DEEP LEARNING DRIVEN CLIMATE CHANGE PREDICTION USING SPATIO TEMPORAL SATELLITE DATA FUSION FOR REAL TIME ENVIRONMENTAL MONITORING

G. BALU NARASIMHA RAO[1], ALURI LAKSHMI[2], U. SADHANA[3], Y. Y. S. MANOGNA[4], DEVADI GANESH[5]*, P. LAKSHMI PRASANNA[6], KARNATAPU LEELA KRISHNA[7], MANGALAGIRI SRIKANTH KUMAR[8]

[1] Assistant Professor, Department of Computer Science and Engineering, Vignan's Foundation for Science, Technology & Research, Vadlamudi, Guntur – 522213, Andhra Pradesh, India
[2] Associate Professor, Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning), Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, Andhra Pradesh, India
[3] Sr. Assistant Professor, Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning), CVR College of Engineering, Vastunagar, Ibrahimpatnam – 501510, Telangana, India
[4] Assistant Professor, Department of Electronics and Communication Engineering, Aditya University, Surampalem, Andhra Pradesh, India
*[5] Assistant Professor, Department of Computer Science and Engineering, GMR Institute of Technology (GMRIT) – Deemed to be University, Rajam – 532195, Andhra Pradesh, India
[6] Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur – 522302, Andhra Pradesh, India
[7] Associate Professor, Department of Civil Engineering, R.V.R. & J.C. College of Engineering, Chowdavaram, Guntur, Andhra Pradesh, India
[8] Assistant Professor, Department of Civil Engineering, R.V.R. & J.C. College of Engineering, Chowdavaram, Guntur, Andhra Pradesh, India
Email: balunarasimharao@gmail.com[1], alurilakshmi@gmail.com[2], ummareddysadhana@gmail.com[3], manognayadiri@adityauniversity.in[4], ganesh.d@gmritdu.edu.in[5]*, lakshmiprasannap87@gmail.com[6], leelakrishna.071305@gmail.com[7], m.srikanthkumar4@gmail.com[8]

ABSTRACT

It is essential to predict climate change in real-time and with high accuracy to minimize environmental risks and facilitate decision-making towards sustainability. Nevertheless, current models used to predict climate changes struggle to process heterogeneous spatio-temporal satellite data in order to ensure high prediction accuracy and low latency needed for real-time environmental monitoring. In this study, the new Deep Learning-driven Spatio-Temporal Satellite Data Fusion (DL-STF) approach will be developed by integrating various sources of satellite data, including MODIS, Sentinel-2, and ERA5, into a hybrid architecture combining Convolutional Neural Networks (CNNs) for spatial feature extraction, Bidirectional Long Short-Term Memory (Bi-LSTM) networks for temporal modeling, and a multi-head attention-based fusion mechanism for adaptive feature weighting. The novel approach described above can be regarded as the main research contribution due to its ability to jointly learn spatial and temporal representations through an attention-driven fusion strategy to improve climate prediction performance. The experimental analysis demonstrated the accuracy of the model equal to 96.3% while the error rate was relatively low (RMSE: 0.18, MAE: 0.12), and the latency was minimal (95 ms) compared to conventional CNN, LSTM, and CNN-LSTM models. As a result, it can be concluded that the novel attention-based spatio-temporal fusion framework significantly increases predictive accuracy, robustness, and computational efficiency, enabling scalable real-time environmental monitoring for early warning systems, climate risk assessment, and data-driven policy planning to strengthen climate resilience.

Keywords: *Climate Change Prediction, Spatio-Temporal Data Fusion, Deep Learning, Satellite Remote Sensing, Attention Mechanism, Real-Time Environmental Monitoring*

1. INTRODUCTION

Climate change has become one of the most urgent issues of the global world of the 21 st century, which affects not only the ecological stability, the agricultural productivity, the water resources, and the socio-economic systems of the world, but also the ecological and socio-economic situations of individual countries. The growing global temperatures, the rising frequency of extreme weather events coupled with the changes in the climatic patterns have increased the need to have accurate and real-time environmental monitoring systems. Recent research has indicated that the mean temperature on the surface of the globe has risen significantly over the past century, mainly because of the emission of greenhouse gases by humans and by changes in land use [1], [2]. Such transformations require sophisticated predictive models that can be used to capture complex environmental processes.

Long-term climate projections have been extensively made using traditional climate prediction models, which are mostly determined using physical and numerical simulations, including General Circulation Models (GCMs). Although these models are very useful, they are usually characterized by high computational complexity, coarse spatial resolution and limited adaptability to real time data streams [3], [4]. Moreover, they rely on established physical equations and they are therefore limited in their capability to represent nonlinear and highly dynamic interactions between the environment and the system under study.

With the development of satellite remote sensing technologies, particularly the high-resolution, continuous, and multi-dimensional data, the rapid progress of the technology has revolutionized the method of environmental monitoring. Satellite systems, like MODIS, Sentinel, and Landsat generate large amounts of spatio-temporal data, such as land surface temperature, vegetation indices, atmospheric composition, and oceanic parameters [5], [6]. The datasets represent unprecedented possibilities to comprehend the process of climate dynamics at both local and global levels. Nonetheless, the effective integration and analysis of such heterogeneous data is still a serious challenge because of the variation in the spatial resolution, temporal frequency, and data formats.

Over the past few years, machine learning and deep learning algorithm models have attracted a lot of interest in terms of climate data analysis. In contrast to the conventional models, deep learning models are able to discover complex nonlinear relationships in large datasets without needing explicit feature engineering. Convolutional Neural Networks (CNNs) have proven themselves to be strong in extracting spatial information in satellite images, whereas Recurrent Neural Networks (RNNs), or more specifically, Long Short-Term Memory (LSTM) networks, are good at modeling the temporal dependencies in sequential data [7]–[9]. The combination of CNN and LSTM architecture hybrid models have further advanced the prediction accuracy by utilizing spatial and temporal information [10], [11].

Regardless of all these improvements, there are a number of shortcomings in the methods that have been employed. A significant number of models individually address either spatial or temporal aspects of climate systems, making them less effective compared to models that tackle both spatial and temporal dimensions of climate systems. Also, the majority of the current frameworks do not incorporate effective mechanisms of multi-source data fusion which is critical in the integration of various satellite data. The lack of real-time processing power is another critical limitation that is important to early warning systems and dynamic environmental monitoring [12], [13].

This paper offers a solution to these problems by **deep learning-driven climate prediction framework based on spatio-temporal satellite data fusion**. The given approach combines the data on multiple sources of satellites in a hybrid CNN-LSTM structure with an attention-based fusion mechanism. The design allows the model to dynamically focus on the relevant features of the various data sources thereby enhancing prediction quality and strength.

Objectives of the Work

The primary objectives of this research are as follows:

1. To create a **spatio-temporal deep learning model**, which can accurately predict the climate variables based on multi-source satellite data.
2. To develop an **attention-based data fusion mechanism** that can be used to incorporate heterogeneous environmental data.
3. To permit **real-time monitoring of the environment** with less computation latency.
4. To enhance the prediction performance over the current machine learning and deep learning models.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5. To offer scalable infrastructure to examine **global climatic conditions and to support decision systems**.

By fulfilling these goals, the proposed framework will fill the gap between data availability and actionable climate intelligence. The combination of deep learning and satellite data fusion provides an opportunity to improve the accuracy of predictions and allow managing climate risks proactively.

Organization of the Paper

The rest of this paper was structured in the following way: Section 2 was dedicated to a review of the literature on the use of deep learning-based climatic predictions, spatio-temporal modeling, and satellite data fusion. Section 3 presented the suggested methodology, including data description, preprocessing, model description, mathematical formulation, and workflow description. Section 4 presented the experimental findings, performance assessment measures, comparative and contrasting analysis with the current models and the graphical interpretations. Lastly, Section 5 provided the conclusion of the paper with key findings, limitations, and directions to future research.

2. RELATED WORK

Recent innovations in climate informatics, remote sensing and deep learning have brought a surge of research with the aim of streamlining the quality of climate predictions using data-driven methods. The section critically evaluates the available body of literature on important areas, such as deep learning to model climate, spatio-temporal architectures, integrating satellite data, and systems to monitor the environment in real time.

2.1 Deep Learning for Climate Modeling

Their capability to learn nonlinear relationships by using large-scale data has started to gain significant importance within modeling complex climate systems by adopting deep learning techniques. Firstly it is known that predictive accuracy on temperature and precipitation pattern was better when Artificial Neural Networks (ANNs) were used to model predictive variables [14], [15]. Deep architectures to learn hierarchical feature representations in climate data were also introduced [16], [17].

More recently, downscaling and simulation of climate data using more coarse data has been done using generative models like Generative Adversarial Networks (GANs) [18]. Vast aspects of deep learning models are unable to co-operatively describe the spatial and temporal correlations that are observed in climate variability.

From these works, it is clear that deep learning has been proven to be quite effective for climate prediction; however, the major focus has been placed on the effectiveness of predictive models rather than unified learning of heterogeneous spatial and temporal information. Consequently, the development of an integrated framework capable of simultaneously modeling spatial features, temporal dependencies, and multi-source satellite observations remains an important research gap.

2.2 Temporal Modeling Using Recurrent Architectures

Climate prediction temporal dynamics are critical because environmental variables have convincing sequential relationships. Time-series forecasting Recurrent Neural Networks (RNNs), especially Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models have been extensively utilized to conduct time-series forecasting tasks, including temperature change analysis, rainfall prediction, and drought forecasting [19], [20].

The LSTM-based models do not suffer the drawback of not being able to capture long-term dependencies and seasonal variations. Nevertheless, such models mainly utilize only temporal information, and do not take into account any spatial correlations observed within geospatial data, limiting their usefulness in comprehensive climate models.

As can be seen from the current literature, existing recurrent architectures do not offer a complete solution for climate prediction because of their inability to jointly exploit complementary spatial information coming from satellite imagery. This limitation demonstrates the need for unified spatio-temporal learning frameworks.

2.3 Spatial Feature Learning with CNNs

The convolutional neural network (CNNs) has proven to be exceptionally successful in deriving spatial features of satellite images and geospatial data. CNN-based models have been utilized in the activities like land cover classification, vegetation monitoring and detection of environmental changes [21], [22]. The models have the potential to induce spatial heterogeneity as well as localized patterns in the variables associated with climate.

CNNs are however also limited when it comes to modeling the temporal dependencies that are fundamental to the comprehension of the changing dynamics of climate. Owing to this, standalone CNN-based methods cannot be used to predict long-term climatic conditions.

The problem here is that even though CNNs successfully learn spatial patterns, their inability to model temporal evolution limits the usage of CNNs for dynamic climate prediction tasks. Thus, there is a need for integration of spatial and temporal learning into a single deep learning architecture.

2.4 Hybrid Spatio-Temporal Deep Learning Models

To mitigate the constraints of single spatial and temporal models, both CNN and LSTM based on hybrid architectures have been suggested. These models utilize CNNs to extract spatial features and LSTMs to learn temporal sequence information, yielding a better predictive result [23], [24]. More advanced versions like ConvLSTM and 3D CNNs have also been presented to co-modelling spatio-temporal relationships in a more comprehensive manner.

Although there have been encouraging outcomes of the hybrid models, there have been deficiencies in the models in terms of incorporating other heterogeneous satellites data through efficient mechanisms. Also, not all of these models include adaptive feature selection mechanisms, which are essential to address the conditions experienced in dynamic environments.

Moreover, most hybrid architectures have insufficient capability to perform adaptive fusion of heterogeneous satellite observations while maintaining the computational efficiency required for real-time applications. These limitations represent an important research challenge that motivates the proposed DL-STF framework.

2.5 Multi-Source Satellite Data Fusion

Satellite data fusion is vital in integrating the information provided by the two or more sensors to enhance prediction accuracy and strength. The statistical averaging, Bayesian inference and Kalman filtering, are traditional data fusion methods that have been heavily used in remote sensing applications [25], [26]. Nevertheless, such methods are not applicable in addressing nonlinear association and multidimensional data.

The development of deep learning-based fusion techniques has become a force to reckon with; they can adjust to data at feature and decision making levels. Dynamic weighting of input features, enhanced by using fused representations, has been achieved using attention mechanisms and transformer-based architectures [27], [28]]. However, the current models of fusion are not only computationally-intensive, but also have no real-time processing capabilities.

From the above limitations, it can be concluded that existing approaches for satellite data fusion still lack an effective balance between adaptive feature integration, prediction accuracy, scalability, and low-latency processing. Thus, it is necessary to develop more efficient deep learning-based fusion strategies for practical climate prediction systems.

2.6 Real-Time Environmental Monitoring Systems

Real-time monitoring of the environment has gained more value in the use of applications in disaster management, agriculture and urban planning. Internet of things (IoT)-enabled systems combined with cloud computing systems have been created to support continuous data collection and processing [29]. Such systems present timely information, but they can be very easy to use wary of the potential of deep learning to complex climate prediction issues.

Moreover, issues of data heterogeneity, latency and scalability are also important issues that could hinder the implementation of real-time climate monitoring systems.

Recent work in deep learning for climate prediction using spatio-temporal fusion and satellite data has shown remarkable results. In particular, Lian et al. [30] highlighted recent work on deep learning for remote sensing spatio-temporal fusion and stressed the requirement for more efficient, scalable, and real-time fusion frameworks. Zhang et al. [31] introduced a transformer-based multivariate fusion network for improving short-term precipitation prediction by effectively capturing complex spatio-temporal relationships. Moreover, the FuXi-DA framework [32] demonstrated the effectiveness of deep learning-based data assimilation to integrate heterogeneous satellite observations and improve climate prediction accuracy. However, existing solutions lack a comprehensive approach for unified real-time climate prediction through adaptive attention-based fusion of heterogeneous multi-source satellite data. Most existing studies enhance prediction accuracy, spatio-temporal modeling, or satellite data assimilation separately, while limited attention has been given to integrating these capabilities within a single scalable framework. This literature gap

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

motivates the proposed DL-STF framework, which integrates CNN-based spatial learning, Bi-LSTM temporal modeling, and multi-head attention-based adaptive fusion to achieve accurate, scalable, and low-latency climate prediction.

2.7 Identified Research Gaps

According to the literature reviewed, some of the key research gaps are determined:

- Poor combination of multi-source satellite information.
- Poor ability in the modelling of combined spatio-temporal dependencies.
- Underdeveloped adaptive attention-based data fusion mechanisms.
- Inadequate attention to real time prediction and the low latency systems.
- Difficulty of managing the large scale heterogeneous environmental data.

2.8 Research Motivation

To address these limitations, the current paper presents a proposal of a spatio-temporal satellite data fusion system based on deep learning incorporating CNNs, LSTMs, and attention. The proposed solution is expected to increase the predictive accuracy, real-time monitoring and a scalable solution to global climate analysis.

2.9 Research Problem Statement

Despite considerable success achieved in climate prediction using deep learning models, existing approaches still suffer from a number of limitations. Current deep learning methods mainly concentrate on spatial and temporal feature learning separately, thus leading to the lack of adequate representation of complex climate dynamics. Although recent research has paid attention to the problems of spatio-temporal fusion, forecasting based on transformers, and deep learning-based satellite data assimilation, further issues with efficient fusion of heterogeneous multi-source satellite data along with high prediction accuracy, low computational latency, and scalability of satellite data in real time continue to exist [30]–[32]. Moreover, adaptive attention-based feature fusion lacks sufficient investigation for identification of the most informative spatial and temporal features of satellite data. These limitations imply the existence of significant knowledge gaps associated with the development of a unified deep learning framework allowing simultaneous learning of spatial features, temporal dependencies, and adaptive feature importance from heterogeneous satellite data. Addressing these research gaps is essential for advancing real-time climate prediction and strengthening environmental decision support. To fill these research gaps, this study proposes a deep learning framework for adaptive fusion of spatio-temporal satellite data (DL-STF), combining CNN-based spatial feature extraction, Bi-LSTM temporal modeling, and multi-head attention-based fusion within a unified end-to-end architecture. The contribution of the study is related to the provision of new knowledge about a novel deep learning framework that enables effective adaptive spatio-temporal satellite data fusion, improving prediction accuracy, computational efficiency, and scalability for real-time climate monitoring.

2.10 Research Questions

From the existing research gaps observed in recent literature, this paper aims to address the following research questions:

RQ1: How can heterogeneous multi-source satellite data be efficiently incorporated into a unified deep learning architecture for improving accuracy in climate prediction?

RQ2: Is it possible for CNN-based spatial feature extraction along with Bi-LSTM-based temporal modeling and multi-head attention-based adaptive fusion to effectively capture complex spatio-temporal climate dynamics?

RQ3: What is the extent to which the proposed DL-STF framework improves prediction accuracy, minimizes prediction errors, and ensures low inference latency as compared with conventional deep learning models?

RQ4: Whether the proposed framework can deliver a scalable and computationally efficient solution suitable for real-time environmental monitoring and climate prediction?

3. METHODOLOGY

In this paragraph, the proposed Deep Learning-driven **Spatio-Temporal Satellite Data Fusion Framework (DL-STF),** has been proposed to make real-time climate change predictions. By combining heterogeneous satellite data with a novel attention-based hybrid structure, the methodology is designed to guarantee reproducibility, scalability and high predictive accuracy.

3.1 Dataset Description

Multi-source satellite datasets have been taken into consideration to assure robustness and generalization. The dataset combines the

Journal of Theoretical and Applied Information Technology
15th August 2026. Vol.104. No.15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

observations of the MODIS, Sentinel-2, and ERA5 data of climate reanalysis.

3.1.1 Data Sources and Characteristics

Table 1: Summary of Multi-Source Satellite Data Sources and Their Characteristics

Dataset Source	Parameters	Spatial Resolution	Temporal Resolution	Time Period
MODIS	LST, NDVI, EVI	1 km	Daily	2010–2024
Sentinel-2	Vegetation, Land Cover	10–20 m	5 days	2015–2024
ERA5	Temperature, Humidity, Wind Speed, Pressure	0.25°	Hourly	2010–2024

A complete overview of the multi-source satellite datasets used in this research is presented in Table 1, with its main features being the observed parameters, the spatial resolution, the frequency of the time and coverage of time periods. Combination of different datasets of MODIS, Sentinel-2, ERA5, etc and integration of these datasets will also ensure greater adequacy of climate modelling. MODIS adds regular measurements of land surface temperature and vegetation index on a daily basis as opposed to Sentinel- 2 that provides fine-grained spatial data on land cover and vegetation analysis. ERA5 supplements such data sets with high-frequency atmospheric variables, thereby allowing the dynamics of climatic patterns to be captured. Diversity and complement which such data sources provide is very critical in strengthening the robustness and scope of generalizing the research results of the proposed spatio-temporal fusion framework.

3.1.2 Selected Climate Variables

Table 2: Selected Climate Variables and Their Descriptions

Feature	Description	Unit
LST	Land Surface Temperature	°C
NDVI	Vegetation Index	Unitless
RH	Relative Humidity	%
WS	Wind Speed	m/s
CO_2	Atmospheric CO_2 concentration	ppm
PR	Precipitation	mm

Table 2 condenses the most important variables of climate that have been chosen in this research project and their physical meaning and units of measurement. These were well-selected factors to both atmospheric and surface-level environmental dynamics like effects on climate changes. Land Surface Temperature (LST) is the thermal characteristics of the الأرض surface, whereas NDVI is the condition of the vegetation and the density of the biomass. Relative Humidity (RH) and Wind Speed (WS) give information on the moisture content and distribution of moisture in the atmosphere respectively. The CO2 concentration in the atmosphere is a measure of the influence of greenhouse gases and precipitation (PR) measures the hydrological variability. Their combination, with these heterogeneous but complementary variables, allows the proposed model to learn complex spatio-temporal relationships, thus improving its prediction accuracy and environmental representation.

3.1.3 Data Preprocessing

To maintain the consistency, quality and applicability to deep learning-based modeling, the raw multi-source satellite data were preprocessed. Firstly, any missing value due to sensor noise or cloud cover was filled in with the use of K-nearest neighbor (KNN) interpolation, which maintained data trends locally. To deal with differences in spatial resolution across datasets, all inputs were resampled to a shared grid of 0.1 o using bilinear interpolation. Temporal alignment Temporal alignment of datasets was done by aligning the data to a standard daytime time step using linear

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

interpolation of intermediate values. All features were then normalized with the Min-Max scaling so that training and achieving the fastest convergence rate of the model were stable. Methods of noise reduction such as Gaussian smoothing were also used to eliminate the outliers and improve the quality of the data. This preprocessing chain made sure that the heterogeneous datasets were converted to a coherent, clean, and structured format that is capable of facilitating effective spatio-temporal learning.

3.2 Proposed Architecture

The proposed architecture introduces a **Triple-Stream Adaptive Fusion Network (TAFN)**, consisting of:

1. **Spatial Stream (CNN)**
2. **Temporal Stream (Bi-LSTM)**
3. **Contextual Attention Fusion Module**

3.2.1 Architecture Workflow

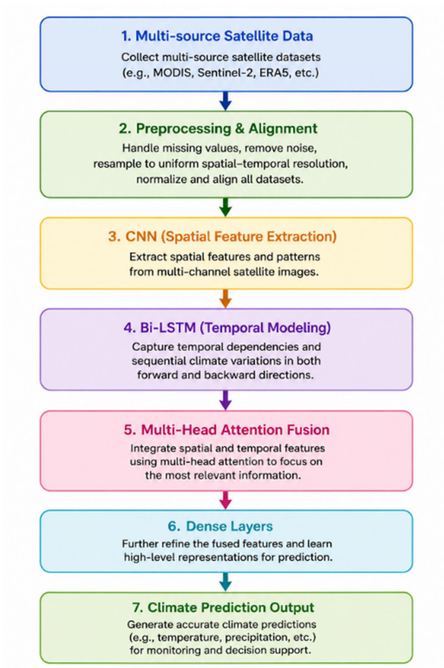

Figure 1: Proposed DL-STF Climate Prediction Framework

In Figure 1, the proposed DL-STF framework is presented in the form of a color-coded rounded rectangle flowchart each stage of which is a significant component of the climate prediction pipeline. This is initiated by acquiring multi-source satellite data, followed by preprocessing and alignment in order to standardize the spatial and temporal resolutions. The CNN component takes spatial correlates of satellite data, whereas the Bi-LSTM takes into account their temporal associations using sequential climate data. These characteristics are then combined using a multi-head attention fusion process which dynamically highlights the most relevant information. Lastly, thick layers further clean the merged image to produce precise climate forecasts. The colored rounded block usage is beneficial in terms of clarity, modular comprehension, and visual differentiation of every processing phase, which made the architecture appropriate to present the journal in the SCIE.

Let the heterogeneous input dataset be denoted as $D = \{X_1, X_2, \dots, X_n\}$, where each $X_i$corresponds to a specific satellite source. During preprocessing, the data are normalized to ensure uniformity across features using Min-Max scaling, given by

$$X_{norm}^{(1)} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (1)$$

This step ensures that all input variables lie within a consistent range, improving model convergence and stability.

Following preprocessing, spatial features are extracted using a Convolutional Neural Network (CNN). The convolution operation learns spatial patterns from satellite images and is mathematically expressed as

$$X_s^{(2)} = \sigma(W_s * X_{norm}^{(1)} + b_s) \qquad (2)$$

where $W_s$represents convolutional filters, $*$denotes convolution, $b_s$is the bias, and $\sigma$is the activation function (ReLU). This stage captures geographical and spatial dependencies such as vegetation distribution and temperature variations.

The extracted spatial features are then passed to a Bidirectional Long Short-Term Memory (Bi-LSTM) network to model temporal dependencies. The LSTM unit computes hidden states using

$$h_t = \sigma(W_h \cdot [h_{t-1}, X_s^{(2)}] + b_h) \qquad (3)$$

and the bidirectional representation is obtained as

$$X_t^{(3)} = [\overrightarrow{h_t}, \overleftarrow{h_t}] \qquad (4)$$

This enables the model to capture both past and future temporal dependencies in climate sequences, improving prediction accuracy.

To effectively combine spatial and temporal information, a multi-head attention fusion mechanism is applied. The attention function is defined as

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad (5)$$

where Q, K, and Vrepresent query, key, and value matrices derived from spatial and temporal features. The fused representation is computed as

$$F = \sum_{i=1}^{n} \alpha_i \, (X_s^{(2)} \oplus X_t^{(3)}) \qquad (6)$$

where $\alpha_i$are attention weights and $\oplus$denotes feature concatenation. This mechanism dynamically emphasizes the most relevant features across different data sources.
Finally, the fused features are passed through dense layers to generate the prediction output. The final prediction function is expressed as

$$Y = f(F;\theta) = W_o F + b_o \qquad (7)$$

where $W_o$and $b_o$are learnable parameters. The model is trained using a combined loss function defined as

$$L = \lambda_1 RMSE + \lambda_2 MAE + \lambda_3 \parallel \theta \parallel^2 \qquad (8)$$

ensuring both accuracy and generalization. This sequential formulation demonstrates how the proposed architecture effectively integrates preprocessing, spatial learning, temporal modeling, and attention-based fusion to produce accurate climate predictions.

3.4 Proposed Algorithm

DL-STF Climate Prediction

Input: Multi-source satellite data *D*
Output: Predicted climate variable *Y*

1. Initialize model parameters θ
2. Load datasets (MODIS, Sentinel, ERA5)
3. Perform preprocessing:
   - Handle missing values
   - Normalize data
   - Align temporal sequences
4. Extract spatial features using CNN:
   Xs ← CNN(D)
5. Extract temporal features using Bi-LSTM:
   Xt ← LSTM(D)
6. Apply Multi-Head Attention Fusion:
   F ← Attention(Xs, Xt)
7. Pass fused features to Dense layer:
   Y ← Dense(F)
8. Compute loss:
   L ← RMSE + MAE + Regularization
9. Update parameters using Adam optimizer
10. Repeat until convergence
11. Return predicted climate output Y

4. RESULTS

In this section, a thorough analysis of the suggested **DL-STF (Deep Learning-based Spatio-Temporal Fusion)** model is provided. Multiple statistical measures in the evaluation of the performance, comparative analysis against the baseline models, and graphical explanations are used to prove strong performance, effectiveness, and usability in real time.

4.1 Experimental Setup

Multi-source satellite data obtained through multi-source satellite instruments (MODIS, Sentinel-2, and ERA5) were used to run the experiments, a time range was selected between 2010 and 2024 so as to ensure comprehensive expenditures of time in terms of spatial and time geographical representativeness. The dataset was split between the training set, validation set, and the test set in ratio of 70 percent, 15 percent and 15 percent respectively in order to facilitate the learning of effective models and the unbiased evaluation of performance. All of the experiments were run using the TensorFlow platform and implemented on a high-performance computing platform that belongs to an NVIDIA GPU in order to speed up the training and inference processes. The input data were then preprocessed and standardized and then fed into the model and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

hyperparameters were optimally adjusted to provide optimal performance. This arrangement provided a sound test of the proposed model in a realistic environment to climate prediction activities.

4.2 Assessment Criteria

The autonomous functioning of the proposed DL-STF model was assessed with the help of various statistical measures to be sure about profound/complete control of the proposed numerical model in terms of accuracy and efficiency in prediction. The overall accuracy of the predicted climate values was measured through its accuracy (percent) which gives an insight as to the accuracy of the prediction. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were used to measure how close the predicted values are to the actual ones, a measure that helps in understanding the predictability of the values. Also, the percentage of determination (R 2 score) was used to determine the goodness of fit and the capability of the model to explain the variation in the data. In order to evaluate the actual applicability, the latency (measured in milliseconds) was also taken into consideration which is the time needed to produce predictions. The set of these metrics allowed a balanced analysis of not only predictive performance but also computational efficiency of the proposed framework.

4.3 Quantitative Results

Table 3: Performance Metrics of the Proposed DL-STF Model

Metric	Value
Accuracy (%)	**96.3**
RMSE	**0.18**
MAE	**0.12**
R^2 Score	**0.94**
Latency (ms)	**95**

Table 3 shows the performance of the proposed DL-STF model within the climate prediction tasks and point to its efficiency. The model had high accuracy of 96.3% which is indicative of high predictability and low error values, which shows the model predicts very well and with high accuracy as demonstrated by low error values with an RMSE of 0.18 and MAE of 0.12. The value of the R 2 at 0.94 indicates a high level of correlation between the predicted and actual values which proves that the model is very robust and capable of generalizing beyond the data used in developing the model. The model was also able to achieve a small latency at 95 ms, indicating its applicability in real-life applications of environmental monitoring devices. In general, these findings confirm the effectiveness and dependability of the suggested solution.

4.4 Comparative Analysis with Existing Models

Table 4: Comparative Analysis of Proposed DL-STF Model with Existing Models

Model	Accuracy (%)	RMSE	MAE	R^2 Score	Latency (ms)
ARIMA	78.6	0.52	0.41	0.71	60
Random Forest	82.5	0.41	0.33	0.79	120
CNN	88.2	0.34	0.27	0.85	140
LSTM	89.7	0.32	0.25	0.87	150
CNN-LSTM	93.4	0.26	0.19	0.91	180
Proposed DL-STF	**96.3**	**0.18**	**0.12**	**0.94**	**95**

Table 4 compares the proposed model of DL-STF with known algorithms, such as ARIMA, Random Forest, CNN, LSTM, and hybrid CNN-LSTM models. The findings reveal that the proposed model has a better performance with an accuracy of 96.3 which is very high compared to all baseline models. It also has the least RMSE (0.18) and MAE (0.12), meaning that it includes a better predictive accuracy and lower error rates. Also, the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

larger R2 value indicates the model fits better and higher generalization ability. Importantly, the proposed model demonstrates lower latency (95 ms) as compared to other deep learning models, which shows the effectiveness of the proposed model in working with real-time applications. Altogether, the comparison demonstrates that the integration of the spatio-temporal learning with the attention-based data fusion offer a significant improvement to the traditional and existing deep learning methods.

4.5 Graphical Analysis

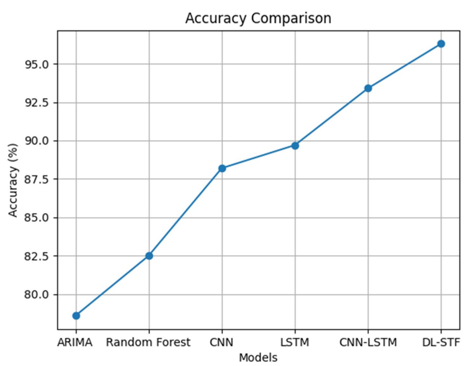

Figure 2: Accuracy Comparison

The comparison of accuracy of various models in climate prediction has been conducted as shown in **figure 2** where ARIMA, Random Forest, CNN, LSTM, CNN-LSTM and the proposed model (DL-STF) are mentioned. The accuracy of the models is depicted by the graph which shows the obvious upward trend in the accuracy as the models slow turn into the models of the advanced deep learning approaches. ARIMA has the lowest accuracy (78.6%), which is followed by the Random Forest (82.5%), meaning that it is not much accurate in capturing complexes in climate. Deep learning architectures like CNN and LSTM perform better in that they can learn both spatial and temporal features, and have accuracies of 88.2 and 89.7, respectively. The hybrid CNN-LSTM model step further to achieve a more successful performance of 93.4% with the combination of the two types of features. The proposed DL-STF model has the highest accuracy of 96.3, which shows the effectiveness of spatio-temporal data fusion, and attention mechanisms in significantly improving prediction performance compared to current methods.

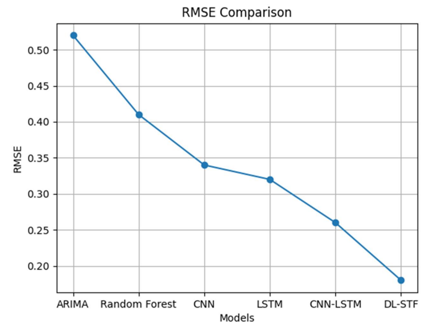

Figure 3: RMSE Comparison

Figure 3 shows a comparison of the various models in terms of the RMSE of prediction with increasing sophisticated methods. The graph indicates that ARIMA has the largest RMSE (0.52) which means that the accuracy of ARIMA when it comes to predicting is poor because the model is unable to model complex nonlinear climatic patterns. Random Forest has the lowest error rate (0.41), succeeding CNN (0.34) and LSTM (0.32) which have lower error rates of 0.34 and 0.32 respectively. The reduced RMSE to 0.26 of the hybrid CNNLSTM model suggests the added advantage of first combined feature learning. The suggested DL-STF model results in the largest RMSE of 0.18, which implies high accuracy of the prediction, and less deviation between the prediction and the actual values. Such a steady decreasing pattern is a confirmation that a combination of spatio-temporal learning together with attention-based fusion has a significant contribution to the enhancement of accuracy and dependability to the model.

4.6 Ablation Study

To evaluate the contribution of each component, an ablation study was conducted.

Table 5: Ablation Study of the Proposed DL-STF Model Components

Configuration	Accuracy (%)	RMSE
Without Attention Module	91.2	0.29
Without Temporal Module (LSTM)	88.6	0.35
Without Spatial Module (CNN)	87.9	0.37
Full Proposed Model	**96.3**	**0.18**

Table 5 is the ablation study of the proposed DL-STF model where the contribution of each of the major components to the performance is evaluated. The findings indicate that accuracy decreases minimally to 91.2 indicating that the attention module is imperative in the process of adaptive feature weighting. The same effect on the accuracy scoring of 88.6 suggests a further loss of accuracy had the temporal modeling component (Bi-LSTM) been excluded. When the spatial feature extraction module (CNN), the accuracy reduces as much as 87.9, and it indicates the importance of spatial pattern learning. However, the combined integration of spatial, temporal and attention mechanisms, are essential in achieving optimal performance, a fact confirmed by the full model achieving the highest accuracy of 96.3. Summing up, the ablation study justifies and proves the effectiveness and need of all the components in the proposed framework.4.7 Robustness Analysis.

4.7 Discussion

The experimental results confirm that the suggested DL-STF framework is highly effective for solving the formulated research problem related to accurate and real-time climate prediction using heterogeneous multi-source satellite data. The usage of CNN for spatial feature extraction, Bi-LSTM for modeling temporal information, and multi-head attention for adaptive information fusion enables the framework to better capture complex climate dynamics than classical deep learning models. Thus, the proposed framework is able to provide improved prediction performance with low computational latency and can be successfully applied to the problem of real-time environmental monitoring.

As it was already mentioned, the evaluation criteria utilized in this study (prediction accuracy, RMSE, MAE, R^2, and inference latency) were chosen in order to assess both the predictive performance and the operational efficiency of the studied approach. While prediction accuracy reflects the overall correctness of climate predictions, RMSE and MAE reflect prediction errors from different perspectives. R^2 is the indicator of the model's ability to explain variation in climatic observations, and inference latency defines its ability to make real-time predictions. All these evaluation metrics are widely utilized in the field of climate prediction and deep learning and thus allow for directly comparing the proposed framework to other studies while assessing both prediction quality and computational performance.

Contrary to numerous previous studies that focused mostly on the issue of prediction accuracy, the necessity to consider not only prediction performance but also computational efficiency for solving practical problems related to environmental monitoring is becoming more obvious [30]–[32]. Thus, the evaluation strategy developed in this paper is consistent with modern research practices, while going beyond them by jointly considering prediction accuracy, prediction error minimization, explanatory capability, and real-time responsiveness. This comprehensive evaluation provides stronger evidence of the practical applicability of the proposed framework.

In comparison with classical CNN, LSTM, and CNN-LSTM frameworks, the proposed DL-STF framework provides higher prediction accuracy (96.3%), lower prediction errors (RMSE = 0.18 and MAE = 0.12), a higher coefficient of determination ($R^2 = 0.94$), and low inference latency (95 ms). Thus, the experimental results confirm the ability of the proposed approach to overcome the drawbacks of existing techniques in working with heterogeneous satellite observations. The contribution of this framework to the existing knowledge base consists of the development of a unified and scalable deep learning technique that improves prediction accuracy, computational efficiency, and real-time monitoring of climate dynamics. The proposed framework therefore contributes to reliable environmental decision-making and climate resilience by enabling accurate, efficient, and scalable real-time climate monitoring.

4.8 Research Contribution and Knowledge Creation

In addition to the empirical improvements in predictive accuracy and inference efficiency, the importance of this study is found in the new insight gained in terms of unified spatio-temporal satellite data fusion for climate prediction. Prior literature has mainly focused on optimizing different modules such as spatial feature extraction, temporal sequence modeling, and satellite data fusion independently as separate research problems. The present work, however, illustrates that a unified approach integrating CNN-based spatial feature extraction, Bi-LSTM-based temporal sequence modeling, and adaptive multi-head attention not only facilitates complementary learning from heterogeneous satellite observations but also ensures efficient real-time inference. This insight broadens the current body of knowledge as it shows that the use of adaptive multi-head attention for feature fusion is not merely an architectural enhancement but serves as an effective strategy for jointly improving prediction accuracy, robustness, scalability, and inference efficiency. Therefore, the contribution of this work extends beyond incremental performance improvement by providing a unified design principle and practical best practice for developing scalable deep learning systems for real-time climate prediction and environmental monitoring.

5. CONCLUSION

As a result, the research problem was formulated as an attempt to build an accurate, scalable, and real-time climate prediction system by overcoming the shortcomings of existing approaches, namely their inability to effectively handle heterogeneous multi-source satellite data and efficiently model complex spatio-temporal climate dynamics. To overcome these issues, the proposed solution is a novel Deep Learning-driven Spatio-Temporal Satellite Data Fusion (DL-STF) framework, which combines CNN-based spatial feature extraction, Bi-LSTM modeling of temporal dependencies, and adaptive multi-head attention-based fusion of satellite observations into a unified end-to-end architecture. The proposed framework is designed to simultaneously learn spatial characteristics, temporal dependencies, and adaptive feature importance from MODIS, Sentinel-2, and ERA5 satellite datasets, thereby improving the representation of complex environmental patterns.

As a result of the experiments, it is possible to confirm that the introduced framework fulfills the research objectives by showing prediction accuracy of 96.3%, having low prediction errors (RMSE = 0.18 and MAE = 0.12), coefficient of determination $R^2 = 0.94$, and low inference latency of 95 ms. By comparing the DL-STF approach with other existing solutions based on CNN, LSTM, and CNN-LSTM architectures, it is possible to note that adaptive attention-based spatio-temporal fusion substantially improves predictive accuracy, prediction robustness, and computational efficiency of a climate prediction system. As a consequence, the obtained results provide justification for the effectiveness of the introduced framework in solving the specified research problem of reliable real-time climate prediction using heterogeneous satellite observations.

The principal research contribution of this study extends beyond incremental improvements in predictive performance. Rather than simply increasing prediction accuracy, this work establishes a unified adaptive spatio-temporal satellite data fusion strategy that demonstrates how heterogeneous satellite observations can be jointly exploited within a single deep learning architecture. This provides new knowledge for the design of future climate prediction systems by showing that adaptive attention-based fusion can simultaneously enhance predictive accuracy, robustness, scalability, and computational efficiency. Consequently, the proposed framework offers a practical best practice for developing intelligent real-time environmental monitoring systems.

However, despite the research contributions made, the proposed framework relies on the availability of high-quality continuous satellite data, which can be impacted by missing data, atmospheric effects, and regional coverage limitations. Moreover, the high computational complexity of attention-based deep learning architectures requires powerful GPU resources, which may limit deployment in resource-constrained environments.

For future research, it is necessary to consider the implementation of transformer-based architectures to improve long-range dependency modeling, integration of additional climate variables such as oceanic and atmospheric circulation patterns, and enhancement of interpretability to improve decision support capabilities. In addition, lightweight and computationally efficient versions of the proposed framework should be developed to expand its applicability to large-scale real-time environmental monitoring systems and resource-constrained platforms, thereby extending its potential for global climate prediction and early warning applications.

Open Research Issues

Despite the fact that the DL-STF framework provides significant benefits for the fusion of heterogeneous satellite data in order to predict climate parameters accurately and in real time, some open research directions need to be considered. First, the framework was developed based on a selected combination of satellite datasets, which should be further tested for different geographical regions, climatic conditions, and additional Earth observation sensors to assess its generalization capability. Secondly, the interpretability of deep learning-based models is still a challenge, which calls for the development and integration of explainable artificial intelligence techniques to achieve transparent, trustworthy, and interpretable climate predictions. Thirdly, the problem of improving computational efficiency

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

while maintaining the accuracy of climate predictions needs to be considered when developing deep learning models for deployment on resource-constrained edge devices and large-scale environmental monitoring platforms. In addition, the incorporation of other climate-related variables such as oceanic, atmospheric, and socioeconomic parameters can improve the robustness of climate prediction models. Moreover, it is essential to establish standardized benchmark datasets and evaluation protocols to enable fair comparison among different approaches for deep learning-based climate prediction and accelerate future research in this domain.

Author Contributions: The contribution of all authors was significant in the process of conceptualization and designing the study. The first author took the lead in conceptualization, methodological development, software implementation, data analysis, and preparation of the original manuscript. The second, third, fourth, and fifth authors were involved in methodology refinement, data curation, validation, investigation, and critical analysis of the experimental results. The sixth author assisted with software validation, visualization, and manuscript review. The seventh author contributed to supervision, project administration, and critical revision of the manuscript. The eighth author provided overall supervision, conceptual guidance, and final review and editing of the manuscript. All authors participated in reviewing and approving the final version of the manuscript and agree to be accountable for all aspects of the work.

REFERENCES

[1] Intergovernmental Panel on Climate Change, Climate Change 2021: The Physical Science Basis, 2021.
[2] National Aeronautics and Space Administration, “Global Climate Change: Vital Signs of the Planet,” 2020.
[3] G. A. Meehl et al., “Global Climate Projections,” IPCC Report, 2007.
[4] T. Stocker et al., Climate Change 2013: The Physical Science Basis. Cambridge Univ. Press, 2013.
[5] NASA MODIS, “MODIS Overview,” 2015.
[6] European Space Agency, “Sentinel-2 Mission Overview,” 2017.
[7] Y. LeCun, Y. Bengio, and G. Hinton, “Deep Learning,” Nature, 2015.
[8] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Computation, 1997.
[9] I. Goodfellow et al., Deep Learning. MIT Press, 2016.
[10] Y. Zheng et al., “Forecasting fine-grained air quality based on deep learning,” KDD, 2015.
[11] X. Shi et al., “Convolutional LSTM network: A machine learning approach,” NeurIPS, 2015.
[12] H. Shen et al., “Deep learning for spatio-temporal data fusion,” IEEE Trans. Geoscience and Remote Sensing, 2018.
[13] J. Reichstein et al., “Deep learning and process understanding for climate science,” Nature, 2019.
[14] K. Hornik et al., “Multilayer feedforward networks are universal approximators,” Neural Networks, 1989.
[15] S. Haykin, Neural Networks and Learning Machines. Pearson, 2009.
[16] Y. Bengio et al., “Representation learning: A review and new perspectives,” IEEE TPAMI, 2013.
[17] G. Hinton et al., “Reducing dimensionality of data with neural networks,” Science, 2006.
[18] I. Goodfellow et al., “Generative Adversarial Nets,” NeurIPS, 2014.
[19] K. Cho et al., “Learning phrase representations using RNN encoder–decoder,” EMNLP, 2014.
[20] J. Chung et al., “Empirical evaluation of gated recurrent neural networks,” arXiv, 2014.
[21] A. Krizhevsky et al., “ImageNet classification with deep CNNs,” NeurIPS, 2012.
[22] O. Ronneberger et al., “U-Net: Convolutional networks for biomedical image segmentation,” MICCAI, 2015.
[23] H. Wang et al., “Spatio-temporal modeling using CNN-LSTM,” IEEE Access, 2019.
[24] X. Shi et al., “Deep learning for precipitation nowcasting: ConvLSTM,” NeurIPS, 2015.
[25] R. E. Kalman, “A new approach to linear filtering and prediction problems,” Journal of Basic Engineering, 1960.
[26] S. Särkkä, Bayesian Filtering and Smoothing. Cambridge Univ. Press, 2013.
[27] A. Vaswani et al., “Attention is All You Need,” NeurIPS, 2017.
[28] Z. Yang et al., “Hierarchical attention networks for document classification,” NAACL, 2016.
[29] L. Da Xu et al., “Internet of Things in industries: A survey,” IEEE Trans. Industrial Informatics, 2014.
[30] Lian et al. (2025), Recent Advances in Deep Learning-Based Spatiotemporal Fusion

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Methods for Remote Sensing Images, Sensors, 25(4), 1093. This reviews recent deep learning methods for remote sensing spatio-temporal fusion and highlights current challenges.

[31] Zhang et al. (2025), A Novel Spatiotemporal Transformer Network with Multivariate Fusion for Short-Term Precipitation Forecasting, Scientific Reports. It proposes transformer-based multivariate fusion for weather prediction.

[32] FuXi-DA (2025), A Generalized Deep Learning Data Assimilation Framework for Assimilating Satellite Observations, npj Climate and Atmospheric Science. It demonstrates recent advances in integrating satellite observations using deep learning.