Anomaly Prediction in Sea Surface Temperature using Time Series Models

Team Number: 9

Rahul Yalavarthi - AM.EN.U4AIE20159
Gayathri P - AM.EN.U4AIE20126
Jayanth S Varma - AM.EN.U4AIE20134

Introduction

- Forecasting oceanic phenomena is critical for maritime safety, coastal management, and disaster preparedness.
- Several sea surface parameters like sea surface height, sea surface temperature are essential for developing ocean forecasting models.
- Utilize satellite measurements for precise, real-time SST data, while employing deep learning for anomaly prediction.
- This proposed model utilizes a Long Short-Term Memory (LSTM) neural network to predict anomalies in SST values.
- Aims to capture both spatial and temporal variations in SST values.
- Various established models like ARIMA, SARIMA, are present, among which one model is compared in this project.

Motivation

- SST changes can have significant impacts on coastal areas, including variations in climate and weather patterns, flooding, habitat loss, erosion, and saltwater intrusion.
- Changes in sea surface temperature are caused by several factors such as solar radiation, oceanic current etc.
- The sea surface temperature has risen at an average rate of 0.14°F per decade from 1901 to 2020. click here for report
- Warmer oceans results in rise in sea levels which pose a significant threat to coastal countries, particularly India, China, Bangladesh, the Netherlands, and others.
- Warmer water in the Indian Ocean is contributing to stronger cyclones, hurricanes, and storm surges. For instance, the Super Cyclone Amphan, 2020
- Motivation : To develop a model for predicting anomalies in SST values which is crucial for assessing these impacts and implementing effective disaster preparedness measures in coastal regions.

Background Study/Related Work

Title & Year	Problem	Contributions	Limitations	Future Work
Title : Multilayer Fusion Recurrent Neural Network for Sea Surface Height Anomaly Field Prediction **Year :** 2022 **Journal :** IEEE transactions on Geoscience and Remote Sensing	Addresses the need for accurate and holistic prediction of sea surface height anomaly (SSHA) in the South China Sea using a Multi-layer Fusion Recurrent Neural Network.	❖ Introduces a novel MLFrnn model that fully captures contextual and time-sequential information from neighboring and remote regions to improve accuracy ❖ Integrates both spatial and temporal features, forecasting the entire SSHA map for the region in a single model. A multilayer fusion cell is designed to effectively combine global and local spatiotemporal features.	1) The MLFrnn's computational complexity could hinder real-time applications. 2) The study focused on the South China Sea; generalization to other regions may require further investigation.	1) Investigate model optimization for faster inference. 2) Extend the MLFrnn to analyze SSHA fields in different ocean regions for broader applicability.
Title : Anomaly Prediction With Hybrid Supervised/Unsupervised Deep Learning for Elastic Optical Networks: A Multi-Index Correlative Approach **Year :** 2022 **Journal :** Journal of Lightwave Technology	Addresses the challenge of predicting network anomalies in complex optical network environments.	❖ Proposes a multi-index anomaly prediction scheme that combines supervised and unsupervised deep learning techniques for elastic optical networks. ❖ The proposed scheme for elastic optical networks (EON) with multiple indicators utilizes correlative prediction and LSTM neural networks to forecast future time series.	1) Limited dataset size may affect the generalization of the model. 2) The scheme's performance on highly dynamic network environments needs further investigation	1) Use data augmentation for improving model.

Background Study/Related Work

Title & Year	Problem	Contributions	Limitations	Future Work
Title : Robust Anomaly Detection for Multivariate Data of Spacecraft Through Recurrent Neural Networks and Extreme Value Theory **Year :** 2021 **Journal :** IEEE ACCESS	Addresses the need for anomaly detection in Spacecrafts which is crucial for avoiding catastrophic failures.	❖ Proposes an unsupervised anomaly detection algorithm combining GRU and EVT. ❖ The proposed method outperforms state-of-the-art approaches in model performance and robustness.	Limited labeled data for precise anomaly identification.	1)The proposed method can be extended to explore transfer learning techniques for improved efficiency and robustness across multiple spacecraft and machines.
Title : Mid-Term Simultaneous Spatiotemporal Prediction of Sea Surface Height Anomaly and Sea Surface Temperature Using Satellite Data in the South China Sea **Year :** 2020 **Journal :** IEEE Geoscience and Remote Sensing Letters	The primary objective of this paper is to tackle the difficulty of forecasting daily sea surface height anomaly (SSHA) and sea surface temperature (SST) in the marine environment for a period.	❖ Proposes a novel data-driven method that combines empirical orthogonal function of multivariate, empirical mode decomposition, and a multilayer perceptron (MEOF-EMD-MLP). ❖ The model takes into account the correlation between SSH and SST, and the temporal and spatial relationship between discrete points allowing more accurate predictions.	1) The proposed model's performance might be specific to the South China Sea region where it was tested. 2) The model may not be well-suited for predicting abrupt oceanic changes caused by external forces like storms, which can significantly impact the marine environment.	1)Developing a data-driven-based ocean-atmosphere coupled model. 2)Extending the model's forecasting capabilities beyond the 30-days period to address long-term predictions. 3)could also focus on incorporating external forces, such as wind conditions.

Persisting Challenges

- **Dataset Size:** the dataset initially selected for the project is only across 2 years and has approximately 700 points. THIS PROBLEM IS SOLVED by finding and taking two new dataset each over 40 years of sst values, one we used as custom testing and other one we used an already used dataset to compare performances.

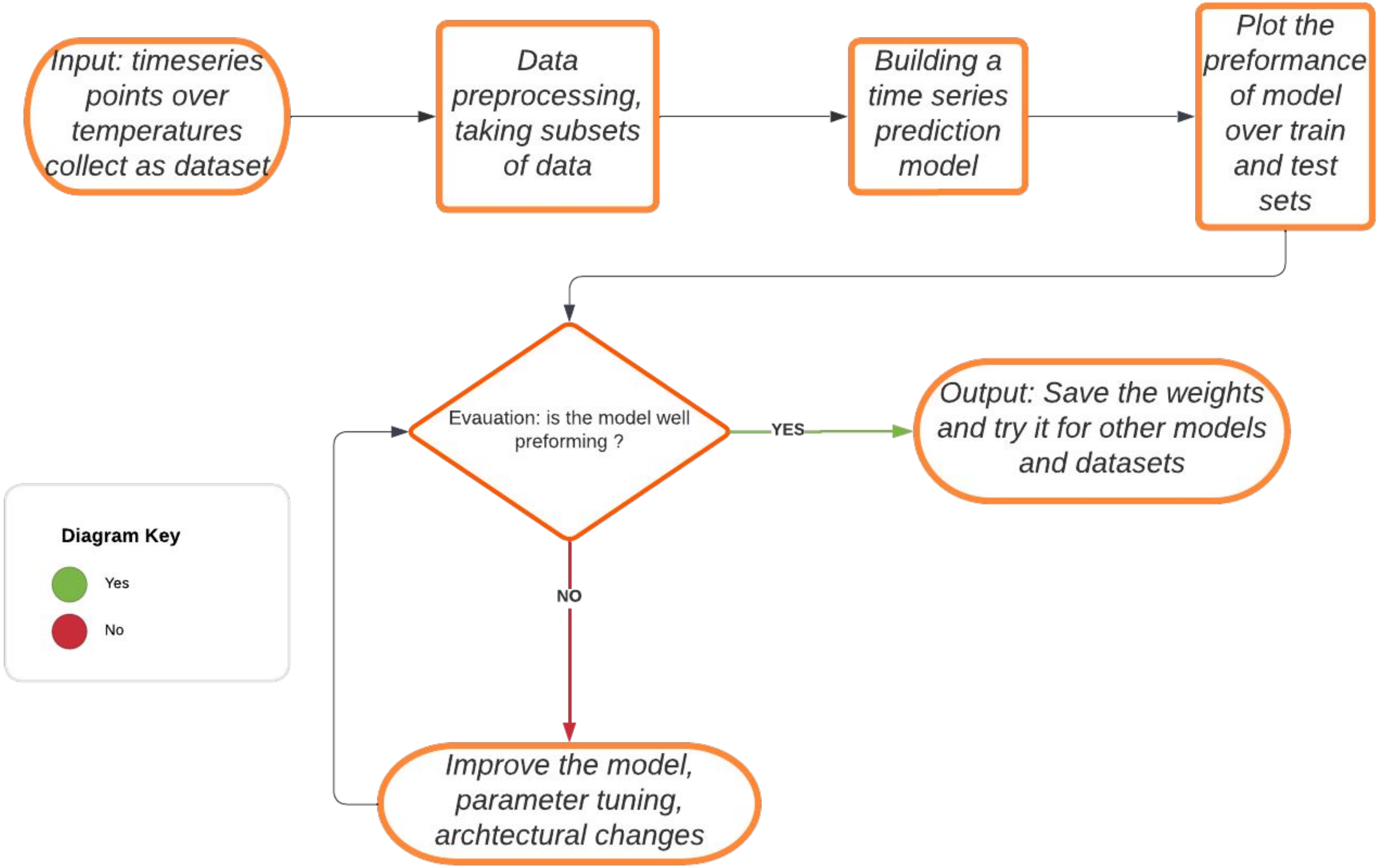
- **Availability of Frameworks:** the frameworks used in some of the base papers are not even published for free public use, example multilayer fusion Rnn. also the existing ConvLSTM doesn't support .nc files as an input. THIS PROBLEM IS ALSO SOLVED by manually extracting values as np arrays along with normalization.

- **Seasonality:** Last review we did not take seasonality into consideration but now we have taken it too in consideration by using models like SARIMA and LSTM.

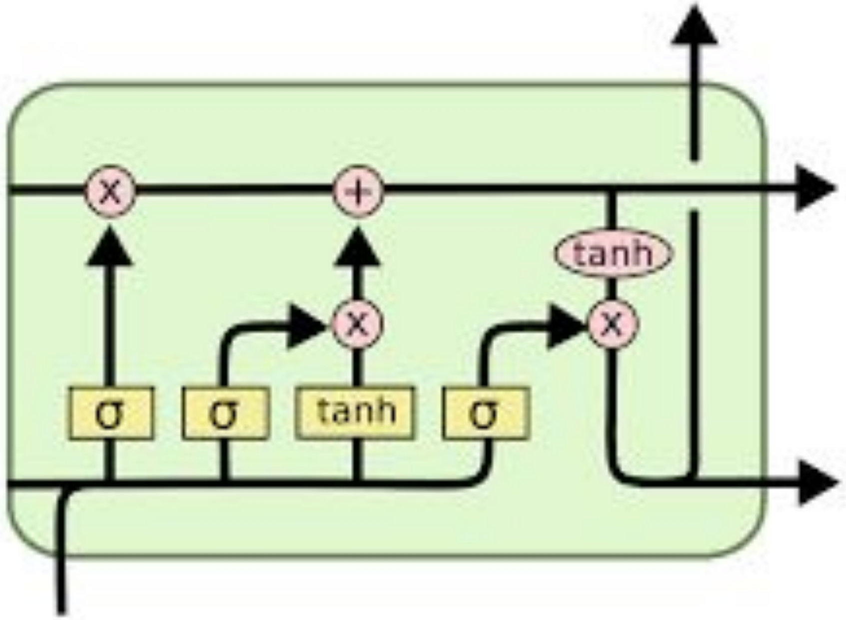
Project Contributions

- An accurate model for predicting anomalies in sea surface temperature, which helps in developing coastal management strategies.
- Crucial for taking precautions in coastal regions during a disaster.
- Researches in “anomaly prediction” is less than in “anomaly detection”.

High Level Design



Algorithms



Algorithm Training LSTM

Input : Training data: (X, Y)
Number of epochs: N
Learning rate: α
LSTM architecture parameters: $(input\_size, hidden\_size, output\_size)$

Output: Trained LSTM model

Initialize LSTM model with random weights;
for *epoch* $\leftarrow 1$ **to** N **do**
    for *each training example* $(x, y) \in (X, Y)$ **do**
        Forward pass:;
            Compute the input gate: $i_t = \sigma(W_{xi}x + W_{hi}h_{t-1} + b_i)$;
            Compute the forget gate: $f_t = \sigma(W_{xf}x + W_{hf}h_{t-1} + b_f)$;
            Compute the output gate: $o_t = \sigma(W_{xo}x + W_{ho}h_{t-1} + b_o)$;
            Compute the cell state: $c_t = f_t c_{t-1} + i_t \tanh(W_{xc}x + W_{hc}h_{t-1} + b_c)$;
            Compute the hidden state: $h_t = o_t \tanh(c_t)$;
        Backpropagate the error and update weights:;
            Compute the loss: $L = \frac{1}{2}(y - h_t)^2$;
            Compute the gradients: $\frac{\partial L}{\partial h_t}, \frac{\partial L}{\partial c_t}, \frac{\partial L}{\partial o_t}, \frac{\partial L}{\partial f_t}, \frac{\partial L}{\partial i_t}$;
            Update weights and biases using gradient descent:;
    end
end

Algorithms

Algorithm ARIMA Model Fitting

Input : Time series data: $y_1, y_2, \ldots, y_n$
ARIMA order: (p, d, q)

Output: Fitted ARIMA model

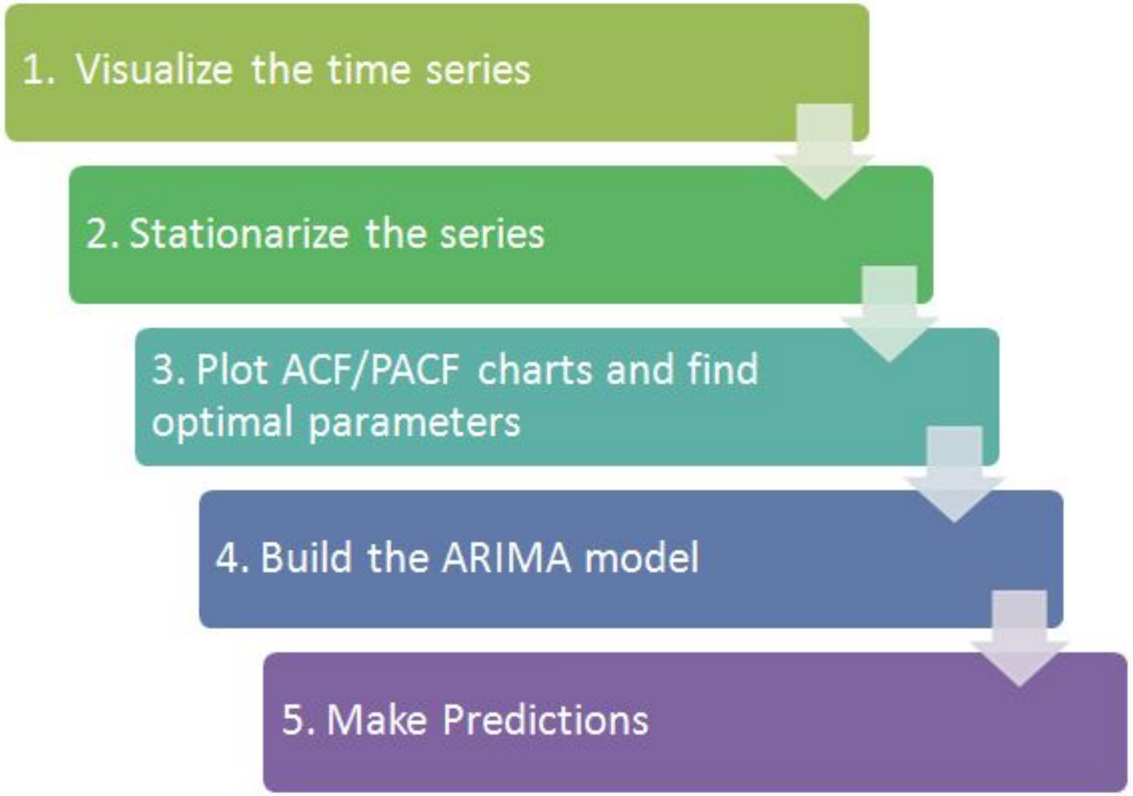
Step 1: Differencing
if $d > 0$ **then**
    for $i \leftarrow 1$ **to** d **do**
        $y_i \leftarrow y_i - y_{i-1}$ `// Difference the data d times`
    end
end

Step 2: Model Identification
Use ACF and PACF plots to determine p and q;

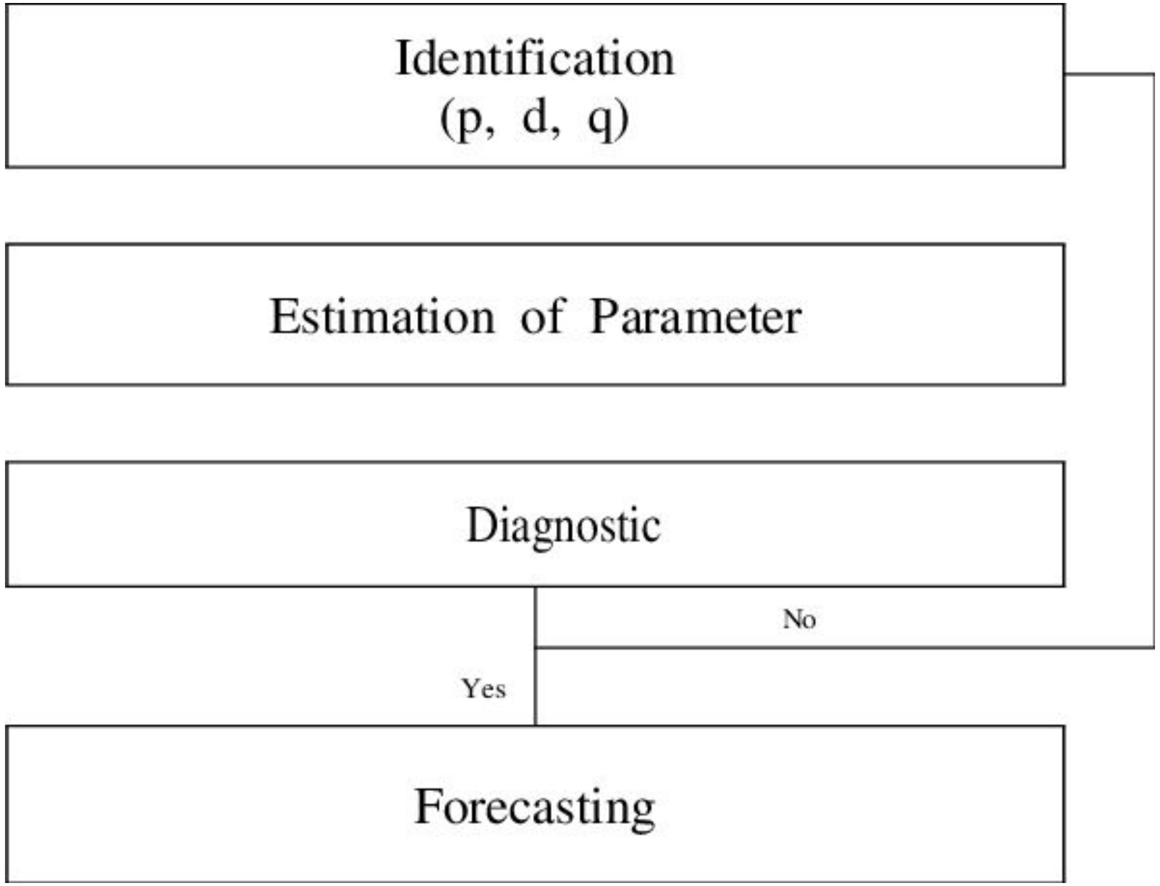
Step 3: Model Estimation
Fit an ARMA(p, q) model to the differenced data using maximum likelihood estimation;

Step 4: Inverse Differencing
if $d > 0$ **then**
    for $i \leftarrow 1$ **to** d **do**
        $y_i \leftarrow y_i + y_{n-i}$ `// Inverse difference the forecasts`
    end
end

Step 5: Forecasting
Use the fitted ARIMA model to make future forecasts;



Algorithms



Algorithm 1 SARIMA Model Fitting Algorithm

Require: A time series y_t and the order of the SARIMA model $(p, d, q)(P, D, Q)_s$.

Ensure: A fitted SARIMA model.

1: **Identify the order of the SARIMA** model This can be done using various methods, such as the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots.

2: **Preprocess the data** This may involve removing outliers and transforming the data to make it stationary.

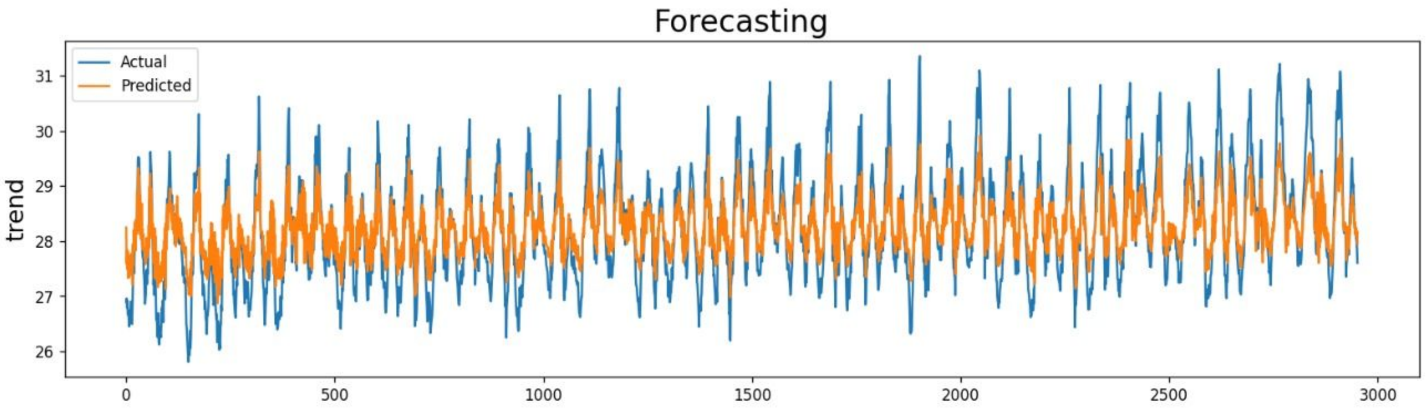
3: **Estimate the parameters of the SARIMA model** This can be done using various methods, such as maximum likelihood estimation.

4: **Check the adequacy of the model** This can be done by examining the residuals of the model and ensuring that they are white noise.

5: **Use the fitted model to forecast future values of the time series**

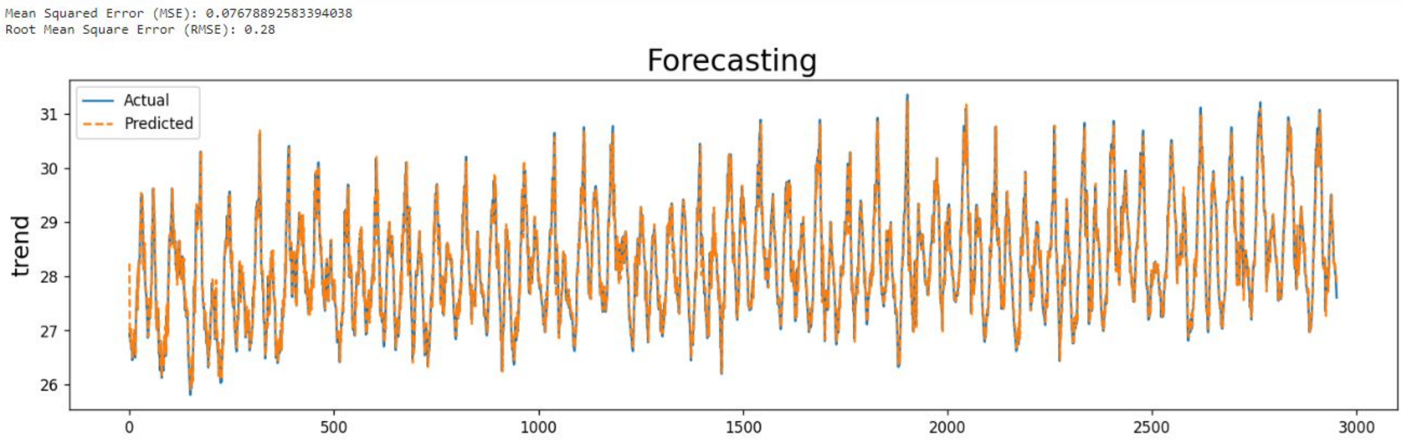
Result Analysis - ARIMA

The work can be found here .



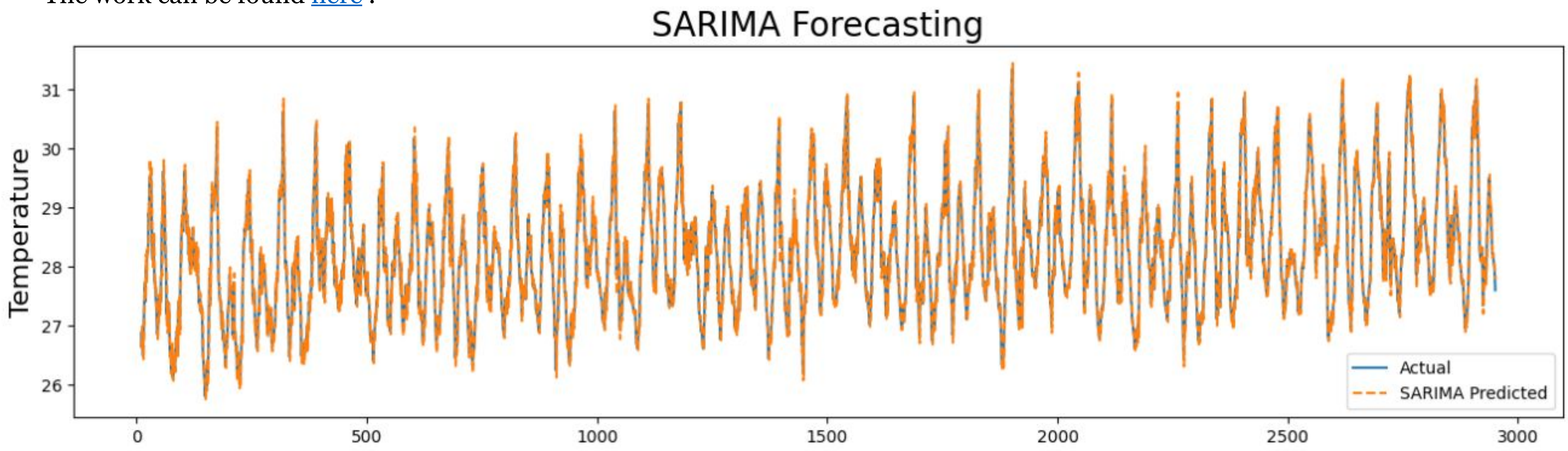
Arima model was significantly improved and provided us with very less mse of 0.07 and rmse of 0.28

Although this looks like a good model, the errors are low but still Arima alone is not able to reach all the peak anomalies within the time series, so we had to try other models too.



Result Analysis - SARIMA

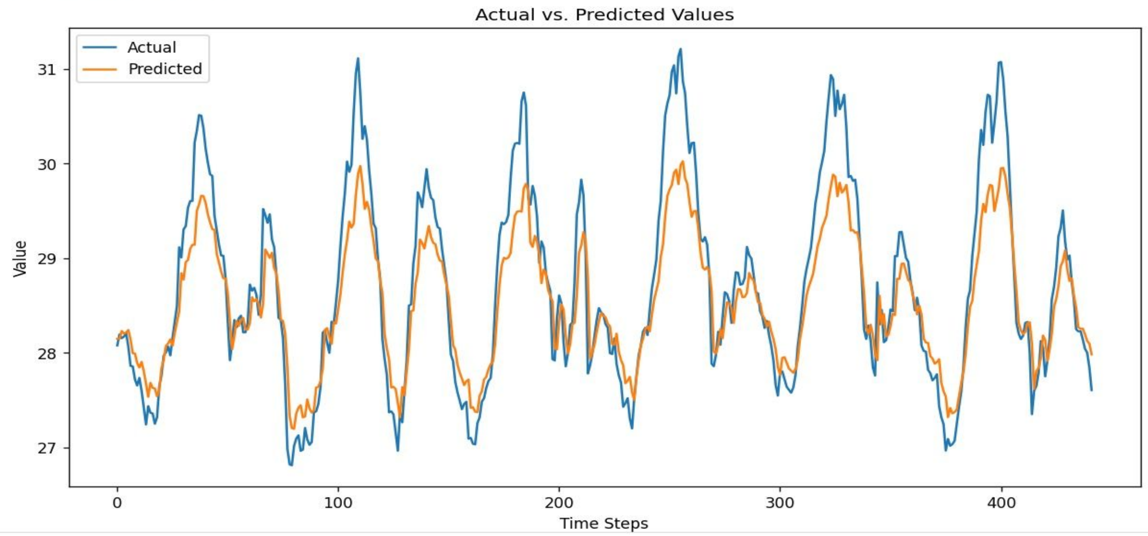
The work can be found [here](#) .



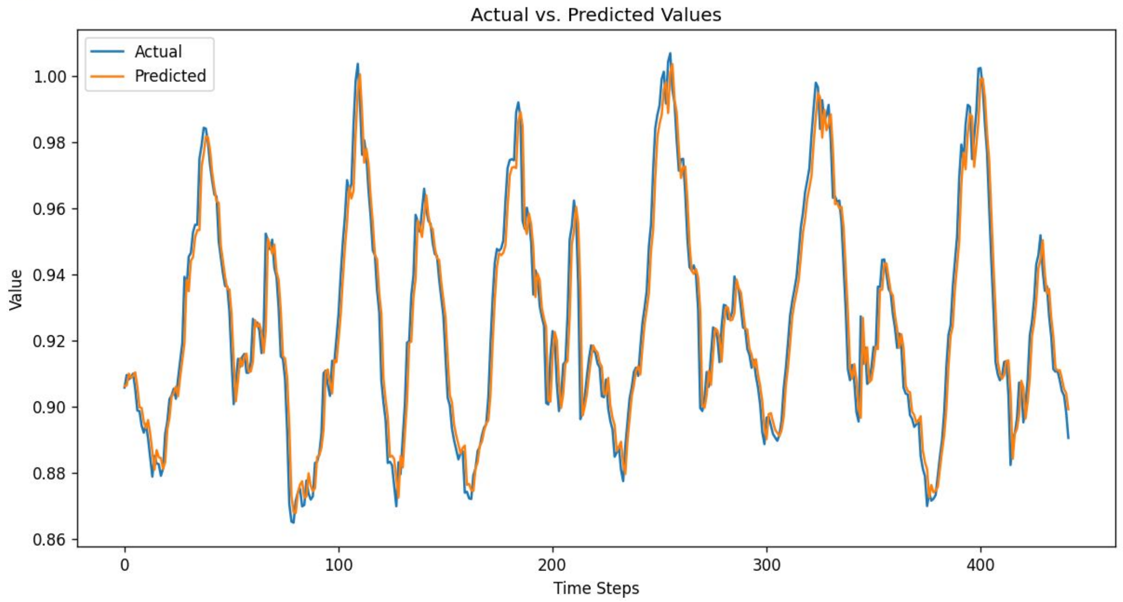

```
Mean Squared Error (MSE): 0.07814550106564634
Root Mean Square Error (RMSE): 0.28
```

Sarima model performed better than Arima, although the mse and rmse are very close to each other, this model managed to reach all the way till the peaks thus giving us a better chance at finding anomaly points.

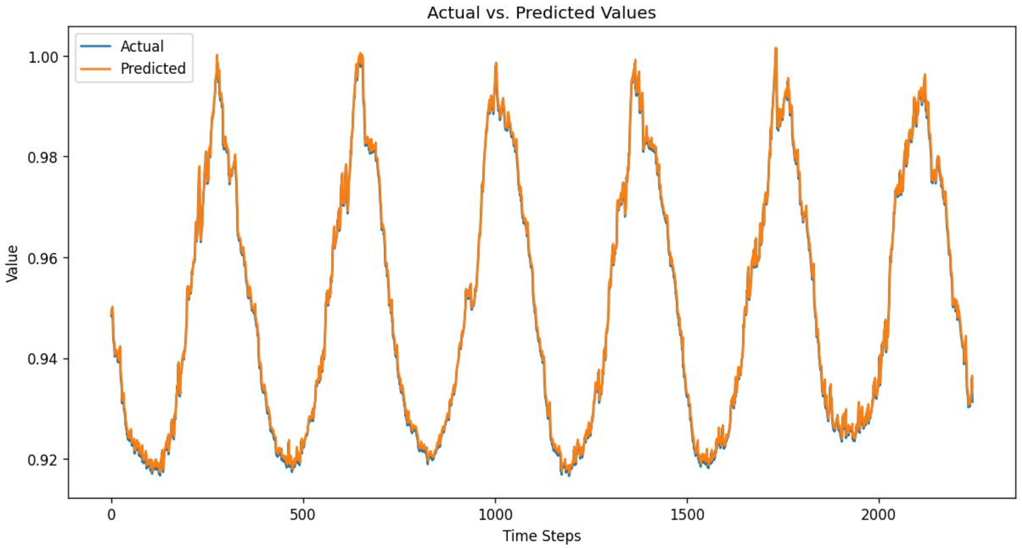
Result Analysis - LSTM



Mean Squared Error (MSE): 7.074711047685766e-05
Standard Deviation: 0.04669551757091632
Root Mean Square Error (RMSE): 0.01



Mean Squared Error (MSE): 2.1608480866302483e-06
Standard Deviation: 0.0353378782755088
Root Mean Square Error (RMSE): 0.00



Result Analysis - Custom Dataset (Gujarat Coast)

The work can be found here .

Features in the dataset :

Year	Month	Day	Mean temperature in kelvin	Mean temperature in Celsius	Mean temperature uncertainty	Fraction of sea ice covered ocean

Performance Evaluation :

Model Name	Mse	Rmse
Arima	0.076	0.28
Sarima	0.078	0.28
LSTM	$7.07 * 10^{-5} = 0.0000707$	0.01

Result Analysis - ERA5 Dataset (Korean Region)

Paper name: Deep-learning model for sea surface temperature prediction near the Korean Peninsula

Features in the dataset :

dimensions(sizes):	longitude(57)	latitude(37)	time(14976)

variables(dimensions):	**float32** longitude(longitude)	**float32** latitude(latitude)	**int32** time(time)	**int16** sst(time, latitude, longitude)

Performance Evaluation

Model Name	RMSE	MAPE
LSTM based SST model 1 day prediction by Authors	0.119	0.362
LSTM based model implemented by us (1 day prediction)	~0.00	0.299

```
Mean Squared Error (MSE): 2.157224326702524e-06
Standard Deviation: 0.03532362813799465
Root Mean Square Error (RMSE): 0.00
Mean Absolute Percentage Error (MAPE): 2.99%
```

Result Analysis - Comparison of LSTM Model

Reason why our model worked better than the authors' model:

- We focused on using normalization to the dataset before utilizing the data in LSTM model.
- Before such normalization the error was too high. The value of temperatures were around 200-280 which account for higher errors.
- Normalizing the values to a range of 0-1 which makes this calculations easier, normally this practice is followed in CNN for pixel values, which was used here the same is proved in next slide.

Result Analysis - Comparison of LSTM Model

Before Normalization

```
Mean Squared Error (MSE): 55.989358691902
Standard Deviation: 7.474080724886077
Root Mean Square Error (RMSE): 7.48
Mean Absolute Percentage Error (MAPE): 2.33%
```

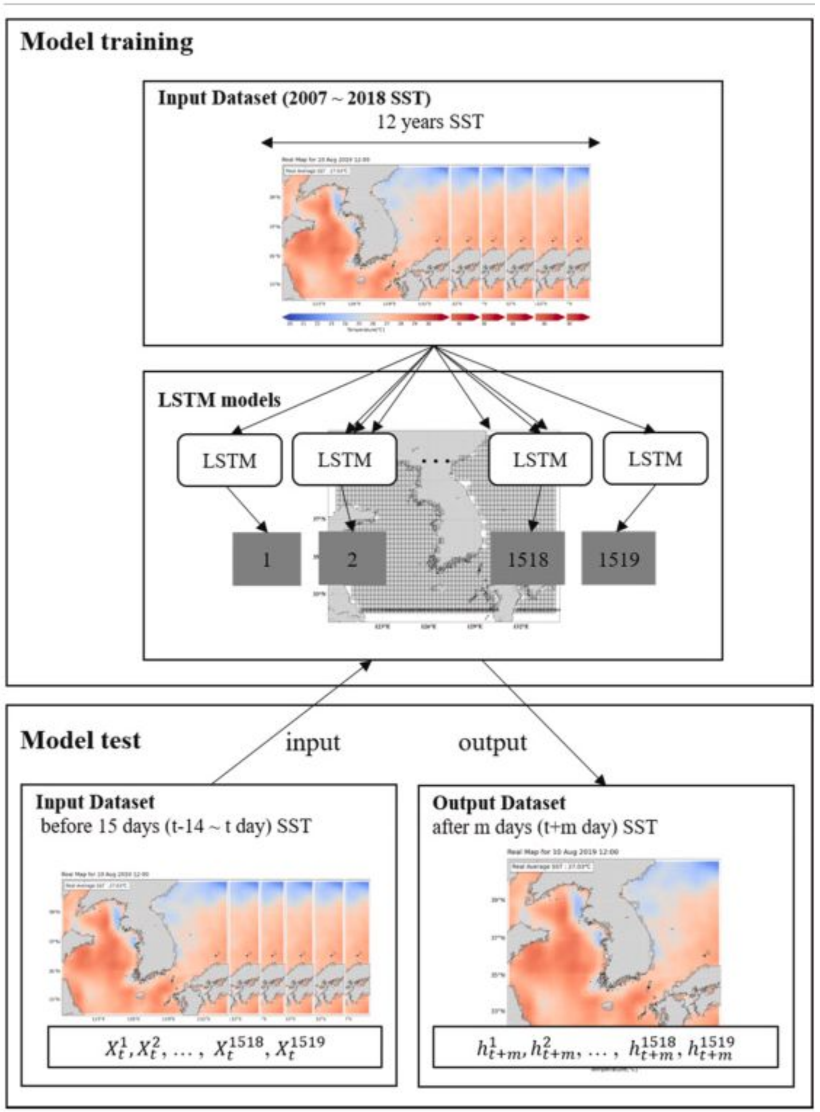
The error is too high and our model did very bad performance

After Normalization

```
Mean Squared Error (MSE): 2.157224326702524e-06
Standard Deviation: 0.035323628137994 65
Root Mean Square Error (RMSE): 0.00
Mean Absolute Percentage Error (MAPE): 2.99%
```

After normalization the model has drastically improved and the error decreased by vast number, thus normalization helped our model to greatest extent.

Architectural Difference of LSTM Models



Author's Usage:
Single lstm gate for each pixel.
results tabulated for best performed pixel.

Our Usage:
Single lstm gate for last point
results tabulated on the same point only
But one crucial point is that the epochs run more faster

Experimented extra:
tried to make the lstm gates more complex by using more layers of lstm gates but the output was not as good as single lstm gate output, also its taking more time to run epochs.

Conclusion and Future Scope

- **Conclusion:** Our project demonstrated the effectiveness of LSTM, ARIMA, and SARIMA models in anomaly detection. These diverse approaches offer valuable insights into time series data, providing a robust foundation for anomaly detection in various domains.

- **Future scope:** Looking ahead, the future scope for this project includes the exploration of hybrid models that combine LSTM, ARIMA, and SARIMA to enhance anomaly detection accuracy. Additionally, further research into optimizing model hyperparameters and streamlining real-time implementation will advance its practical applications.

References

[1] Y Zhou, C Lu, K Chen, X Li, "Multilayer fusion recurrent neural network for sea surface height anomaly field prediction", *IEEE transactions on geoscience and remote sensing*, vol. 60, pp. 1-11, 2022, 2022.

[2] H Yang, Y Wan, Q Yao, B Bao, C Li, Z Sun, H Wang, J Zhang, M Cheriet, "Anomaly Prediction With Hybrid Supervised/Unsupervised Deep Learning for Elastic Optical Networks: A Multi-Index Correlative Approach", *Journal of Lightwave Technology*, Vol. 40, Issue 14, pp. 4502-4513, Jul 2022.

[3] G Xiang, R Lin, "Robust Anomaly Detection for Multivariate Data of Spacecraft Through Recurrent Neural Networks and Extreme Value Theory", *IEEE Access*,Vol. 9, pp. 167447-167457, 2021.

[4] Q Shao, W Li, G Hou, G Han, X Wu, "Mid-term simultaneous spatiotemporal prediction of sea surface height anomaly and sea surface temperature using satellite data in the South China Sea", *IEEE Geoscience and Remote Sensing Letters*, Vol. 19, pp. 1-5, 2020.

Thank You