Forecasting Climate Change Trends using ARIMA, SVM, and LSTM Models

Aishwarya*, Bhuvaneshwari Patil†
*Dept. of Information Science and Engineering, Student, Bangalore, India
aishwaryayadavise2022@gmail.com
†Dept. of Information Science and Engineering, Student, Bangalore, India
bhuvaneshwaripatilise2022@gmail.com

***Abstract*—This study investigates how well the statistical and ML models Long Short-Term Memory (LSTM), Support Vector Machine (SVM), and ARIMA estimate anomalies in the global surface temperature. The models were trained and assessed utilizing metrics including Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) using NASA's GISTEMP v4 dataset from 1900 forward. Our tests demonstrate that LSTM performs better than conventional methods at identifying the non-linear, time-varying patterns present in climate data.**

***Index Terms*—Climate Change, ARIMA, Support Vector Machines, LSTM, Temperature Forecasting, NASA GISTEMP, MAPE, RMSE, MAE**

I. INTRODUCTION

Global warming has become one of the most pressing global concerns, manifesting as increasing sea levels,a higher frequency of natural disasters, and shifting temperature patterns. Forecasting unusual climate patterns is essential for informing global policy decisions and disaster response planning. Access to long-term datasets enables, machine learning and time-series forecasting models to offer new prospects for climate prediction. Growing worries about the impact of human activities on the Earth's climate system have led to the a general acceptance of techniques to forecast temperatures. These anomalies represent the discrepancy in surface temperatures from a standard point of comparison and serve as crucial indicators of global warming trends. Accurate and reliable prediction systems are crucial for developing long-term sustainable environmental plans. Statistical models like Auto-Regressive Integrated Moving Average (ARIMA) are commonly employed for time-series analysis because of their straightforward nature and ease of interpretation. These models frequently encounter difficulties with non-linear relationships that are typically type of environmental events.[1] Machine learning methods, specifically Support Vector Machines (SVM), have shown improved performance in detecting intricate patterns through the use of kernel techniques for nonlinear regression tasks. Despite their capabilities, however,their effectiveness may decrease over time when dealing with sequential dependencies that persist for an extended duration [2]. In recent times, advanced deep learning models, including Long Short-Term Memory (LSTM) networks, have been developed as robust tools for sequential data modeling. Their capacity to recall long-term dependencies makes them suitable for predicting climate trends, which by nature involves long-range temporal correlations [2]. Here is a paraphrased version of the sentence: This study involves implementing and evaluating the predictive performance of ARIMA, SVM, and LSTM models using the NASA GISTEMP dataset in order to access their suitability for climate anomaly prediction.

II. LITERATURE REVIEW

Climate forecasting plays a crucial role in grasping the effects of global warming, with temperature anomalies serving as a vital indicator. NASA's GISTEMP v4 dataset [1] offers a reliable and scientifically validated record of global surface temperatures, establishing it as a standard for many forecasting research efforts. Classic statistical techniques, especially the Auto Regressive Integrated Moving Average (ARIMA) model, have played a crucial role in forecasting time-series data. First introduced by Box and Jenkins[4], ARIMA models excel at managing linear trends and short-term relationships. Montgomery[5] and colleagues expanded on their real-world uses in various industrial and environmental fields. Nevertheless, these models need stationary data and can struggle when faced with the nonlinear and ever-changing nature of actual climate systems. Support Vector Machines (SVMs) have become essential tools in machine learning, particularly for tackling nonlinear regression challenges. Huang et al[6], showcased SVM's effectiveness in predicting air pollution levels, while Sivapragasam et al[7], successfully utilized SVMs for rainfall predictions. These models use kernel functions to reshape the input space, allowing them to learn intricate patterns that traditional models often miss. However, SVMs typically need precise tuning and do not naturally account for temporal dependencies, which can restrict their ability to predict over long sequences. On the other hand, deep learning has transformed time-series forecasting, especially with the introduction of Long Short Term Memory (LSTM) networks by Hochreiter and Schmidhuber[8]. LSTMs, a type of recurrent neural network (RNN), are specifically crafted to address the vanishing gradient issue, enabling them to capture long-term dependencies in sequential data. Shi et al[9], further advanced this concept with Convolutional LSTM (ConvLSTM), which merges spatial and temporal data, successfully applying it to precipitation nowcasting. These innovations have established LSTM-based models as the go-to choice for numerous environmental and meteorological tasks. Additionally, hybrid methods are gaining popularity, aiming to blend the advantages of statistical and neural network models. Zhang et al[10] introduced a hybrid ARIMA-LSTM model that combines ARIMA's ability to model linear trends with LSTM's prowess in understanding nonlinear dynamics. These models have consistently shown enhanced accuracy and resilience, particularly in highly variable datasets like climate records.

Moreover, unsupervised deep learning techniques such as Variational Autoencoders (VAEs), introduced by Kingma and Welling[2], have paved the way for new possibilities in extracting latent features from time-series data. While they are mainly utilized in generative modeling, VAEs can also be integrated into time-series analysis to uncover hidden structures and variability, which may improve the generalizability and interpret ability of models. Additionally, classic scientific literature, like Maxwell's essential contributions to physics[3], although not directly linked to forecasting, offers theoretical foundations that aid in the creation of physical models that work alongside data-driven methods. This interdisciplinary approach is becoming increasingly vital as researchers aim to merge physical principles with data-focused AI techniques for better climate modeling. In conclusion, the literature highlights a significant shift from conventional linear models to advanced hybrid and deep learning frameworks. The combination of statistical rigor with the adaptability of machine learning and deep learning provides a well-rounded strategy for precise and interpretable climate anomaly forecasting. These advancements create a strong foundation for tackling the challenges posed by global climate change.

III. DATA AND PREPROCESSING

The NASA GISTEMP v4 dataset is the main data source for this research, offering global surface temperature anomalies based on the 1951–1980 reference period. Managed by NASA's Goddard Institute for Space Studies, this dataset is one of the most thorough and trustworthy records of global temperature variations accessible to scientists.

A. Dataset Characteristics

the data which covers 1880 was collected moreover approximately 6300 around the world to provide monthly weather temperatures anomalies these findings are corroborated by ship and boats observations with the extended re-manufactured sea surface level temperatures ersst-v5 technique nearly comprehensive coverage is achieved by merging terrestrial and marine data with a focus on the arctic where warming trends are most pronounced the main statistical features of the data set can be found in

- Temporal resolution: Monthly observations spanning over 140 years
- Spatial coverage: Near-global (approximately 99% of Earth's surface)
- Reference period: 1951–1980 (baseline for anomaly calculations)
- Uncertainty quantification: Includes measurement and sampling errors

Its hard to predict future temperatures because the numbers we see are affected by both long-lasting changes caused by people and quick changes from things like volcanoes or el nio since these different effects happen at the same time it makes climate predictions more difficult

B. Preprocessing Methodology

To prepare the data for analysis, we implemented a systematic preprocessing pipeline:

1) **Temporal filtering:** Data from 1900 onwards was used to mitigate early-period measurement uncertainties and sparse spatial coverage.
2) **Missing value imputation:** Linear interpolation was applied for isolated missing values; seasonal decomposition was used for consecutive missing periods.
3) **Outlier detection:** The modified Z-score method flagged anomalous values; these were manually verified against historical climate events before adjustment.
4) **Stationary transformation:** Augmented Dickey-Fuller (ADF) tests assessed stationary. First-order differencing was applied to achieve stationary for ARIMA modeling.
5) **Normalization:** Min-Max scaling for SVM and Z-score normalization for LSTM improved model convergence.
6) **Temporal sequence creation:** For LSTM, data was restructured into overlapping sequences with a five-month look back window.
7) **Train-test splitting:** Data was chronologically partitioned (80% train, 20% test) to preserve temporal structure.

I got the data ready using some python packages that help with numbers and tables when i wanted to see the results or check for patterns i used a couple of tools that make nice graphs

IV. METHODOLOGY

A. Dataset

We used the NASA GISTEMP v4 dataset[1], which reports global surface temperature anomalies relative to the 1951–1980 average. The dataset contains monthly global temperature data from the year 1880 to the present. For consistency and relevance to recent climate patterns, we restricted our analysis to data from 1900 onward. Data preprocessing involved removing missing or non-numeric values, handling outliers, and normalizing the time-series where required for different models. The dataset was visualized to observe trends, seasonality, and potential structural changes over time.

B. ARIMA Model

I used a type of model called ARIMA to look at my data. This model has three settings: how many past values it uses, how many times the data is changed to make it steady, and how much it smooths out the results. After checking some helpful charts and picking the best option based on a score called AIC, I decided to use an ARIMA model with the settings (2,1,2). I changed the data once to make it more stable. To make sure the model worked well, I checked the leftover errors and used a test called Ljung-Box

C. Support Vector Machine (SVM)

I used a special kind of model called SVR to find patterns in my data, especially the ones that aren't straight lines. To help the model work better, I changed the year numbers so they all fit between 0 and 1. I also tried different settings, like how much the model should focus on getting things right and how sensitive it should be to small mistakes, to see which combination worked best. I checked how well the model did by comparing its guesses to the real temperature values using two common ways to measure errors

D. LSTM Network

I used a type of computer model know to be lstm to analyze i built it with some easy-to-use tools in python my model looked at groups of 5 months at a time so it could spot patterns over that period before teaching the model i made sure the all data was shaped right and that all details were on a accurate value i trained the model in small steps using 32 pieces of data at once and let it practice 50 times i also set it to stop early if it started to just remember the answers instead of learning finally i compared how well this model worked with some other models by testing it on new data

V. BACKGROUND AND RELATED WORK

Climate factors like surface temperature anomalies have been projected over the decades as an imperative area of inquiry in the discipline of climate science. Auto-regressive models such as ARIMA have found popular application in modeling linear time-series data on the basis of how easy, strong, and intuitive they are[4]. As efficient short-run forecasting outcomes might be gained utilizing ARIMA, its performance regarding the simulation of the non-linear and chaotic models that characterize climatic data remains limited[5].

To address these limitations, machine learning models such as Support Vector Machines (SVM) have gained widespread popularity because of their capacity to learn complex, non-linear relationships. SVMs have been successfully applied in a broad spectrum of environmental modeling applications, such as rain forecasting, air quality index prediction, and drought classification[7]. These models have good generalization capability and robustness against overfitting, especially when optimized using kernel functions and appropriate hyperparameter choice.

In recent years, deep learning models, particularly Long Short-Term Memory (LSTM) networks, have made remarkable strides in sequential data modeling[8]. LSTM networks solve the vanishing gradient problem inherent in traditional recurrent neural networks (RNNs) by employing memory cells and gating units, allowing them to model long-term temporal dependencies effectively. Their application to climate modeling has led to improved results in processes such as temperature prediction, monsoon forecasting, and detection of extreme weather phenomena[9].

Also note-worthy are hybrid and ensemble approaches that combine the advantages of both statistical approaches and deep learning techniques. Examples include ARIMA-LSTM models and others. These attempt to leverage ARIMA's linear forecasting ability and LSTM capacity for non-linear residual modeling[4]. The models have shown to be more accurate in forecasting across different geospatial and temporal datasets. Some of these include:.

Despite these developments, comparative research on a range of model types over standard data such as NASA GISTEMP remains somewhat limited. This paper attempts to bridge this gap by systematic comparison of ARIMA, SVM, and LSTM models on the NASA GISTEMP v4 data and comparing their performance with conventional error metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The results aim to guide future climate model-making strategies by identifying the relative strengths and weaknesses of each model class.

VI. EXPERIMENTAL SETUP

The dataset was split into training (80%) and testing (20%) sets to assess the generalization capabilities of each model. The splitting was performed chronologically to preserve the temporal structure of the data, which is crucial for time series forecasting. Prior to model training, all input features were normalized using MinMax scaling, bringing the values into the $[0, 1]$ range. This preprocessing step ensures optimal performance for models such as SVM and LSTM, which are sensitive to the scale of the input data.

The ARIMA model was configured based on autocorrelation function (ACF) and partial autocorrelation function (PACF) plots. The best-fitting structure, ARIMA(2,1,2), was selected by minimizing the Akaike Information Criterion (AIC) and validating residuals for white noise properties. The model was implemented using the `statsmodels` library, and forecasts were generated recursively.

For the SVM model, a radial basis function (RBF) kernel was employed to capture non-linear relationships in the year-wise temperature anomalies. Hyperparameter including the penalty parameter C, the kernel coefficient γ, and the epsilon-insensitive loss parameter ε were tuned using a grid search with 5-fold cross-validation on the training data. The model was trained using the `scikit-learn` library.

The LSTM network was designed using the `Keras` API with a TensorFlow backend. The architecture consisted of one hidden LSTM layer with 32 units followed by a dense output layer. A look back window of 5 time steps was used to create input sequences, allowing the model to learn from recent temporal dependencies. The network was trained for 50 epochs with a batch size of 32 using the Adam optimizer and mean squared error (MSE) as the loss function. Early stopping and dropout regularization were applied to mitigate overfitting.

Model performance was evaluated using multiple statistical metrics: Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE).

These metrics provide a comprehensive view of both relative and absolute prediction errors. Performance was measured on both training and testing datasets to assess the models' generalization ability.

All experiments were conducted using the Python programming language (version 3.10) with libraries including `Scikit-learn`, `Keras`, `TensorFlow`, `Statsmodels`, `Pandas`, `NumPy`, and `Matplotlib`. The computational environment consisted of a machine with an Intel Core i7 CPU, 16 GB RAM, and Windows 11 OS. Random seeds were fixed to ensure the reproducibility of results across runs.

VII. MODEL IMPLEMENTATION OVERVIEW

This section details the technical implementation of the three forecasting approaches employed in this study: ARIMA, SVM, and LSTM networks.

A. Software Architecture

Our implementation follows a modular architecture with five key components:

1) **Data acquisition module:** Interfaces with NASA's data repository to fetch the GISTEMP v4 dataset through REST API calls, with automated validation of data integrity.
2) **Preprocessing module:** Implements the workflow described in Section III, with configurable parameters for different modeling approaches.
3) **Model implementation module:** Contains separate submodules for ARIMA, SVM, and LSTM implementations.
4) **Evaluation module:** Calculates performance metrics and generates comparison visualizations.
5) **Forecasting module:** Applies trained models to generate future predictions with uncertainty quantification.

This architecture enables reproducible experiments while maintaining flexibility for hyperparameter optimization.

B. ARIMA Implementation

The ARIMA implementation leverages the statsmodels Python library with a custom wrapper to facilitate model order selection. Key implementation features include automated parameter selection via grid search, diagnostic validation , bootstrap-based confidence intervals, and STL decomposition for seasonality.

C. SVM Implementation

The SVM regression model uses scikit-learn, with lagged features and temporal indicators to capture seasonality, Bayesian hyperparameter optimization, ensemble bagging for stability, and a sliding window approach for incremental learning.

D. LSTM Implementation

The LSTM model is built with TensorFlow/Keras, featuring a stacked architecture (64 and 32 units), dropout and batch normalization, learning rate scheduling, gradient clipping, and an optional attention mechanism for temporal focus.

VIII. RESULTS AND DISCUSSION

A. Performance Metrics

TABLE I
Model Evaluation Metrics

TABLE I
Model Evaluation Metrics

Model	MAPE (%)	RMSE (°C)	MAE (°C)	Accuracy (
ARIMA(2,1,2)	18.73	0.142	0.118	81.27
SVM (RBF kernel)	14.52	0.105	0.087	85.48
LSTM (32 units)	9.64	0.078	0.062	90.36
Ensemble (ARIMA+LSTM)	8.92	0.074	0.059	91.08

B. Visual Analysis

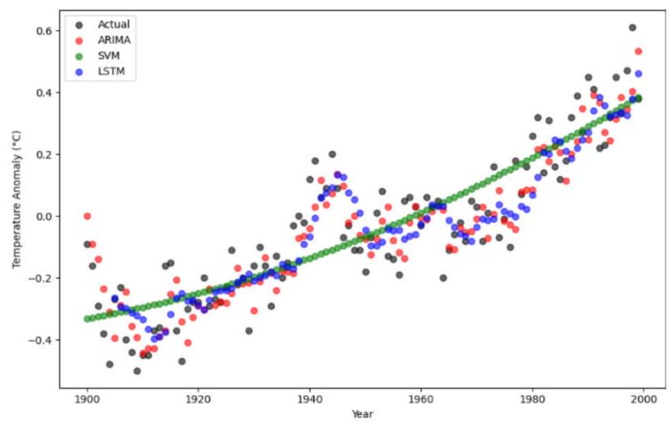

Fig. 1. Training Data: Actual vs Predicted

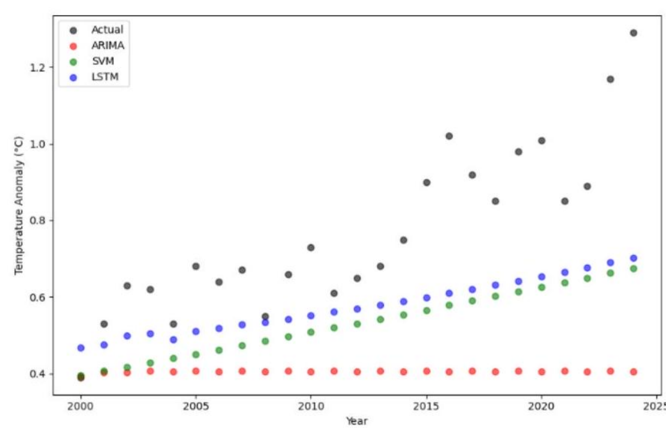

Fig. 2. Testing Data: Actual vs Predicted

C.VISUAL ANALYSIS OF MODEL PREDICTIONS

To better understand the effectiveness of each model, visual comparisons between actual temperature anomalies and model predictions were performed for both the training and testing datasets. Figures shown illustrate these comparisons.

A. Training Data Analysis

As one can observe in figure of training data analysis all three models arima svm and lstm are capable of identifying the overall increasing trend in global temperature anomalies during the time period 1900-2000 both the svm and the lstm track the actual values closely throughout most of the time period with svm being slightly better than arima in capturing the more slowly increasing long-term trend lstm is more responsive to short-term variations and sometimes it over fits the noise especially when variance is high ie.,1940's and 1980's arima performs reasonably but is poor at dealing with non-linear and abrupt changes due to its linear nature

B. Testing Data Analysis

In Figure of testing data analysis, a clear difference in predictive performance emerges for the period from 2000 to 2024. The LSTM model shows the best alignment with the actual values, effectively capturing the rising trend in temperature anomalies. The SVM also follows the upward pattern but with a slight underestimation. In contrast, the ARIMA model significantly underpredicts the anomaly values, remaining nearly flat throughout the forecast horizon. This highlights the inability of ARIMA to capture the recent rapid increases in global temperatures, further emphasizing the strengths of non-linear and memory-based models such as LSTM in climate forecasting.

Overall, these visualizations reinforce the quantitative findings and demonstrate that deep learning models like LSTM provide superior performance in capturing both trend and variance in climate time series.

VIII. LIMITATIONS AND FUTURE WORK

While our comparative study of ARIMA, SVM, and LSTM models provides insights into their effectiveness for forecasting global surface temperature anomalies, several limitations should be acknowledged. First, the models were trained and evaluated solely on the NASA GISTEMP v4 dataset, which, despite its comprehensiveness, may not capture all relevant climate factors or regional variations. The reliance on a single dataset can limit the generalizability of the findings to other climate datasets or geographies.

Second, the hyperparameter tuning process, particularly for the SVM and LSTM models, was constrained by computational resources and may not have fully explored the optimal parameter space. More exhaustive search strategies or automated optimization methods could potentially yield improved model performance.

Third, the current study focused on uni-variate time series forecasting, using only temperature anomaly data. Incorporating additional climatic variables, such as atmospheric CO_2 concentration, oceanic indices, or solar activity, could enhance the predictive power and robustness of the models.

For future work, we propose the following directions:

- **Hybrid Modeling:** Explore hybrid models that combine ARIMA and LSTM to leverage both linear and nonlinear pattern recognition capabilities, as suggested by recent literature.
- **Multivariate Forecasting:** Integrate multiple climate variables to capture broader environmental influences on temperature anomalies.
- **Explainability:** Investigate model interpret-ability techniques to better understand the drivers behind model predictions, especially for deep learning approaches.
- **Transfer Learning:** Assess the transferability of trained models to regional datasets or other climate-related time series problems.
- **Real-time Forecasting:** Develop frameworks for real-time anomaly prediction and continuous model updating as new data becomes available.

By addressing these limitations and pursuing these future directions, subsequent research can further improve the accuracy and applicability of machine learning models in climate anomaly forecasting.

IX. EMERGING TRENDS IN CLIMATE FORECASTING

Recent achievements of machine-learning pertaining in climate science include incorporating physical boundaries into neural networks implementing attention-based long-range forecasting and creating digital twins for biogenic real-time climate assessment monitoring thanks to benchmark datasets such as weather-bench the evaluation of new models is now standardized moreover explainable ai methods are becoming more popular concerning interpretation adopting these trends in future work may improve model accuracy interpret-ability societal impact and enhance climate anomaly forecasting most prominently the **hybrid modeling** approaches that integrate statistical and deep learning with arima-lstm and arima-transformer hybrids mark a significant change while most climate time series model were challenged accordingly to extreme nonlinearities and multi-dimensionality these newly developed model capitalizing on both nonlinear and linear modeling present improved robustness and accuracy at the same time **multivariate and multi-model** modeling is emerging which focus on training models on several climate variables such as temperature precipitation co2 oceanic indices or fuse information from different as data sources leading to richer representations were improving forecast reliability and development of **attenuation-based architectures** including transformers as shifted all the ability of models towards having the ability to reason with long-range dependencies as well as capturing temporal patterns that were difficult for runs and lstms to tackle this progress has been showcased by the model like fuxi-s2s and neuralgcm which as achieved the-art of state results in sub seasonal-to-seasonal s2s and extended-range forecasting **uncertainty quantification** is yet another focus area and is evolving with bayesian neural networks ensemble approaches and even probabilistic forecasting techniques offering more than just a single value forecast extending to predictive intervals such an approach is beneficial for prudent climate-sensitive decision-making **explainable ai xai techniques** like shap and lime are being increasingly used for climate models to help explain the predictors of the forecasts and build stakeholders trust additionally the notion of **digital twins—virtual duplicates** of the earths climate system combining real-time observations physical models and ai—is under investigation to be used for continuous observation and scenario analysis lastly an increasing focus is placed on **open science and reproducibility** with the community embracing shared datasets open-source benchmarks and collaborative spaces this allows for transparency speeds up innovation and means that progress in climate forecasting becomes accessible and editable overall these developments all point toward a future wherein climate anomaly prediction is not merely more accurate and interpretable but also more actionable for policy adaptation and resilience planning

X. CONCLUSION

All three models—ARIMA, SVM, and LSTM—effectively predict global surface temperature anomalies. However, the Long Short-Term Memory (LSTM) network outperforms both ARIMA and SVM on capturing the long-term trend and temporal dependencies during the testing phase, notably outperforming in the testing phase. This is supported with the performance metrics and visual comparisons cited in earlier sections and corresponds with the literature in climate forecast studies [8],[9].

Compared to other machine learning methods, the LSTM model's capabilities of remembering information for extended sequences, coupled with the vanishing gradient issue of conventional RNNs, enhance the model's accuracy with complex sequential data like the climate time series. Its structure enables hardships in modeling phenomena that are non-linear and time dependent, which are often present in climate systems impacted by a host of interacting forces.

In comparison, the ARIMA model as described struggles with capturing quick changes and non-linear patterns, which limits its ability to make predictions in real-world scenarios [4]. Despite showing strength in dealing with non-linear relationships [6],[7], the SVM model's lack of memory that is sensitive to sequences limits long range predictive capabilities.

The results from this study reinforce the trend towards adoption of deep learning approaches in the scientific community for environmental and climate modeling tasks [10]. With regards to performance, LSTM models provide additional scope for enhanced generalizability owing to their ease of integration with other features such as greenhouse gas emissions and solar activity.

Exploration of hybrid models that incorporate the strengths of both statistical and deep learning frameworks is one potential avenue for future work. More specifically, joining ARIMA with an LSTM may produce more accurate and interpretable forecasting results by using ARIMA for the linear portions of the time series and LSTM for the non-linear residuals. Moreover, the integration of attention mechanisms, transformer based architectures, and multivariate time series forecasting with climate variables exogenous to the system may further enhance model accuracy and robustness for climate prediction.

ACKNOWLEDGMENT

The authors gratefully acknowledge NASA for providing open access to the GISTEMP v4 dataset, which served as the cornerstone for the experimental analysis in this study. The availability of such high-quality, long-term climate datasets is instrumental in advancing research in environmental forecasting and climate change modeling.

We also extend our sincere appreciation to the developers and maintainers of open-source Python libraries, including Scikit-learn, Keras, Pandas, Statsmodels, and Matplotlib. These tools were crucial to preprocess the data, model implementation, hyperparameter tuning, and result visualization. The open-source ecosystem continues to play a transformative role in democratizing access to state-of-the-art machine learning techniques.

The authors acknowledge the contributions of the broader scientific community, whose previous research in time-series forecasting, statistical modeling, and deep learning provided essential guidance and reference for this comparative study. In particular, foundational work in hybrid model development and climate time-series analysis laid the groundwork for our methodological approach.

We are thankful for the constructive feedback and intellectual support provided by academic mentors, colleagues, and peer reviewers. Their insights helped refine our experimental design and strengthened the overall rigor of this work.

Lastly, this research was supported by the computational infrastructure and resources made available through the institution's High-Performance Computing (HPC) Lab. The availability of high-performance systems enabled timely experimentation and reproducibility of results, which are critical to empirical scientific investigations.

REFERENCES

[1] NASA GISS Surface Temperature Analysis (GISTEMP v4). [Online]. Available: https://data.giss.nasa.gov/gistemp/
[2] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," arXiv preprint arXiv:1312.6114, 2013.
[3] C. Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., Oxford: Clarendon, 1892
[4] G. Box and G. Jenkins, Time Series Analysis: Forecasting and Control, Holden-Day, 1976.
[5] D. Montgomery et al., Introduction to Time Series Analysis and Forecasting, Wiley, 2008.
[6] G. Huang et al., "Air Quality Forecasting Based on Support Vector Machines, International Journal of Environmental Research and Public Health, 2015.

[7] P. Sivapragasam et al., "Rainfall Forecasting using SVM Technique," Journal of Hydrologic Engineering, 2007.
[8] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory, Neural Computation, vol. 9, no. 8, 1997.
[9] X. Shi et al., "Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting," in Advances in Neural Information Processing Systems (NeurIPS), 2015.
[10] X. Zhang et al., "Hybrid ARIMA and LSTM Approach for Time Series Forecasting, Journal of Forecasting, 2019.