Dynamic optimization using long short-term memory and genetic algorithms for predicting marine data

Mukhlis[1,2], Indra Jaya[3], Sri Nurdiati[4], Karlisa Priandana[1], Irman Hermadi[1]
[1]Computer Science Study Program, School of Data Science, Mathematics, and Informatics, IPB University, Bogor, Indonesia
[2]Department of Informatics, Faculty of Computer Science, Bhayangkara Jaya University, Jakarta, Indonesia
[3]Department of Marine Science and Technology, Faculty of Fisheries and Marine Sciences, IPB University, Bogor, Indonesia
[4]Mathematics Study Program, School of Data Science, Mathematics, and Informatics, IPB University, Bogor, Indonesia

Article Info

Article history:

Received Mar 12, 2025
Revised Feb 6, 2026
Accepted Apr 22, 2026

Keywords:

Dynamic deep learning optimization
Evolutionary hyperparameter tuning
LSTM-GA hybrid
Oceanographic time-series
Prediction

ABSTRACT

This study aims to develop an accurate and efficient ocean data prediction model to tackle the challenges posed by climate change and complex oceanographic dynamics. The main goal is to use long short-term memory (LSTM) networks along with genetic algorithms (GA) to predict four key ocean factors at once: sea surface temperature (SST), sea surface height (SSH), sea surface salinity (SSS), and chlorophyll-a (Chl-a). An experimental quantitative approach is employed, utilizing satellite data from the Banda Sea region. This approach involves time series modeling using LSTM, which is optimized by GA for hyperparameters such as the number of neurons and batch size. The results show that the combined LSTM-GA model greatly improves prediction accuracy and successfully identifies seasonal trends and irregular changes in all variables, even when there is a lot of noise. Tests reveal that the optimal configuration varies for each variable, and the GA optimization process can expedite model convergence by as little as 10 epochs. These findings underscore the effectiveness of integrating evolutionary techniques in training deep learning (DL) models for ocean data. The implications of this research include potential applications in adaptive ocean monitoring systems, early warning initiatives, and data-driven planning in marine resource management.

Corresponding Author:

Indra Jaya
Department of Marine Science and Technology, Faculty of Fisheries and Marine Sciences, IPB University
Bogor 16680, Indonesia
Email: indrajaya@apps.ipb.ac.id

1. INTRODUCTION

Ocean parameters such as sea surface salinity (SSS), sea surface temperature (SST), sea surface height (SSH), and chlorophyll-a (Chl-a) concentrations play a pivotal role in shaping marine ecosystems, regulating global climate systems, and influencing human activities. Accurate prediction of these parameters is critical for sustainable marine resource management, disaster mitigation, and effective maritime planning [1]. However, the temporal complexity and nonlinear interdependencies among these parameters make long-term and seasonal prediction challenging when using conventional statistical or physical models. This complexity necessitates the development of a robust predictive framework capable of capturing the dynamic and nonlinear nature of ocean data [2], [3].

Variations in oceanographic dynamics—such as SST, ocean currents, and salinity—have substantial implications for global climate systems and the management of marine resources. Consequently, there is an increasing demand for adaptive and high-precision prediction models. In recent years, machine learning

(ML) and deep learning (DL) techniques have demonstrated great potential in modeling nonlinear and time-dependent environmental systems. Among these, the long short-term memory (LSTM) neural network has gained prominence for its ability to capture long-term temporal dependencies in sequential data, while optimization algorithms such as the genetic algorithm (GA) have been widely adopted to enhance model performance through adaptive parameter tuning [4], [5]. Several studies have explored ML- and DL-based models for ocean data prediction. For example, the authors [6], [7] utilized LSTM to predict SST and reported superior performance compared to traditional autoregressive models. Similarly, Chen *et al.* [8] demonstrated that integrating LSTM with GA improves predictive accuracy for sequential data. Other studies have extended LSTM applications to predict coastal climate variables with improved accuracy over longer periods [9], [10], while hybrid models combining LSTM with metaheuristic algorithms have been developed to accelerate convergence and enhance the stability of ocean weather predictions [11]. Furthermore, the authors [12], [13] successfully applied LSTM to predict ocean current patterns with higher accuracy than conventional models.

Despite these advancements, the integration of LSTM and GA for multidimensional marine parameter prediction remains limited. For instance, Alizadeh and Nourani [14] employed GA to optimize a sequential model for predicting water temperature; however, its application to complex, high-dimensional marine datasets has not been fully investigated. Stajkowski *et al.* [15] discussed the use of LSTM for blockchain transaction cost prediction, yet the integration of GA and its adaptation to marine data remains unexplored, indicating a significant research gap in this domain. Most previous studies have focused on standalone implementations of LSTM or other ML approaches without incorporating optimization techniques such as GA, even though parameter optimization plays a crucial role in enhancing model generalization and predictive accuracy [16]. Additionally, many studies analyzed only one or two oceanic variables independently, neglecting the intricate interactions among SSS, SST, SSH, and Chl-a. Determining the optimal number of hidden layers and neurons in an LSTM model constitutes an non-deterministic polynomial-hard (NP-hard) problem [17]. Although GA does not guarantee a global optimum, it efficiently explores the search space to produce near-optimal solutions, thereby mitigating the computational burden of manual hyperparameter tuning [18].

Hybrid architectures such as convolutional neural network (CNN)–LSTM and spatiotemporal deep models have shown improved predictive capabilities for ocean data. Nevertheless, these models still exhibit limitations in handling spatiotemporal uncertainties and dynamic dependencies [19]–[21]. While the authors in [22]–[24] emphasized the relevance of evolutionary algorithms in optimizing LSTM networks, few studies have systematically explored the integration of GA and LSTM for spatiotemporal marine data prediction. This research gap underscores the necessity for further investigation into how temporal and spatial variations influence the robustness and adaptability of hybrid predictive models in ocean science.

This study proposes an integrated framework combining LSTM and GA to predict multiple marine parameters—SSS, SST, SSH, and Chl-a—simultaneously. The GA is employed to adaptively optimize LSTM parameters and hyperparameters, thereby improving the model's capacity to learn complex inter-variable relationships and nonlinear temporal-spatial dynamics. Unlike prior studies focusing on single-variable predictions, the proposed approach emphasizes multivariate integration and dynamic interaction modeling. The GA performs optimization through an evolutionary process inspired by natural selection, enabling efficient exploration of near-optimal configurations [25]. Recent developments in predictive modeling confirm that neural network architectures, particularly LSTM, are powerful tools for analyzing nonlinear and irregular marine data patterns [19], [26]. Nevertheless, manually setting model parameters or relying on conventional heuristics often results in suboptimal convergence and limited generalization [27], [28]. Therefore, combining LSTM with an evolutionary optimization strategy such as GA offers a promising avenue to enhance training efficiency and predictive reliability.

The present study aims to develop a more accurate and adaptive model for predicting marine parameters by leveraging the synergy between LSTM and GA. The expected contributions of this work include advancing the state-of-the-art in ocean prediction modeling. The study also supports evidence-based decision-making for marine resource management, disaster preparedness, and maritime operations [29].

2. METHOD

2.1. Data collection and preprocessing

Marine data are collected from sources such as Google Earth Engine and Copernicus Marine Service. Initial preprocessing includes standardizing data types (e.g., float and datetime) and handling missing values through interpolation, imputation, or removal. The dataset is then split into 80% training–validation and 20% testing. Training data are used for model learning, while testing data evaluate the final performance. An LSTM model optimized with GA is trained to handle time-series variables (SST, SSS, SSH, and Chl-a), where GA determines optimal parameters such as neuron count, learning rate, epoch, and batch size. From the training portion, validation data support parameter evaluation and prevent overfitting. GA iteratively generates parameter configurations, trains the LSTM, and evaluates performance; if results are suboptimal, new configurations are produced via selection, crossover, and mutation. Once the best configuration is found, the model is retrained and validated, then tested using the unseen 20% data to compute mean absolute error (MAE), root mean square error (RMSE), and other metrics. The overall process from data collection to testing is illustrated in Figure 1.

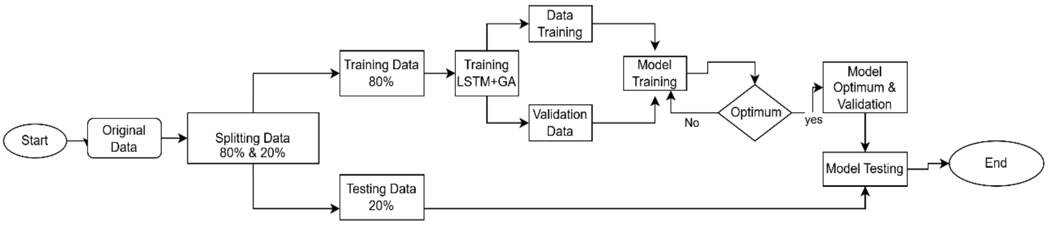
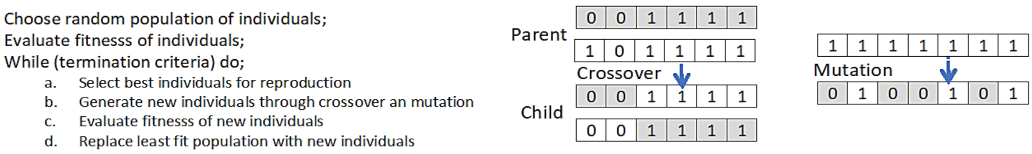


Figure 1. Research stages

2.2. Ocean data prediction with long short-term memory and genetic algorithms

LSTM optimization using GA begins with population initialization, where each individual encodes configurations such as neuron count and time-step length. Each configuration is built into an LSTM model and evaluated using metrics like RMSE or MAE [2]. Individuals with the best fitness values undergo crossover and mutation to explore improved solutions and avoid local optima. After several generations, the best-performing configuration is selected and retrained using marine data to ensure predictive accuracy [9]. The GA operations—including crossover and mutation—are illustrated in Figure 2, which explains the basic concept of a GA, an optimization method inspired by the process of biological evolution. In Figure 2(a) stages of the GA process and Figure 2(b) examples of genetic operations: crossover and mutation.



(a) (b)

Figure 2. GA flow and mechanism of (a) stages the GA process and (b) GA crossover and mutation operations

The model prediction results are compared with the target value using RMSE or MAE to assess the overall accuracy of the marine data, as shown in the LSTM–GA optimization process in Figure 3. After the fitness stage, the GA performs a crossover process to generate new individuals based on the combination of two selected parents [30]. Using the DEAP library, this process is carried out through a gene merging mechanism according to the crossover type used, such as uniform (genes are randomly selected from each parent), one-point (gene exchange after one random point), and two-point (gene exchange between two random points). This crossover plays an important role in exploring solutions and determining the optimal configuration for the marine data prediction model.

2.3. Genetic algorithm formation and process

The GA operates with a population size of 10, five generations, and a chromosome length of 15. The process begins with the random initialization of 10 individuals, each represented by a 15-gene chromosome. Each individual is then evaluated using an objective function to measure its fitness. The selection stage chooses the fittest individuals to become parents for the next generation. During crossover, selected parents exchange parts of their genes—for example, splitting at the 8th gene position—to produce new offspring. Mutation is applied to maintain genetic diversity by randomly altering certain genes, such as changing a bit from 0 to 1 or vice versa. After crossover and mutation, a new population of 10 individuals is formed. This cycle continues for five generations, after which the best individual across all generations is taken as the optimal solution.

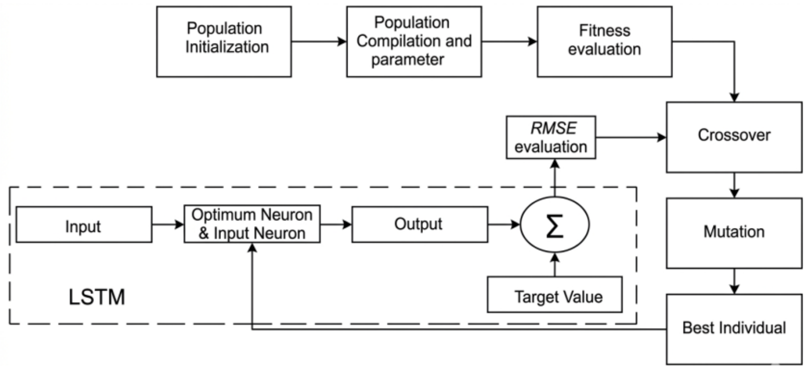


Figure 3. Flowchart of LSTM optimization with GA

2.4. Population size, generations, and gene length

In GA, population size, number of generations, and gene length are key parameters for solution optimization [2], [31]. A population size of 10 means each generation contains 10 candidate solutions. A generation size of 5 indicates that the GA runs for five iterations before terminating, with selection, crossover, and mutation conducted in each cycle. A gene length of 15 represents that each individual is encoded as a chromosome with 15 genes, where each gene contributes to the solution. This gene length corresponds to the number of parameters the GA optimizes [32].

In the context of GA, population, generation, and gene length are essential parameters for optimizing the solution [31]. Here is an explanation of each based on the given provisions. Population size =10 means that each generation has 10 individuals (potential solutions) to be evaluated. Each individual in the population represents a possible solution to the problem being solved. At the beginning of the algorithm, this population will be initialized randomly. Generation size =5 is the number of generations, indicating how many cycles or iterations GA will carry out. In this case, after 5 generations, the algorithm stops, and the best solution found will be considered the optimal solution. The selection process, crossover (gene exchange), and mutation will be carried out in each generation to produce a new population. Gene length =15, is each individual in the population represented by a chromosome consisting of a series of genes. The gene length indicates how many variables or bits are in the chromosome. In this case, each chromosome has a length of 15 genes, meaning 15 values in each individual affect the solution. The length of this gene can be interpreted as the number of parameters optimized by the GA [32].

2.5. Long short-term memory optimization equation using genetic algorithms

GA is used to optimize parameters such as weight W and bias b in both hidden layers [33]. The LSTM model predicts marine data with (1). Where $\hat{y}_{marine}$ is the predictive output; X_{marine} is the input data, $W$and b, is the GA-optimized parameters.

$$\hat{y}_{marine} = LSTM\ (X_{marine}, W, b) \tag{1}$$

2.6. Main steps of optimization with genetic algorithms

Population initialization is that each individual in the GA population represents a complete set of weights W and bias b in both hidden layers. The optimized weight matrix according to (2) and the optimized bias according to (3).

$$w_f^{(1)}, w_i^{(1)}, w_C^{(1)}, w_0^{(1)}, w_f^{(2)}, w_i^{(2)}, w_C^{(2)}, w_0^{(2)} \tag{2}$$

$$b_f^{(1)}, b_i^{(1)}, b_C^{(1)}, b_0^{(1)}, b_f^{(2)}, b_i^{(2)}, b_C^{(2)}, b_0^{(2)} \tag{3}$$

The fitness function is used to assess how well each individual (a set of weights and biases) performs, typically using error measures such as MAE, as shown in (4), or RMSE. The GA aims to minimize this error. MAE is selected as the primary metric in studies such as dynamic optimization using LSTM and GA for predicting oceanographic data for several technical and practical reasons. First, MAE is robust to outliers because it calculates the average absolute difference between predicted and actual values, making it more stable than RMSE, which is highly sensitive to extreme values. This is important for marine data, which often contains natural anomalies caused by storms or sudden ocean dynamics.

Second, MAE is easier to interpret since it retains the same unit as the original data, allowing researchers and practitioners to understand the average error directly (e.g., “an average error of 0.5 meters”).

This is beneficial for real-time marine applications that rely on practical error magnitudes rather than squared penalties. Third, MAE aligns well with the optimization objective. When used as the objective function in GA, MAE supports a more linear and stable search for solutions, preventing the algorithm from overemphasizing extreme errors as RMSE does. This makes it more suitable when the aim is to reduce overall global error rather than focus solely on severe deviations. Fourth, MAE is appropriate for oceanographic time-series data, where errors are often irregular and non-Gaussian [34]. Unlike RMSE, MAE does not rely on assumptions about error distribution. During selection, the best individual is chosen based on its fitness, using methods such as tournament selection. In the crossover stage, two parent individuals exchange parts of their weights and biases to produce new offspring, as described in (5).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - x| \quad (4)$$

Where n is the number of data points, and x_i and x denote the forecasted and actual values, respectively.

$$W_{child} = \alpha W_{parent1} + (1 - \alpha)W_{parent2} \quad (5)$$

After crossover, some weights and biases may undergo small mutations to maintain diversity in the population according to (6).

$$W_{mutasi} = W_{child} + random\ noise \quad (6)$$

GA continues to run selection, crossover, and mutation until it reaches a satisfactory result or a certain number of iterations. The LSTM and GA methods can be concluded using the parameters of the number of neurons in LSTM, using selection, crossover, and mutation to find the best configuration with the configuration, the size of the individual has 2 parameters, namely the number of neurons in two layers of LSTM. Population size, 10 individuals in each generation. Number of generations, 5 is used to find the best configuration, which aims to optimize the number of neurons in LSTM.

3. RESULTS AND DISCUSSION

3.1. Sea surface temperature prediction with long short-term memory and genetic algorithms

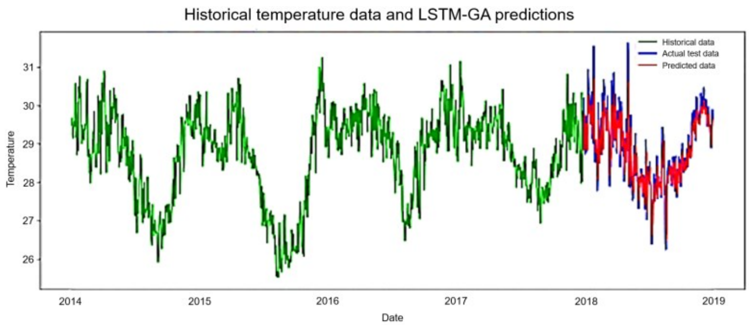
The SST prediction using the LSTM model optimized with a GA was carried out on the 2014–2018 time-series dataset. The model uses two hidden layers with 99 and 44 neurons and was trained for 10 epochs. The prediction results, illustrated in Figure 4, compare historical data, actual test data, and the LSTM-GA predicted values. In Figure 4(a), historical SST data, actual test data, and LSTM-GA prediction results. In Figure 4(b), a comparison between actual SST values and predicted SST values using the LSTM-GA model.

Overall, the LSTM-GA model is able to capture the main temperature patterns, especially the general upward and downward trends. The model also replicates major peaks and troughs reasonably well. However, it responds slightly late to rapid changes and occasionally deviates from the actual values, leading to minor overestimation or underestimation during sharp fluctuations. The use of only 10 epochs may limit performance for such highly variable data, and additional training could enhance accuracy. The two-layer architecture with 99 and 44 neurons helps the model learn both long-term and short-term patterns, but further refinement is needed for better adaptation to complex SST dynamics.

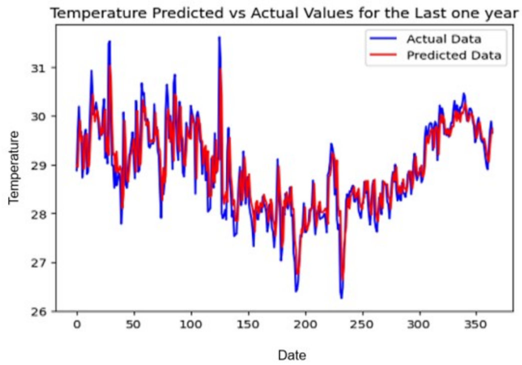
3.2. Sea surface salinity prediction with long short-term memory and genetic algorithms

The SSS prediction uses an LSTM model trained on 2014–2018 time-series data, with two hidden layers containing 87 and 99 neurons. These neuron numbers were selected experimentally to balance model complexity, avoid underfitting from too few neurons, and prevent overfitting from too many neurons. The second hidden layer has more neurons because it processes more complex feature representations. The model was trained for 10 epochs, meaning the entire dataset was passed through the network ten times to optimize weights while avoiding under- or overfitting.

Using a GA to optimize the model, the LSTM-GA predictions were compared with historical, actual, and predicted SSS data, as shown in Figure 5, Figure 5(a) shows the predicted SSS results using the LSTM-GA model based on time series data from 2014–2018. Figure 5(b) shows a comparison between the actual data and the model predictions, indicating that the model follows the main salinity patterns quite well. The model generally follows the main patterns of the actual SSS, capturing seasonal trends, although it struggles with sharp fluctuations, particularly in late 2018, where predictions lag slightly behind rapid changes. In more stable periods (2018–2019), the model performs well, with only minor deviations in peak and trough points.

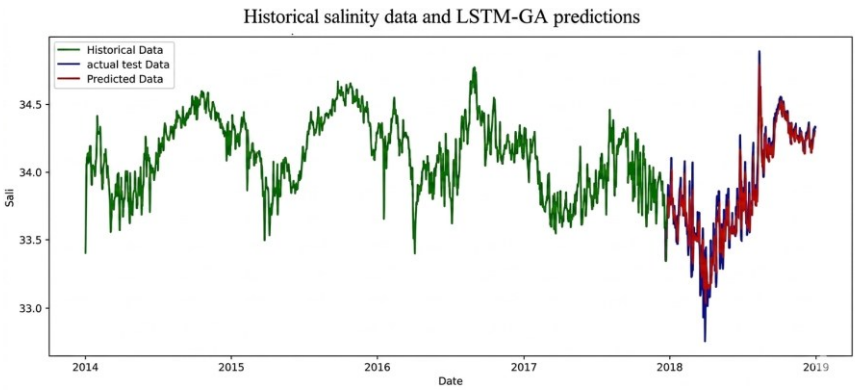


(a)

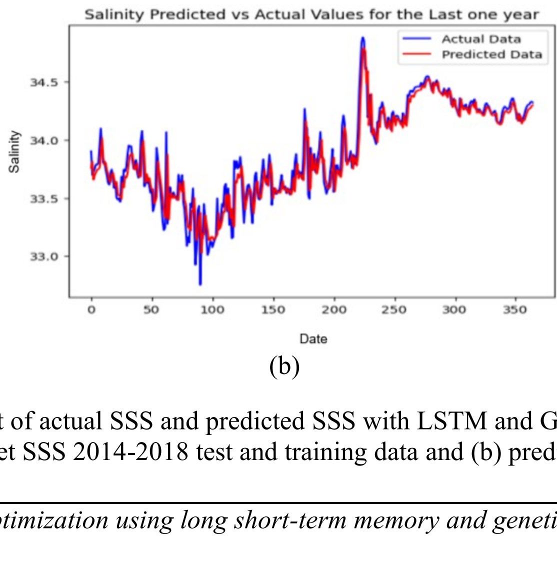


(b)

Figure 4. Comparison graph of actual SST and predicted SST with LSTM and GA prediction for one year of (a) original SST dataset 2014-2018 test and training data and (b) predicted SST test data



(a)



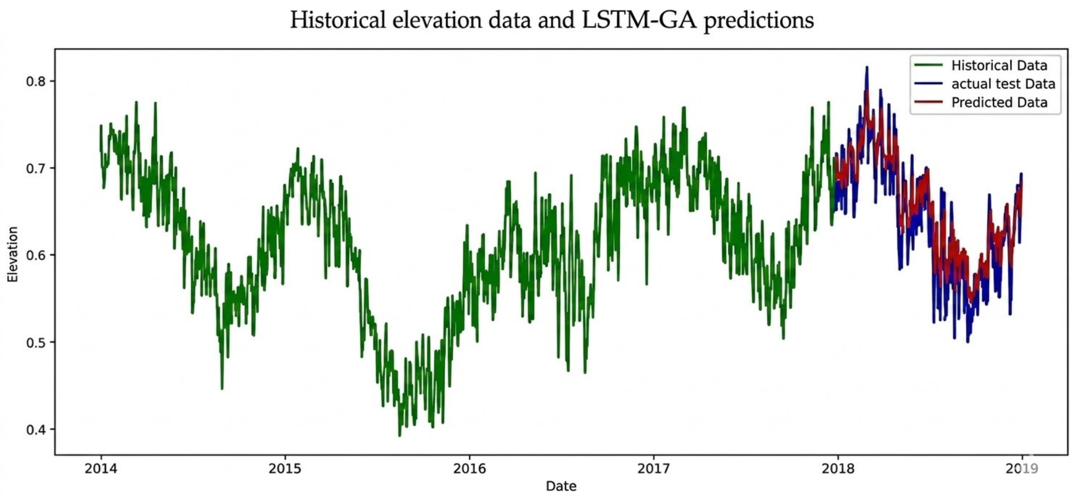
(b)

Figure 5. Comparison chart of actual SSS and predicted SSS with LSTM and GA prediction for one year of (a) original dataset SSS 2014-2018 test and training data and (b) predicted SSS test data

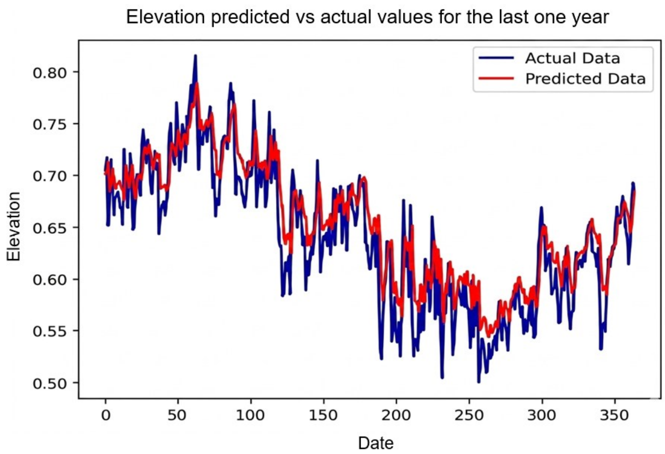
3.3. Sea surface height prediction with long short-term memory and genetic algorithms

SSH timeseries input data 2014–2018, hidden layer 1 =90 neurons, hidden layer 2 =98 neurons, epoch =10, using GA for neural network optimization. Elevation prediction using the LSTM model is optimized with GA by comparing historical, actual, and predicted data. Performance analysis of LSTM and GA in SSH prediction. With LSTM and GA, the performance analysis of SSH is the global pattern suitability in general; the LSTM model successfully follows the significant development of test data. The model successfully replicates long-term elevation fluctuations. The model is good at capturing general patterns similar to actual test data. Prediction results of the LSTM model optimized with GA for elevation data, by comparing historical data, actual data from the test set, and predicted data from the LSTM and GA models, SSH prediction, with LSTM and GA, are shown in Figure 6. Historical, actual, and predicted SSH elevation data using the LSTM model optimized with GA, as shown in Figure 6(a). Performance analysis of the LSTM-GA model for SSH elevation prediction by comparing actual data and predicted data, as shown in Figure 6(b).

The difference in hourly fluctuations is that around 2018, several sharp fluctuations in the actual test data (blue line) were challenging to follow by the model prediction (red line). This often happens in time-series predictions when the model cannot fully capture fast dynamics. The stable part in 2018–2019 is at the end; the LSTM prediction is more stable and closer to the test data. This shows that the model can capture long-term developments quite well after going through sharp initial fluctuations.



(a)

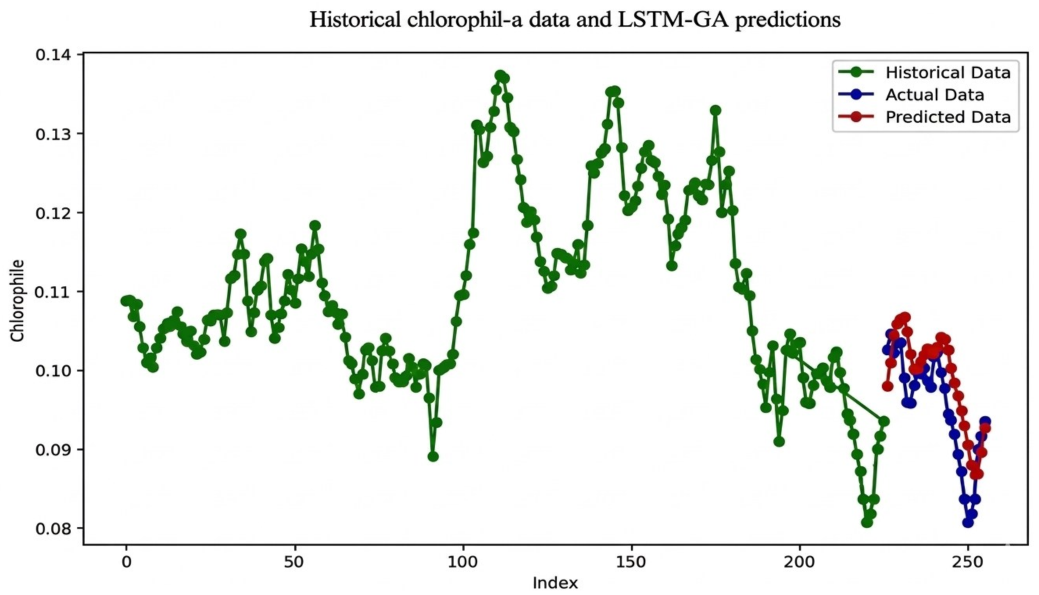


(b)

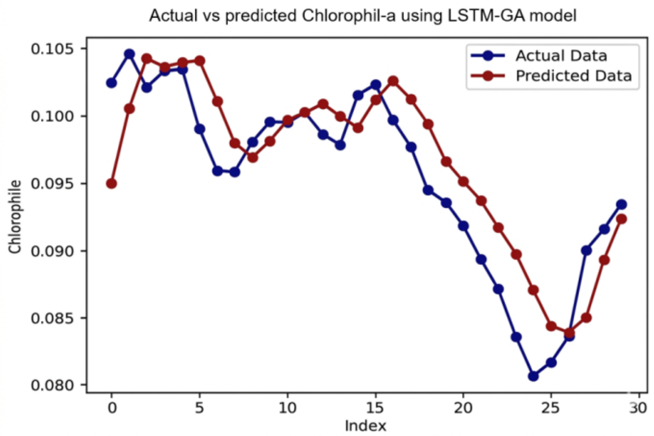
Figure 6. Comparison chart of actual SSH and predicted SSH with LSTM and GA prediction for one-year of (a) original SSH dataset 2014-2018 training and test data and (b) SSH test data

3.4. Marine chlorophyll-a data prediction with long short-term memory and genetic algorithms

Marine Chl-a prediction using an LSTM model optimized with a GA was conducted on Chl-a timeseries data from 2014–2018, using two hidden layers of 83 neurons each and 10 training epochs. The GA-based optimization improved the network's ability to model Chl-a concentration dynamics. As shown in Figure 7, the model's predicted values closely follow both the historical dataset (Figure 7(a)) and the actual test data (Figure 7(b)), demonstrating good predictive capability despite several minor deviations. These discrepancies may be attributed to data noise or the need for additional training epochs, more neurons, or extended historical data in future experiments. Overall, the LSTM–GA model successfully captures the fluctuation patterns of Chl-a concentrations, particularly the sharp variations appearing near the end of the historical record.



(a)



(b)

Figure 7. Comparison graph of actual Chl-a and predicted Chl-a with LSTM and GA for one-year prediction of (a) original Chl dataset 2014-2018 training and test data and (b) Chl-a test data

3.5. T-test

The t-test comparison between the LSTM and LSTM-GA predictive models shows no statistically significant performance difference, as indicated by the p-value of 0.720 in Table 1—well above the 0.05 threshold. This result demonstrates that incorporating a GA into the LSTM model does not yield a statistically meaningful improvement. Although LSTM-GA performs reasonably well, it still exhibits higher errors during abrupt oceanographic fluctuations, such as sudden salinity changes or extreme weather events. These findings suggest that GA integration alone is insufficient for substantial enhancement, and future studies may require additional exogenous variables (e.g., wind or currents) to improve prediction accuracy under highly dynamic conditions.

Table 1. T-test results

T-test results	LSTM	LSTM + GA
LSTM		0.0720
LSTM + GA	0.072	LSTM+GA

3.6. Prediction performance and comparison of prediction models based on mean absolute error

The performance of marine data prediction using LSTM-GA was evaluated through MAE across different train–test splits (70:30, 80:20, and 90:10) for SSS, SSH, SST, and Chl-a. As shown in Table 2, the MAE values remain consistently low across all configurations, indicating that the LSTM-GA model maintains stable predictive accuracy even with reduced training data proportions [35]. A comparison of autoregressive integrated moving average (ARIMA), LSTM, and LSTM-GA models using MAE, presented in Table 3, shows that ARIMA performs poorly due to its linear design and inability to capture long-term temporal dependencies in marine time-series data. LSTM significantly improves prediction accuracy by learning nonlinear sequential patterns. LSTM-GA further refines performance through parameter optimization, although it requires caution to avoid overfitting. Overall, LSTM and LSTM-GA outperform ARIMA, demonstrating their suitability for complex oceanographic variable prediction.

Table 2. Performance of marine data prediction with LSTM and GA algorithms based on MAE values

LSTM + GA	Training and testing data configuration		
	70-30%	80-20%	90-10%
SSS	0.0024	0.0026	0.0026
SSH	0.0064	0.0048	0.0042
SST	0.0059	0.0053	0.0042
Chl-a	0.0085	0.0048	0.0032

Table 3. Comparison of prediction models based on MAE

Variables	MAE		
	ARIMA	LSTM	LSTM-GA
SSS	0.781	0.026	0.024
SSH	0.344	0.061	0.048
SST	0.516	0.053	0.050
Chl-a	0.093	0.014	0.048

4. CONCLUSION

The t-test results indicate that the combination of LSTM-GA has not produced a statistically significant difference compared to pure LSTM (p-value =0.720). Despite this, incorporating the GA still offers advantages in model optimization. GA is a method that searches widely to improve the settings of the LSTM model, such as the number of neurons, learning rate, and batch size. This optimization process aims to enhance model accuracy and mitigate the risk of getting trapped in local minima during training. The advantages of integrating GA become more evident in complex data scenarios or when conducting repeated experiments that require thorough parameter tuning. GA facilitates a more comprehensive exploration of the solution space and is not reliant on a manual trial-and-error approach, thereby increasing efficiency in the model development process.

ACKNOWLEDGMENTS

Support for this research was provided by the School of Data Science, Mathematics and Informatics, IPB University, and the Faculty of Computer Science, Bhayangkara Jaya University, which has given the author the opportunity to work and produce this paper.

FUNDING INFORMATION

The authors declare that this research was conducted independently and did not receive any specific grant, financial support, sponsorship, or external funding from public, commercial, or non-profit organizations. All stages of the study, including research design, data collection, data analysis, interpretation of results, and manuscript preparation, were carried out solely by the authors without any external financial assistance. The authors also confirm that no funding agency had any involvement in the conduct of the research or the decision to publish the findings presented in this study.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Mukhlis	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	
Indra Jaya	✓	✓		✓		✓	✓			✓	✓			
Sri Nurdiati	✓		✓	✓	✓					✓				
Karlisa Priandana	✓	✓			✓	✓				✓	✓	✓	✓	✓
Irman Hermadi	✓	✓	✓	✓			✓	✓		✓	✓	✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The author declares that there are no known conflicts of interest associated with this publication. To the best of the author's knowledge, no financial, personal, professional, or institutional relationships exist that could have influenced the design, conduct, interpretation, or reporting of the work in any inappropriate manner. The author also confirms that this manuscript was prepared independently and objectively, without any external pressure or interest that might compromise its integrity.

INFORMED CONSENT

Prior to their participation in this study, informed consent was obtained from all individuals included, after they had been given a clear explanation of the study objectives, procedures, potential risks, and benefits, as well as assurance that their participation was voluntary and that they could withdraw at any time without any consequence.

ETHICAL APPROVAL

This study did not involve human participants or animal subjects, and no experiments, observations, measurements, or interventions were performed on human beings throughout the research process. Therefore, the authors did not seek approval from an institutional review board or ethics committee, as ethical review was not applicable to this research and no procedures were conducted that required such authorization.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [IJ], upon reasonable request.

REFERENCES

[1] M. Kim, H. Yang, and J. Kim, "Sea surface temperature and high water temperature occurrence prediction using a long short-term memory model," *Remote Sensing*, vol. 12, no. 21, pp. 1–21, 2020, doi: 10.3390/rs12213654.
[2] H. Cen *et al.*, "Applying deep learning in the prediction of chlorophyll-a in the East China Sea," *Remote Sensing*, vol. 14, no. 21, 2022, doi: 10.3390/rs14215461.

[3] X. Chen *et al.*, "A neural speech decoding framework leveraging deep learning and speech synthesis," *Nature Machine Intelligence*, vol. 6, no. 4, pp. 467–480, 2024, doi: 10.1038/s42256-024-00824-8.
[4] R. Teixeira, A. Cerveira, E. J. S. Pires, and J. Baptista, "Enhancing weather forecasting integrating LSTM and GA," *Applied Sciences*, vol. 14, no. 13, pp. 1–23, 2024, doi: 10.3390/app14135769.
[5] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May. 2015, doi: 10.1038/nature14539.
[6] L. Wei and L. Guan, "Seven-day sea surface temperature prediction using a 3DConv-LSTM model," *Frontiers in Marine Science*, vol. 9, Dec. 2022, doi: 10.3389/fmars.2022.905848.
[7] U. Tangke and B. Senen, "Distribution of sea surface temperature and chlorophyll-a concentration its correlation with small pelagic fish catch in Dodinga Bay," in *IOP Conference Series: Earth and Environmental Science*, 2020, vol. 584, no. 1, pp. 1–8, doi: 10.1088/1755-1315/584/1/012020.
[8] H. Chen, T. Lu, J. Huang, X. He, and X. Sun, "An improved VMD–EEMD–LSTM time series hybrid prediction model for sea surface height derived from satellite altimetry data," *Journal of Marine Science and Engineering*, vol. 11, no. 12, 2023, doi: 10.3390/jmse11122386.
[9] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, "Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: comparison with machine learning approaches," *Energies*, vol. 11, no. 7, 2018, doi: 10.3390/en11071636.
[10] C. Butler and M. Crane, "Blockchain transaction fee forecasting: a comparison of machine learning methods," *Mathematics*, vol. 11, no. 9, pp. 1–26, 2023, doi: 10.3390/math11092212.
[11] J. Wang, Y. Zhou, L. Zhuang, L. Shi, and S. Zhang, "A model of maritime accidents prediction based on multi-factor time series analysis," *Journal of Marine Engineering and Technology*, vol. 22, no. 3, pp. 153–165, 2023, doi: 10.1080/20464177.2023.2167269.
[12] B. Shi *et al.,* "Sea surface temperature prediction using ConvLSTM-based model with deformable attention," *Remote Sensing*, vol. 16, no. 22, 2024, doi: 10.3390/rs16224126.
[13] D. Menaka and S. Gauni, "Prediction of dominant ocean parameters for sustainable marine environment," *IEEE Access*, vol. 9, pp. 146578–146591, 2021, doi: 10.1109/ACCESS.2021.3122237.
[14] M. J. Alizadeh and V. Nourani, "Multivariate GRU and LSTM models for wave forecasting and hindcasting in the southern Caspian Sea," *Ocean Engineering*, vol. 298, pp. 117–193, 2024, doi: 10.1016/j.oceaneng.2024.117193.
[15] S. Stajkowski, D. Kumar, P. Samui, H. Bonakdari, and B. Gharabaghi, "Genetic-algorithm-optimized sequential model for water temperature prediction," *Sustainability*, vol. 12, no. 13, 2020, doi: 10.3390/su12135374.
[16] Z. Niu *et al.*, "Recurrent attention unit: a new gated recurrent unit for long-term memory of important parts in sequential data," *Neurocomputing*, vol. 517, no. 1, pp. 1–9, 2023, doi: 10.1016/j.neucom.2022.10.050.
[17] S. Gu and Y. Yang, "A deep learning algorithm for the max-cut problem based on pointer network structure with supervised learning and reinforcement learning strategies," *Mathematics*, vol. 8, no. 2, 2020, doi: 10.3390/MATH8020298.
[18] V. Tomar, M. Bansal, and P. Singh, "Metaheuristic algorithms for optimization: a brief review," *Engineering Proceedings*, vol. 59, no. 1, pp. 1–16, 2023, doi: 10.3390/engproc2023059238.
[19] F. Xie, H. Yan, Y. Long, H. Guo, H. Liu, and P. Yu, "Weather prediction based on multivariate LSTM neural network model," in *Advances in Transdisciplinary Engineering*, 2024, vol. 47, pp. 298–303, doi: 10.3233/ATDE231201.
[20] L. Wei, L. Guan, L. Qu, and D. Guo, "Prediction of sea surface temperature in the China seas based on long short-term memory neural networks," *Remote Sensing*, vol. 12, no. 17, pp. 1–20, 2020, doi: 10.3390/RS12172697.
[21] D. Shen, S. Bao, L. J. Pietrafesa, and P. Gayes, "Improving numerical model predicted float trajectories by deep learning," *Earth and Space Science*, vol. 9, no. 9, pp. 1–13, 2022, doi: 10.1029/2022EA002362.
[22] L. Wang *et al.*, "An effective algorithm for offshore air temperature prediction with LSTM neural network and wavelet decomposition and reconstruction," *2022 3rd International Conference on Signal Processing and Computer Science*, 2022, doi: 10.1088/1742-6596/2414/1/012016.
[23] Y. Fu, J. Song, J. Guo, Y. Fu, and Y. Cai, "Prediction and analysis of sea surface temperature based on LSTM-transformer model," *Regional Studies in Marine Science*, vol. 78, 2024, doi: 10.1016/j.rsma.2024.103726.
[24] M. Kotyrba, E. Volna, H. Habiballa, and J. Czyz, "The influence of genetic algorithms on learning possibilities of artificial neural networks," *Computers*, vol. 11, no. 5, 2022, doi: 10.3390/computers11050070.
[25] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: the state of the art," *International Journal of Forecasting*, vol. 14, no. 1, pp. 35–62, 1998, doi: 10.1016/S0169-2070(97)00044-7.
[26] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
[27] J. H. Holland, *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*, Cambridge, Massachusetts: MIT Press, 1992.
[28] R. C. Staudemeyer and E. R. Morris, "Understanding LSTM -- a tutorial into long short-term memory recurrent neural networks," *arXiv:1909.09586*, Sep. 2019.
[29] A. Miller and J. I. Virmani, "Advanced marine technologies for ocean research," *Deep-Sea Research Part II: Topical Studies in Oceanography*, vol. 212, 2023, doi: 10.1016/j.dsr2.2023.105340.
[30] A. Hassanat, K. Almohammadi, E. Alkafaween, E. Abunawas, A. Hammouri, and V. B. S. Prasath, "Choosing mutation and crossover ratios for genetic algorithms-a review with a new dynamic approach," *Information*, vol. 10, no. 12, 2019, doi: 10.3390/info10120390.
[31] G. Moges, A. Alemu, and J. Ejepu, "Structural analysis of gravity and magnetic data: implication for groundwater study in the Ziway, Abijata and Langano lakes corridor and surroundings, central main Ethiopian rift," *Geophysical Journal International*, vol. 240, no. 3, pp. 1505–1522, Jan. 2025, doi: 10.1093/gji/ggae461.
[32] E. R.-Tamariz, M. A. Z.-Garcia, and R. Batres, "Optimization of a drum boiler startup using dynamic simulation and a micro-genetic algorithm," *Energy Reports*, vol. 6, pp. 410–416, 2020, doi: 10.1016/j.egyr.2019.11.095.
[33] J. Li, H. Fu, K. Hu, and W. Chen, "Data preprocessing and machine learning modeling for rockburst assessment," *Sustainability*, vol. 15, no. 18, 2023, doi: 10.3390/su151813282.
[34] D. Häfner, J. Gemmrich, and M. Jochum, "FOWD: a free ocean wave dataset for data mining and machine learning," *Journal of Atmospheric and Oceanic Technology*, vol. 38, no. 7, pp. 1305–1322, 2021, doi: 10.1175/JTECH-D-20-0185.1.
[35] A. B.-Folgmann, R. Roscher, S. Wenzel, B. Uebbing, and J. Kusche, "Sea level anomaly prediction using recurrent neural networks,"in *Proceedings of the 2017 conference on Big Data from Space*, 2017, pp. 1-4.

BIOGRAPHIES OF AUTHORS

Mukhlis, S.Kom., M.T. is a permanent lecturer in Informatics at Universitas Bhayangkara Jakarta Raya and is currently completing a doctorate in Computer Science at IPB University. He actively teaches and guides students in Informatics Engineering. He is recognized in the AD scientific index with an H-index of 10 and 467 citations in engineering and technology. He can be contacted at email: clisfadhlimukhlis@apps.ipb.ac.id or mukhlis@dsn.ubharajaya.ac.id.

Prof. Dr. Indra Jaya, M.Sc. is a professor of Underwater Acoustics at IPB University. He is recognized as a highly productive and innovative researcher. He earned his doctorate in Underwater Acoustics/Marine Studies from the University of Delaware, United States, in 1996, followed by a one-year postdoctoral program at the Department of Mathematics, Rensselaer Polytechnic Institute (RPI), New York. He can be contacted at email: indrajaya@apps.ipb.ac.id.

Sri Nurdiati earned her bachelor's degree in Statistics from IPB University, her master's degree in Computer Science from the University of Western Ontario, Canada, and her doctoral degree in Applied Mathematics from Twente University, Netherlands. She is currently a professor of mathematical science at the Mathematics Study Program, Faculty of Mathematics and Natural Sciences, IPB University, with extensive research contributions in applied mathematics and network science. She can be contacted at email: nurdiati@apps.ipb.ac.id.

Karlisa Priandana is a senior member of IEEE. She received her bachelor's degree from Institut Teknologi Bandung in 2008, her master's degree from Tokyo Institute of Technology in 2010, and her Ph.D. degree from Universitas Indonesia in 2017. She is an associate professor in the Computer Science Study Program at IPB University. Her research interests include robotics and artificial intelligence. She is an active member of IEEE Indonesia. She can be contacted at email: karlisa@apps.ipb.ac.id.

Irman Hermadi received his bachelor's degree in Computer Science from IPB University in 1999, his master's degree in Computer Science from King Fahd University of Petroleum and Minerals in 2004, and his Ph.D. degree in Computer Science from the University of New South Wales, Australia, in 2015. He has been a faculty member of the Computer Science Study Program at IPB University since 2004. His research interests include software engineering, computational intelligence, and information systems. He can be contacted at email: irmanhermadi@apps.ipb.ac.id.