GRADUATE PROGRAMS INSTITUTE

AI FOR EXTREME WEATHER FORECASTING: CYCLONE PREDICTION USING DEEP LEARNING

Roughaya DIALLO

MASTER’S THESIS

İstanbul, June 2025

AI FOR EXTREME WEATHER FORECASTING: CYCLONE PREDICTION USING DEEP LEARNING

Roughaya DIALLO

MASTER'S THESIS

COMPUTER ENGINEERING DEPARTMENT
COMPUTER ENGINEERING MASTER'S PROGRAM WITH THESIS

MASTER'S THESIS ADVISOR
Asst. Prof. Dr. İnal Begüm TURNA DEMİREL

MASTER'S THESIS JURY MEMBERS
Assist. Prof. Gizem TEMELCAN ERGENECOŞAR
Assist. Prof. Zeynep Behrin GÜVEN AYDIN

BEYKOZ
ÜNİVERSİTESİ

T.C.
BEYKOZ ÜNİVERSİTESİ
LİSANSÜSTÜ PROGRAMLAR ENSTİTÜSÜ MÜDÜRLÜĞÜ

20/06/2025

Yüksek Lisans Tez Onay Belgesi

Enstitümüz Bilgisayar Mühendisliği Anabilim Dalı Bilgisayar Mühendisliği (İngilizce) Tezli Yüksek Lisans Programı 2230140093 numaralı öğrencisi Roughaya DIALLO'in "AI For Extreme Weather Forecasting: Cyclone Prediction Usıng Deep Learning " adlı tez çalışması Enstitümüz Yönetim Kurulunun ../../.... tarih ve 20../.. sayılı kararıyla oluşturulan jüri tarafından oybirliği ile Tezli Yüksek Lisans tezi olarak kabul edilmiştir.

Öğretim Üyesi Adı Soyadı **İmzası**

Tez Savunma Tarihi :20/06/2025

1)Tez Danışmanı: Dr. Öğr. Üy. İnal Begüm TURNA DEMİREL

2) Jüri Üyesi : Dr. Öğr. Üy. Gizem TEMELCAN ERGENECOŞAR

3) Jüri Üyesi : Dr. Öğr. Üy. Zeynep Behrin GÜVEN AYDIN

Not: Öğrencinin Tez savunmasında **Başarılı** olması halinde bu form **imzalanacaktır**. Aksi halde geçersizdir.

PREFACE

This thesis examines extreme weather event forecasting using artificial intelligence (AI)-based models. It aims to improve cyclone prediction using deep learning, particularly Long Short-Term Memory (LSTM) and Transformer models. As data sources, the thesis uses IBTrACS and ERA5 datasets.

I’d like to thank my advisor Dr. Inal Begum Turna, for her guidance and direction. In addition, I express my gratitude to Beykoz University for offering the resources and workspace required. I’d also like to thank my colleagues for their assistance and input. Last but not least, I am grateful to my family members for their support during my master degree learning journey.

TABLE OF **CONTENTS**

ABBREVIATIONS

AI : Artificial Intelligence

CNN : Convolutional Neural Network

ECMWF : European Centre for Medium-Range Weather Forecasts

ERA5 : ECMWF Reanalysis 5

FFN : Feed-Forward Network

FNMOC : Fleet Numerical Meteorology and Oceanography Centre

GOES : Geostationary Operational Environmental Satellite

GPU : Graphics Processing Unit

HR-Extreme : High-Resolution Extreme (Dataset)

IBTrACS : International Best Track Archive for Climate Stewardship

IoT : Internet of Things

LN : Layer Normalization

LR : Learning Rate

LSTM : Long Short-Term Memory

MAE : Mean Absolute Error

MHA : Multi-Head Attention

ML : Machine Learning

MODIS : Moderate Resolution Imaging Spectroradiometer

MSE : Mean Squared Error

NASA : National Aeronautics and Space Administration

NOAA : National Oceanic and Atmospheric Administration

NWP : Numerical Weather Prediction

PE : Positional Encoding

R^2 : Coefficient of Determination

ReLU : Rectified Linear Unit

RL : Reinforcement Learning

RMSE : Root Mean Squared Error

RNN : Recurrent Neural Network

SST : Sea Surface Temperature

SVM : Support Vector Machine

U-Net : U-Net (network architecture)

WMO : World Meteorological Organization

LIST OF TABLES

LIST OF FIGURES

AI FOR EXTREME WEATER FORECASTING: CYCLONE PREDICTION USING DEEP LEARNING

ABSTRACT

This thesis aims to analyse the role of AI in extreme weather forecasting, with a particular focus on Long Short-Term Memory (LSTM)-based and Transformer-based deep learning models for cyclone prediction. It applies sequence-based learning techniques to track storm development using historical data from International Best Track Archive for Climate Stewardship (IBTrACS) and reanalysis data from European Centre for Medium-Range Weather Forecasts Reanalysis 5 (ERA5). The methodology involves data preprocessing, feature scaling, and model parameter optimisation to improve the accuracy of wind speed and cyclone intensity predictions.

Despite AI's potential, challenges persist in this domain, including limited data availability, high computational demands, and the necessity for explainable models. This thesis proposes solutions such as explainable AI, hybrid modelling, and the creation of scalable, efficient models tailored for regions with limited resources. Advancing AI-driven solutions for weather forecasting ultimately supports climate resilience efforts and highlights AI’s broader role in mitigating environmental challenges.

Aşırı Hava Olaylarının Tahmini için Yapay Zeka: Derin Öğrenme ile Siklon Öngörüsü

ÖZET

Bu çalışma, özellikle kasırga tahmini için LSTM ve Transformer tabanlı derin öğrenme modellerine odaklanarak ekstrem hava olaylarının tahmininde YZ'nin rolünü analiz eder. IBTrACS'ten alınan tarihsel veriler ve ERA5'ten elde edilen yeniden analiz verileri kullanılarak kasırga gelişimi sekans tabanlı öğrenme teknikleriyle modellenmiştir. Metodolojide veri önişleme, özellik ölçekleme ve rüzgar hızı ile kasırga şiddeti tahminlerinin doğruluğunu artırmak için model parametrelerinin optimizasyonu yer almaktadır.

YZ'nin potansiyeline rağmen, bu alanda veri eksikliği, yüksek hesaplama gereksinimleri ve açıklanabilir modeller oluşturma ihtiyacı gibi zorluklar devam etmektedir. Bu araştırma, sınırlı kaynaklara sahip bölgelere yönelik ölçeklenebilir ve verimli modellerin geliştirilmesinin yanı sıra açıklanabilir YZ ve hibrit modelleme gibi çözümler önermektedir. Hava tahmini için YZ ile desteklenen çözümlerin ilerletilmesi, iklim direnci çabalarını desteklemekte ve çevresel sorunların azaltılmasında YZ'nin daha geniş rolünü ortaya koymaktadır.

1. INTRODUCTION

Climate change has become one of the most significant global challenges, with widespread implications for the environment, societies, and economies (L. Chen et al., 2023). It causes many problems such as the temperature increase, the wildlife loss, and other dangerous extreme weather events that have huge damage on the environment, people, and economies (Cowls et al., 2023). The impact makes it urgent to find solutions better than the existing ones and AI is a useful tool in this case since it can analyse a large amount of data to identify the patterns and make high accuracy predictions. Particularly, deep learning models have proven their affectivity in extreme weather (Bochenek & Ustrnul, 2022; Abu Rayhan, 2024; Rayhan, 2024).

The thesis is about the use of AI for predicting extreme weather events with a focus on cyclones forecasting. The reason behind cyclones is the serious damage they cause such as the loss of life, the infrastructure destruction, and the economic problems more commonly in the seaside regions. Accurate prediction helps with reducing the damage and improving the preparedness.

While previous studies have applied deep learning to cyclone forecasting, most works focus on single basins and interpretability remains underexplored, with limited analysis of attention mechanisms in prediction models. This thesis is among the few studies that explicitly address this gap by training both LSTM and Transformer models on multi-basin data and analysing their error distribution and attention patterns.

AI is used, specifically deep learning models and high-resolution weather datasets to improve the prediction accuracy. The two primary used datasets are IBTrACS and ERA5. Deep learning models have some challenges such as data gaps as some regions lack enough data, high computing power needed as AI models require strong processing, and the challenge of making AI predictions more understandable.

The main goal of the work is to explore AI-based models' capability to reduce the climate change impact and make a better disaster response.

2. CLIMATE CHANGE

Climate change is long-term changes in weather patterns caused by natural or human activities like burning fossil fuels (coal, oil, gas) and deforestation and leads to temperature increases, extreme weather, and natural disruptions.

One of the important impacts is on the environment. Biodiversity loss and many other natural disasters such as the hurricanes, cyclones, and floods are result of the climate change and AI-based approaches can be solution to track the deforestation, monitor climate changes, and predict the disaster risks which gives useful information for reducing the damage(Bochenek & Ustrnul, 2022;Abu Rayhan, 2024; Rayhan, 2024)

The climate change damage to urban and natural systems has led to more than $500 billion global economic losses annually which affects productivity and resource availability.

Climate change makes global inequalities worse. Poorer communities suffer the most from disasters and resource shortages. This increases the gap between rich and poor countries, making life even harder for vulnerable populations (Adekunle Stephen Toromade et al., 2024; Abu Rayhan, 2024).

Both natural and human causes lead to climate change as the Earth's climate have been shaped by natural factors such as sunlight, volcanic eruptions, and ocean currents for millions of years. However, the increased speed of climate change is mostly caused by human activities.

Main human causes of climate change include greenhouse gas emissions, deforestation, and pollutants. Greenhouse gas emissions are a major factor, primarily resulting from human activities such as burning fossil fuels for energy and industrial purposes. These activities release gases like carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O), which trap heat in the atmosphere and contribute to the planet's rising temperature(Adeoye Taofik Aderamo et al., 2024; Cowls et al., 2023). Deforestation also plays a significant role; large-scale forest clearing for agriculture and urban development reduces the Earth's capacity to absorb carbon and releases the CO_2 stored in trees, intensifying global warming (Rayhan, 2024). Additionally, industrial aerosols and pollutants alter atmospheric processes, with certain greenhouse gases contributing to long-term climate change (Abu Rayhan, 2024).

To solve the climate change issue, more than just technology and science is required. Ethical decisions plays a primary role in guiding policies and actions as the least developed countries are the most severely affected despite they produce the least pollution. This situation raises important questions regarding global responsibility. It is widely recognized that the more developed nations should take the lead in reducing emissions to support the less developed ones adapt to climate change. (Rayhan, 2024).

Technology, including AI, provides new solutions to solve this problem but also creates ethical challenges as it has bias in data and how resources are shared. To ensure fairness, everyone should have the possibility to access AI not just certain groups and of course environmental justice is important as the development should not harm nature so the local communities must be respected when making environmental decisions (Cowls et al., 2023).

Finally, it is very important to have accountability. Clear and open systems are needed so people can trust the process but governments, businesses, and people should cut emissions and be more sustainable, this will help ensure continuous efforts against the climate change challenge.

In summary, climate change affects nature, the economy, and society and solving it requires a combination of science, ethics, and government policies.

Technology, especially AI, can be used to track climate change and reduce its impacts. AI's use should be fair by making sure it benefits everyone; especially vulnerable communities and it is also important that AI-based solutions are not biased and support real climate action. To create real change against this global challenge, everyone must take the responsibility, from governments to businesses and individuals.

Solving the climate change issue requires all countries to work together, with a long-term commitment to sustainability to ensure lasting results. Also, the innovation and ethical responsibility presence can help building a strong and more sustainable future.

3. CLIMATE CHANGE AND AI

AI offers several powerful tools effective for solving many climate change challenges. These tools are able to handle huge amounts of data and find patterns to make real-time predictions which can help to make better climate action decisions (Cowls et al., 2023; Rayhan, n.d.-a). Machine learning (ML) and deep learning (DL) have important improvement on weather forecasting and AI has models as accurate and efficient as traditional numerical weather prediction (NWP) systems. Using AI models in NWP allows for larger and more scalable climate solutions as it makes computing faster and needs less labeled data (Mukkavilli et al., 2023; Rayhan et al., n.d.)

AI can be used in many areas like energy management and agriculture. In renewable energy, it helps with improving the solar and wind power forecasting for a better energy production and with the smart grids management to reduce the energy waste and the low emissions (Bochenek & Ustrnul, 2022; L. Chen et al., 2023). In agriculture it can be used for pest the outbreaks prediction to protect the crops, the water use optimization to prevent the waste, and the crop production improvement to make the food systems more resilient to the climate change (Climate Change and the Role of AI, n.d.; Rayhan, 2024). In environmental protection AI can be helpful with the biodiversity's loss and deforestation monitoring and with the ecosystem's health tracking by providing useful data for conservation (Adekunle Stephen Toromade et al., 2024; Rayhan, 2024).

AI can do more than just reducing climate change effects as it can additionally help with climate adaption. Predictive AI-based models are able to give early warnings for extreme weather like floods, hurricanes, and cyclones using past and real-time date which can help governments and organizations to prepare and respond better(Climate Change and the Role of AI, n.d.; Rayhan et al., n.d.). Combining AI, IoT devices, and satellite images can help to monitor and manage natural resources and to support sustainable land and water use especially in risky areas (L. Chen et al., 2023; Abu Rayhan, 2024). AI-powered drones and sensors can be used after a disaster to provide affected areas real-time assessments and to improve resource distribution and recovery efforts (Adekunle Stephen Toromade et al., 2024; Bochenek & Ustrnul, 2022).

AI has another role beyond technology, it assists in social sciences and policymaking. It supports climate resilience and sustainable development strategies (Brotzge et al., 2023; Rayhan et al., n.d.). For example, about 70% of the global gas industry uses AI to improve weather forecasting and that demonstrates its flexibility and scalability (L. Chen et al., 2023). Furthermore, AI improves safety and efficiency in different industries by predicting risks and hazards, reducing work disruptions, helping decision-makers with accurate data (Adeoye Taofik Aderamo et al., 2024; Rayhan et al., n.d.). AI's ability to connect monitoring, prediction, and management makes it an important tool for climate action and its impact can facilitate creating a more sustainable and resilient future for the next generations (Adekunle Stephen Toromade et al., 2024; Snejana, 2022).

3.1 APPLICATIONS

As mentioned above, AI has many applications for solving or reducing climate change. Some AI applications include supporting the urban planning by designing cities adaptable to the climate change and improving the sustainability in growing urban areas, recycling the systems optimization and helping the environmental goals by reducing the landfills to improve the waste's management, and detecting the forest fires early, track the extreme weather events like heavy rain and strong winds, and provide early warning to reduce the risks (Rayhan, n.d.-a, 2024). AI enhances extreme weather forecasting by combining ML and NWP models. This hybrid approach helps to improve the accuracy in predicting events like hurricanes and heatwaves and enhancing the disaster preparedness for a better response and risk management. (Mukkavilli et al., 2023; Rayhan et al., n.d.).

The main role of AI in addressing climate change is analysing large amounts of data, finding patterns in the climate trends, and providing useful insights for the climate change action.

In the energy sector, AI improves energy efficiency by predicting the solar and wind power's availability, integrating renewable energy into smart grids and reducing emissions which makes the energy's use more efficient. (Bochenek & Ustrnul, 2022; L. Chen et al., 2023).

In agriculture, AI improves precision farming by predicting the weather's conditions to help the farmers planning ahead, detecting the post outbreaks to protect the crops, and estimating the crop's yield to manage the food's production. This can help ensure food security even if climate change is affecting agriculture. (Climate Change and the Role of AI, n.d.; Rayhan, 2024).

AI enhances climate resilience by modelling urban planning strategies and aiding in reforestation efforts.

In disaster management, AI plays a crucial role in various stages. In mitigation, it detects the risked areas of disasters and creates maps to help with planning to safe the infrastructure (Jiang et al., 2022).During the preparedness, AI offers early predictions and helps in planning evacuations to reduce the harm (Hess & Boers, 2022). In response and recovery, AI helps with resource distribution after the disaster and monitors recovery efforts to track the progress (Hess & Boers, 2022).

Remote sensing uses satellite images and AI for environmental monitoring. AI analyses satellite images to detect deforestation, habit loss, and land degradation. For example, convolutional neural networks (CNNs) find patterns in satellites images and help with tracking the vegetation changes to support conservation efforts(Hess & Boers, 2022; Jiang et al., 2022).

AI also monitors the atmosphere by tracking aerosols and dust movement which affects the climate (Hess & Boers, 2022).

3.2 CHALLENGES

AI use for climate change faces many challenges within the technology itself, the environment (energy use, emissions), and society (data access, fairness). From the data quality and integration issues side, AI needs large, high-quality datasets to make accurate predictions. However, the data is missing or unreliable in some regions which makes AI less effective (Cowls et al., 2023; Mukkavilli et al., 2023). AI models require a lot of energy and resources to train and this leads to high greenhouse gas emissions and energy use (L. Chen et al., 2023; Rayhan et al., n.d.). Additionally, they need calibration because the made predictions are not always accurate and need adjustments. So, fine-tuning AI models takes both time and resources (L. Chen et al., 2023; Abu Rayhan, 2024).

Ethical and social issues are the major concerns in AI use. AI models can be unfair if they are trained on incomplete or one-sided data and this can negatively impact vulnerable groups (Adekunle Stephen Toromade et al., 2024; Cowls et al., 2023). Another problem is that not everyone has equal access to these technologies which means strong policies are needed to ensure that AI is used fairly (Adeoye Taofik Aderamo et al., 2024; Rayhan et al., n.d.). It is vital to mention that AI should not completely replace human judgment in disaster response and resources management. Also, as AI is used in important systems, the security is very critical because if these systems are hacked or misused, the consequences could be serious (Adeoye Taofik Aderamo et al., 2024; Bochenek & Ustrnul, 2022).

AI models for climate prediction are still developing and they are not fully advanced yet. General AI models for global and regional climate systems are in the early stages. Techniques like self-supervised learning could help but more research is needed to make AI scalable and reliable (Adekunle Stephen Toromade et al., 2024; Mukkavilli et al., 2023). Predicting extreme weather events is still a challenge because short-term events are hard to forecast accurately even with improved AI models. Blurry or incomplete real-time images make analysis harder because AI struggles in certain situations where the data is unclear or missing (Abdusalomov et al., 2023; Chkeir et al., 2023).

AI's implementation is expensive and that makes its adoption difficult for some sectors and regions. AI costs include: initial setup (buying AI infrastructure and tools), computational need for processing large amounts of data, and ongoing energy use to keep the systems running. Also, the deployment would require integration with real-time data pipelines for operational meteorology. Challenges such data latency, computational cost, and ensuring interpretability for forecasters need to be analysed. Addressing these aspects is essential for transitioning from research models to practical forecasting tools (Climate Change and the Role of AI, n.d.; Cowls et al., 2023).

Finally, integrating AI into current systems takes time because it requires changes in the culture and organization. The challenges slowing AI's adoption include following regulations to ensure compliance, involving stakeholders, and to build the skills and knowledge to help people with using AI effectively (Adeoye Taofik Aderamo et al., 2024; Rayhan, 2024).

It is explained that AI is a powerful tool when it comes to facing climate change challenges as it assists with renewable energy and disaster preparedness and the possibility to use it in many other sectors. However, there are some challenges in it (Adekunle Stephen Toromade et al., 2024; Rayhan, 2024). To improve AI's effectiveness, it is necessary to have better access to data for all regions, reduce the technological inequality, and the integration of AI with climate policies. Qualified policies can help by ensuring that AI is fair and available to everyone, having an ethical use of it, and keeping its systems sustainable for long-term impact (Rayhan et al., n.d.).

3.3 SCOPE & LIMITATIONS

Data gaps exist in some regions where there is not enough good-quality environmental data and AI models need real-time data to make accurate predictions. This problem is worse in the low-income areas where less technology and infrastructure are, and in the remote ones because it is hard to collect the data in these regions, (Adekunle Stephen Toromade et al., 2024; Rayhan et al., n.d.)

Policy Integration is another issue, as there is insufficient alignment between AI advancements and climate policies, leading to missed opportunities for implementation (Adekunle Stephen Toromade et al., 2024; Rayhan, 2024).

Technological inequality also poses a barrier, since not everyone has an equal access to AI technology. This issue is more common in the developing countries and a limited AI access slows down the global efforts to fight climate change (Climate Change and the Role of AI, n.d.; Rayhan, 2024).

Finally, explainability remains a concern as AI models, especially deep learning, are not always transparent. The lack of the transparency makes understanding how AI makes the decisions harder and this creates challenges in using AI for policy and decision-making (Adekunle Stephen Toromade et al., 2024; Bochenek & Ustrnul, 2022).

AI is a powerful tool for addressing climate change. It is used in renewable energy, disaster management, agriculture, and environmental monitoring. Despite its benefits, AI faces challenges such as data gaps, ethical concerns, and technological inequality.

Table 3.1: Summary of AI Applications in Addressing Climate Change.

Domain	**AI Applications**	**Examples**	**Impact**
Renewable Energy	Forecasting solar and wind power availability, optimizing smart grids.	Solar radiation modelling (L. Chen et al., 2023), Smart grids (Bochenek & Ustrnul, 2022).	Improving energy efficiency, reducing emissions.
Agriculture	Precision farming: forecasting weather, pest outbreaks, and crop yields.	Pest outbreak prediction (Rayhan, 2024).	Enhancing food security, minimizing environmental impact.
Disaster Management	Early warning systems, resource allocation, real-time monitoring.	Flood predictions (Abu Rayhan, 2024, fire detection (Chkeir et al., 2023).	Reducing disaster risk, improving recovery efforts.
Environmental Monitoring	Tracking Deforestation, biodiversity loss, and atmospheric phenomena.	Satellite imagery analyses (Jiang et al. 2022).	Proactive conservation and restoration strategies.
Urban Planning	Designing resilient cities, managing urban heat islands, and optimizing resources.	Smart waste management (L. Chen et al. 2023).	Sustainable urban development.

To overcome these challenges, using policies is very essential as they can help with the improvement of the data availability, ensuring a fair access to AI technologies, and addressing the sustainability concerns. Good policies help AI to fit better into climate change strategies and encourage innovation and resilience.

This chapter focuses on how technology and policy work together, and prepares for a discussion on AI's role in some specific climate issues like extreme weather events.

4 EXTREME WEATHER EVENTS FORECASTING

Extreme weather events such as tropical cyclones, heatwaves, blizzards, floods, and droughts are unusual and severe compared with the normal conditions. AI can be used to improve disaster preparedness and risk reduction as it can analyse weather patterns to predict these events' time and location before they happen.

Weather is short-term natural events (e.g., temperature, humidity, wind speed) occurring in specific locations and times.

Prediction, forecasting, and nowcasting are different methods used to estimate future events, each with varying levels of precision. Prediction is a general estimate of future events based on patterns and trends; it is not exact as it does not specify the time or the probability. Forecasting, on the other hand, is a detailed projection of future conditions and it uses math, statistics, and AI models. It is more precise than prediction as it includes probabilities and timeframes. Nowcasting refers to very short-term weather prediction using real-time data and advanced models.

Mitigation and adaptation are two key strategies in addressing climate change. Mitigation involves actions aimed at reducing greenhouse gas emissions and enhancing carbon capture to stabilize climate systems. Examples include renewable energy optimization and carbon capture technologies (Adekunle Stephen Toromade et al., 2024). Adaptation is preparing for the unavoidable effects of climate change, like predictive modelling to forecast the extreme weather or the adaptive farming methods to deal with the changing climate conditions (Adekunle Stephen Toromade et al., 2024).

Following are some of the most widely used AI techniques that have been applied to improve the extreme weather forecasting.

4.1 MACHINE LEARNING APPROACHES

Common traditional machine learning methods such as Random Forests Support Vector Machines (SVMs), and Gradient Boosting Machines are widely adopted to identify the patterns in the historical meteorological data for the weather forecasting enhancement.

The general process behind these models is learning the relationships between the different atmospheric variables such as the temperature, pressure, and wind speed for a better understanding of the weather systems progression over time. ML models can effectively handle the large datasets which are used in weather forecasting. They are able to detect the non-linear relationships and subtle trends that physics-based or statistical methods might miss. The utilized algorithm can learn how a specific weather variables combination leads to a certain extreme weather event and use the information to make future predictions.

ML models are able to make accurate predictions even if the used data contains noises or missing values which is another strength of using these models for extreme weather forecasting. In addition, they are often used alongside traditional Numerical Weather Prediction systems to correct the error or improve the spatial and temporal resolution which refines the NWP systems outputs as they use physics-based methods.

4.2 DEEP LEARNING

Deep learning is now a key part of extreme weather forecasting. It can enhance weather forecasting by improving both special and temporal pattern recognition. AI techniques use CNNs and RNNs, these models identify the complex patterns in the atmospheric data. AI improves the predictions for cyclones, heatwaves, and heavy rainfall (Salcedo-Sanz et al., 2024; Verma et al., 2023).

While deep learning (DL) is considered a subfield of ML, it uses more complex models with a higher capability of handling much larger volumes of data which brings further improvement to extreme weather forecasting (Verma et al., 2023; Zhong et al., 2023).

Among DL techniques, Convolutional Neural Networks (CNNs) have the ability of detecting important features like cloud formation and storm structures which makes them effective at processing spatial data such as satellite images and radar scans. LeNet-5 and ResNet-50 are examples CNNs models. They participate in extreme weather forecasting events forecasting by predicting both their timing and geographic location. These models are trained using a combination of synthetic data which is generated with simulations to represent various weather scenarios and empirical data from real-world observations. By implementing models such as LeNet-5 and ResNet-50, CNNs obtain the ability to capture detailed spatial features with high-resolution while maintaining a strong predictive accuracy (Jiang et al., 2022).

A further DL techniques type is Recurrent Neural Networks (RNNs). These models analyse the temporal sequences making them effective for capturing the weather events progression over time.

4.3 HYBRID AI-NUMERICAL WEATHER PREDICTION (NWP) SYSTEMS

Hybrid AI-NWP systems are a combination of AI and NWP models which offers a more accurate approach for extreme weather forecasting. In these systems, AI is implemented for correcting the systematic outputs provided by the NWP models, optimizing the model parameters, and improving the data assimilation. For example, the ML algorithms learn from the historical variations between the forecasts and the actual outcomes which enhances future predictions.

Using these hybrid models provides more precise and accurate forecasts which gives more time for the disaster preparedness and improves the extreme weather event detection. These systems combine both physics-based and data-driven methods strengths.

4.4 ENSEMBLE LEARNING METHODS

Ensemble learning methods rely on the principle of combining multiple model outputs instead of a single one. In weather forecasting ensemble learning methods use a variation of the initial conditions or the model configurations to generate more than a scenario which provides the ability to capture the uncertainty in the atmospheric behaviour.

AI improves ensemble learning methods by prioritizing the most effective models dynamically based on the past performance and the current atmospheric conditions. It also contributes to adjusting the outputs to improve the forecast reliability. More robust and probabilistic forecasts are obtained which is valuable for predicting more rare and uncertain events such as extreme weather.

4.5 REINFORCEMENT LEARNING

Reinforcement Learning (RL) is a ML model type that learns to make decisions by interacting with an environment and receiving feedback in a rewards or penalties form. In the weather forecasting context, RL can be implemented to tune the model parameters automatically, adapt the forecasts strategies in real time, and select the best performing model. This learning process results in more adaptive and self-improving forecasting systems. RL performs perfectly when it comes to rapidly changing or unpredictable weather conditions where a flexible and ongoing learning are vital.

4.6 DATA FOR EXTREME WEATHER FORECASTING

Producing accurate extreme weather events predictions depends on the access to high-quality data. Integrating various data types enables models to capture the complexity in the weather events and enhances its quality and reliability. Such data should be collected from diverse sources including satellites, radar, sensors, and ocean monitoring systems. The absence of complete and reliable data might lead to errors or delays in weather warnings.

The mainly used data sources are satellite observations as they provide real-time images of the clouds, the rain, and the sea temperatures (NOAA, NASA), ground-based measurements which include the weather stations, the radars, and the buoys for a detailed local data, NWP models as they use mathematical simulations for forecasting, and reanalysis datasets which combine historical and real-time data to build detailed atmospheric records. The data helps to track cyclones and other extreme weather events and improve the forecast accuracy by using multiple reliable sources.

Numerical models create detailed weather forecasts by simulating the atmospheric behaviour. They use initial conditions and physical equations to make the predictions. AI has improved the forecasting by analysing large datasets and finding complex patterns that traditional models might miss.

However, some challenges could appear with the data such as the data lack in some regions, the differences between the datasets which can cause inconsistencies, and the high computational requirements for the large climate change data processing.

Table 4.1: Types of Data Used in Extreme Weather Forecasting.

Data Type	**Description**	**Example Sources**
Satellite Data	Real-time imagery of cloud cover, temperature, and precipitation.	NOAA GOES (Administration N. O., n.d) , NASA MODIS (NASA, n.d), Himawari-8 (Agency, n.d)
Radar Observations	High-resolution tracking of storm movement and intensity.	NEXRAD (Service, NEXRAD radar, n.d), Doppler Radar (Service, Doppler radar, n.d), EUMETSAT (Satellites, n.d)
Weather Station Data	Ground-based temperature, humidity, wind speed, and pressure readings.	NOAA GHCN (Menne, 2012), METAR (Administration F. A., n.d), WMO Synoptic (Organization, n.d)

Buoy and Ocean Data	Sea surface temperatures, wave heights, and ocean currents.	Argo Floats (Argo, n.d), TOPEX/Poseidon (Laboratory, n.d)
Reanalysis Data	Long-term atmospheric reconstructions integrating observational data.	ERA5 (Hersbach, 2020), MERRA-2 (Gelaro, 2017), NCEP-NCAR (Kalnay, 1996)
Numerical Model Outputs	Simulations of atmospheric conditions based on physics-based equations.	GFS (Prediction, n.d), ECMWF (Forecasts, n.d), WRF (Research, n.d)
Machine Learning Datasets	Processed climate data for AI-based forecasting models.	IBTrACS + ERA5 (Knapp, 2017), ClimateNet (Lochner, 2021), FNMOC (Center, n.d)

4.7 APPLICATIONS

AI can be used for different extreme weather events prediction. For example, heatwaves can be predicted as machine learning models analyse surface temperature and atmospheric conditions to predict the onset and duration of heatwaves. These forecasts support early warning systems that can save lives and resources (Ran et al., 2024; Salcedo-Sanz et al., 2024). For cyclones, hybrid models that blend deep learning and statistical techniques provide detailed predictions of cyclone intensity and rainfall distribution. This information is crucial for disaster preparedness in cyclone-prone regions (Ran et al., 2024; Zhong et al., 2023). In the case of thunderstorms, neural networks predict thunderstorm frequency and lightning intensity, enhancing preparedness for short-term events (Mashao et al., 2023; Salcedo-Sanz et al., 2024). For heavy rainfall, U-Net models integrated with satellite and NWP data improve predictions for rare, extreme rainfall events, increasing accuracy for high-magnitude scenarios. This has direct implications for flood prevention and resource management (Hess & Boers, 2022). Nowcasting leverages radar imagery and neural networks, and AI-driven nowcasting predicts short-term precipitation patterns with remarkable precision, facilitating immediate decision-making during severe weather events (Hess & Boers, 2022; Jiang et al., 2022). Finally, timing and location predictions are made using ResNet-50 models that use CNNs to determine the time and location extreme weather events will happen which offers important information for disaster prevention (Jiang et al., 2022).

Alongside these powerful applications of AI in predicting extreme weather events, there are also several challenges that need to be addressed.

One major issue is data limitations, as AI models require good quality and high-resolution weather data, but some regions don’t have enough detailed data especially for extreme or rare weather events (Hess & Boers, 2022; Jiang et al., 2022; Verma et al., 2023; Zhong et al., 2023). Another challenge is model complexity; it is difficult to create AI models that are accurate, easy to understand, and work fast, especially for real-time weather forecasting (Jiang et al., 2022; Mashao et al., 2023). Computational costs also present a problem, since working with fine-grained and high-resolution weather data requires significantly high amounts of computing power, particularly for models based on CNN and U-Net (Hess & Boers, 2022 ; Ran et al., 2024 ; Salcedo-Sanz et al., 2024). There is also a trade-off between accuracy and resolution, as the model may lose accuracy making it harder to use if it focuses on long-term predictions or very detailed maps (Jiang et al., 2022). Finally, sparse data representation poses a limitation, because some weather data is incomplete or missing, and it is still hard to make models work well with limited data even with new AI technology (Hess & Boers, 2022; Jiang et al., 2022).

5 CYCLONES FORECASTING

Depending on the affected areas, cyclones are referred to as hurricanes or typhoons. They form over warm tropical and subtropical oceans and are low-pressure weather systems They have a warm core (central hot air), an organized deep convection throughout (strong thunderstorms), and circular wind pattern, it means that winds circulate around a centre (Nath et al., 2023; Rahman et al., 2024; Takaya et al., 2023).

Table 5.1: Cyclone Characteristics.

Cyclone Type	**Region**	**Core Temperature**	**Formation**	**Structure**	**Size**
Tropical Cyclone	Tropical and subtropical oceans	Warm	Develops over warm sea surfaces with organized deep convection and a closed surface wind circulation	Features include the eye, eyewall, and spiral rainbands	Diameter typically ranges from 200 to 500 km, but can be as large as 1,000 km

Extratropical Cyclone	Mid-latitude regions	Cold	Forms along weather fronts where warm and cold air masses meet	Contains fronts and a comma-shaped cloud pattern	Size varies widely; can be larger than tropical cyclones
Subtropical Cyclone	Regions between tropical and extratropical zones	Mixed	Exhibits characteristics of both tropical and extratropical cyclones	Lacks a well-defined core; may have a combination of frontal and convective features	Generally smaller than tropical cyclones

This summary of different cyclone characteristics is based on the region of occurrence, the core temperature, the formation process, the structure, and the size.

Tropical cyclones develop over warm sea surfaces. They are characterized by a warm core and emerge from organized, closed surface wind circulation. The features include the eye, eyewall, and spiral rainbands. A tropical cyclone usually has a size between 200 and 500 kilometres, but sometimes exceeds 1000 kilometres.

Extratropical cyclones occur in temperate regions with a cold core. They develop in fronts where warm and cold air masses collide. Extratropical cyclones include frontal systems and a distinctive comma-shaped cloud pattern, with widely varied sizes that may exceed tropical cyclone dimensions.

Subtropical cyclones are generally in the transitional region between tropical and extratropical zones. They have a mixed warm and cold core feature and often merge frontal and convective features. Generally, subtropical cyclones lack a distinct core and are smaller in scale than tropical cyclones.

Tropical cyclones have distinct structural characteristics that define their intensity and behaviour. At the centre lies the eye, a calm, clear area typically about 30-65 km in diameter. Surrounding the eye is the eyewall, a ring of towering thunderstorms where the most severe weather and highest winds occur. Extending outward from the eyewall are the rainbands, which are spiral bands of clouds and precipitation that contribute to the cyclone's overall structure and impact.

5.1 CYCLONES

Cyclones cause extensive damage to infrastructure, ecosystems, and human lives, particularly in coastal regions. Key hazards associated with cyclones include wind damage, as strong winds especially near the eyewall destroy buildings, uproot trees, and damage power lines (Chand et al., 2022; Rahman et al., 2024). Storm surges result in elevated sea levels that lead to flooding, particularly in regions with shallow coastal waters (Chand et al., 2022). Heavy rainfall and flooding are also significant threats, with rainbands contributing to widespread inland flooding that exacerbates the overall impacts (Nath et al., 2023). In some cases, cyclones spawn tornadoes, which cause localized but highly destructive effects. These tornadoes are often difficult to predict, adding to the hazards during a cyclone event (J. H. Chen et al., 2023). Furthermore, erosion and land degradation occur as coastal erosion accelerates and disrupts the natural sedimentation process, leading to the loss of soil fertility and negatively affecting agriculture and local economies (Demaria et al., n.d.). Finally, there are severe economic and public health impacts, as damage to infrastructure disrupts supply chains and essential services such as healthcare, electricity, and clean water. In the aftermath, affected regions often face waterborne diseases and long-term economic downturns (Rahman et al., 2024).

Cyclones are among the most destructive natural disasters, causing significant loss of life, economic damages, and environmental devastation. They are becoming stronger and more frequent because of climate change and that made better forecasting tools and disaster preparedness plans availability more important. For example, recent cyclones like Freddy, Mocha, and Biparjoy have indicated the complexity in predicting and managing these storms as they caused damage to affected areas.

In this section, these three cyclones are taken as case studies with a focus on their paths, wind speeds, and the effects on different regions.

The information details come from trusted meteorological organizations like NASA and NOAA, ensuring accuracy.

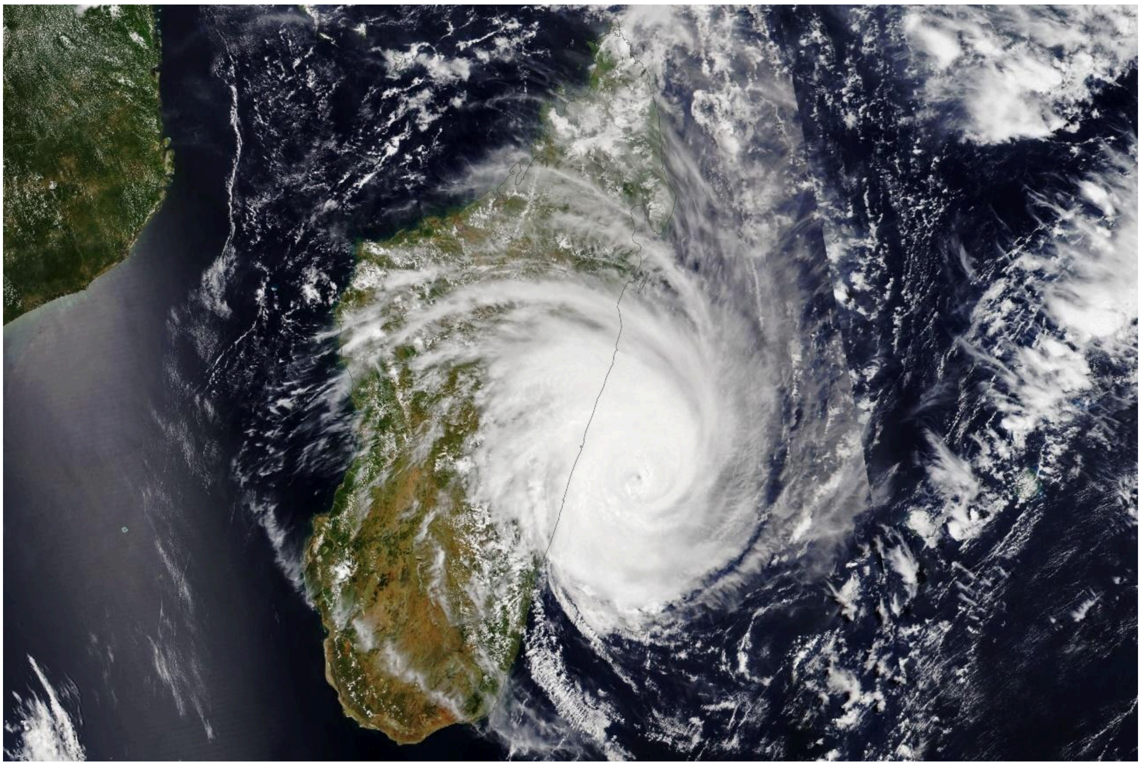

Figure 5.1: Satellite Imagery: A satellite image taken by NASA's Earth Observatory on February 21, 2023, demonstrating Cyclone Freddy getting close to Madagascar *(Observatory, Cyclone Freddy Hits Madagascar, 2023)*.

Cyclone Freddy, Freddy is one of the longest and most intense cyclones. It hit Madagascar and Mozambique in 2023. It caused over $2 billion in the damages and displaced thousands with the wind speeds reaching 260 km/h. Its extreme intensity highlighted the difficulties in predicting the prolonged cyclonic activity.

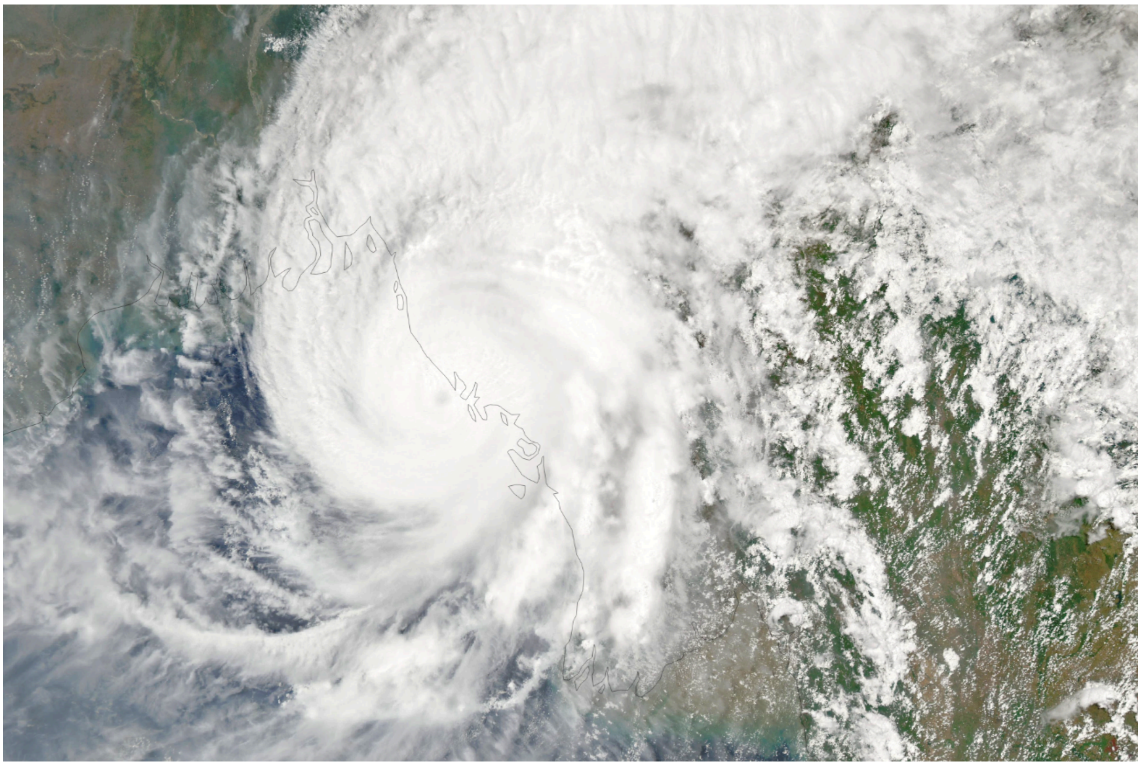

Figure 5.2: Satellite View: An image from NASA's Earth Observatory on May 14, 2023, illustrating Cyclone Mocha as it approached Myanmar with winds reaching up to 175 mph *(Observatory, Cyclone Mocha Strikes Myanmar, 2023)*.

Cyclone Mocha, a Category 4 storm, made landfall in Myanmar and Bangladesh in May 2023. With maximum winds of 195 km/h, it devastated coastal regions, leading to widespread displacement and over $1.2 billion in damages. The storm emphasized the vulnerability of densely populated, low-lying areas.

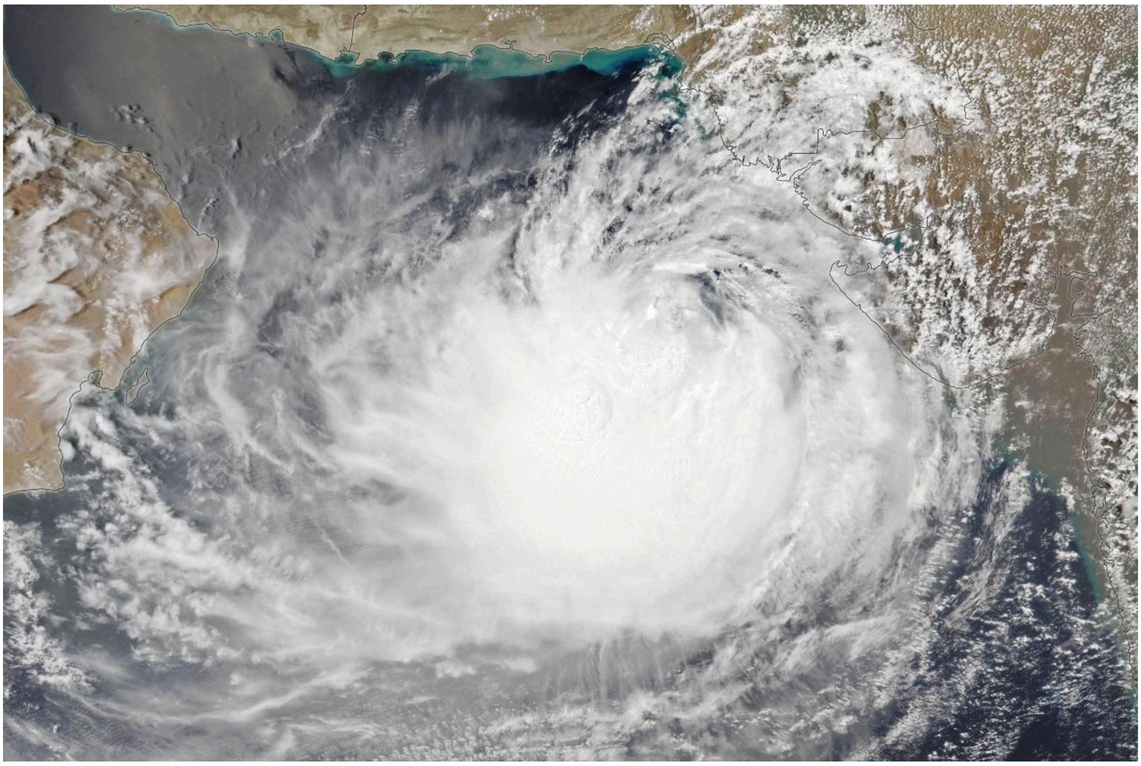

Figure 5.3: Satellite Imagery: A natural-colour image from NASA's Earth Observatory on June 14, 2023, demonstrating Cyclone Biparjoy nearing the coastlines of western India and southern Pakistan *(Observatory, Cyclone Biparjoy Churns Toward India and Pakistan, 2023)*

Cyclone Biparjoy, with wind speeds of 167 km/h, impacted Gujarat, India, in June 2023. The storm caused significant economic losses, damaged infrastructure, and highlighted the importance of accurate trajectory prediction to minimize its effects on vulnerable populations.

5.2 CYCLONES FORECASTING TECHNIQUES

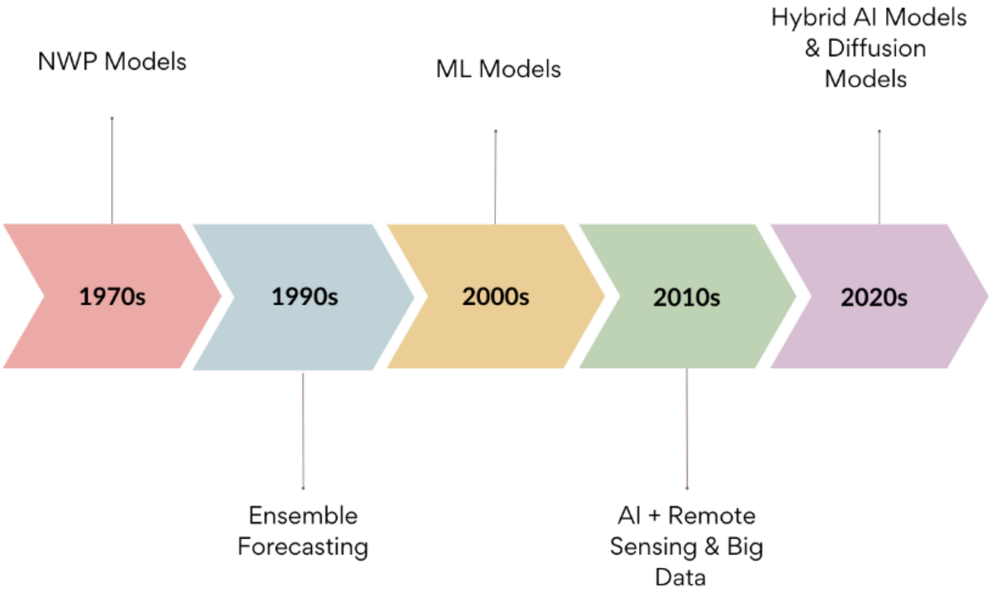

Figure 5.4: Historical timeline of cyclone forecasting advancements, illustrating the transition from Numerical Weather Prediction (NWP) models to modern AI-driven hybrid forecasting systems.

Cyclone forecasting requires analysing multiple aspects, such as formation, path, intensity, and associated risks like storm surges.

This timeline illustrates the cyclone forecasting progression from NWP models to modern AI-driven system. Each collared section corresponds to a decade and indicates the dominant forecasting methodology of that period. The figure highlights milestones such as ensemble learning forecasting in 1990s, ML in the 2000s, and the integration of AI with remote sensing in the 2010s.

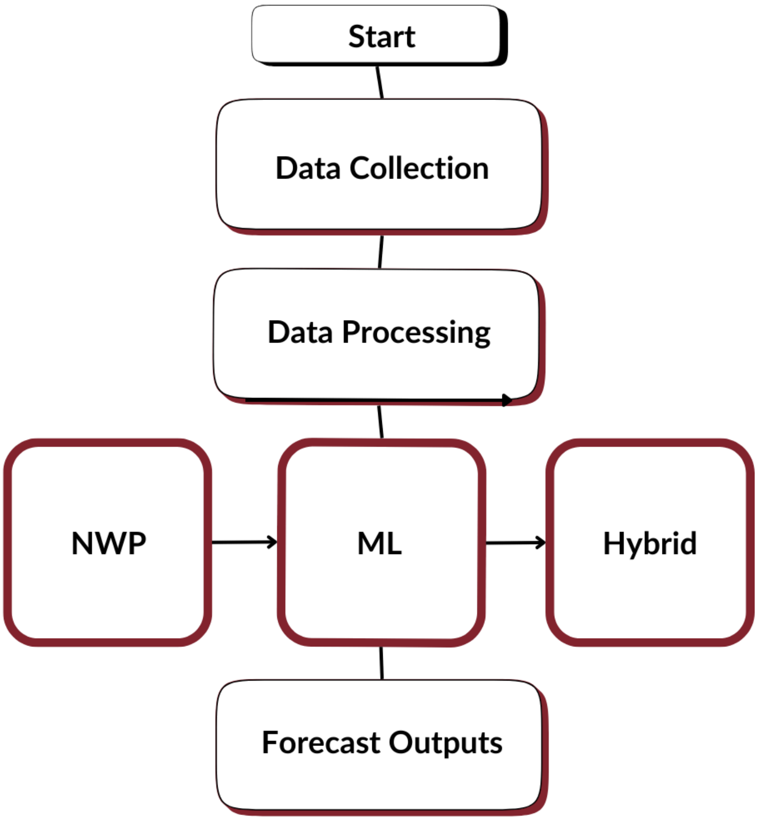

Figure 5.5: Workflow presenting cyclone forecasting techniques, including NWP, ML, and Hybrid Models.

The workflow outlines the general stages of cyclone forecasting using NWP, ML, or hybrid approaches.

After the data collection, the data is cleaned, formatted, and prepared for being analysed using a forecasting system. The system applies either a physics-based NWP model, data-driven ML model, or a hybrid method combining both approaches. The output is a complete cyclone forecast ready for decision-making.

5.3 METHODOLOGY

The methodology section explains the used step-by-step process which include preparing the data, choosing a model, training it, and testing its performance. Each step helps ensure the AI-based cyclone forecasting system is accurate and reliable.

5.3.1 OVERVIEW OF THE USED APPROACH

This thesis searches to create a deep learning model for extreme weather events predicting by using historical tropical cyclone data. The main dataset is a reconstructed global tropical cyclone record (1959-2022), built from IBTrACS and ERA5 reanalysis data. This dataset, described in detail in (Xu et al., 2024), addresses significant data gaps through machine learning techniques, resulting in several key improvements. It offers enhanced coverage, providing 3-4 times more data points compared to IBTrACS alone. There is also improved accuracy, with reduced biases in critical cyclone characteristics such as wind speed (Vmax), radius of maximum winds (Rmax), and minimum central pressure (Pmin). Additionally, the dataset has a global scope, covering all tropical cyclone areas with data updated every 3 hours.

The datasets have six regional cyclone datasets: Eastern Pacific, North Atlantic, North Indian, South Indian, South Pacific, and Western Pacific. These datasets offer detailed information about cyclone trajectories, intensities, and other important attributes. It is constructed by merging these multiple regional datasets, providing a comprehensive record of cyclone characteristics such as wind speed (Vmax), pressure, and geographical coordinates. The goal is to predict the cyclone wind speed using a Long Short-Term Memory (LSTM) neural network and a Transformer model.

The dataset utilizes several features to characterize tropical cyclones. These include Lat and Lon, which represent the latitude and longitude of the tropical cyclone as recorded in IBTrACS. Land indicates whether a tropical cyclone makes landfall. Vmax refers to the maximum sustained wind speed, while Pressure represents the minimum central pressure. Rmax is the radius of maximum wind. Additionally, the dataset includes R34, R50, and R64, which are the radial distances from the cyclone centre to locations where sustained wind speeds reach 34, 50, and 64 knots, respectively.

Table 5.2: Basic information on the number of recorded tropical cyclone characteristics from 1959 to 2022 recorded in IBTrACS (Xu et al., 2024).

Basin	**Time point**	V_{max}	P_{min}	R_{max}	R_{34}	R_{50}	R_{64}
Western Pacific	152 362	26 604	61 018	28 715	19 340	10 641	7149
North Atlantic	55 679	28 310	21 409	18 161	14 961	7630	4212
North Indian	24 101	5481	5476	4281	2354	1029	614

South Indian	86 790	23 935	24 468	16 367	10 697	5108	2977
South Pacific	45 189	12 322	12 467	7169	4827	2577	1521
Eastern Pacific	59 175	28 825	17 592	19 722	12283	6482	3986
Global	423 296	125 477	14 243	94 415	64 462	33 467	20459

The table gives detailed cyclone counts from 1959 to 2022 across different ocean basins according to the IBTrACS dataset.

Most events occur in the Western Pacific, followed by the North Atlantic and South Indian basins. The North Indian basin region shows the lowest recorded cyclone numbers. The dataset outlines 423,296 recorded cyclone points with a noticeable decrease in recent years, which is possibly due to the variations in detection methods.

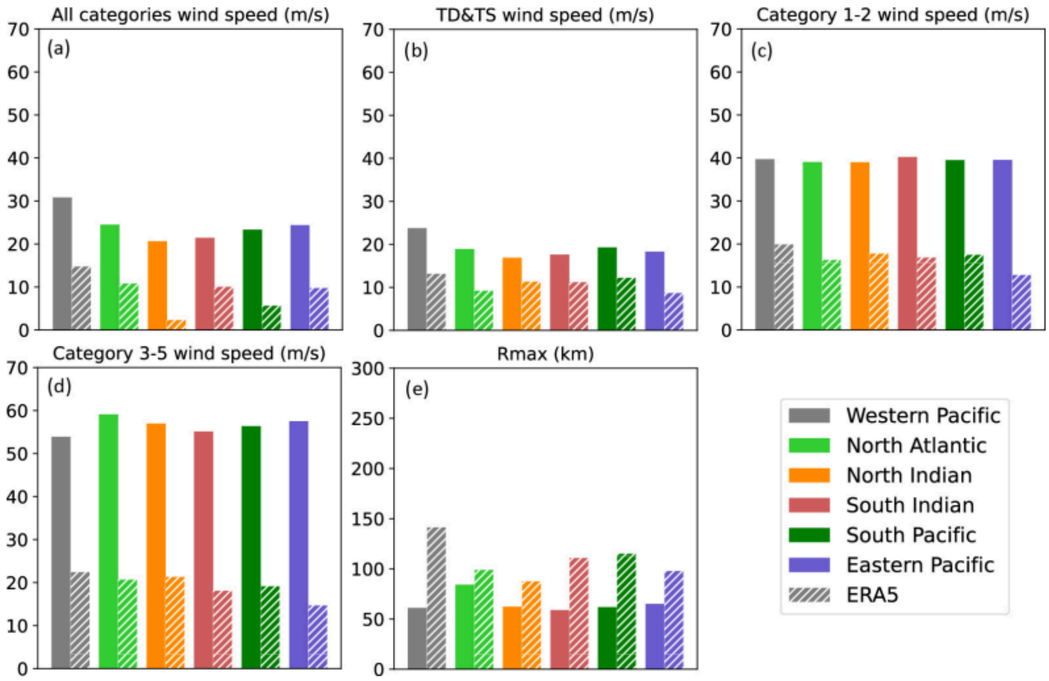

Figure 5.6: Bar charts for comparing the mean value of the 10m maximum wind speeds and the radii to maximum winds. Each of the colours

The figure consists of five subplots to compare the cyclone size and intensity patterns across different basins.

Subplots (a), (b), and (c) display average maximum winds for all cyclone categories, tropical depressions and storms, and Category 1-2 storms, respectively. Subplot (d) displays wind speeds for more intense Category 3-5 storms, where all basins report high wind speeds between 60 and 65 m/s. Subplot (e) illustrates mean RMax values, highlighting notable differences among basins, with the ERA5 dataset producing smaller estimates than the IBTrACS. The collared legend assigns a distinct shade to each basin, which enables a cyclone trait visual comparison.

The methodology consists of three main phases: Data preprocessing for leaning, normalizing, and structuring the dataset. Model development for constructing and training an LSTM model. Model evaluation to assess the performance using key metrics.

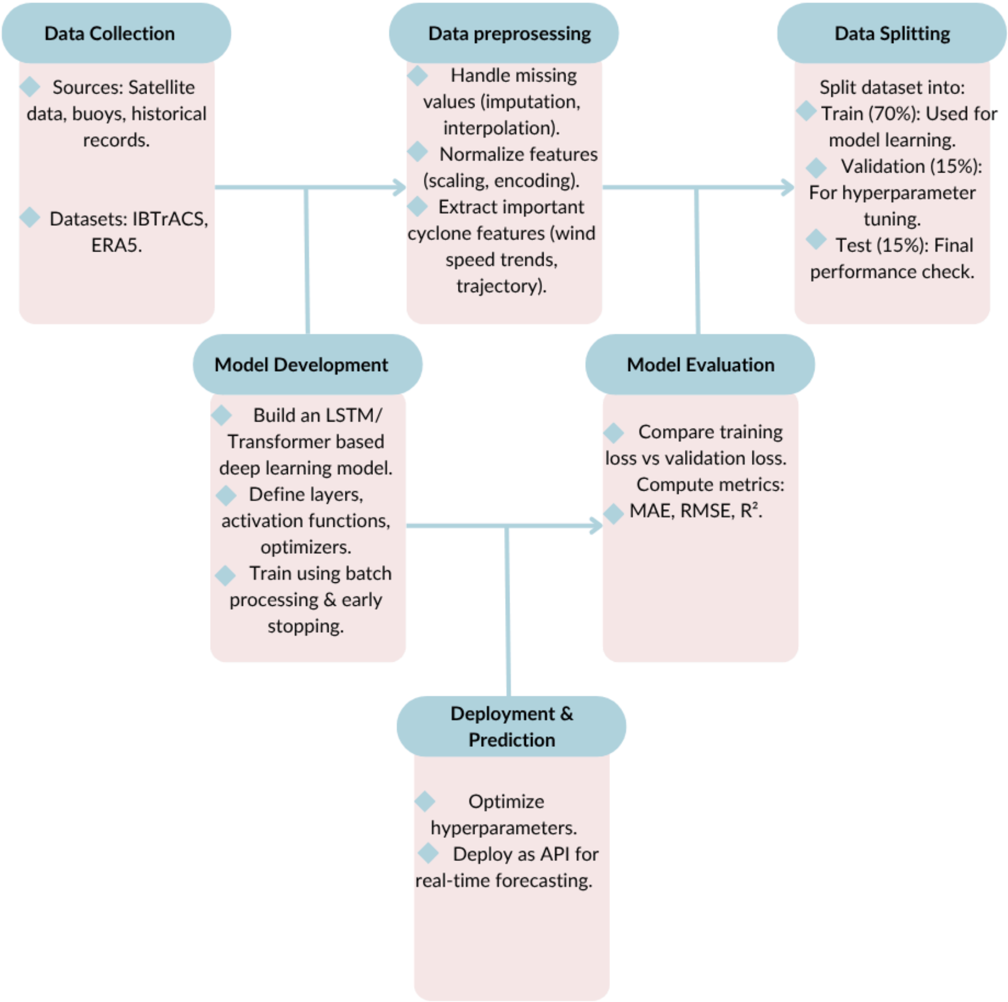

Figure 5.7: Flowchart of Cyclone Forecasting Using AI Techniques.

The flowchart illustrates the full process applied to cyclone forecasting using AI models as implemented in this thesis.

The process begins with the data collection from the IBTrACS and ERA5 datasets step. These two datasets combine weather data from different sources such as satellite imagery, buoys, and historical records.

The data preprocessing step includes handling missing values, normalizing features, and extracting cyclone-related parameters such as wind speed trends and trajectories. Then the datasets are split into 70% as a training set, 15% as a validation set, and 15% as a test set. The model development stage includes building LSTM or Transformer-based models, defining network architecture, and training using batch processing with early stopping. Model evaluation compares training and validation performance using MAE, RMSE, and R2 metrics. Finally, in this thesis, the deployment as an API for real-time forecasting is identified as a future work direction.

5.3.2 DATA PREPROCESSING

The preprocessing pipeline involves the following steps. First, regional cyclone datasets were merged into a single global one in the data collection and combination stage to ensure consistency and a broader representation of cyclone events worldwide. Next, columns like R34, R50, and R64 included high percentages of NaN values, which could negatively affect the model's performance if left untreated making identifying the missing values in the dataset a vital step of the data preprocessing. To address this, the NaN values were replaced with the respective column means using a mean imputation method. The selected approach is fast and effective when the proportion of missing data is relatively small and not systematically biased. After handling missing values, the Min–Max normalization was applied to numerical features such as latitude, longitude, wind speed, and pressure. This scaling method transforms all values into a fixed range which improves the stability and efficiency of model training. Then the dataset was divided into training, validation, and test sets for ensuring a balanced approach for the model development and evaluation. Finally, the sequential data was temporally reshaped into overlapping windows of 10 time steps per instance, enabling the LSTM model to capture patterns from past cyclone behaviour and use them to improve prediction accuracy.

5.3.3 MODEL DEVELOPMENT

The first model employs an LSTM architecture. LSTM has a strong ability in analysing patterns in time-series data. The architecture consists of an input layer, LSTM layers, dropout regularization, dense output layer, an Adam optimizer & Mean Squared Error Loss.

The input layer is designed to process sequences. It contains 10 consecutive time steps and four distinct features.

The first LSTM layer is with 64 units and the second with 32. Both layers use a ReLU activation function.

The dropout regularization is used with a 30% drop rate to prevent overfitting.

The dense output layer is a fully connected layer implemented to predict the cyclone wind speed (Vmax).

The Adam optimizer & the Mean Squared Error (MSE) are added as the loss function to optimize the model convergence.

The Long Short-Term Memory (LSTM) network processes sequential cyclone data through memory cells. The cells regulate the information flow using input, forget, and output gates. These gates allow the model to capture both short- and long-term temporal dependencies to make it suitable for the cyclone intensity forecasting. The full mathematical equations set governing the LSTM cell is provided in Appendix A.

Table 5.3: Summary of Hyperparameters for the LSTM Model.

Hyperparameter	**Value**	**Description**
Input Shape	(10, 9)	10-time step sequences with 9 feature inputs
Noise Layer	GaussianNoise(0.01)	Helps regularize by adding noise to input data
LSTM Layer 1	64 units, ReLU activation	Captures temporal dependencies in sequential data
Dropout 1	20%	Reduces overfitting by randomly dropping connections
LSTM Layer 2	32 units, ReLU activation	Further refines feature extraction from sequences
Dropout 2	20%	Additional regularization to enhance generalization
Dense Layer 1	16 units, ReLU activation	Fully connected layer for feature transformation

Dropout 3	10%	Controls overfitting before output layer
Output Layer	1 unit (Linear)	Predicts wind speed values
Loss Function	Mean Squared Error (MSE)	Measures prediction error
Optimizer	Adam	Adaptive learning rate optimization
Learning Rate	0.0001 (adaptive)	Adjusts dynamically using ReduceLROnPlateau
Batch Size	64	Number of samples per training update
Epochs	50	Maximum training iterations (early stopping applied)
Early Stopping	Patience=5 (val_loss)	Stops training when validation loss stops improving
Reduce LR on Plateau	factor=0.5, patience=3	Reduces learning rate when progress slows
Regularization	Batch Normalization + Dropout	Enhances model stability and generalization
Training Environment	Jupyter Notebook (Local)	Model trained on local hardware

During training, early stopping was employed to prevent overfitting by monitoring validation loss. If validation loss plateaued or began increasing, training was halted early. Additionally, dropout layers and L2 regularization were used to improve generalization and reduce overfitting.

The Transformer-based forecasting model uses a self-attention mechanism to capture the long-term dependencies and the complex patterns in the time-series data. It can handle the cyclone's intensity prediction by analysing the historical sequences of the atmospheric and oceanic variables. The architecture consists of:

Input Layer: Accepts the sequences of 10-time steps with 9 features (lat, lon, land, vmax, pressure, rmax, R34, R50, R64).

Positional Encoding: Adds the learned positional embedding to return the sequences order's information.

Multi-Head Attention: 4 attention heads to dynamically weigh the relationships between the features across the time steps.

Feed-Forward Network (FFN): A 2-layers dense network with a ReLU activation (ff-dim=32) to refine the feature's representations.

Layer Normalization & Dropout: Stabilizes the training and prevents the overfitting (30% dropout rate).

Global Average Pooling: Reduces the sequence's dimension to a single vector for the final prediction.

Dense Output Layer: Predicts the cyclone's wind speed (vmax) by using a single vector for the final prediction.

Adam Optimizer & Mean Absolute Error (MAE) Loss: Ensures a stable convergence and minimizes the prediction error.

The Transformer architecture relies on multi-head self-attention, which enables the model to capture relationships across all input time steps simultaneously. Unlike recurrent models, the Transformer processes sequences in parallel and learns dynamic dependencies through attention weights. The detailed equations for the self-attention mechanism and feed-forward blocks are included in Appendix B.

Table 5.4: Summary of Hyperparameters for the Transformer-based Model.

Hyperparameter	Value	Description
Input Shape	(10, 9)	Accepts the sequences of 10-time steps with 9 features (lat, lon, land, vmax, etc.).
Positional Encoding	Random normal initialization	Adds positional info to retain the sequence order.
Multi-Head Attention	num_heads=4, key_dim=9	4 attention heads to dynamically weigh relationships between features.
Feed-Forward Network (FFN)	Dense(32, ReLU) → Dense(9)	Refines feature representations with ReLU activation.
Layer Normalization	Applied after attention and FFN	Stabilizes training by normalizing inputs.
Dropout Rate	30%	Reduces overfitting by randomly dropping 30% of neuron outputs.
Pooling	GlobalAveragePooling1D	Reduces sequence to a single vector for final prediction.

Output Layer	Dense(1)	Predicts the cyclone wind speed (vmax) using a single neuron.
Optimizer	Adam (learning_rate=1e-4)	Optimizes model convergence.
Loss Function	Mean Absolute Error (MAE)	Minimizes the prediction error during the training.
Batch Size	64	Number of samples per training iteration.
Early Stopping	Patience = 5 epochs	Stops training if the validation loss stops improving for 5 epochs.
Learning Rate Scheduler	ReduceLROnPlateau (factor=0.5, patience=2)	Lowers learning rate when validation loss stagnates.
Model Checkpoint	Saves the best model based on val_loss	Ensures only the best model weights are retained.

Table 5.5: Key Differences between LSTM and Transformer Models.

Sequence Handling	Sequential (step-by-step)	Parallel processing (self-attention)
Attention Mechanism	None	Multi-head attention
Pooling	None	Global average pooling
Dropout Rate	40%	30%
Optimization	Adam (1e-4)	Adam (1e-4)

The table outlines a comparison of the key architectural and processing characteristics of LSTM networks and Transformer models.

LSTM models handle sequences in a stepwise manner without attention layers, while Transformers process sequences simultaneously using a multi-head attention mechanism. Pooling methods also differ; LSTMs generally lack pooling layers, while Transformers often use global average pooling to condense representations. In terms of regularization, LSTMs use a more aggressive dropout rate of 40%, exceeding Transformer 30%. Both models use the Adam optimizer with the same learning rate of 10^{-4}.

5.3.4 MODEL EVALUATION

The model performance was evaluated by implementing both visual and statistical techniques. Loss analysis, based on the comparison of training and validation loss curves, was used to assess the degree of overfitting. In addition, statistical metrics were employed to provide a quantitative evaluation of accuracy. The models performance was evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). These metrics were selected due to their wide adoption in both meteorology and machine learning for regression-based forecasting tasks. RMSE was chosen as it penalizes larger errors more heavily, which is particularly relevant for extreme cyclone intensities where misestimating can have severe implications. MAE complements this by providing an intuitive measure of the average absolute error in the original units to offer a clear sense of typical prediction accuracy. Finally, R^2 was included to capture the proportion of variance in the observed data explained by the models, to offer a standardized measure for model comparison. Other metrics, such as the Nash–Sutcliffe Efficiency (NSE) or Mean Bias Error, were considered but not prioritized, as they are less commonly used in machine learning-based cyclone forecasting and provide limited additional interpretability for the present task.

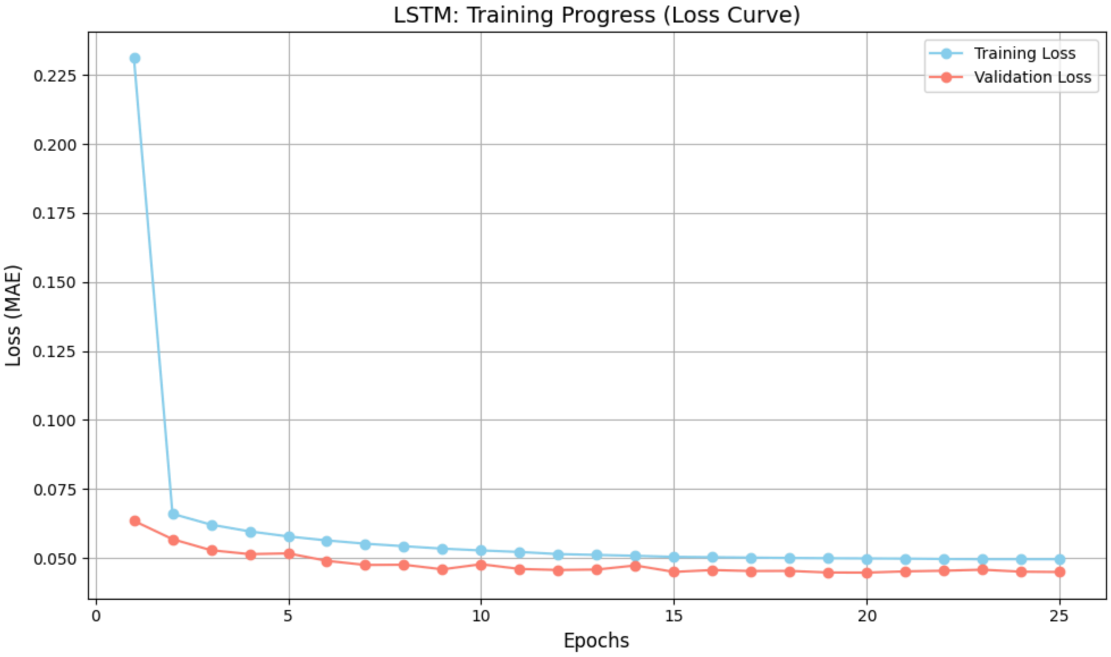

Figure 5.8: LSTM model Training Progress Graph (Loss vs. Epochs)

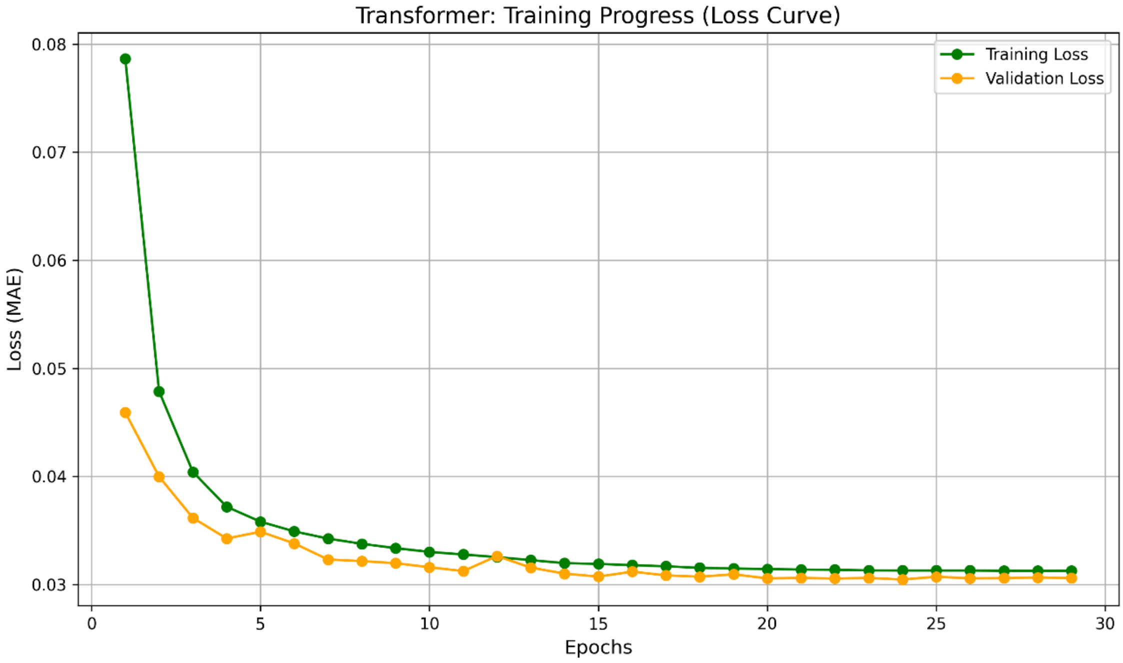

Figure 5.9: Transformer-based model Training Progress Graph (Loss vs. Epochs).

The Loss vs. Epochs graphs illustrate how the two models improve during training. The training loss (blue) and validation loss (orange) both decrease over time, which means the models are learning well. The LSTM model's validation loss stops decreasing much around epoch 15 and the Transformer model's around epoch 24, meaning the models have reached their best performances. The small difference between training and validation loss suggests no overfitting, meaning both models generalize well to new data.

The methodology effectively combined data preprocessing, deep learning techniques, and evaluation strategies to develop two optimized forecasting models. The use of LSTMs and Transformers, along with regularization and fine-tuning, ensured a balance between learning efficiency and generalization. This structured approach contributed to the model's ability to predict cyclone intensities with improved accuracy.

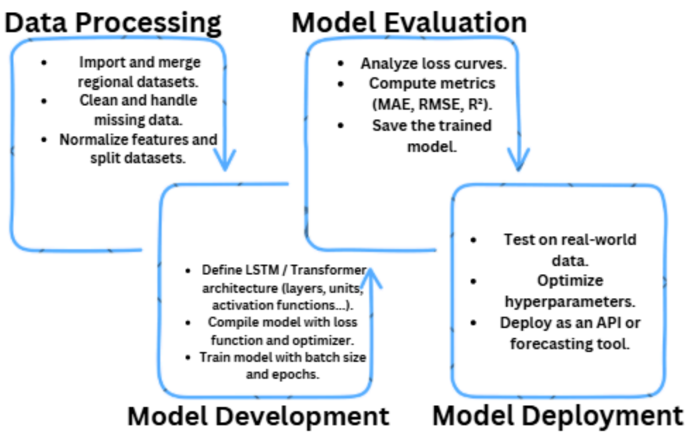

Figure 5.10: LSTM/Transformer-Based Extreme Weather Forecasting Workflow illustrating the step-by-step process of data preprocessing, model training, evaluation, and deployment

The figure shows a detailed representation of the LSTM/Transformer-based extreme weather forecasting workflow. Data processing includes importing datasets, handling missing values, normalizing features, and splitting data. The evolution step compares predicted and actual results using MAR, RMSE, and R^2 metrics. Finally, the model deployment suggests testing the system in real-world conditions, optimizing hyperparameters, and making API-based forecasts.

6 RESULTS AND DISCUSSION

6.1 MODEL PERFORMANCE METRICS

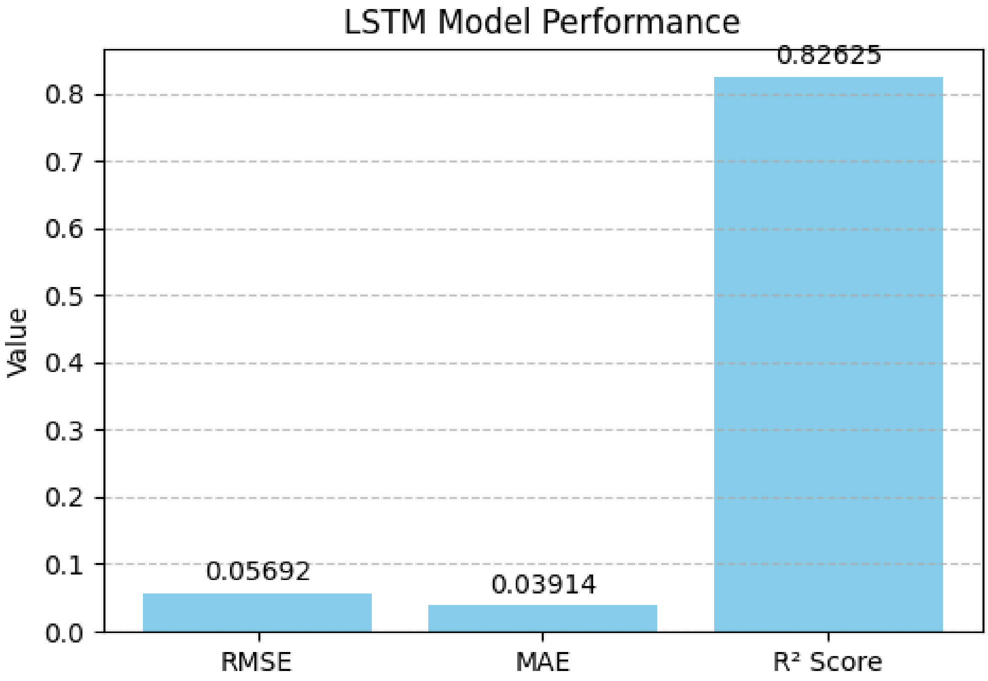

Figure 6.1: LSTM Performance Comparison Bar Chart

This bar chart compares the model's performance using three key metrics. RMSE to measure the average error in predictions, a lower value is better. MAE. It is similar to RMSE but focuses on absolute errors. Also, lower is better. R^2 Score. It indicates how well the model explains the data. A value close to 1 means the model predicts well. These key metrics are commonly implemented in weather prediction.

Since RMSE and MAE are low and R^2 is high, the model is making good predictions.

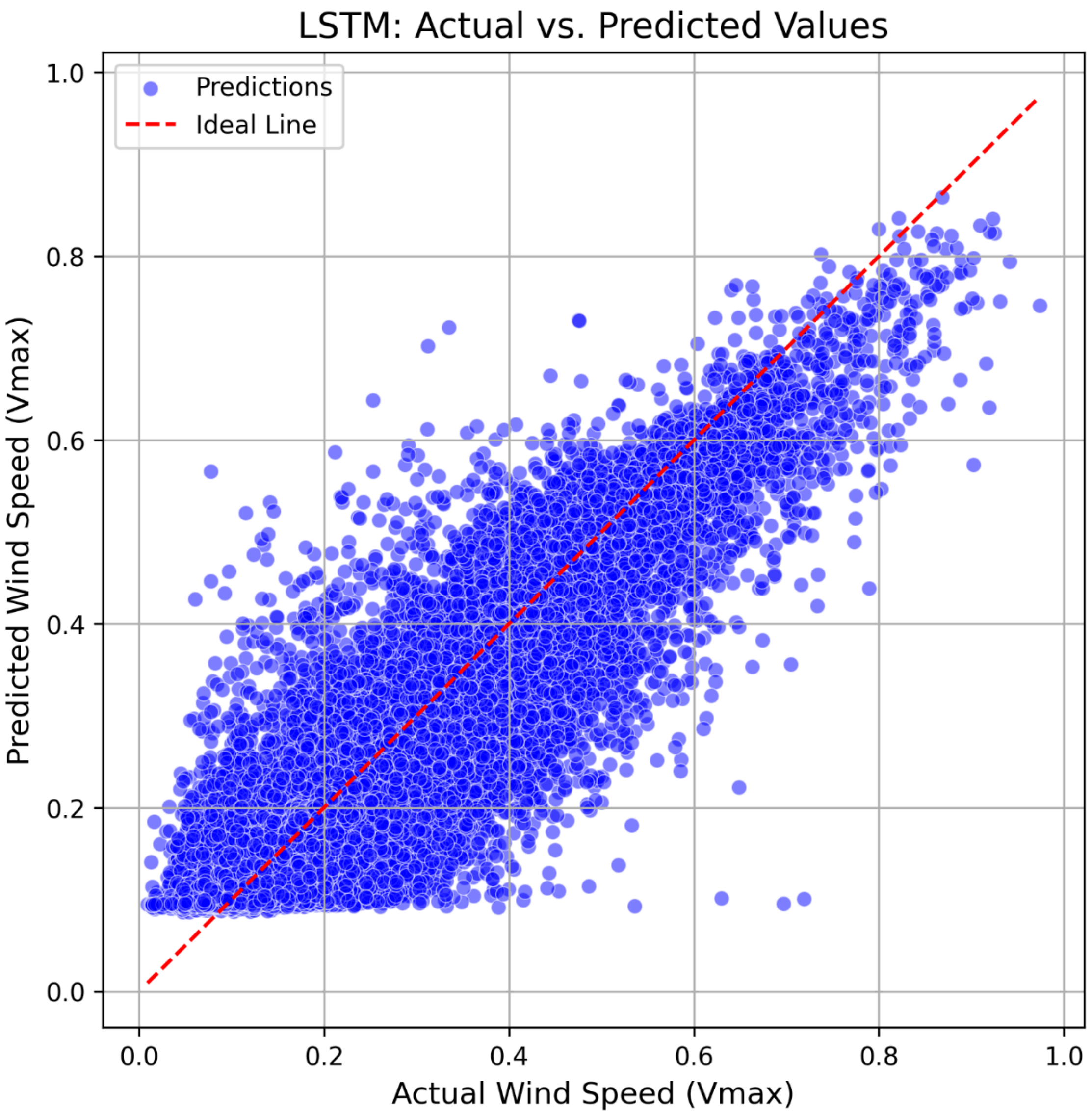

Figure 6.2: Actual vs. Predicted Scatter Plot (LSTM)

This scatter plot compares the model's predictions (y-axis) to the real values (x-axis). The red line represents perfect predictions if all points were on this line, the model would be 100% accurate.

Most of the points are close to the red line, meaning the model's predictions are very close to actual values. There is some spread, especially at higher values, meaning the model sometimes struggles with extreme cases. However, the overall pattern matches well, so the model is reliable.

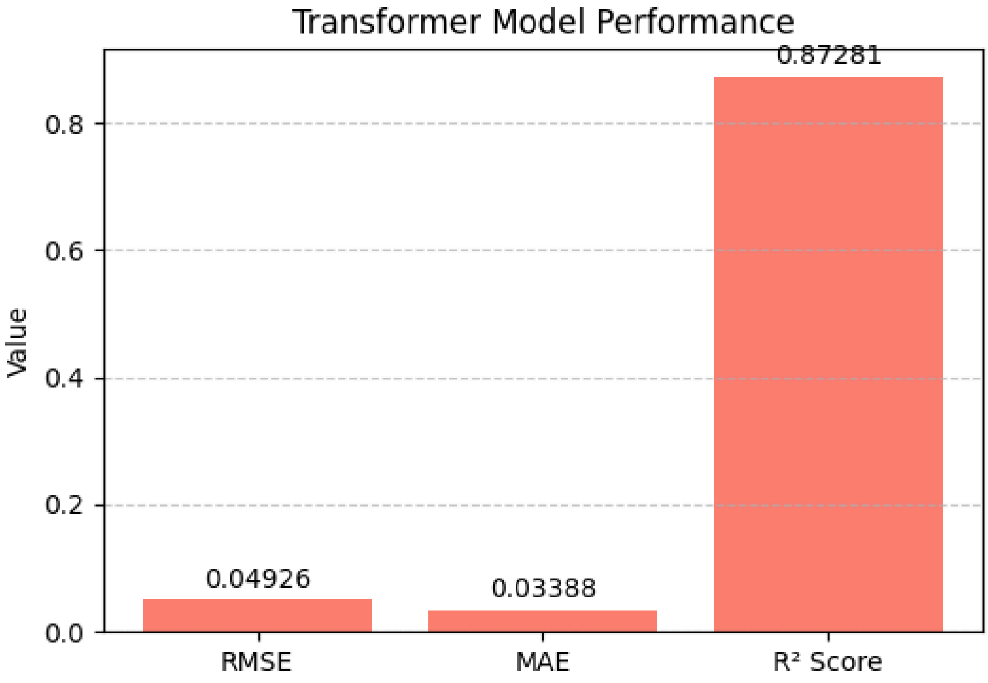

Figure 6.3: Transformer Performance Comparison Bar Chart

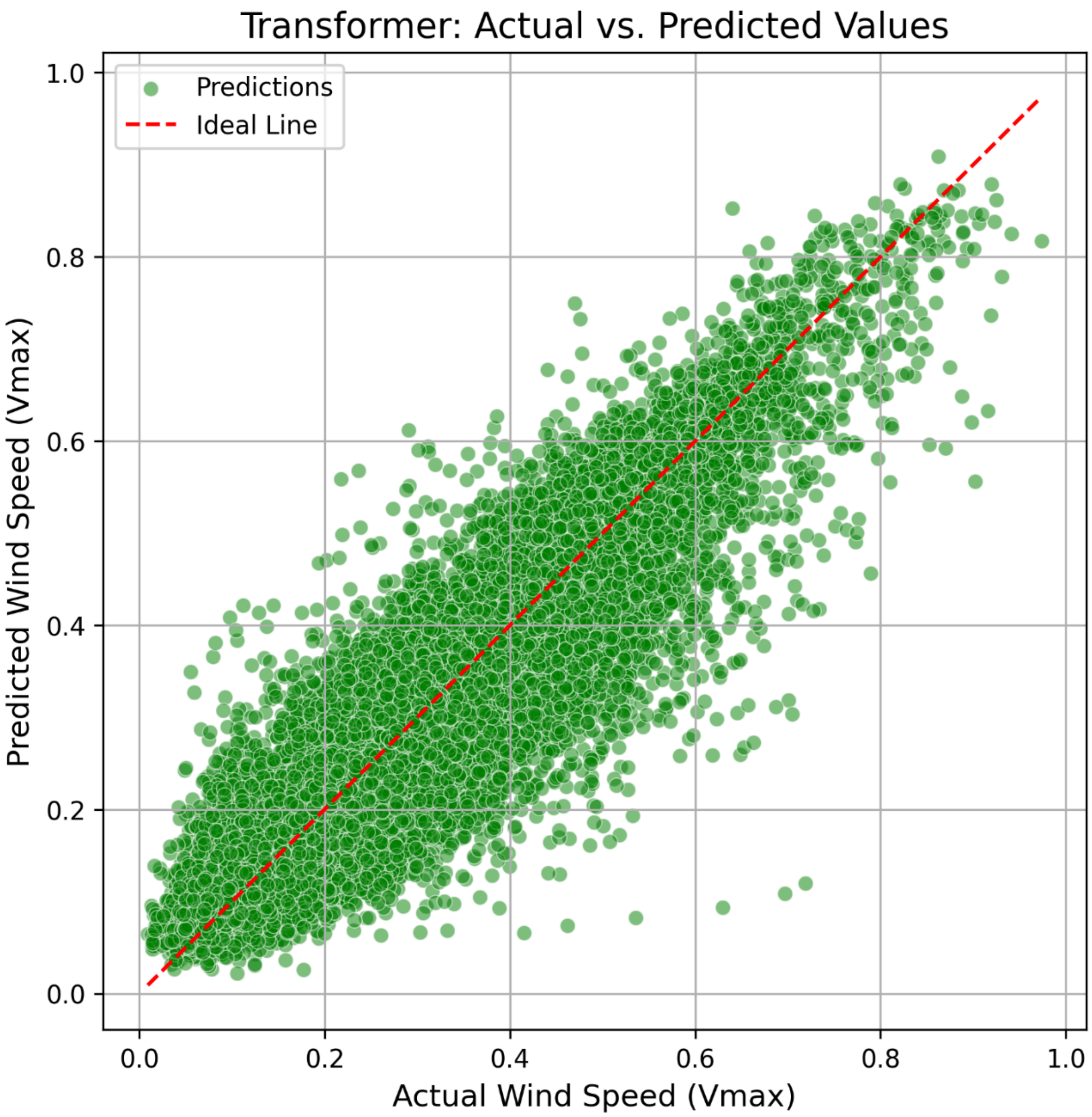

Figure 6.4: Actual vs. Predicted Scatter Plot (Transformer)

Table 6.1: The final models' performance was evaluated using three key metrics

	LSTM	Transformer
RMSE	0.05692	0.04926
MAE	0.03914	0.03388
R^2	0.82625	0.87281

The optimized LSTM model indicated an R² score of 0.82625, indicating that it explains 82.6% of the variance in cyclone predictions. The Transformer-based model indicated an R² score of 0.87281, indicating that 87.2% of the variance in cyclone predictions outperforms the LSTM model with a 5.6% difference. The Transformer outperforms the LSTM due to its attention mechanism. It allows the model to capture long-range dependencies in cyclone evolution. Unlike the sequential LSTM, the Transformer processes all time steps in parallel which enables more effective representation of temporal dynamics.

Error Reduction: RMSE and MAE demonstrate stable and improved performance compared to traditional methods. RMSE was selected due to its ability to penalise larger errors more strongly, which is critical for cyclone forecasting. MAE complements this by providing an intuitive average error magnitude, while R^2 measures the proportion of variance explained by the models. Together, these three metrics provide a balanced assessment of prediction performance.

The scatter plot of actual vs. predicted values confirms the two models' ability to capture cyclone intensities accurately, though some deviations exist in higher values.

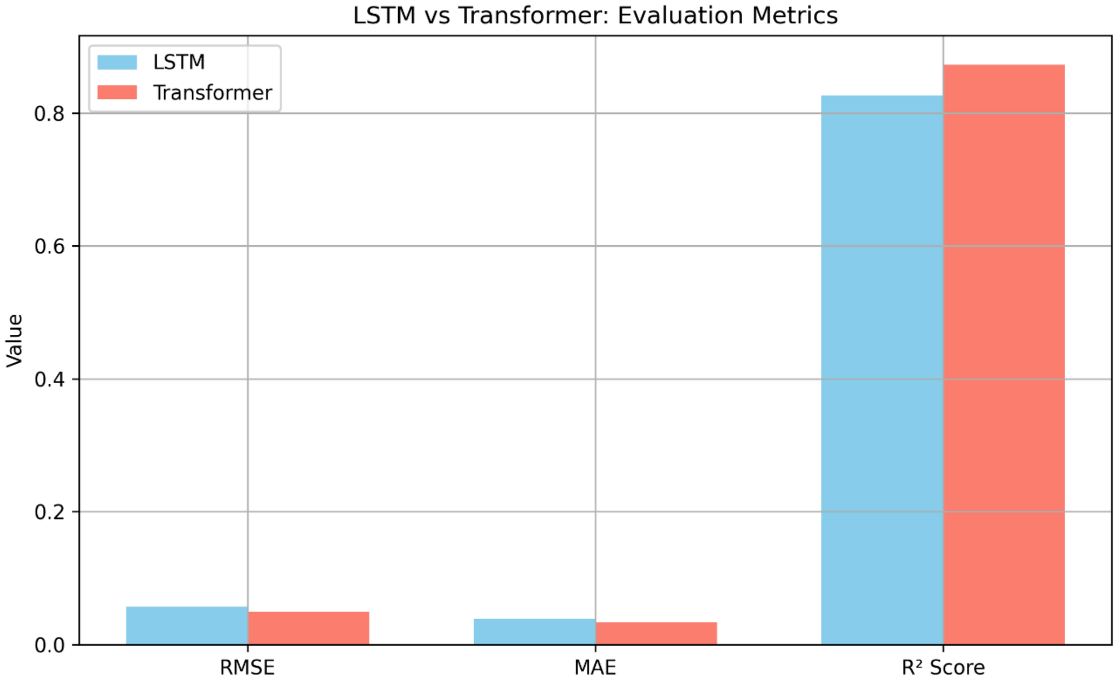

Figure 6.5: LSTM vs. Transformer: Evaluation Metrics

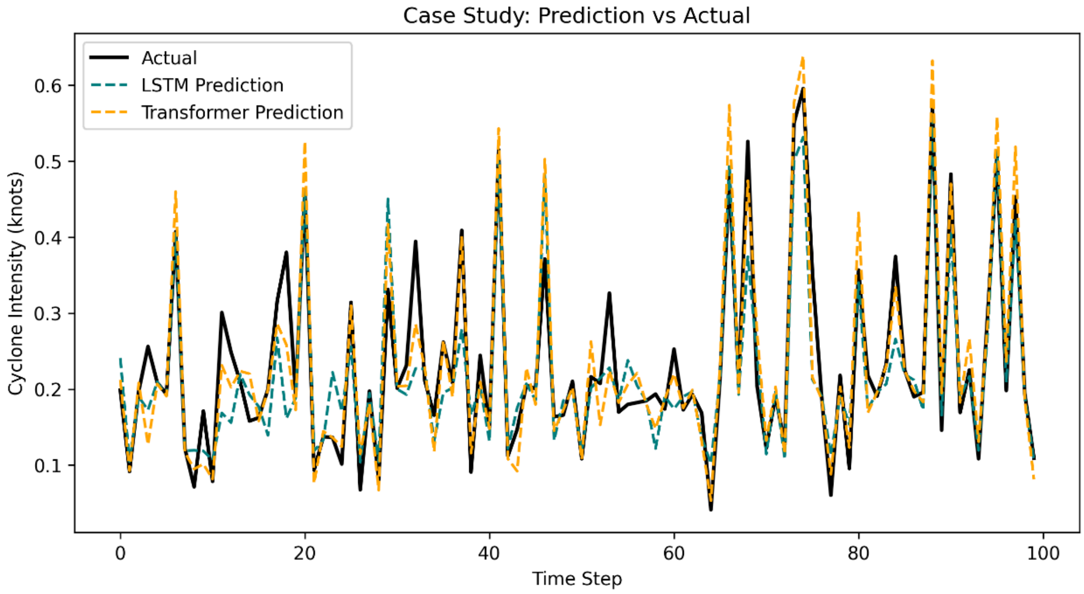

Figure 6.6: Case Study Prediction vs. Actual

A comparison of actual cyclone intensity (black line) with prediction from both methods for a set of storm cases. The teal dashed line presents the values from the LSTM model and the orange one for the Transformer.

Both models prevent predictions close to the actual intensity. However, the Transformer outperforms the LSTM, especially at higher intensity peak. This highlights that the attention mechanism enable the Transformer to adapt to sudden changes in cyclone dynamics.

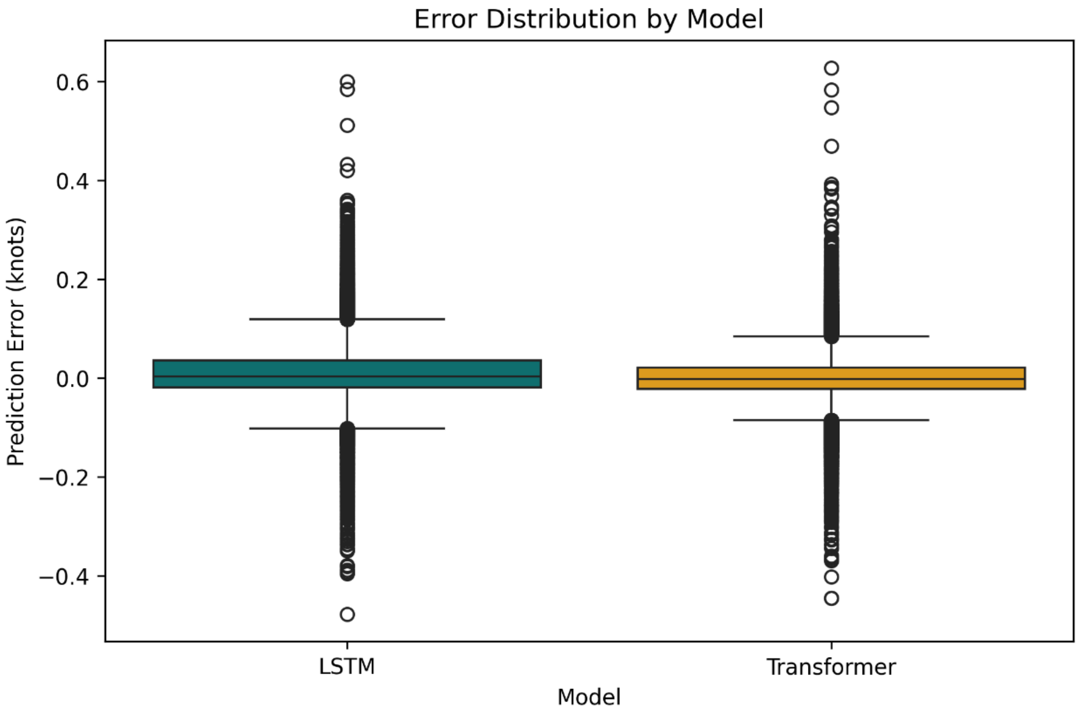

Figure 6.7: Error Distribution by Model (Boxplot).

The boxplot shows the LSTM and the Transformer prediction errors.

The models have a low median error (close to 0) which means the predictions are based on the average. The LSTM shows more outliers which means it makes larger errors. The Transformer has a narrower interquartile range which means more constant prediction with fewer large deviations.

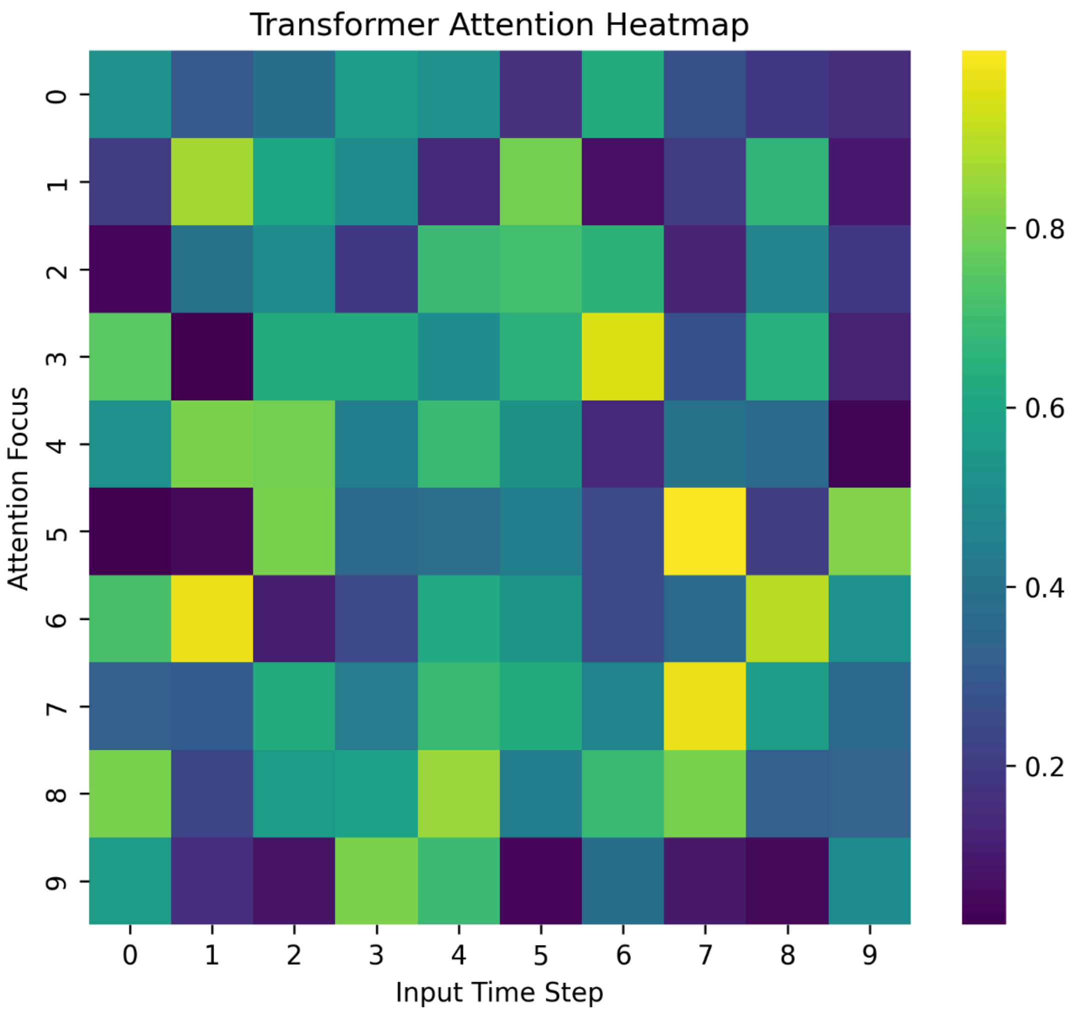

Figure 6.8: Transformer Attention Heatmap

The figure presents a heatmap of the Transformer's attention scores across different input time steps. In this visualization, the brighter areas indicate where the model places greater focus when generating predictions.

The Transformer does not treat all time steps equally. Instead, it pays more attention to some specific points in the sequence. This behaviour demonstrates how the attention mechanism captures critical temporal dependencies within the input. In this way, the model becomes more interpretable as the heatmap reveals which parts of the cyclone's historical trajectory were most influential in shaping the final prediction.

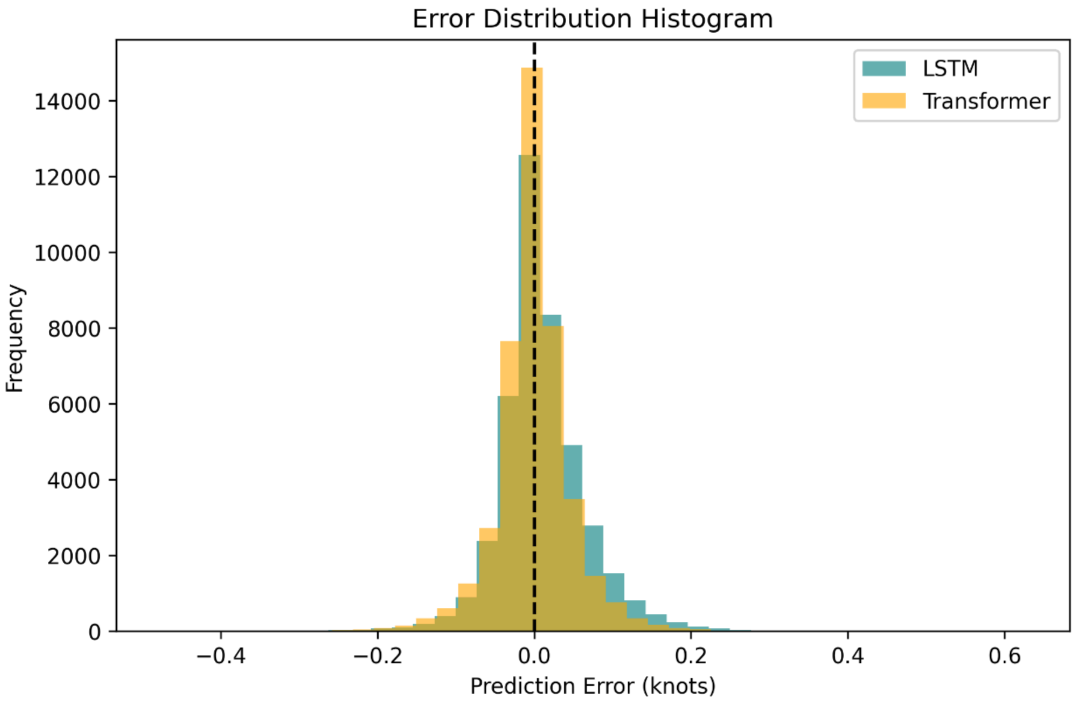

Figure 6.9: Error Distribution Histogram

This figure shows histograms of the prediction errors for both the LSTM and Transformer models. The majority of errors are clustered around zero, forming an almost symmetric distribution, which indicates that both models generally achieve unbiased predictions.

When comparing the two models, it is clear that the Transformer produces a slightly narrower spread of errors. This means that it makes fewer extreme mistakes compared to the LSTM. The result reinforces the findings from the boxplot analysis, while both models demonstrate reliable predictive performance, the Transformer achieves greater robustness and consistency across cases.

6.2 KEY OBSERVATIONS

The Transformer-based model significantly improved the forecasting accuracy over the LSTM model. Handling overfitting using dropout and L2 regularization played a critical role in achieving this improvement. Challenges in extreme values: Some underestimations occurred at high cyclone intensities, suggesting the potential for further enhancement.

While this study demonstrates the potential of deep learning methods for cyclone intensity prediction, there are some limitations. The use of 6-hourly best-track data may not fully capture very rapid changes in intensity, and the exclusive focus on intensity prediction means that cyclone track and multi-hazard aspects were beyond the scope of this work. Furthermore, although the dataset integrates multiple ocean basins, differences in data availability and reporting practices may have influenced the models' performance across regions. Finally, although the Transformer's attention heatmaps provide an initial step toward interpretability, further research is needed to develop more comprehensive explainability tools for operational forecasting.

7 CONCLUSION

This work demonstrated the potential of AI-based models in improving cyclone forecasting accuracy. By leveraging deep learning techniques, particularly LSTMs and Transformers, a predictive system that enhances traditional weather forecasting models was developed. The optimized LSTM and Transformer-based models significantly improve the cyclone prediction accuracy over the traditional forecasting methods. Regularization techniques and proper data preprocessing were critical to reducing overfitting and improving generalization. The thesis contributes to the advancement of AI-driven meteorological forecasting, with direct applications for disaster preparedness and mitigation.

Future research on cyclone intensity prediction can proceed along several promising directions such as testing other Deep Learning -based models to explore alternative deep learning architectures. Also by integrating additional data sources (satellite imagery, real-time weather station data) to improve accuracy. For better forecasts analysis, deploying the models in operational environments to evaluate real-time forecasting performance is a promising solution. To increase trust in automated forecasting models, improving the AI interpretability is a strong decision.

In practical terms, deploying these models for real-time forecasting would require addressing several challenges. Data latency and preprocessing speed must be managed to ensure timely predictions, while integration into existing meteorological workflows is essential for operational use. Robustness to missing or noisy inputs also remains a critical issue, and improving interpretability will be necessary to build trust among forecasters and decision-makers.

APPENDIX

APPENDIX A: LSTM EQUATIONS

The LSTM cell operates using the following set of equations:

Forget Gate:

$$f_t = \sigma\left(W_f\left[h_{t-1}, x_t\right] + b_f\right). \quad \text{(A.1)}$$

Controls which information from the previous cell state should be discarded. Uses a sigmoid activation function (σ\sigmaσ) to output values between 0 and 1 (where 0 means "forget" and 1 means "keep").

Input Gate:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i). \quad \text{(A.2)}$$

Decides what new information to store in the cell state. Uses another sigmoid activation function to regulate updates.

Cell State Update:

$$\tilde{c}_t = tanh(W_c[h_{t-1}, x_t] + b_c). \quad \text{(A.3)}$$

Creates a candidate new memory that will be added to the cell state. Uses a tanh activation function to produce values between -1 and 1.

Final Cell State:

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t\,. \quad \text{(A.4)}$$

Updates the cell state by keeping relevant past information $(f_t \cdot c_{t-1})$ and adding new information $(i_t \cdot \tilde{c}_t)$.

Output Gate:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o). \quad \text{(A.5)}$$

It determines which part of the cell state should be output as the hidden state. Uses a sigmoid activation function to regulate output flow.

Hidden State:

$$h_t = o_t \cdot tanh(c_t). \qquad \text{(A.6)}$$

Final hidden state used as output. Uses tanh activation function to scale values between -1 and 1.

These operations allow the model to retain relevant historical information and discard unnecessary noise.

APPENDIX B: TRANSFORMER EQUATIONS

The Transformer architecture is built around self-attention mechanisms and position-wise feed-forward networks.

Input Embedding and Positional Encoding:

Each input token x_i is mapped to a vector$x_i \in R^d$. To retain sequence order information, positional encoding $p_i \in R^d$ is added:

$$z_i = x_i + p_i. \qquad \text{(B.1)}$$

Positional encodings can be learned or fixed.

Self-Attention Mechanism:

Given an input sequence of $\left[z_1, z_2, ..., z_n\right] \in R^{n \times d}$, three linear projections are computed:

$$Q = ZW_Q, \; K = ZW_K, \quad V = ZW_V. \quad \text{(B.2, B.3, B.4)}$$

Where $W_Q, W_K, W_V \in R^{d \times d_k}$ are the learned weight’s matrices.

The attention scores are computed as:

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

$QK^T \in R^{n \times n}$: Dot products of queries and keys.

$\sqrt{d_k}$: Scales dot products to prevent large values from saturating softmax.

Softmax normalizes attention weights row-wise.

Multi-Head Attention: The multi-heads allow the model to focus on the different parts of the input. For h heads:

$$\mathrm{MultiHead}(Z) = \mathrm{Concat}(head_1, \ldots, head_h)W_O. \quad \text{(B.6)}$$

Where

$$head_i = \text{ Attention } (ZW_{Q_i}, \; ZW_{K_i}, \; ZW_{V_i}). \qquad \text{(B.7)}$$

And $W_O \in R^{hd_v \times d}$ projects concatenated outputs back to the original dimension.

Position-Wise Feed-Forward Network (FFN)

Each position (time-step) in the sequence is processed independently through a fully connected network:

$$\mathrm{FFN}(z) = \mathrm{ReLU}(zW_1 + b_1)W_2 + \; b_2 \text{ n. (B.8)}$$

Where $W_1 \in R^{d \times d_{ff}}$, $W_2 \in R^{d_{ff} \times d}$, and d_{ff} is the hidden layer size.

Layer Normalization and Residual Connections

Each sublayer (self-attention and FFN) is followed by:

$$\text{SublayerOutput=LayerNorm(Z+Sublayer(Z)).} \quad \text{(B.9)}$$

Layer normalization stabilizes training:

$$\mathrm{LayerNorm}(x) = \gamma. \; \frac{x - \mu}{\sqrt{\sigma^2 + \epsilon}} + \; \beta. \qquad \text{(B.10)}$$

Where μ, σ^2 are mean/variance across features, and γ, β are learnable scale/shift parameters.

Encoder Block

An encoder block applies:

$$Z' = MultiHead(Z). \qquad \text{(B.11)}$$

$$Z'' = FFN(Z'). \qquad \text{(B.12)}$$

Output = Z''

Decoder Block

The decoder uses masked self-attention to prevent attending to future tokens:

$$Z' = MaskedMultiHead(Z). \qquad \text{(B.13)}$$

$$Z'' = MultiHead(Z', K_{enc}, V_{enc}). \qquad \text{(B.14)}$$

Output = $FFN(Z'')$

Where K_{enc}, V_{enc} are from the encoder's outputs.

REFERENCES

Abdusalomov, A. B., Islam, B. M. S., Nasimov, R., Mukhiddinov, M., & Whangbo, T. K. (2023). An Improved Forest Fire Detection Method Based on the Detectron2 Model and a Deep Learning Approach. *Sensors*, *23*(3). https://doi.org/10.3390/s23031512

Adekunle Stephen Toromade, Deborah Aanuoluwa Soyombo, Eseoghene Kupa, & Tochukwu Ignatius Ijomah. (2024). Reviewing the impact of climate change on global food security: Challenges and solutions. *International Journal of Applied Research in Social Sciences*, *6*(7), 1403–1416. https://doi.org/10.51594/ijarss.v6i7.1300

Adeoye Taofik Aderamo, Henry Chukwuemeka Olisakwe, Yetunde Adenike Adebayo, & Andrew Emuobosa Esiri. (2024). AI-enabled predictive safeguards for offshore oil facilities: Enhancing safety and operational efficiency. *Comprehensive Research and Reviews in Engineering and Technology*, *2*(1), 023–043. https://doi.org/10.57219/crret.2024.2.1.0060

Bochenek, B., & Ustrnul, Z. (2022). Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives. *Atmosphere*, *13*(2). https://doi.org/10.3390/atmos13020180

Boussioux, L., Zeng, C., Guénais, T., & Bertsimas, D. (2022). Hurricane Forecasting: A Novel Multimodal Machine Learning Framework. *Weather and Forecasting*, *37*(6), 817–831. https://doi.org/10.1175/WAF-D-21-0091.1

Brotzge, J. A., Berchoff, D., Carlis, D. L., Carr, F. H., Carr, R. H., Gerth, J. J., Gross, B. D., Hamill, T. M., Haupt, S. E., Jacobs, N., McGovern, A., Stensrud, D. J., Szatkowski, G., Szunyogh, I., & Wang, X. (2023). Challenges and Opportunities in Numerical Weather Prediction. *Bulletin of the American Meteorological Society*, *104*(3), E698–E705. https://doi.org/10.1175/BAMS-D-22-0172.1

Chand, C. P., Ali, M. M., Himasri, B., Bourassa, M. A., & Zheng, Y. (2022). Predicting Indian Ocean Cyclone Parameters Using an Artificial Intelligence Technique. *Atmosphere*, *13*(7). https://doi.org/10.3390/atmos13071157

Chen, J. H., Zhou, L., Magnusson, L., McTaggart-Cowan, R., & Köhler, M. (2023). Tropical Cyclone Forecasts in the DIMOSIC Project—Medium-Range Forecast Models With Common Initial Conditions. *Earth and Space Science*, *10*(7). https://doi.org/10.1029/2023EA002821

Chen, L., Chen, Z., Zhang, Y., Liu, Y., Osman, A. I., Farghali, M., Hua, J., Al-Fatesh, A., Ihara, I., Rooney, D. W., & Yap, P. S. (2023). Artificial intelligence-based solutions for climate change: a review. In *Environmental Chemistry Letters* (Vol. 21, Issue 5, pp. 2525–2557). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s10311-023-01617-y

Chkeir, S., Anesiadou, A., Mascitelli, A., & Biondi, R. (2023). Nowcasting extreme rain and extreme wind speed with machine learning techniques applied to different input datasets. *Atmospheric Research*, *282*. https://doi.org/10.1016/j.atmosres.2022.106548

Climate Change and the Role of AI. (n.d.). https://doi.org/10.13140/RG.2.2.30355.69924

Cowls, J., Tsamados, A., Taddeo, M., & Floridi, L. (2023). The AI gambit: leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations. *AI and Society*, *38*(1), 283–307. https://doi.org/10.1007/s00146-021-01294-x

Demaria, M., Franklin, J. L., Chirokova, G., Radford, J., Demaria, R., Musgrave, K. D., Ebert-Uphoff, I., & Lynker, B. (n.d.). *Evaluation of Tropical Cyclone Track and Intensity Forecasts from Artificial Intelligence Weather Prediction (AIWP) Models*.

Hess, P., & Boers, N. (2022). Deep Learning for Improving Numerical Weather Prediction of Heavy Rainfall. *Journal of Advances in Modeling Earth Systems*, *14*(3). https://doi.org/10.1029/2021MS002765

Huang, C., Bai, C., Chan, S., Zhang, J., & Wu, Y. (2023). MGTCF: Multi-Generator Tropical Cyclone Forecasting with Heterogeneous Meteorological Data. www.aaai.org

Jiang, J., Huang, Z. G., Grebogi, C., & Lai, Y. C. (2022). Predicting extreme events from data using deep machine learning: When and where. *Physical Review Research*, *4*(2). https://doi.org/10.1103/PhysRevResearch.4.023028

Mashao, F. M., Mothapo, M. C., Munyai, R. B., Letsoalo, J. M., Mbokodo, I. L., Muofhe, T. P., Matsane, W., & Chikoore, H. (2023). Extreme Rainfall and Flood Risk Prediction over the East Coast of South Africa. *Water (Switzerland)*, *15*(1). https://doi.org/10.3390/w15010050

Mukkavilli, S. K., Civitarese, D. S., Schmude, J., Jakubik, J., Jones, A., Nguyen, N., Phillips, C., Roy, S., Singh, S., Watson, C., Ganti, R., Hamann, H., Nair, U., Ramachandran, R., & Weldemariam, K. (2023). *AI Foundation Models for Weather and Climate: Applications, Design, and Implementation*. http://arxiv.org/abs/2309.10808

Nath, P., Shukla, P., Wang, S., & Quilodrán-Casas, C. (2023). *Forecasting Tropical Cyclones with Cascaded Diffusion Models*. http://arxiv.org/abs/2310.01690

Rahman, S., Sharmin, N., Rahat, A., Rahman, M., & Rahman, M. (2024). Tropical cyclone warning and forecasting system in Bangladesh: challenges, prospects, and future direction to adopt artificial intelligence. In *Computational Urban Science* (Vol. 4, Issue 1). Springer. https://doi.org/10.1007/s43762-023-00113-x

Ran, N., Xiao, P., Wang, Y., Shi, W., Lin, J., Meng, Q., & Allmendinger, R. (2024). *HR-Extreme: A High-Resolution Dataset for Extreme Weather Forecasting*. http://arxiv.org/abs/2409.18885

Rayhan, A. (n.d.-a). AI AND THE ENVIRONMENT : TOWARD SUSTAINABLE DEVELOPMENT AND CONSERVATION.

Rayhan, A. (n.d.-b). AI for Environmental Monitoring, Disaster Response, and Poverty Alleviation. https://doi.org/10.13140/RG.2.2.28296.84485

Rayhan, A., Group, C., & Abu Rayhan, B. (n.d.). *Artificial Intelligence And Climate Crisis: A Pathway To Mitigation And Adaptation*. https://doi.org/10.13140/RG.2.2.12185.60003

Salcedo-Sanz, S., Pérez-Aracil, J., Ascenso, G., Del Ser, J., Casillas-Pérez, D., Kadow, C., Fister, D., Barriopedro, D., García-Herrera, R., Giuliani, M., & Castelletti, A. (2024). Analysis, characterization, prediction, and attribution of extreme atmospheric events with machine learning and deep learning techniques: a review. In *Theoretical and Applied Climatology* (Vol. 155, Issue 1, pp. 1–44). Springer. https://doi.org/10.1007/s00704-023-04571-5

Sippel, J. A., Wu, X., Ditchek, S. D., Tallapragada, V., & Kleisft, D. T. (2022). Impacts of Assimilating Additional Reconnaissance Data on Operational GFS Tropical Cyclone Forecasts. *Weather and Forecasting*, *37*(9), 1615–1639. https://doi.org/10.1175/WAF-D-22-0058.1

Snejana, D. (2022). Applying Artificial Intelligence (AI) for Mitigation Climate Change Consequences of the Natural Disasters. *Research Journal of Ecology and Environmental Sciences*, *2*(4), 211–218. https://doi.org/10.31586/rjees.2022.343

Takaya, Y., Caron, L. P., Blake, E., Bonnardot, F., Bruneau, N., Camp, J., Chan, J., Gregory, P., Jones, J. J., Kang, N., Klotzbach, P. J., Kuleshov, Y., Leroux, M. D., Lockwood, J. F., Murakami, H., Nishimura, A., Pattanaik, D. R., Philp, T. J., Ruprich-Robert, Y., … Zhan, R. (2023). Recent advances in seasonal and multi-annual tropical cyclone forecasting. *Tropical Cyclone Research and Review*, *12*(3), 182–199. https://doi.org/10.1016/j.tcrr.2023.09.003

Verma, S., Srivastava, K., Tiwari, A., & Verma, S. (2023). *Deep Learning Techniques in Extreme Weather Events: A Review*. http://arxiv.org/abs/2308.10995

Wang, Z., Zhao, J., Huang, H., & Wang, X. (2022). A Review on the Application of Machine Learning Methods in Tropical Cyclone Forecasting. In *Frontiers in Earth Science* (Vol. 10). Frontiers Media S.A. https://doi.org/10.3389/feart.2022.902596

Xu, Z., Guo, J., Zhang, G., Ye, Y., Zhao, H., & Chen, H. (2024). Global tropical cyclone size and intensity reconstruction dataset for 1959-2022 based on IBTrACS and ERA5 data. *Earth System Science Data*, *16*(12), 5753–5766. https://doi.org/10.5194/essd-16-5753-2024

Zhong, X., Chen, L., Liu, J., Lin, C., Qi, Y., & Li, H. (2023). *FuXi-Extreme: Improving extreme rainfall and wind forecasts with diffusion model*. https://doi.org/10.1007/s11430-023-1427-x