Study of Sea Surface Temperature Prediction and Oceanographic Exploration using Deep Learning

1st Biswaraj Choudhury
Department of Computer Science and Engineering, TSSOT
Assam University
Silchar-788011, Assam, India
biswaraj.choudhury@aus.ac.in

2nd Kunal Chakraborty
Department of Computer Science and Engineering, TSSOT
Assam University
Silchar-788011, Assam, India
kunal.chakraborty@aus.ac.com

3rd U Poirainganba Singha
Department of Computer Science and Engineering, TSSOT
Assam University
Silchar-788011, Assam, India
poirainganbasinghaofficial@gmail.com

4th Debojyoti Kuri
Department of Computer Science and Engineering, TSSOT
Assam University
Silchar-788011, Assam, India
debojyotikuri832@gmail.com

5th Mousum Handique
Department of Computer Science and Engineering, TSSOT
Assam University
Silchar-788011, Assam, India
mousum.smit@gmail.com

6th Neha Sharma
Tata Consultancy Services (TCS)
Pune, India
nvsharma1975@gmail.com

***Abstract*—The integration of data science and marine science into a single platform has led to a revolution in the understanding of oceanographic processes. Sea Surface Temperature (SST) prediction plays a vital role in various fields, namely marine ecology, climate change studies, and environmental forecasting. This paper delves into the most recent advancements in SST prediction techniques and their impact on oceanographic exploration. Moreover, it presents a novel model aimed at addressing the limitations of previous methodologies. The utilisation of advanced Deep Learning and Machine Learning architectures has significantly improved the accuracy of the SST forecasts, surpassing the less accurate results previously obtained through numerical models. Modern techniques can capture spatial correlations and temporal dependencies in SST data. This enables predicting SST values more reliably and accurately. These cutting-edge discoveries provide valuable insights into oceanographic phenomena, aiding in the enhanced understanding of the ocean and bolstering our capacity to predict and comprehend significant and captivating climate events. This study underscores the importance of leveraging the critical role of harnessing the vast advancements in SST prediction to advance marine science and facilitate informed decision-making across diverse sectors related to the marine realm.**

***Index Terms*—Sea Surface Temperature(SST), Marine Data Science, Transformer, LSTM, CNN, Attention Mechanism, Machine Learning, Oceanography, ConvLSTM**

I. Introduction

Oceanography delves into the scientific examination of Earth's oceans, covering a wide range of aspects, including their physical properties, chemical composition, biological components, and geological characteristics. Advancing our understanding of Earth's oceans and their profound influence on the planet stands as the fundamental goal of oceanography. Within this realm, the essentiality of SST cannot be overstated. SST plays a pivotal role in unravelling the intricate interactions between the ocean and the Earth's atmosphere, thereby holding immense significance. It serves as a critical component with diverse applications, including the examination of marine ecosystems, weather prediction, and the analysis and modelling of complex climate patterns. A comprehensive study conducted by Collin et al. [1] further substantiates the significance of SST in this particular context.

Predicting SST with utmost precision is an extremely challenging endeavour due to the complex nature of heat radiation and flux and uncertainty in wind patterns over the sea surface.

In recent years, various data-driven approaches have been explored, including the Long Short Term Memory (LSTM) model as examined by Xiao et al. [2], Deep Learning Neural Networks (DNNs) along with ensembling of Stacked DNNs as investigated in the works in [3] [4] and [5], Attention Mechanisms and Graph Neural Networks (GNNs) These methods are the types of Artificial Neural Networks (ANNs) demonstrated by Tripathi et al. [6] that have gained significant popularity for SST prediction. The inspiration for employing ANNs stems from the brain, our most significant organ that endows humans with cognitive capabilities and numerous other remarkable attributes. This feature enables the model to seamlessly handle non-linearity and aptly fit random data as shown in [6]. Novel advancements in SST prediction have witnessed significant progress in the application of Deep Learning (DL) techniques, particularly the integration of Convolutional Neural Networks (CNNs) and LSTM models. These advancements have improved correctness and enhanced capabilities in capturing spatial and temporal dependencies within SST data.

Several advancements have been made in the field of ocean temperature prediction. One notable development is the Multilayer ConvLSTM model proposed by Zhang et al. [7]. It combines CNN, LSTM, and layer stacking to predict ocean temperature in a 3D space, taking into account horizontal and vertical temperature variations across different depths. This approach has shown improved reliability, particularly in the upper layers over time. Another approach worth mentioning is the combination of 3D CNN and LSTM with an attention mechanism, as explored [8]. It considers spatial correlation and temporal dependency of SST data for a specific geographical area, leveraging XGBoost for feature extraction. Experimental results have demonstrated enhanced prediction accuracy and reduced complexity. Additionally, LSTM networks have proven effective in predicting SST, outperforming traditional neural networks, as demonstrated by Sarkar et al. in [9]. Integration of DL techniques with numerical estimators and the use of GNNs with attention mechanisms have also shown promising results in SST prediction as portrayed by the work done in [8] [9] and [10].

These recent developments have paved the way for the development of models capable of being more reliable and accurate in predicting the SST and providing precious insights for various applications such as climate research, oceanography, and environmental monitoring. Specifically, in the realm of SST prediction, there is an inherent need for advancements in our models and methodologies to ensure greater accuracy and to pave the way for further growth and progress in this direction.

Existing methods struggle to properly capture the spatial and temporal dependencies in SST data. To overcome this, we propose the ConvLSTM-Transformer, a novel model that combines Convolutional layers, LSTM, and Transformer architecture. This integration captures spatial, sequential, and global interactions, improving prediction accuracy and generalisation. Our model's ability to handle long-term dependencies has the potential to revolutionise SST forecasting and enhance understanding of oceanic conditions.

The main contributions of this paper are as follows:

(i) A comprehensive analysis of various research studies focusing on the forecasting of SST highlights the advancements achieved and the challenges encountered by current models.

(ii) An innovative model called ConvLSTMTransformer, improves the precision of predicting SST by integrating spatial and temporal information. Moreover, this model effectively overcomes the limitations of existing methods.

The paper is organised in the following manner: Section II discusses the details of preliminaries related to SST prediction and oceanographic exploration. In Section III several pertinent studies that apply to the present work are presented and analysed. Section IV is reported to our proposed model and is extensively discussed. Finally, Section V presents conclusive findings and engages in a comprehensive discussion on the future prospects of the study.

II. Preliminaries

In the introductory section, the concise definitions of key terms and relevant topics that are significant to the research article are provided. The aim of this section is to offer brief information that helps in understanding each topic.

1) **Spatial Correlation and Temporal Dependency:** Spatial correlation refers to the statistical relationship or the dependency among the observations at various locations in a spatial domain. It captures the essence of the similarity or relatedness of the values of a particular variable across neighbouring locations. In the context of the paper, spatial correlation is relevant in the analysis and modelling of SST data. While temporal dependency pertains to the correlation or interdependence between observations at various time points within a dataset. It defines that the value of a variable at a particular time is influenced by the values that are observed in previous time steps leading to that. Capturing and interpreting temporal dependency is crucial for producing accurate time series forecasting.
2) **ARIMAX:** The acronym ARIMAX refers to the model known as AutoRegressive Integrated Moving Average with Exogenous Variables as explored by the work in [11]. This forecasting model is widely used and builds upon the traditional ARIMA model by incorporating external variables. By incorporating autoregressive and moving average components, as well as differencing for capturing temporal patterns in time series data, the ARIMAX model further enhances its forecasting capabilities. Moreover, it considers the influence of exogenous variables, leading to improved overall accuracy in SST prediction.
3) **Layer Stacking, XGBoost and Attention Mechanisms:** Layer stacking involves merging multiple layers of models into a unified platform to enhance the predictive capability of the overall model. XGBoost, a popular machine learning algorithm, utilises a gradient-boosting framework to combine multiple weak predictive models, often decision trees, into a powerful ensemble model that exhibits superior performance. Attention mechanisms, on the other hand, are components frequently employed in neural network models. They enable selective weighting and focusing on distinct segments of input data, allowing the model to prioritise and emphasise the most pertinent and significant elements within the input sequence.
4) **CNN and 3D CNNs:** These are deep learning architectures commonly utilised for processing visual and graphical data. CNNs excel in image and object recognition tasks by leveraging convolutional filters to automatically learn hierarchical representations in the input data. These filters enable the model to easily detect local patterns in the input image, which are then combined and transformed in subsequent layers to capture more complex and abstract features. Expanding upon the success of CNNs, 3D CNNs extend this concept to volumetric data distributed across a 3D space. By incorporating additional dimensions, these models can capture spatial and temporal dependencies within the data. Through the application of 3-dimensional convolutional filters, 3D CNNs excel in learning spatiotemporal features.
5) **LSTM & GRU Encoder-Decoder (GED):** LSTMs and GED models are recurrent neural network architectures specifically designed to capture and recognize sequential dependencies in time series data. LSTMs overcome the vanishing gradient problem commonly faced by traditional Recurrent Neural Networks (RNNs) by incorporating a gating mechanism that allows the model to

capture long-term dependencies and retain information over extended periods [12]. The GED model extends the LSTM architecture by incorporating gating mechanisms in both the encoding and decoding components. During the encoding phase, the input sequence is processed upon, resulting in the generation of a fixed-length context vector. The decoder utilises this contextual vector to produce an output sequence that suits the provided input.

6) **Graph Neural Networks (GNNs):** These are specialised neural network architectures that work on data that are structured in the form of graphs. GNNs take account of the relationships and connection parameters which exist in the graph to perform computation and to learn how the nodes and edges in the graph are represented. By the process of iteratively aggregating information from the surrounding nodes, these neural networks can capture both local and global graph structures and relationships, making these structures well-suited for tasks that involve graph data, such as social networks, recommendation engines, and molecular chemistry.
7) **Multilayer Convolutional LSTM (M-convLSTM) and Transformers:** Models that combine the strength of multiple neural networks as well as incorporating advanced architectures are known as M-convLSTMs and Transformers. M-convLSTM combines the features of the spatiotemporal modelling capability of CNNs and LSTM networks to identify complex patterns and dynamics within sequential data along with spatial information. Transformers, on the other hand, are based on self-attention models that excel at capturing long-range dependencies in sequential data without actually needing the recurrent connections. These have gained massive attention in natural language processing tasks as they prove to be a suitable candidate for such, but could also be applied to various other sequential data domains and are very capable in order to produce paramount levels of accuracy in the output.

These combinations of definitions provide a very concise and easy to dive into perspective into the wide field of neural networks and oceanographic exploration, and will further help the reader to understand the paper with a better outlook towards related terms and the key characteristics of those.

III. Related Work

Several existing works in different domains related to the prediction of SST have been carefully studied and analysed. These papers employ several models and techniques to enhance the accuracy and improve the efficiency of SST prediction. The paper by Zhang et al. [7] introduces the M-convLSTM model, which combines CNN, LSTM and layer stacking to predict 3 Dimensional ocean temperature. The model used in the paper considered the horizontal and vertical temperature variations across the varying depths of the ocean, which in turn, led to better accuracy in upper layers with increasing time steps. Another paper by Qiao et al. [8] proposes a novel approach that combines 3D CNN, LSTM with attention mechanisms, and XGBoost for feature extraction to predict SST. The experimental results from the paper highlight improved prediction accuracy while comparing them to the existing models, along with lower complexity and higher training efficiency. The deep learning neural networks are used as explored by the work in [9], where LSTM networks outperform the traditional neural networks, which proved LSTM are a better alternative in SST prediction. This fact can be verified by the model achieving high correlation values almost close to 1.0 for location-specific SST forecasts in varying time domains. The effectiveness of LSTM for use in the prediction of SST on long, medium and short time scale reflected in [13] is very high, specifically around the coastal seas. Also, usage of an ensemble of stacked DNNs is laid out and studied in [14] to incorporate air temperature and water temperature into the predictive model, resulting in an improved amount of accuracy in SST forecasting. The CFCC-LSTM model is implemented and understood and studied in [15], where this paper addresses the various limitations of previous models by the method of combining both the spatial and temporal information and then bringing out promising results that outperform traditional as well as the previous implementations of LSTM models. The GED model with SST codes, as well as dynamic influence link (DIL), has been introduced to us by the works in [16], where the paper highlights that the model is able to capture static information, along with solving the long-scale dependency problem in SST prediction. Finally, the authors in [10] propose a Global Spatiotemporal Graph Attention Network, which leverages the GNNs and attention mechanisms to properly analyse and interpret the spatiotemporal dependencies in the SST data and outperforms existing methods in terms of prediction accuracy.

A. *Shortcoming and Limitations of Existing Work*

Despite the recent advancements made in SST prediction, there are still some voids and limitations in the existing works. These are mentioned as below:

(i) Limited focus on the incorporation of various other physical oceanographic features as well as environmental variables which in turn could improve upon the accuracy of the models as highlighted in [14], where the incorporation of air and water temperature into a single model resulted in very accurate SST forecasts by the model.
(ii) Potential long term issues with generalizability and transferability of models across different geographical regions and datasets. Since, each of the models cater to one and only one particular geographical area.
(iii) The need for further development and exploration into the interpretability of the model and understanding the underlying physical processes that govern SST variations.
(iv) Challenges in handling missing or sparse data need to be dealt with more extensively.

Table I presents the contemporary analysis of eight papers focusing on Sea Surface Temperature (SST) prediction, providing a comprehensive assessment of the current research landscape in this field.

B. *Key Findings and Achievements*

As per the discussion of the literature review, we have observed the following key findings and achievements in terms of SST prediction and oceanographic exploration.

(a) The M-convLSTM model [7] takes into consideration, both the horizontal as well as vertical temperature varia-

TABLE I
CONTEMPORARY LITERATURE ANALYSIS ON STUDIES ON A FEW SST PREDICTION TECHNIQUES

Domain	Serial Number	Title of Paper	Year	Objective	Dataset	Algorithms and Techniques Used
Oceanography	1	Prediction of Sea Surface Temperature Using Long Short-Term Memory [13]	2017	To predict sea surface temperature using LSTM on short-term and long-term scales	Coastal seas of China	Long Short-Term Memory (LSTM)
	2	A CFCC-LSTM Model for Sea Surface Temperature Prediction [15]	2017	To predict sea surface temperature by combining temporal and spatial information	China Ocean and Bohai Sea datasets	CFCC-LSTM
	3	An Adaptive Scale Sea Surface Temperature Predicting Method Based on Deep Learning With Attention Mechanism [16]	2019	To predict SST using a GRU encoder-decoder model with attention mechanism and dynamic influence link	NOAA OISST grid data	GRU, LSTM, Attention Mechanism
	4	Prediction of sea surface temperatures using deep learning neural networks [9]	2020	To compare deep learning LSTM networks with other techniques for SST prediction	Numerical model products, in situ measurements, and satellite observations	Deep learning neural networks, LSTM
	5	Prediction of 3-D Ocean Temperature by Multilayer Convolutional LSTM [7]	2020	To predict 3-D ocean temperature using a model called M-convLSTM with CNN, LSTM, and layer stacking	ARGO data	Multilayer Convolutional LSTM, CNN, LSTM
	6	Sea Surface Temperature Prediction Approach Based on 3D CNN and LSTM with Attention Mechanism [8]	2021	To predict sea surface temperature using 3D CNN, LSTM with attention mechanism, and XGBoost for feature extraction	SST data from NOAA, selected SST data from Bohai and South China Sea	3D CNN, LSTM, Attention Mechanism, XGBoost
	7	Sea Surface Temperature Forecasting With Ensemble of Stacked Deep Neural Networks [14]	2021	To improve SST prediction accuracy using an ensemble of stacked deep neural networks	NOAA and Argo datasets	Deep Neural Networks (DNNs)
	8	Global Spatiotemporal Graph Attention Network for Sea Surface Temperature Prediction [10]	2023	To improve SST prediction accuracy by combining GNNs with attention mechanisms	NOAA OISST data, datasets from the Bohai Sea and the South China Sea	Graph Attention Network, Graph Neural Networks, Gated Temporal Convolutional Networks

tions in the ocean, leading to the improvement in accuracy in 3D ocean temperature predictions.

(b) The approach used in [8] combines 3D CNN as well as LSTM with the use of an attention mechanism and also XGBoost for feature extraction, resulting in improved accuracy, training efficiency, and lower complexity.
(c) Deep learning LSTM networks in [9] have taken a lead over traditional neural networks after achieving very high correlation values of around 1.0 for location-specific forecasts.
(d) LSTM-based models, as proposed in [13], effectively highlight the temporal relationship in SST data as well as improving the prediction accuracy.
(e) The incorporation of air and water temperature data into the model in [14] has led to better prediction accuracy, suggesting that increasing the number of factors incorporated into the model could improve accuracy further.
(f) The CFCC-LSTM model introduced in [15] combines temporal and spatial data, addressing the limitations of previous plain LSTM models and outperforming traditional and numerical models.
(g) The GED model that incorporated SST codes as well as DIL in [16] effectively captures the static information present in the dataset and solves the long-scale dependency problem, surpassing previous models in accuracy.
(h) The Global Spatiotemporal Graph Attention Network proposed in [10] utilizes GNNs and attention mechanisms to capture spatiotemporal dependencies within the SST data, resulting in higher accuracy compared to previous and traditional implementations and approaches.

IV. PROPOSED MODEL

The proposed model aims to utilise the Hybrid Convolutional-LSTM with Transformer (ConvLSTM-Transformer) model in order for it to predict SST accurately. This model combines the strengths of the convolutional layers, LSTM and Transformer architecture to effectively integrate the spatial and temporal information, while addressing the identified gaps and limitations within the previous implementations. A diagram of a standard LSTM model is depicted in Fig. 1 [17].

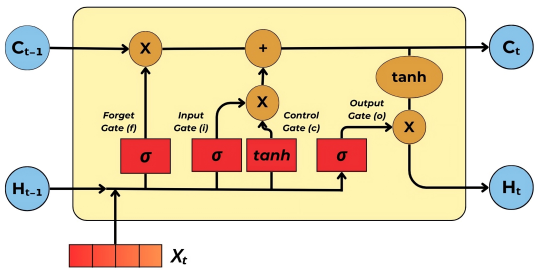

Fig. 1. Standard Diagram of LSTM

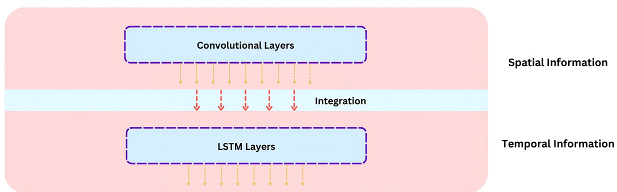

Fig. 2. Integration of Spatial and Temporal Information in the ConvLSTM-Transformer Model

A. Methodologies and Algorithms used

1. **Convolutional Layers:** The model begins with convolutional layers that extract the spatial features from the input SST data stream. These layers help in capturing the local patterns as well as correlations within the data.
2. **LSTM Layers:** The spatially extracted features are then fed forward into the LSTM layers, which then model the temporal dependencies in the SST data. Also in this phase, LSTM can capture sequential patterns as well as long-term dependencies. The integration of the Convolutional and LSTM layers, as shown in Fig. 2.
3. **Transformer Layers:** In addition to the Conv-LSTM layers, the proposed model incorporates Transformer layers. Transformers are known to leverage self-attention

979-8-3503-0219-6/23/$31.00 ©2023 IEEE

mechanisms to capture the global patterns and long-range dependencies within the data. The attention mechanisms in this phase enable the model to focus more on the important spatial and temporal contexts.

4. **Integration of Spatial and Temporal Information:** The ConvLSTM-Transformer architecture integrates the spatial and temporal information with the help of the connections between the convolutional layers, LSTM layers, and the Transformer layers. This tight integration enhances the model's overall ability to capture both local as well as global patterns in the SST data. The Architecture of the proposed ConvLSTM model is depicted in Fig. 3.

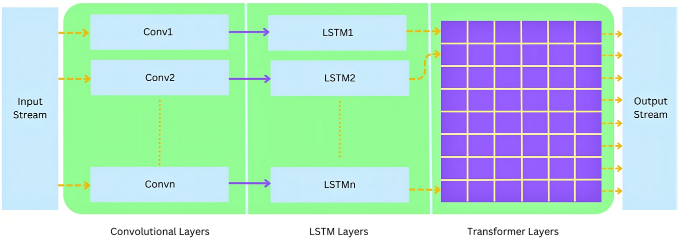

Fig. 3. Architecture of the ConvLSTM-Transformer Model

B. Benefits over Previous Implementations

1. **Enhancement in Prediction Accuracy:** The ConvLSTM-Transformer model combines the strengths of convolutional layers, LSTM, and Transformer architecture to capture the spatial as well as temporal dependencies more effectively than ever before. By integrating both spatial and temporal information, the proposed model has the potential to revolutionarily improve upon the accuracy of SST predictions compared to the previous implementations.
2. **Comprehensive Modeling of SST Data:** The integration of convolutional layers, LSTM, and Transformers allows the model to capture spatial, sequential, and global interactions in SST data. This comprehensive modeling approach leads to a more holistic representation of data, resulting in the improvement of prediction performance.
3. **Handling Long-Term Dependencies:** The ConvLSTM-Transformer model addresses the limitations of its predecessors associated with long-term predictions. The LSTM layers and Transformer components are known to excel in capturing long-range dependencies, enabling this model to make accurate predictions over extended periods.
4. **Improved Generalization:** With the conglomeration of convolutional layers, LSTM, and Transformer, the proposed model can usually generalise well into unseen SST data. The spatial and temporal information captured by the model facilitates the better adaptation of the model to varying oceanic conditions in diverse geographical regions, leading to improved forecasting capabilities.

V. Conclusion and Future Aspects

The review of relevant literature has enabled the dive into the depths of recent developments in the field of oceanography and to gain valuable insights into the existing work on SST prediction. Several models, such as M-convLSTM, CNN-LSTM with attention, LSTM-based models, CFCC-LSTM, GED, and Global Spatiotemporal Graph Attention Network, have been proposed to improve SST predictions by considering spatial and temporal dependencies, incorporating additional factors, and leveraging advanced techniques like attention mechanisms and GNNs.

The key findings highlight the effectiveness of these models in improving prediction accuracy over time, capturing spatial and temporal patterns, addressing long-scale dependencies, and outperforming traditional and numerical models. The incorporation of additional factors, such as air and water temperature, has shown promise in enhancing prediction accuracy. Despite these achievements, some limitations persist in the existing work. These include the limited incorporation of various physical oceanographic features and environmental variables, challenges in generalizability across different geographical regions and datasets, the need for model interpretability and understanding of underlying physical processes, and sparse data handling.

To address these limitations, the Hybrid Convolutional-LSTM with Transformer (ConvLSTM-Transformer) model for SST prediction is proposed. The model, as depicted in Fig. 2 and Fig. 3, integrates spatial and temporal information by combining convolutional layers, LSTM, and Transformer architecture. This approach enhances prediction accuracy, handles long-term dependencies, and improves generalisation. The proposed model overcomes the limitations of previous implementations by incorporating additional factors, providing a comprehensive modelling approach, and leveraging advanced techniques.

In summary, the proposed ConvLSTM-Transformer model offers a promising solution for accurate SST prediction by addressing the limitations of existing models. This research opens up avenues for further advancements in SST prediction, with implications for improved oceanic forecasting, climate studies, and environmental monitoring. Future directions include further exploration of incorporating various physical oceanographic features and environmental variables to enhance accuracy, developing models with improved generalizability, enhancing interpretability and understanding of underlying processes, and addressing challenges in handling missing or sparse data in historical SST datasets.

References

[1] A. Collin, P. Prot, H. Bouazzaoui, and S. Bnikkou, "Spatiotemporal patterns of the sea surface temperature using statistical anomaly detection across the 18-year-old global scale," in *OCEANS 2021: San Diego – Porto*, 2021, pp. 1–7.

[2] C. Xiao, N. Chen, C. Hu, K. Wang, J. Gong, and Z. Chen, "Short and mid-term sea surface temperature prediction using time-series satellite data and lstm-adaboost combination approach," *Remote Sensing of Environment*, vol. 233, p. 111358, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425719303773

[3] D. Qi, J. Li, and X. Yuan, "An ensemble learning approach for oceanographic data analysis and prediction," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016.

[4] Z. Şen, M. F. Demircan, and Y. C. Ozdemir, "Ensemble machine learning for oceanographic data analysis and forecasting," *IEEE Transactions on Geoscience and Remote Sensing*, 2014.

[5] B. Qiao, Z. Wu, L. Ma, Y. Zhou, and Y. Sun, "Effective ensemble learning approach for sst field prediction using attention-based predrnn," *Frontiers of Computer Science*, vol. 17, no. 1, p. 171601, 2023.

[6] K. C. Tripathi, M. L. Das, and A. Sahai, "Predictability of sea surface temperature anomalies in the indian ocean using artificial neural networks," *Indian Journal of Marine Sciences*, vol. 35, pp. 210–220, 09 2006.

[7] K. Zhang, X. Geng, and X.-H. Yan, “Prediction of 3-d ocean temperature by multilayer convolutional lstm,” *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 8, pp. 1303–1307, 2020.

[8] B. Qiao, Z. Wu, Z. Tang, and G. Wu, “Sea surface temperature prediction approach based on 3d cnn and lstm with attention mechanism,” in *2021 23rd International Conference on Advanced Communication Technology (ICACT)*, 2021, pp. 342–347.

[9] P. P. Sarkar, P. Janardhan, and P. Roy, “Prediction of sea surface temperatures using deep learning neural networks,” *SN Applied Sciences*, vol. 2, no. 8, p. 1458, 2020. [Online]. Available: https://doi.org/10.1007/s42452-020-03239-3

[10] Z. Gao, Z. Li, J. Yu, and L. Xu, “Global spatiotemporal graph attention network for sea surface temperature prediction,” *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1–5, 2023.

[11] S. B. Mahongo and M. C. Deo, “Using artificial neural networks to forecast monthly and seasonal sea surface temperature anomalies in the western indian ocean,” *The International Journal of Ocean and Climate Systems*, vol. 4, no. 2, pp. 133–150, 2013.

[12] C. Olah. (2015) Understanding LSTM networks. [Online]. Available: http://colah.github.io/posts/2015-08-Understanding-LSTMs/

[13] Q. Zhang, H. Wang, J. Dong, G. Zhong, and X. Sun, “Prediction of sea surface temperature using long short-term memory,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1745–1749, 2017.

[14] M. Jahanbakht, W. Xiang, and M. R. Azghadi, “Sea surface temperature forecasting with ensemble of stacked deep neural networks,” *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[15] Y. Yang, J. Dong, X. Sun, E. Lima, Q. Mu, and X. Wang, “A cfcc-lstm model for sea surface temperature prediction,” *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 207–211, 2018.

[16] J. Xie, J. Zhang, J. Yu, and L. Xu, “An adaptive scale sea surface temperature predicting method based on deep learning with attention mechanism,” *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 5, pp. 740–744, 2020.

[17] F. Deng, Z. Chen, Y. Liu, S. Yang, R. Hao, and L. Lyu, “A novel combination neural network based on convlstm-transformer for bearing remaining useful life prediction,” *Machines*, vol. 10, no. 12, 2022. [Online]. Available: https://www.mdpi.com/2075-1702/10/12/1226