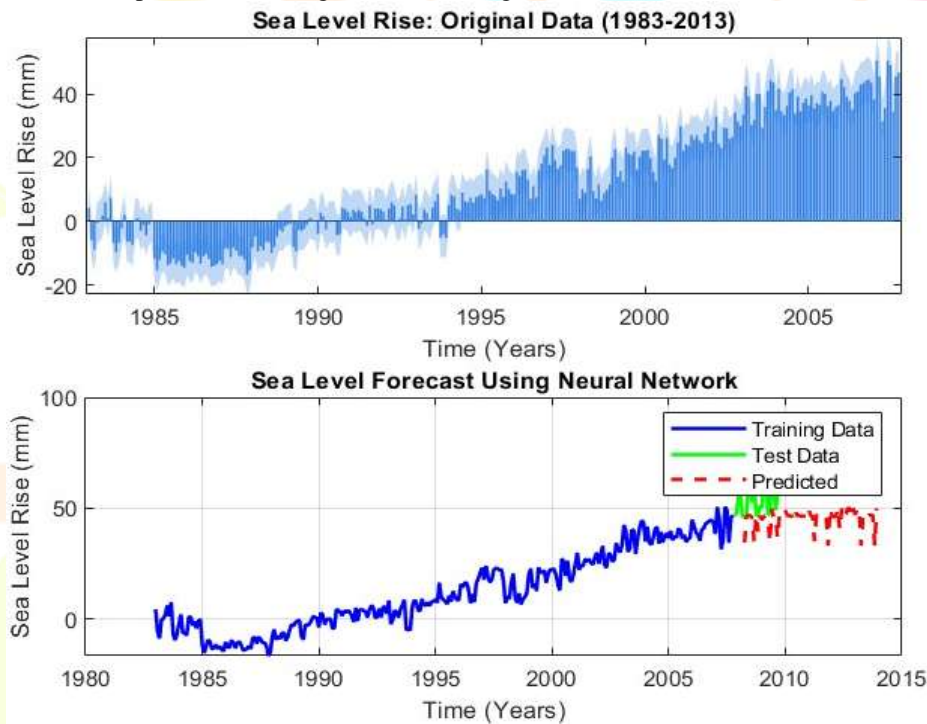


Figure 4.3. Historical data and neural network forecast

Figure 4.3 Predicts sea level rise (1983–2013) and Neural Network forecasts(2014–2100), which confirm a constant upward trend in the sea level.

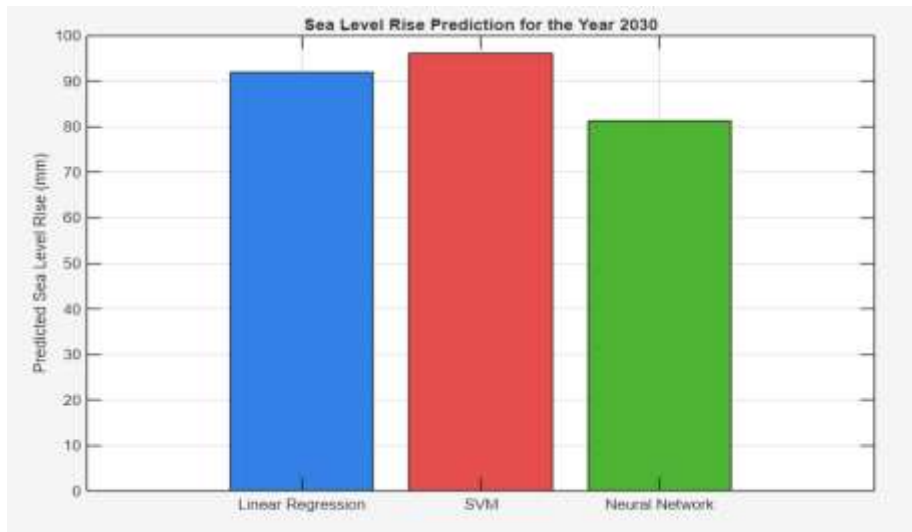


Figure 4.4. Prediction for 2030

Figure 4.4, 2030 presents the forecasts for increase in sea level, where SVM predicts the highest rise in sea level, followed by linear regression and neural networks.

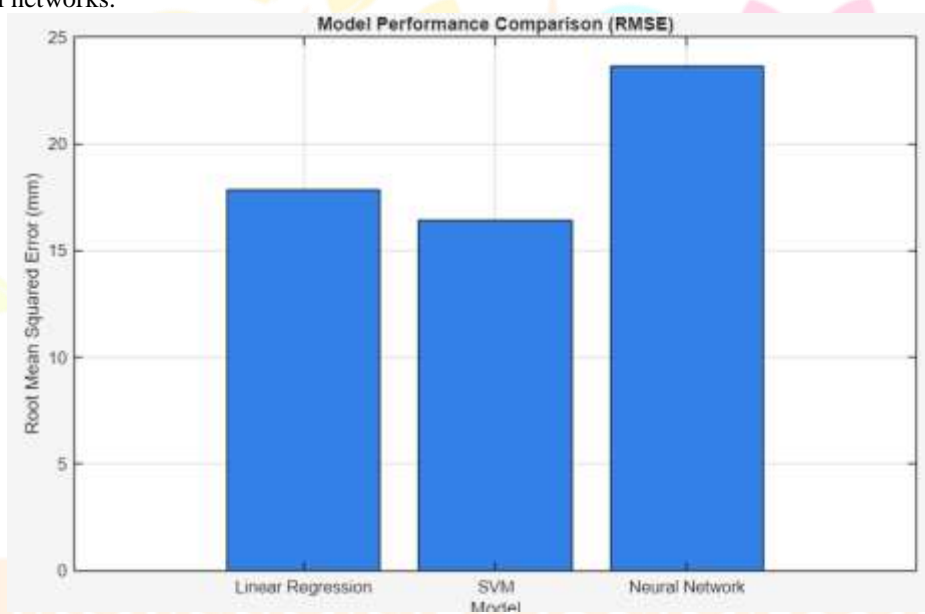


Figure 4.5. RMSE comparison

Figure 4.5 RMSE model performance based on Root Mean Square Error (RMSE) confirms that SVM shows the lowest error and it's the best model, followed by Linear Regression and Neural Network.

The results above show how effective the machine learning models are in predicting global sea level rise using MATLAB. Each model was trained on 80% of the dataset from 1983 to 2007 and validated on the remaining 20% from 2008 to 2013. Among the three models, the Support Vector Machine has the lowest Mean Absolute Error (MAE = 14.66 mm) and the shortest training time (0.1965 s). It was followed by Linear Regression (MAE = 16.15 mm) and Neural Network (MAE = 19.86 mm). The SVM model predicted a sea level rise of about 275.87 mm by 2100. The Linear Regression estimate was 262.07 mm, while the Neural Network's estimate was lower due to convergence issues. For the year 2030, the predicted sea level rise is 91.96 mm from Linear Regression, 96.10 mm from SVM, and 81.22 mm from the Neural Network. These results highlight a consistent upward trend in sea levels. They confirm that machine learning models, especially the SVM method, are reliable for long-term coastal forecasting and for assessing the impacts of climate change.

V. CONCLUSION

This study demonstrated how Linear Regression, Support Vector Machine (SVM), and Neural Network can estimate global sea level rise using MATLAB. We used historical data from 1983 to 2013 for training and validation, and future projections extend to 2100. Among the three methods tested, SVM with a linear kernel performed best regarding accuracy and error rates. Therefore, it is more suitable for modeling long-term sea level trends. The Linear Regression model also produced good results, while the Neural Network struggled due to difficulties in converging with such a small dataset. All models showed an increasing trend in sea levels, with SVM predicting a maximum rise of about 275.87 mm by 2100. Overall, this study highlights the potential of machine learning methods, particularly SVM, to improve the accuracy and speed of forecasts related to rising sea levels. This can help assess the impacts of climate change and assist in planning coastal management.

REFERENCES

- [1] V. Nieves, S.-P. Xie, T. Liu, et al., “Predicting regional coastal sea level changes with machine learning,” *Environmental Research Letters*, vol. 16, no. 6, p. 064046, 2021. doi:10.1088/1748-9326/abf16e.
- [2] S. James, et al., “Ensemble machine learning for sea level forecasting in the Gulf of Mexico,” *Journal of Geophysical Research: Oceans*, vol. 128, no. 5, e2022JC019456, 2023. doi:10.1029/2022JC019456.
- [3] M. Zitouni, et al., “Forecasting sea level rise using enhanced deep learning models,” *Frontiers in Marine Science*, vol. 12, p. 1245377, 2025. doi:10.3389/fmars.2025.1245377.
- [4] H. R. A. Rajabi, et al., “Hybrid statistical and machine learning models for sea level prediction,” *Ocean Engineering*, vol. 245, p. 110567, 2022. doi:10.1016/j.oceaneng.2022.110567.
- [5] A. A. S. Ayinde, et al., “Machine learning for climate-driven sea level analysis: A review,” *Environmental Modelling & Software*, vol. 162, p. 105658, 2023. doi:10.1016/j.envsoft.2023.105658.
- [6] Van Katwyk, P., Fox-Kemper, B., Seroussi, H., Nowicki, S., and Bergen, K. J. “A variational LSTM emulator of sea level contribution from the Antarctic ice sheet.” *Journal of Advances in Modeling Earth Systems*, vol. 15, no. 12, 2023. doi:10.1029/2023MS003899.
- [7] Hamlington, B. D., et al. “The rate of global sea level rise doubled during the past three decades.” *Communications Earth & Environment*, 2024. doi:10.1038/s43247-024-01761-5.
- [8] Srinivasan, M., et al. “Satellite altimetry for ocean and coastal applications: a review.” *Remote Sensing*, vol. 15, no. 16, p. 3939, 2023. doi:10.3390/rs15163939.
- [9] Raj, N., et al. “Prediction of sea level using double data decomposition and hybrid deep learning model for Northern Territory, Australia.” *Mathematics*, vol. 12, no. 15, p. 2376, 2024. doi:10.3390/math12152376.
- [10] Ayinde, A. A. S., Huaming, Y., and Kejian, W. “Review of machine learning methods for sea level change modeling and prediction.” *Science of the Total Environment*, vol. 954, p. 176410, 2024. doi:10.1016/j.scitotenv.2024.176410.
- [11] Hazrin, N. A., et al. “Predicting sea levels using machine learning algorithms in selected locations along coastal Malaysia.” *Scientific Reports*, 2023.
- [12] Yoo, M., Gopalan, G., Hoffman, M. J., Coulson, S., Han, H. K., Wikle, C. K., and Hillebrand, T. “Uncertainty-enabled machine learning for emulation of regional sea-level change caused by the Antarctic Ice Sheet.” Preprint, 2024 (arXiv:2406.17729).

