Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer
Vol. 24, No. 1, November 2024, pp. 179∼194
ISSN: 2476-9843, accredited by Kemenristekdikti, Decree No: 200/M/KPT/2020
DOI: 10.30812/matrik.v24i1.4032

Integration of Deep Learning and Autoregressive Models for Marine Data Prediction

Mukhlis [1], Puput Yuniar Maulida[3], Achmad Mujib[2], Adi Muhajirin[1], Alpi Surya Perdana[4]
[1]Universitas Bhayangkara Jakarta Raya, Jakarta, Indonesia
[2]IPB University, Bogor, Indonesia
[3]Universitas Pendidikan Indonesia, Bandung, Indonesia
[4]Universitas Majalengka, Majalengka, Indonesia

Article Info

Article history:

Received May 06, 2024
Revised September 16, 2024
Accepted November 25, 2024

Keywords:

Algorithm
Correlation
Deep Learning
Mean Absolute Error

ABSTRACT

Climate change and human activities significantly affect the dynamics of the marine environment, making accurate predictions essential for resource management and disaster mitigation. Deep learning models such as Long Short-Term Memory excel at capturing non-linear temporal patterns, while autoregressive models handle linear trends to improve prediction accuracy. This aim study predicts sea surface temperature, height, and salinity using deep learning compared to Moving Average and Autoregressive Integrated Moving Average methods. The research methods include spatial gap analysis, temporal variability modeling, and oceanographic parameter prediction. The relationship between parameters is analyzed using the Pearson Correlation method. The dataset is divided into 80and 20and Autoregressive models. The results show that Long Short-Term Memory performs best with a Root Mean Squared Error of 0.1096 and a Mean Absolute Error of 0.0982 for salinity at 13 sample points. In contrast, Autoregressive models produce a Root Mean Squared Error of 0.193 for salinity, 0.055 for sea surface height, and 2.504 for sea surface temperature, with a correlation coefficient 0.6 between temperature and sea surface height. In conclusion, the Long Short Term Memory model excels in predicting salinity because it is able to capture complex non-linear patterns. Meanwhile, Autoregressive models are more suitable for linear data trends and explain the relationship between parameters, although their accuracy is lower in salinity prediction. This approach shows great potential in improving the accuracy of complex ocean parameter predictions.

Corresponding Author:

Mukhlis, +62858-9269-1149
Faculty Computer Science
Universitas Bayangkara Jakarta Raya, Jakarta, Indonesia / IPB University, Bogor, Indonesia
Email: clisfadhlimukhlis@apps.ipb.ac.id

How to Cite:
M. Mukhlis, P. Maulidia, A. Mujib, A. Muhajirin, and A. Perdana, "Integration of Deep Learning and Autoregressive Models for Marine Data Prediction", *MATRIK: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer*, Vol. 24, No.1, pp. 179-194, November, 2024.

1. INTRODUCTION

The LSTM (Long Short Term Memory) architecture was developed to predict the behavior of nonlinear time series, i.e., various architectures are available and have been used in different application areas such as image processing, manufacturing, or autonomous systems. Recurrent neural networks and exceedingly Long short-term memory (LSTM) have been investigated intensively in recent years due to their ability to model and predict nonlinear time-variant system dynamics. The present paper delivers a comprehensive overview of existing LSTM cell derivatives and network architectures for time series prediction [1]. Time Series Forecasting (TSF) is used to predict the target variables at a future time point based on the learning from previous time points. To keep the problem tractable, learning methods use data from a fixed-length window from the past as explicit input. In this paper, we study how the performance of predictive models changes as a function of different look-back window sizes and different amounts of time to predict the future. Transformer models, which have succeeded in the image and natural language processing domains [2]. Extensive and high-dimensional data processing methods such as machine learning and deep learning have rapidly developed in the technological era—spatial-temporal linkages in Big Data. The characteristics of the time and space domains are diverse and highly dimensional, which leads to more significant cpagelenges in Big Data analysis, especially marine data. In the Big Data era, data assimilation could integrate all these sources of Big Earth Data to achieve a more spatiotemporally and physically consistent representation of the Earth system. The estimation scheme consists of an LSTM model in a Generative Adversarial Network (GAN) configuration, which is sampled multiple times and yields a single trajectory prediction with uncertainty. The performance of the estimation scheme is demonstrated and compared against two other commonly used methods, showing that the probabilistic heat map provides valuable information compared to the baseline methods [3]. Ocean circulation, including ocean currents and systems of ocean currents, such as ocean gyres and the meridional overturning circulation, play a vital role in the climate system by redistributing heat, freshwater, carbon, and ecosystem-relevant quantities. Some of these systems of ocean currents are on a large spatial scale and of global climate relevance. Remote sensing data's temporal and spatial continuity is flawed due to cloudiness, sensor malfunction, or atmospheric pollution. Different methods have been presented to estimate missing values in remote sensing data. In this study, we evaluate the performance of a spatiotemporal gap-filling algorithm. This algorithm is exciting and worthy of further evaluation because it achieves high accuracy while maintaining a considerably low computational complexity.

Forecasting will never be 100% accurate because the future has a problem of uncertainty. However, using the right method can give forecasting a low error rate value to provide a good forecast for the future. This study aims to determine the effect of increasing the number of hidden layers and neurons on the performance of the long short-term Memory (LSTM) forecasting method. In various architectural scenarios, LSTM performance measurement is done by root mean square error (RMSE). The LSTM algorithm can handle long-term dependencies on its input and can predict data for a relatively long time. Based on research conducted from all models, the best results were obtained with an RMSE value of 0.699 obtained in model 1 with the number of hidden layers 2 and 64 neurons [4]. The method chosen in this study is LSTM recurrent neural network as one of the best algorithms that perform better in predicting time series. The LSTM models in this study were used to compare the performance between modeling using meteorological factors and without meteorological factors [5]. Time Series Forecasting (TSF) is used to predict the target variables at a future time point based on the learning from previous time points. There are many variants of RNN, such as RNN itself, long-short-term memory (LSTM), and gated recurring unit, so it is often debated which algorithm from the RNN family has the most efficiency and optimal computation time. When developing a prediction system, sequential or time series data is needed to make accurate predictions. Sequential or time series data involves data arranged in a time sequence, such as weather data, financial data, carbon emission data, and traffic data recorded over time. The research aims to predict the speed of ocean surface currents and their directions using LSTM. There are many prediction methods, one of which is Long Short-term Memory (LSTM). The working principle of LSTM is to process information from previous memory through three gates, namely the forget gate, input gate, and output gate, to produce an output that will be input to the next process. Based on trials with several parameters including Hidden Layer, Learning Rate, Batch Size, and Learning rate drop period, the smallest MAPE value was obtained in the U component and V component of 14.15% and 8.43% with hidden layers parameters 50, Batch size 32 and Learn rate drop 150 [6]. Accurately forecasting energy metrics is essential for efficiently managing renewable energy generation. Given the high variability in load and renewable energy power output, this represents a crucial area of research to pave the way for increased adoption of low-carbon energy solutions. While the impact of different neural network architectures and algorithmic approaches has been researched extensively, the impact of utilizing additional weather variables in forecasts has received far less attention. This article demonstrates that weather variables can significantly influence energy forecasting and presents methodologies for using these variables within a long short-term memory (LSTM) architecture to improve forecasting accuracy. A variety of LSTM architectures and hyperparameters were investigated. Sea level was predicted using both techniques up to the end of 2023. Performance indicators, including Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE), were employed to assess the quality of both prediction models. The result shows that the ARIMA (1,1,4) model is more effective in predicting sea level than the LSTM. The

MAE, MAPE, and RMSE values for ARIMA (1,1,4) are 7.19, 4.86%, and 10.35, respectively [7]

One approach used for marine data prediction is deep learning with the Long Short-Term Memory (LSTM) algorithm. The prediction results using LSTM show better accuracy than the ARIMA method's. The key to LSTM design is incorporating non-linear control and data dependency into the RNN cells, which are trained to ensure that the gradient of the objective function does not vanish. LSTM addresses the vanishing gradient problem when the gradient value is close to or equal to zero. LSTM uses a gating mechanism to handle information stored or removed from memory. Each LSTM cell processes input from the previous state (h(t-1)) and the current input (x(t)). LSTM combines the last state, the current memory, and the new input, making it very efficient in capturing long-term dependencies, Sea surface temperature (SST) prediction has received increasing attention in recent years due to its paramount importance in the various fields of oceanography. Studies have shown that neural networks are particularly effective in making accurate SST predictions by efficiently capturing spatiotemporal dependencies in SST data. Among different models, the ConvLSTM framework is notably prominent. This model skillfully combines convolutional neural networks (CNNs) with recurrent neural networks (RNNs), enabling it to capture spatiotemporal dependencies within a single computational framework simultaneously. To overcome the limitation that CNNs primarily capture local spatial information, we propose a model named DatLSTM that integrates a deformable attention transformer (DAT) module into the ConvLSTM framework, thereby enhancing its ability to effectively process more complex spatial relationships. Specifically, the DAT module adaptively focuses on salient features in space, while ConvLSTM further captures the temporal dependencies of spatial correlations in the SST data. This way, DatLSTM can adaptively capture complex spatiotemporal dependencies between the preceding and current states within ConvLSTM. To evaluate the performance of the DatLSTM model, we conducted short-term SST forecasts in the Bohai Sea region with forecast lead times ranging from 1 to 10 days. We compared its efficacy against several benchmark models, including ConvLSTM, PredRNN, TCTN, and SwinLSTM. Our experimental results show that the proposed model outperforms all of these models in terms of multiple evaluation metrics for short-term SST prediction. The proposed model offers a new predictive learning method for improving the accuracy of spatiotemporal predictions in various domains, including meteorology, oceanography, and climate science [8]. Sea surface temperature (SST) prediction has received increasing attention in recent years due to its paramount importance in the various fields of oceanography. Studies have shown that neural networks are particularly effective in making accurate SST predictions by efficiently capturing spatiotemporal dependencies in SST data. Among different models, the ConvLSTM framework is notably prominent. This model skillfully combines convolutional neural networks (CNNs) with recurrent neural networks (RNNs), enabling it to capture spatiotemporal dependencies within a single computational framework simultaneously. To overcome the limitation that CNNs primarily capture local spatial information, we propose a model named DatLSTM that integrates a deformable attention transformer (DAT) module into the ConvLSTM framework, thereby enhancing its ability to effectively process more complex spatial relationships. Specifically, the DAT module adaptively focuses on salient features in space, while ConvLSTM further captures the temporal dependencies of spatial correlations in the SST data. This way, DatLSTM can adaptively capture complex spatiotemporal dependencies between the preceding and current states within ConvLSTM. To evaluate the performance of the DatLSTM model, we conducted short-term SST forecasts in the Bohai Sea region with forecast lead times ranging from 1 to 10 days. We compared its efficacy against several benchmark models, including ConvLSTM, PredRNN, TCTN, and SwinLSTM. Our experimental results show that the proposed model outperforms all of these models in terms of multiple evaluation metrics for short-term SST prediction. The proposed model offers a new predictive learning method for improving the accuracy of spatiotemporal predictions in various domains, including meteorology, oceanography, and climate science. Prediction of oceanographic variables such as sea surface temperature (SST), sea surface salinity (SSS), sea surface height (SSH), and chlorophyll-a (Chl-a) concentrations is of great importance in oceanography and environmental research. These variables influence marine ecosystems, global weather, and climate patterns, and a better understanding of their temporal changes can provide essential insights for weather prediction, fisheries management, and climate change mitigation. However, the prediction of oceanic data is cpagelenging due to its complex, multivariate, and often nonlinear characteristics. The data is usually heavily influenced by external factors such as wind patterns, tidal cycles, and global climate phenomena such as El Niño and La Niña. This is where Long Short-Term Memory (LSTM) becomes essential [8] Chlorophyll-a concentration in area of research ranging from 0,22 mg/m^3- 1,15 mg/m^3. Chlorophyll-a concentration value each month fluctuates follow wind of progress. Maximum value of chlorophyll-a concentration happening in wesh season and minimum value occurs in transitional season 2. Layang scad, banyar fish and eastern little tuna has a negative response to sea surface temperature especially in east season [9].

To support this, we propose a new timeTSFmachine-learning framework for the short-term forecasting of data oceanography conditions to support critical decision-making associated with marine operations. In contrast, Arlindo Flores transport variability is dominated by the annual period (AV); the Indonesian Throughflow (ITF) is an inter-ocean Pacific-Indian current system that passes Indonesian Seas, such as via the primary path of Makassar Strait to Flores Sea and via the secondary path of Lifamatola Strait to Banda Sea. Temporal data has a meaningful order. For example, in oceanographic data, the sea surface temperatures on day 1, day 2, and so on have related patterns. "Temporal context" refers to the relationship between observations made at different times—for example, the temperature on one day may be influenced by the temperature on the previous day [10]. LSTM is a neural network that handles sequential data, such as time data. LSTM can retain information from the past (for example, observations from days or months ago) and use this information to make predictions or understand the current context. We apply climate attribution techniques to sea surface temperature time series from five regional North Pacific ecosystems to track the growth in human influence on ocean temperatures over the past seven decades (1950–2022). Using Bayesian estimates of the Fraction of Attributable Risk (FAR) and Risk Ratio (RR) derived from 23 global climate models, we show that human influence on regional ocean temperatures could first be detected in the 1970s and grew until 2014–2020 temperatures showed overwhelming evidence of human contribution [11]. Research Gap: while previous studies have demonstrated the potential of LSTM and its variants in predicting time series data across various domains, few have focused on oceanographic variable prediction specifically addressing multivariate interactions between Sea Surface Temperature (SST), Sea Surface Salinity (SSS), Sea Surface Height (SSH), and chlorophyll-a (chl-a).

Furthermore, limitations remain in managing these variables' temporal and spatial continuity, particularly in sparse or inconsistent data scenarios. Prior research lacks a comprehensive examination of the effects of different prediction window sizes, an attention-based long short-term memory (LSTM) neural network approach is used to learn the short-term temporal patterns from in situ observations. This is then integrated with an existing, low computational cost spatial nowcasting model to develop a complete framework for spatiotemporal forecasting. The entire spatiotemporal forecasting system is demonstrated using a case study based on independent observation locations near the southwest coast of the United Kingdom. Results are validated against in situ data from two wave buoy locations within the domain and compared to operational physics-based wave forecasts from the Met Office (the United Kingdom's national weather service). For these two example locations, the spatiotemporal estimates have an accuracy of R2 = 0.9083 and 0.7409 in forecasting 1-h-ahead significant wave height and R2 = 0.8581 and 0.6978 in 12-h-ahead forecasts, respectively. Importantly, this represents respectable accuracy levels comparable to traditional physics-based forecast products but requires only a fraction of the computational resources [12]. The experimental fishing method for data collection, image analysis, statistical analysis, and interpretation is expected to answer the research objective for optimized wild fishery production. The research finds that simultaneously, sea surface temperature and chlorophyll-a significantly influence pelagic fish catch with an Fhit value of 90.403 and significance of 0.000.

In contrast, partial surface temperature and chlorophyll-a are the oceanographic factors that contribute to the distribution of small pelagic fish in Dodinga Bay, with an r-square value of 0.8525 for sea surface temperature and 0.6966 for chlorophyll-a [13]. Habitat suitability modeling is essential in the planning process of marine fisheries resource management in the future. The environmental conditions of the waters affect the habitat of skipjack tuna. This study aims to evaluate the habitat suitability model of skipjack tuna in the southern waters of West Java - Banten. Sea surface temperature and chlorophyll-a concentration data used in this modeling were obtained from MODIS satellite imagery from December 2018 to November 2019. Skipjack tuna catch location data were obtained from catch reports. This modeling was run using MaxEnt software. Sea surface temperature affects the model in the 28.3-29.2^{o}C range for the west season and 24.0-28.5^{o}C for the east season [14]. The modeling results show excellent performance with an accuracy level of 0.84 for the west season and 0.92 for the east season. To mitigate these losses, predicting such outbreaks to prevent or respond to them as early as possible is necessary. In the present study, we propose an HWT prediction method that applies sea surface temperatures (SSTs) and deep-learning technology in a long short-term memory (LSTM) model based on a recurrent neural network (RNN). The LSTM model predicts time series data for the target areas, including the coastal region from Goheung to Yeosu, Jeollanam-do, Korea, which has experienced frequent HWT occurrences in recent years. To evaluate the performance of the SST prediction model, we compared and analyzed the results of an existing SST prediction model for the SST data, as well as additional external meteorological data [15]. Pearson correlation helps identify significant relationships between parameters, aiding in developing machine learning-based predictive models. Pearson's correlation coefficient (PCC), traditionally used in statistics, was used to analyze the correlations among the eight data parameters. PCC is a numerical value quantifying the correlation between data X and Y. It has a value between $+1$ and -1, with a positive correlation for a positive value, a negative correlation for a negative value, and no correlation for a zero value. Generally, it is judged that there is a meaningful linear correlation with values of $+0.5$ or more or -0.5 or less [16] The data were analyzed using the moving average method and pearson-correlation. The results of this study indicated that the fishing season for skipjack, tuna mackerel, mackerel, and sailfish generally occurred in the east season. There was a close relationship between the production of skipjack and tuna mackerel on monthly fluctuations in sea surface temperature and chlorophyll-a values (p-value$<$0.05). An increase in the concentration of chlorophyll-a, and a decrease of sea surface temperature values would be followed by an increase in the production of skipjack and tuna mackerel in Palabuhanratu waters. The correlation index of skipjack to chlorophyll-a was in the strong category (r = 0.697), while that of tuna mackerel was in the medium category (r = 0.485).

This study utilizes a time series dataset of oceanographic data (sst, sss, ssh, and chl-a), focusing on the "Date" and "Value" features. It evaluates the performance of a spatiotemporal gap-filling algorithm, which is notable for its high accuracy and low computational complexity. The research examines how predictive model performance varies with different look-back window sizes and prediction timeframes. Data preparation includes preprocessing steps such as handling missing values, normalization, clustering points for prediction, correlating variables, implementing the LSTM model, and evaluating its RMSE performance compared to the ARIMA method for accuracy in oceanographic variable prediction. In this study, we hope to contribute to science and stakeholders in the marine and fisheries sector by identifying oceanographic data conditions—such as chlorophyll-a, sea surface temperature, sea level, and salinity—as potential fish habitats, using Landsat/MODIS satellite imagery and combined data as environmental information. In addition, they provide guidelines for related parties on the activities and resources needed to determine oceanographic parameters, which aim to predict future oceanographic variable areas through the RNN-LSTM approach and its optimization.

2. RESEARCH METHOD

The primary data needed for predictions is oceanographic parameter data collected in original data in files with the extension nc from the data-marine-Copernicus website. The data is cleaned using Ocean Data View and converted in the form of a Txt file into a Microsoft Excel file according to the location of the longitude and latitude. Then, the training and testing data will be selected before searching for neurons in the input layer, and modeling will be done using Recurrent Neural Network-Long short-term memory.

The research stages, according to Figure 1, start from taking environmental data sample points taken at each location, totaling 188 sample points for daily and weekly data using the Modis satellite, Landsat, and the Hycom application to extract the appropriate sample points from the imagery on the Google Earth Engine server. The following data collection is time series data on chlorophyll-a and sea surface temperature, salinity, and sea level height for 2014-2018, obtained via Google Earth Engine and www.oceancolor.gsfc.nasa.com. The data format chosen is NonConformance (.nc). The data was then extracted using the QGIS 3.10.2 application and cropped according to the coordinates of the research area.

Research Methods Used: 1) Quantitative Approach: (a) Numerical data is processed for prediction using machine learning methods, specifically Long Short-Term Memory (LSTM), which is an algorithm based on recursive neural networks (RNN). (b) The analysis is based on measuring the accuracy of the prediction results using numerical metrics such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). 2) Quantitative Research Stages, (a) Data collection: Taking time-series environmental data from Modis satellite, Landsat, Hycom, and the Google Earth Engine site. (b) Data processing: Cleaning data using Ocean Data View, QGIS, and converting data from .nc to .txt and Excel formats. (c) Data analysis: Splitting data into training data (80%) and testing (20%) for model testing. 3). Specific Methods: Algorithm, LSTM is used for time-series data analysis and prediction, which includes: (a) Creating a supervised learning model. (b) Determine the number of neurons in the hidden layer. (c) Determine the number of epochs. (d) Perform data prediction and accuracy evaluation. This quantitative study uses the LSTM algorithm to process numerical data and produce accurate predictions through a structured and data-driven process.

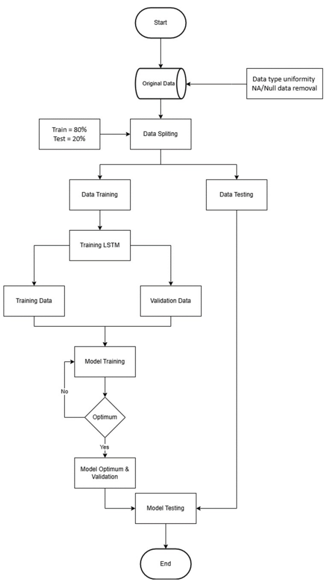

Figure 1. Research Stages

The research steps are as follows: 1) Dataset Determination: The dataset consists of three main variables, namely SST, SSH, and SSS, obtained in the 2014-2018 time span. Each variable is processed in a preprocessing stage, including data cleaning and division into training and test data. 2) Prediction Method: Two approaches are used to predict these oceanographic variables: (a) ARIMA (Auto Regression Integrated Moving Average): Used as a comparative method, ARIMA has long been used in time series forecasting. However, ARIMA has limitations in capturing complex seasonal patterns and nonlinear relationships between variables. (b) Long Short Term Memory): As a neural network-based model, LSTM can capture long-term relationships and seasonal patterns, which is expected to improve accuracy compared to ARIMA. 3) Model Performance Measurement: Model performance is measured using the Root Mean Square Error (RMSE), which provides an average measure of the model's prediction error in actual data units. A lower RMSE value indicates better model performance [17].

2.1. Data Acquisition

The Banda Sea is a sea in the Maluku Islands of Indonesia, connected to the Pacific Ocean but surrounded by hundreds of islands and the Pagemahera and Seram Seas. It is about 1000 km (600 miles) from east to west and about 500 km (300 miles) from north to south. The processes contributing to the mesohaline structure are studied using data from three ARLINDO cruises in 1993, 1994, and 1996. An inverse-model analysis using salinity and CFC-11 data is applied to a vertical section along the main flow path from the Makassar Strait to the Flores Sea and Banda Sea. The model reproduces the throughflow's seasonal and interannual variability and shows flow reversals in the vertical structure. This study examines the vertical structure of the Banda Sea circulation using the Hybrid Coordinate Ocean Model+Navy Coupled Ocean Data Assimilation Global Analysis (GLBa0.08) products.

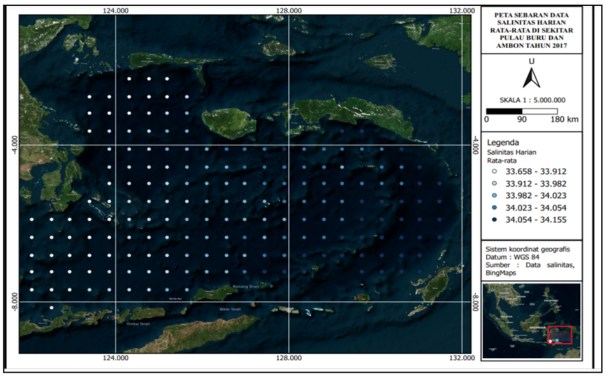

Figure 2. Salinity of sample points from 188 locations in the Banda Sea

The vorticity distribution shows a four-layer structure in the Banda Sea circulation, more robust in the upper 700 m than in the deeper layers. Circulation during the northwesterly monsoon is cyclonic in the surface layer and anticyclonic in the upper layer, opposite to its counterpart during the southeasterly monsoon [18]. With coordinates 6^{o}S 127^{o}E, starting Latitude -7.0999996 Longitude 131.916663 or -7^{o}05' 60. 00 "S 131^{o}54'59.99 "E. Environmental data sampling points were taken at each location as many as 188 sample points for weekly data using the Modis satellite. Meanwhile, the time of capture point data from the capture point data was carried out in 2014 and 2018. With measurement data of chlorophyll-a (chl-a) per 500 m, sea surface temperature (SST) per 500 m, and current salinity (sss) 0.080 (degrees) and sea level height (ssh) 0.080 (degrees), with a measurement of 1 degree = 111 km so 0.08 is less than 1000 m, which is around 888 m. Sample points for oceanographic salinity data were taken at each location, totaling 188 sample points for weekly data using the Modis satellite, as shown in Figure 2.

2.2. Data Preprocessing

Aqua MODIS satellite is one of the satellite methods that can measure sea surface temperature (SST) spatially and temporally. The purpose of this study is to determine changes in sea surface temperature and determine the correlation of SPL with the El Nino Southern Oscillation (ENSO) and Indian Ocean Dipole (IOD) phenomena for 20 years from 2003 to 2022 in the Flores Sea Waters and surrounding areas. The data processing consists of preprocessing, processing, and spatial and temporal analysis. The processing results will be presented temporally and spatially to observe monthly, seasonal, and annual fluctuations. The study was used to determine the relationship between sea surface temperature and ENSO, and IOD used the Pearson correlation equation. The highest monthly average SST occurred in December at 30.63^{o}C, while the lowest occurred in August at 27.67^{o}C. The lowest monthly average SPL in 20 years occurred in August 2006 at 26.86^{o}C [19]—for example, Noviar et al. Test the spectral response using a spectrometer and then analyze the spectral pattern at each age of rice. The research results showed that the NIR band's reflectance value (750-900mm) was higher in the generative phase than in the vegetative phase. In the vegetative phase, the NIR reflectance value is generally around 35%, while in the generative phase, it is around 40%. Irregular or inaccurate reflectance can be caused by field measurement interference or Landsat-8 image interference. Landsat-8 standard level (level 1T) data received by users still in digital form can be used directly for land cover/land use mapping. Activities in preprocessing in connection with research into developing deep learning methods and long short-term Memory for predicting marine data from satellite images, namely Cloud Removal and Clipping. (1) The cloud-removing method removes clouds and cloud shadows in optical satellite images. This method eliminates the cloud display by displaying image data from different times (multi-temporal). The principle used in this cloud-removing process is overlapping cloudy image data with clean image data, then localizing the cloud area and covering it with a clean image with algorithm formulation facilities applied to each band. Things that must be considered in the cloud-removing process to obtain optimal results are: a) The difference in image data acquisition time between cloudy and covered image data is not too significant; the closer, the better. b) The geometry of both image data must be the same and have a uniform appearance. (i) The difference in recording time in the images to be combined is vital because objects on the earth's surface change very quickly; for example, residential development means that land cover data will change quickly. The pixel values will differ significantly if the merged images are not from adjacent recordings. (ii) The geometry of the two image data must be the same so that a particular Object at a location will combine with the Object in the complementary image. This cloud removal process can be carried out after previously carrying out a map-to-map rectification/orthorectification process, where one of the satellite image data, which has been geometrically appropriately corrected, is used as a reference for other satellite image data so that later the objects between the primary satellite image data with replacement satellite image data not shifting. (2). Clipping and Erasing are essential geoprocessing techniques in ArcGIS. Its primary function is to produce maps that focus more on specific areas. It should be noted that the condition of the image data before the analysis process stage contains several errors, including (a) Systematic errors, Errors caused by measurement bias so that the results are different from the actual situation, for example, systematic errors due to the influence of the curvature of the earth (systematic geographic error), sensor/scanner damage, atmospheric disturbances (b) Unsystematic/(Random) errors, (i) Variations in the height of the satellite (platform) and the angle of data collection (ii) Weather disturbances: clouds.

Sensors called long wave (L-band) microwave radiometers or optical sensors such as the Sentinel-3 satellite are usually used to obtain data about salinity, sea surface temperature, and sea level height from satellites. Longwave microwave radiometers make it possible to measure sea surface temperature, which in turn can be used to estimate salinity because salinity correlates with the thermal conductivity of seawater. The Sentinel-3 satellite, for example, is equipped with a sensor called the Sea and Land Surface Temperature Radiometer (SLSTR), which can provide information on sea surface temperature. This sea surface temperature can then be used together with other data, such as data on temperature, pressure, and humidity, to calculate salinity using hydrographic models. Processing salinity data from satellites involves a series of steps: (1) Data collection: Satellites collect data through sensors installed on them. (2) Data calibration: Data received from sensors must be calibrated to ensure accuracy and consistency. (3) Data processing: The calibrated data is then processed to obtain sea surface temperature information. (4) Validation: The data processing results are then validated using other sources or field observation data. (5) Estimation of salinity: Sea surface temperature information and other parameters are used to estimate salinity using hydrographic models. With a comparison with data from https://www. Copernicus. Eu/en/copernicus-satellite-data-access and Google Earth Engine (GEE) are cloud computing platforms that allow users to process, analyze, and visualize extensive geospatial data- magnitude. GEE provides access to a variety of data sources, including satellite data, and provides tools for analyzing that data.

3. RESULT AND ANALYSIS

This study applied the Long Short Term Memory (LSTM) deep learning method to analyze the seasonal trend of oceanographic data collected from 1 January 2014 to 30 December 2018. The LSTM method was chosen because of its ability to handle sequential data and capture long-term patterns, especially in oceanographic data with complex seasonal and fluctuation characteristics. The neural network architecture used in the LSTM (Long Short-Term Memory) model, both SST, SSH, and SSS, in the graph below, is the prediction process; there are two hidden layers and one dense layer, and the previous method using the Adam optimizer is often used to adjust model parameters during training and is a machine learning/deep learning library such as TensorFlow, PyTorch, and Keras which are part of python

3.1. Salinity Banda Sea

Meanwhile, the results of processing the original data set produced an image of the 2018 Banda Sea salinity map, which showed salinity values of 34.034027 to 34.31760 PSU, according to previous research. The Banda Sea, with a depth of more than 300 bar (1 bar = 10 meters), has the same salinity as the Flores Sea, namely around 34.6 ppm, shown in Figure 3. The distribution of raw satellite data from 2014 to 2018, shown in Figure 4, comes from the North Pacific. The increased salinity in the Banda Sea is carried to the lowest layer of the thermocline in the Flores Sea through the water mass circulation of the Makassar Strait. Salinity is a measure of the concentration of salt in water. The results of iteration and prediction of salinity data using LSTM produced RMSE = 0.190064402 and MAE = 0.0982284, with a total of 3,871 parameters and epoch iterations starting at a slowdown point starting below the value 0.004, as shown in Figure 5 and 6. The image above shows a map of the distribution of salinity data in the Banda Sea 2018. As indicated in the legend, salinity values are demonstrated in PPT (Parts Per Thousand), ranging from 34,031 to 34,237. The blue gradient on the map depicts variations in salinity, with lighter areas having higher salinity. The map uses the WGS 84 geographic coordinate system with a scale 1:5,000,000 for the large study area. This data is essential for analyzing ocean dynamics, such as the Banda Sea's current circulation and water mass distribution. The map also utilizes BingMaps as a background to provide a clear geographic context., shown in Figure 3. The Figure 4. above shows the raw salinity data as a time series. This graph depicts the fluctuation of salinity values over time, with the horizontal axis representing time and the vertical axis representing salinity values. The data pattern shows seasonal variations or periodic changes, which may be influenced by environmental factors such as ocean current circulation, evaporation, and rainfall. Sharp fluctuations in the graph may indicate anomalies or significant changes in environmental conditions. This data is essential as an initial step for further analysis, such as predictive modelling using algorithms such as LSTM. The increased salinity in the Banda Sea is carried to the lowest layer of the thermocline in the Flores Sea through the water mass circulation of the Makassar Strait. The figure below shows the loss curve during the training process of the Long Short-Term Memory (LSTM) model. At the beginning of training, the loss value is very high, indicating that the model has a significant error in predicting the output. As the number of epochs increases, the loss decreases sharply, indicating that the model has successfully learned the data pattern and improved its prediction accuracy. After several epochs, the loss curve flattens, indicating that the model has reached a point of convergence or stability. At this stage, the decrease in loss becomes very small, and the model tends not to experience significant improvement anymore. This curve is essential to ensure the model does not experience overfitting or underfitting during training.shown in Figure 5 and Figure 6, is a graph of historical data (Green Line): this line shows the salinity data from 2014 to 2017 that was used to train the LSTM model. The salinity pattern during this period is what the model learns to make predictions. After the model is trained, data from around 2017 to 2018 will be used as the test data during the prediction and test period. This is where we compare the actual data (blue line) and the model predictions (red line). Blue vs Red Line (Actual Data vs Predictions): in the period after 2017, the blue and red lines compare the actual salinity values with those predicted by the model. Ideally, the red line will close to or overlap the blue line if the LSTM model performs wel.

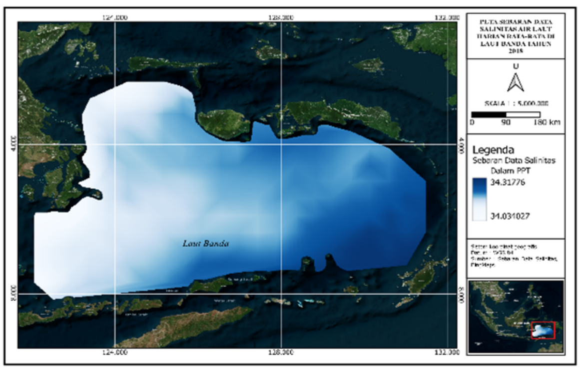

Figure 3. Banda sea salinity map 2018

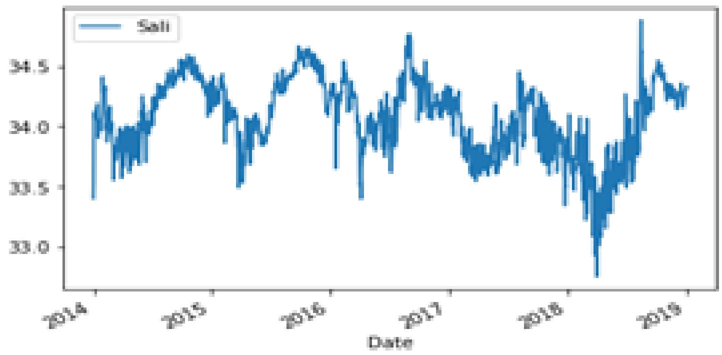

Figure 4. Banda sea salinity distribution year 2014-2018

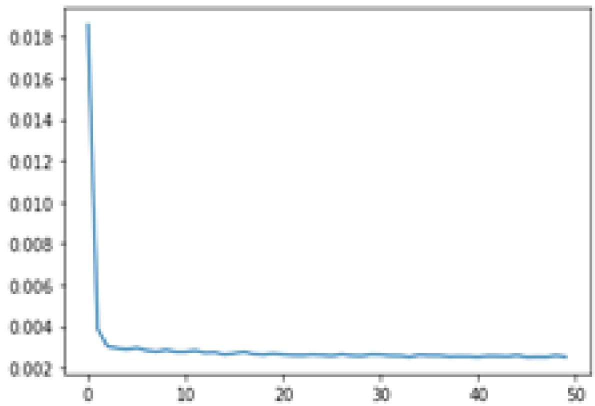

Figure 5. Epoch salinity iteration with LSTM

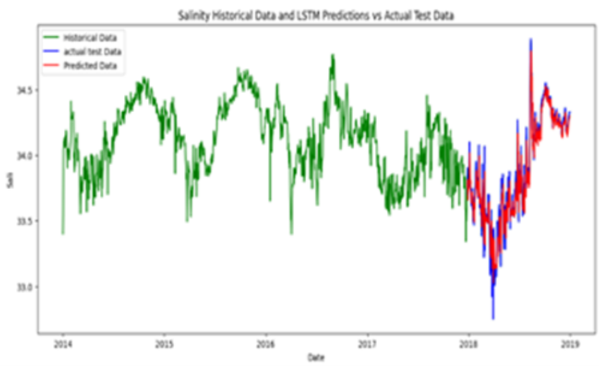

Figure 6. Graph of salinity prediction with LSTM

3.2. Elevation Banda Sea

The image above is a map of the average daily sea level elevation in the Banda Sea in 2018, with the range of elevation values indicated by the color gradation in the legend (e.g., from purple to yellow). The following is a more detailed analysis. The area covered is the Banda Sea, surrounded by large islands such as Sulawesi to the northwest and the Maluku Islands to the east. The Banda Sea is one of the deepest seas in Indonesia, and it has complex oceanographic dynamics as shown in 7. Shows a periodic fluctuation pattern, possibly following a particular season or seasonal cycle (for example, the influence of monsoon winds or tides). Research data from the daily average sea level height map of time series data from 2014 to 2018. Location in the Banda Sea coordinates 6 ° S 127 ° E, starting Latitude -7.0999996 Longitude 131.916663 or -7 ° 05'60.00 "S 131 ° 54'59.99" E. The following is the distribution of sea level in 2018; the highest was 0.655605 m, and the lowest was 0.576662 m, shown in Figure 2.7. Nonstationary is a statistic such as the average and variance change over time, seen from the amplitude of fluctuations that are not constant. Nonlinearity is a relationship between data values at various times that is not simple or can be represented linearly as shown in 8. Furthermore, The figure above shows the loss curve during the model training process through epoch iterations. The horizontal axis represents the number of epochs, while the vertical axis shows the loss value. The loss value is very high at the beginning of training but decreases sharply in the first few epochs, indicating that the model is rapidly learning from the data. After about ten epochs, the loss value tends to stabilize with a slower decrease, indicating convergence toward the optimal solution. This graph illustrates the effectiveness of model training in reducing errors and increasing prediction accuracy. as shown in 9. Historical Data (Green Line): the green line shows the elevation pattern from 2014 to 2017, used as the basis for training the LSTM model. This pattern includes fluctuations in elevation that the model learns. Prediction and Test Period: After the model is trained with historical data, data from around 2017 to 2019 will be used as test data. The model predictions (red line) are compared to the actual data (blue line) during this period. Blue vs Red Line (Actual Data vs Prediction), from 2017 to 2018, the comparison between the blue and red lines shows how accurate the model predictions are to the actual data. Similarities between the blue and red lines indicate that the model performs well, while differences indicate prediction errors. as shown in 10. The LSTM results graph shows that although it is not yet precise, it tends to follow the raw data trend. The LSTM method produces.

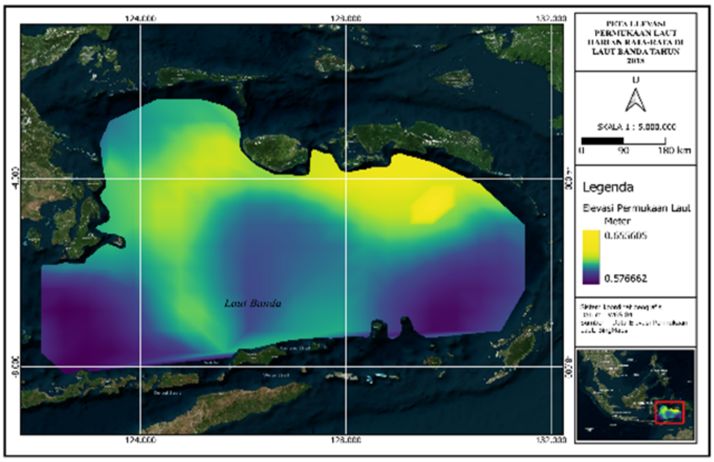

Figure 7. Map of sea level height (elevation) Banda Sea in 2018

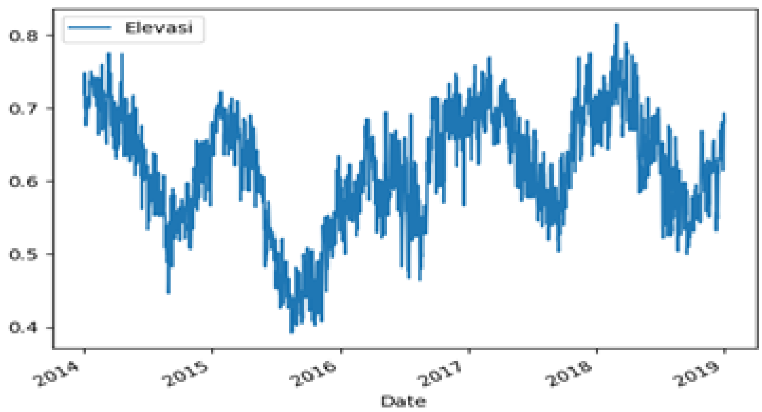

Figure 8. Daily distribution of elevation Banda Sea in 2014-2018

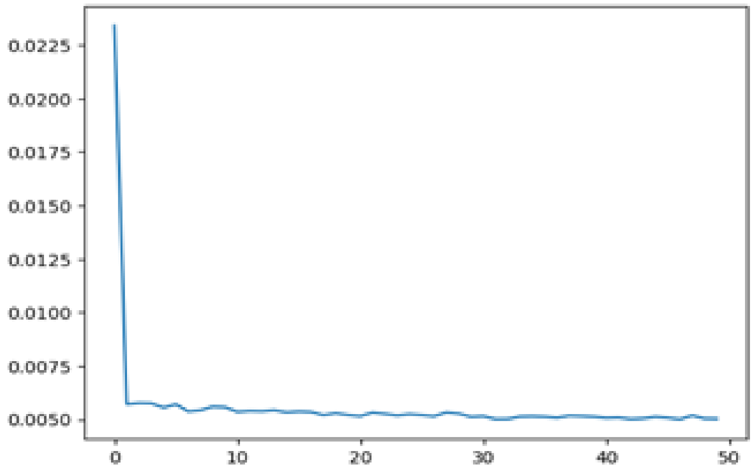

Figure 9. Elevation epoch iteration with LSTM

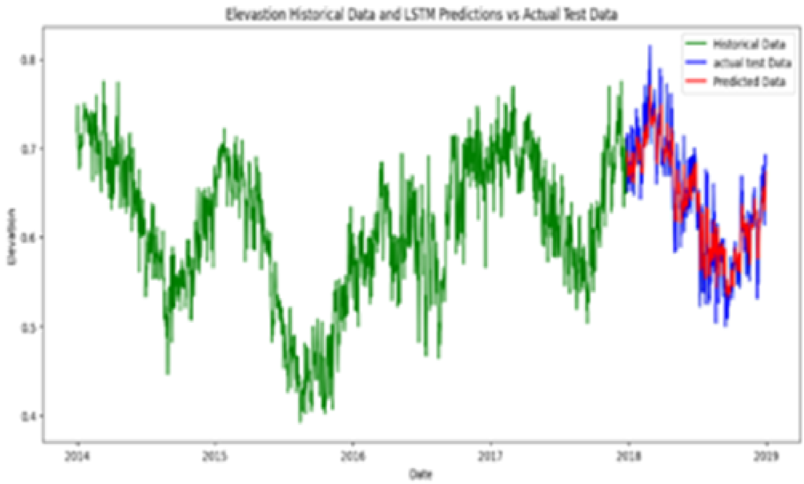

Figure 10. Elevation prediction graph with LSTM

3.3. Temperature Banda Sea

Location, showing the average daily sea surface temperature (SST) map in the Banda Sea, Indonesia, for 2018. The area covered includes the waters of the Banda Sea and its surroundings. The research data of the temperature distribution map, the value of the sea surface temperature distribution in 2018, is shown in Figure 2.1. Scale: the map uses a scale of 1:5,000,000 with a north orientation at the top. Color Palette: The red indicates the area with the highest temperature (around 29.31 °C). White indicates the area with the lower temperature (around 28.50 °C). The data source is temperature data obtained from satellite imagery and uses a geographic coordinate system with a WGS84-based map projection., as shown in 11. Graph Description: The Horizontal Axis (X-axis) represents time in annual format (2014 to 2018); each point reflects the time of temperature data collection. The Vertical Axis (Y-axis) represents sea surface temperature in degrees Celsius, ranging from 26°C to 31°C. The highest temperature is around 31°C, and the lowest temperature is around 26°C. is shown in 12. The temperature epoch iteration tends to decrease sharply, starting at 0.020 seconds. At 0.008 seconds, it slopes slightly along with the increase in epochs until it approaches 100, as well as the trend in temperature prediction with LSTM. following temperature data, as seen in 13 and 14. In this graph, the last year of SST from the 2014–2018 dataset shows an increase to around 34°C in April and May 2018, which coincides with the end of the rainy season. The temperature decreases to around 26°C in June and August, corresponding to the beginning of the rainy season. In particular, the highest temperatures are recorded between the 40th and 140th days, and between the 180th and 240th days, the lowest temperatures are seen. In this period, roughly from June to February and April/May, temperatures exceeding 30°C occur more frequently.

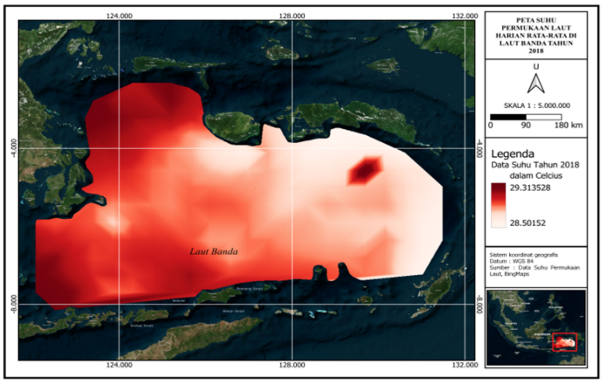

Figure 11. Sea Surface Temperature Map (temperature/sst) Banda Sea 2018

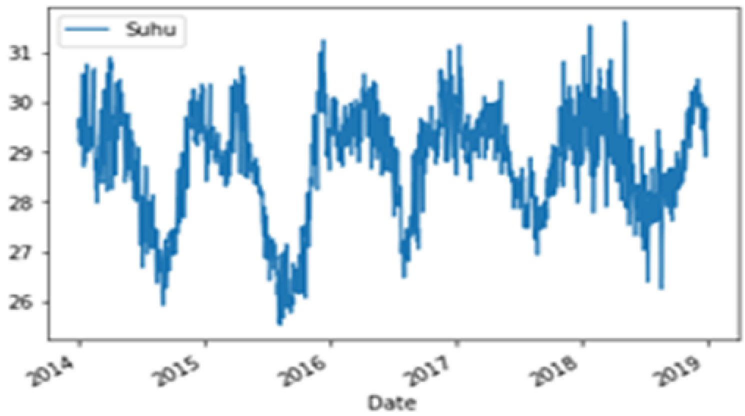

Figure 12. Sea Surface Temperature Distribution Banda Sea 2014-2018

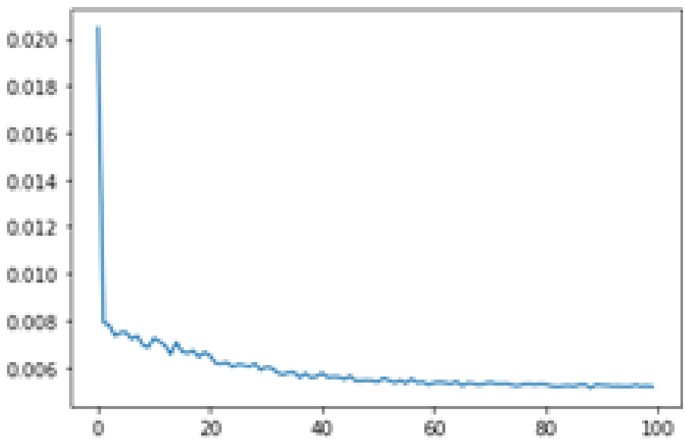

Figure 13. Sea surface temperature epoch iteration with LSTM

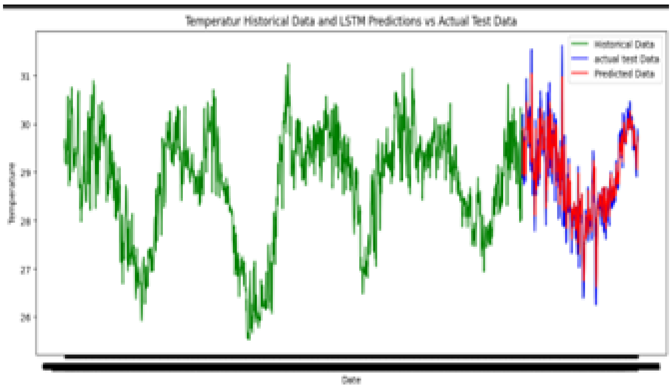

Figure 14. Sea surface temperature prediction graph with LSTM

3.4. Correlation plot

Correlation results without the Chloropil-a (Chl-a) parameter, the correlation between temperature parameters (sst) and sea surface height/elevation (ssh) is 0.6. The correlation results between variables are depicted in Figure 15 is the relationship between variables using Pearson correlation, with the approach of each variable, meaning the relationship between parametric test variables to test the relationship between 2 variables with a numerical measurement scale (interval-ratio or ratio-interval). The Pearson Correlation Test is carried out if the assumption of at least one of the variables is normally distributed. which shows the correlation between temperature parameters (sst) and sea level height/elevation (ssh) of 0.6.

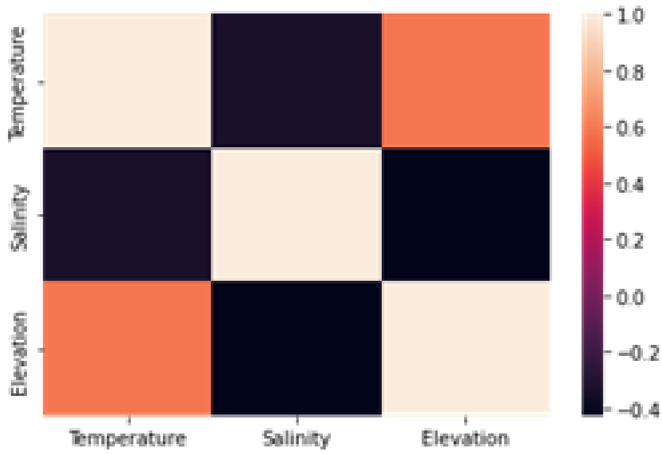

Figure 15. Correlation plot between variables

Research using correlation compared to previous research is a Deeper Understanding of Marine Variable Interactions. Interactions between marine variables are crucial to understanding complex ocean dynamics. For example, there is a strong relationship between SST and chl-a in predicting marine biological productivity and between SSH and SSS in identifying phenomena such as upwelling or ocean circulation. By correlating these variables, the model can capture deeper patterns of interaction and not just predict each variable separately. This approach provides more accurate and realistic predictions because it considers how changes in one variable can affect the other. Research that correlates marine variables can produce more robust models by including physical relationships between variables, providing richer insights and prediction results that align with actual ocean conditions. The image above is a correlation matrix, with dark and light blocks indicating the strength and direction of correlation between pairs of variables. In the context of oceanographic variables, this matrix likely shows correlations between variables such as Sea Surface Temperature (SST), Sea Surface Salinity (SSS), Sea Surface Height (SSH), and chlorophyll-a (chl-a) concentrations. The color gradient on the right side of the matrix shows that higher values (darker colors) indicate a strong positive correlation. In comparison, lower values (lighter colors) indicate a weak or negative correlation.

Here's how our research approach compares to previous research, focusing on the issue of correlation: 1) Deeper Understanding of Variable Interactions: Unlike the last research that might predict each variable separately, your model takes advantage of the relationships between variables. For example, the strong positive correlation between SST and chl-a suggests that higher temperatures may increase phytoplankton growth, which impacts ocean productivity. Recognizing these relationships helps the model make more informed predictions. 2) Improved Predictive Power: By incorporating correlations, models do not simply make individual predictions for SST, SSH, SSS, or chl-a. Instead, models capture patterns of interactions that provide context for how fluctuations in one variable can affect the others. For example, if SST increases, models can predict changes in chl-a concentrations, improving the accuracy of forecasting ecological events. 3) Capturing Complex Ocean Phenomena: Correlations, such as between SSH and SSS, can indicate dynamic processes such as upwelling or circulation changes, which are often missed in single-variable models. By correlating these variables, your research can provide a more realistic representation of oceanographic events. Data Differences from Previous Research, (a) Variable Interactions: Previous research may have analyzed variables indepen- dently and thus missed their combined effects (e.g., SST affecting chl-a). This research provides a more layered understanding by considering correlations critical for complex systems like the ocean. (b) Robustness of Predictions: Models in previous research may need to have the ability to adapt to changing relationships between variables, resulting in less reliable predictions. By correlating variables, our research can produce more robust models that better accommodate natural variations in ocean conditions. Research results: The extended Short-Term Memory model performs best with a Root Mean Squared Error of 0.1096 on test one and a Mean Absolute Error of on test two 0.0982 for salinity at 13 sample points. The table below shows the RMSE values obtained from each model (ARIMA and LSTM) for each predicted variable. The results are as follows:

Table 1. RMSE measurement results of Arima and LSTM statistical models (processed results)

Parameter	RMSE by ARIMA	RMSE by LSTM
Salinity	0.193	0.110
Elevation	0.055	0.050
Sea surface temperature	2.504	0.399

Table 1 shows that LSTM provides a lower RMSE value compared to ARIMA for all variables, indicating that LSTM has better prediction accuracy. This difference is most striking in predicting sea surface temperature (SST), where the ARIMA RMSE of 2.504 drops to 0.399 in the LSTM model. The results of this study indicate that the LSTM model is superior in predicting oceanographic variables compared to ARIMA. Previous research by [20] showed that neural network-based models such as LSTM are more effective in capturing complex patterns and nonlinear relationships between variables in sequential data than traditional methods such as ARIMA. The significant decrease in RMSE indicates that LSTM can capture seasonal patterns and long-term trends, which is essential in forecasting ocean conditions. For example, LSTM successfully predicts sea surface temperature fluctuations with much higher accuracy than ARIMA, which helps analyze climate change and marine ecosystems.

4. CONCLUSION

This study concluded that LSTM can capture complex and nonlinear patterns in time series data. This means that LSTM can better capture the complex relationship between ocean salinity variables and other variables that may be nonlinear. Flexibility in modeling time series data: LSTM can learn and remember information from the past that is relevant over a more extended period, thanks to its long-term and short-term memory features. This allows LSTM to handle time series data with varying levels ofocean Adaptation to data variations: LSTM can adapt well to variations in time series data, including changes in trends, seasonality, and other patterns that may emerge over time. This makes LSTM more flexible in handling unstable or changing time series data. Dependence on parameters: The ARIMA method requires the determination of model parameters such as autoregressive order (p), difference order (d), and moving average order (q). The selection of these parameters can be cpagelenging, but if they are appropriately chosen, ARIMA can correctly capture patterns in the data. Different data: The performance of LSTM and ARIMA can vary depending on the characteristics of the data used. LSTM is better suited for complex and nonlinear time series data, while ARIMA is better suited for more straightforward and linear patterns. While LSTM can provide better results in some cases, it is essential to remember that this advantage does not necessarily apply in all situations. The selection of the best method must consider the specific characteristics of the data and the desired prediction goals. Further research is needed to optimize the neural network weights in the LSTM algorithm to produce a more accurate RMSE.

Further research with the LSTM approach requires optimizing the neural network and iterative optimization to produce more accurate predictions. From the results of each variable, a correlation plot is then carried out to produce the relationship between one variable and another. This is done statistically using multivariate analysis, which is often used to understand the complex relationships between various variables in a dataset and can provide deeper insights than bivariate analysis.

5. ACKNOWLEDGEMENTS

Thank you to all colleagues who have taken the time to collect data on researchers' research and to those who have helped in the research process.

6. DECLARATIONS

AUTHOR CONTIBUTION

All four authors contributed significantly to the research's conception, design, analysis, and interpretation. Authors 1 and 2 were involved in data collection, while Author 3, Author 4, and Author 5 contributed to the data analysis. All authors participated in critically drafting and revising the manuscript for important intellectual content. Each author has approved the final version of the manuscript and takes responsibility for the accuracy and integrity of the work.

FUNDING STATEMENT

This research was self-funded, and the authors received no external financial support for the study's design, data collection, analysis, or interpretation. All expenses related to this research were borne by the authors personally.

COMPETING INTEREST

All authors reported this article with no competing financial interests or personal conflicts.

REFERENCES

[1] B. Lindemann, T. Müller, H. Vietz, N. Jazdi, and M. Weyrich, "A survey on long short-term memory networks for time series prediction," *Procedia CIRP*, vol. 99, no. July 2020, pp. 650–655, 2021, https://doi.org/10.1016/j.procir.2021.03.088.

[2] J. Shi, M. Jain, and G. Narasimhan, "Time Series Forecasting Using Various Deep Learning Models," *World Academy of Science, Engineering and TechnologyInternational Journal of Computer and Systems Engineering Vol:16, No:6, 2022*, no. June, 2022, https://doi.org/10.48550/arXiv.2204.11115.

[3] J.-J. F. A. Andreu-Francisco Javier Francisco Javier López, Morales-Juan Antonio López, Guillen-Zaida Hernández, Rodrigo-Juan Antonio Carrero, Alcaraz-Marta Sánchez, "Deep Learning-Based Time Series Forecasting Models Evaluation for the Forecast of Chlorophyll a and Dissolved Oxygen in the Mar Menor," *Journal of Marine Science and Engineering*, vol. 11, no. 7, 2023, https://doi.org/10.3390/jmse11071473.

[4] A. W. Saputra, A. P. Wibawa, U. Pujianto, A. B. Putra Utama, and A. Nafalski, "LSTM-based Multivariate Time-Series Analysis: A Case of Journal Visitors Forecasting," *ILKOM Jurnal Ilmiah*, vol. 14, no. 1, pp. 57–62, 2022, https://doi.org/10.33096/ilkom.v14i1.1106.57-62.

[5] E. M. C. Wattimena, A. Annisa, and I. S. Sitanggang, "CO and PM10 Prediction Model based on Air Quality Index Considering Meteorological Factors in DKI Jakarta using LSTM," *Scientific Journal of Informatics*, vol. 9, no. 2, pp. 123–132, 2022, https://doi.org/10.15294/sji.v9i2.33791.

[6] I. I. Zulfa, D. C. R. Novitasari, F. Setiawan, A. Fanani, and M. Hafiyusholeh, "Prediction of Sea Surface Current Velocity and Direction Using LSTM," *IJEIS (Indonesian Journal of Electronics and Instrumentation Systems)*, vol. 11, no. 1, p. 93, 2021, https://doi.org/10.22146/ijeis.63669.

[7] G. D. A. N. Yanuar Nurul Hilal, , and F. K. Sabilla Hamda Syahputri, "COMPARISON OF ARIMA AND LSTM METHODS IN PREDICTING JAKARTA SEA LEVEL Yanuar Nurul Hilal, Gibson Daniel Andrianto Nainggolan, Sabilla Hamda Syahputri, and Fitri Kartiasih *," *Jurnal Ilmu dan Teknologi Kelautan Tropis*, vol. 16, no. August, 2024, https://doi.org/10.29244/jitkt.v16i2.52818.

[8] H. H. Benyun Shi, Chonghui Ge, Howang Lin, Qi Tan, Yue Peng, "Sea Surface Temperature Prediction Using ConvLSTM-Based Model with Deformable Attention," *Remote Sensing*, vol. 16, no. 22, p. 4126, 2024.

[9] E. Putra, J. L. Gaol, and V. P. Siregar, "Hubungan Konsentrasi Klorofil-a Dan Suhu Permukaan Laut Dengan Hasil Tangkapan Ikan Pelagis Utama Di Perairan Laut Jawa Dari Citra Satelit Modis," *Jurnal Teknologi Perikanan dan Kelautan*, vol. 3, no. 2, pp. 1–10, 2017, https://doi.org/10.24319/jtpk.3.1-10.

[10] 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, https://doi.org/10.3390/rs12213654.

[11] M. A. Litzow, M. J. Malick, T. Kristiansen, B. M. Connors, and G. T. Ruggerone, "Climate attribution time series track the evolution of human influence on North Pacific sea surface temperature," *Environmental Research Letters*, vol. 19, no. 1, 2024, https://doi.org/10.1088/1748-9326/ad0c88.

[12] J. Chen, I. G. C. Ashton, E. C. C. Steele, and A. C. Pillai, "A Real-Time Spatiotemporal Machine Learning Framework for the Prediction of Nearshore Wave Conditions," *Artificial Intelligence for the Earth Systems*, vol. 2, no. 1, pp. 1–14, 2022, https://doi.org/10.1175/aies-d-22-0033.1.

[13] U. Tangke and B. Senen, "Distribution of sea surface temperature and chlorophyll-a concentration its correlation with small pelagic fish catch in Dodinga Bay," *IOP Conference Series: Earth and Environmental Science*, vol. 584, no. 1, pp. 1–8, 2020, https://doi.org/10.1088/1755-1315/584/1/012020.

[14] M. Suhermat, "Pemodelan Kesesuaian Habitat Ikan Cakalang Menggunakan Penginderaan Jauh di Perairan Selatan Jawa Barat-Banten," *Jurnal Ilmu Lingkungan*, vol. 22, no. 3, pp. 667–671, 2024, https://doi.org/10.14710/jil.22.3.667-671.

[15] J. Sun, H. Dong, Y. Gao, Y. Fang, and Y. Kong, "The Short-Term Load Forecasting Using an Artificial Neural Network Approach with Periodic and Nonperiodic Factors: A Case Study of Tai'an, Shandong Province, China," *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1–8, 2021, https://doi.org/10.1155/2021/1502932.

[16] K. S. Kim, J. B. Lee, M. I. Roh, K. M. Han, and G. H. Lee, "Prediction of ocean weather based on denoising autoencoder and convolutional LSTM," *Journal of Marine Science and Engineering*, vol. 8, no. 10, pp. 1–24, 2020, https://doi.org/10.3390/jmse8100805.

[17] T. Dimri, S. Ahmad, and M. Sharif, "Time series analysis of climate variables using seasonal ARIMA approach," *Journal of Earth System Science*, vol. 129, no. 1, pp. 129–149, 2020, https://doi.org/10.1007/s12040-020-01408-x.

[18] A. C. Dompeipen, D. R. Rumohoira, A. S. Atmadipoera, Nabil, Apriansyah, D. Nugroho, and N. M. Natih, "Ocean Circulation and Variability in the Seram and Northern Banda Seas," *IOP Conference Series: Earth and Environmental Science*, vol. 1350, no. 1, pp. 1–15, 2024, https://doi.org/10.1088/1755-1315/1350/1/012004.

[19] I. G. Putu, W. Paramartha, I. D. Nyoman, N. Putra, I. W. Gede, and A. Karang, "Analisis Perubahan Suhu Permukaan Laut di Wilayah Laut Flores dan Sekitarnya," *JOURNAL OF MARINE RESEARCH AND TECHNOLOGY*, vol. 7, no. 2, pp. 170–176, 2024, https://doi.org/10.24843/JMRT.2024.v07.i02.p14.

[20] S. Siami-Namini, N. Tavakoli, and A. Siami Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*, pp. 1394–1401, 2018, https://doi.org/10.1109/ICMLA.2018.00227.