Predicting Tropical Cyclone Formation and its Landfall's Characteristics

by

Sandeep Kumar
(Roll No. - PhD16010)

Under the supervision of

Dr. Ashish Kumar Pandey

INDRAPRASTHA INSTITUTE OF INFORMATION TECHNOLOGY DELHI
NEW DELHI- 110020

October, 2022

Predicting Tropical Cyclone Formation and its Landfall's Characteristics

A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

SANDEEP KUMAR

(Roll No. - PhD16010)

Department of Computer Science and Engineering

INDRAPRASTHA INSTITUTE OF INFORMATION TECHNOLOGY DELHI

NEW DELHI- 110020

October, 2022

CERTIFICATE

This is to certify that the thesis titled **Predicting Tropical Cyclone Formation and its Landfall Characteristics**, submitted by **Mr. Sandeep Kumar**, to the Indraprastha Institute of Information Technology Delhi, for the award of the degree of **Doctor of Philosophy**, is a bonafide record of the research work done by him under my supervision. The contents of this thesis, in full or in parts, have not been submitted to any other institute or university for the award of any degree or diploma.

Dr. Ashish Kumar Pandey
(Thesis Supervisor)
Assistant Professor
Department of Mathematics
Indraprastha Institute of Information Technology Delhi
New Delhi - 110020

Place: New Delhi
Date: October 18, 2022

DECLARATION

I hereby declare that the work presented in this thesis titled "Predicting Tropical Cyclone Formation and its Landfall Characteristics", submitted to the Department of Computer Science, Indraprastha Institute of Information Technology Delhi, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Computer Science, is an authentic proof of my own work carried out under the supervision Dr. Ashish Kumar Pandey. The work in this thesis has not been submitted in any form for another degree or diploma at any university or another institute.

Sandeep Kumar

(PhD16010)

October 18, 2022

ABSTRACT

Disaster risk reduction is integral to social and economic development as per the 2030 Agenda of Sustainable Development. Among all the natural disasters, storms and floods have significant contributions in terms of their frequency, the number of affected people, and economic loss. In the tropical and subtropical regions of the world, tropical cyclones are the primary cause of storms and floods in the coastal areas. In the last 50 years, about 1942 disasters happened due to tropical cyclones that killed about 0.8 million people and caused a financial loss of $US\$1407.6$ billion. Any tropical cyclone-related prediction task is challenging as the atmospheric and oceanographic causal factors are multidimensional in nature and have a complex non-linear relationship among them. Tropical cyclone-related research mainly focuses on predicting a cyclone formation, track, intensity, storm surge, and associated rainfall. The existing operational models are primarily numerical and statistical in nature. The numerical methods are computationally involved and time-consuming. The statistical techniques are too simple to capture complex non-linear relationships between a large number of causal factors with spatial and temporal dimensions. Recently various deep learning studies appeared that successfully answer various tropical cyclone-related prediction problems.

This research work tries to answer various prediction problems related to a cyclone's different development phases. Starting from the formation (genesis) of a cyclone, the first work proposes a deep-learning model that detects the formation of a cyclone well advance in time in six ocean basins across the world. If a cyclone dies over the sea, one can simply ignore it, as it will not cause significant damage. But if it crosses the ocean and moves over to the land (known as Landfall), it causes a colossal disaster. Therefore, in our second work, we propose a deep learning classification model that answers the fundamental question of whether a cyclone will make a Landfall or not. The extent of disaster caused by a cyclone is determined by the location, intensity, and time of its Landfall. We propose a model that predicts the intensity, location, and time of Landfall for a cyclone across six ocean basins of the world.

In the first work, the proposed deep learning model forecasts the tropical cyclone

formation in six ocean basins of the world. It achieves a 5-fold accuracy in the range of $91.7\% - 97.7\%$, $96.4\% - 99.3\%$, $95.4\% - 99.1\%$, and $86.9.7\% - 92.9\%$ at lead times of 24h, 36h, 48h, and 60h respectively using only 12h of data. The second proposed model tries to answer the fundamental question whether a cylcone will make the landfall or not? It achieves an accuracy in the range of $96.4\% - 99.2\%$ and $93.0\% - 98.7\%$ using 12h or 24h of data (during initial 72h of cyclones progress) respectively across all six ocean basins of the world. The third work focuses on predicting the landfall's characteristics in the form of its location, time and intensity across six ocean basins of the world. The first model in this direction achieves a 5-fold cross-validation MAE of $4.24(\pm 0.40)$knots, $4.5(\pm 0.58)$h, and $51.7(\pm 1.20)$KM for landfall's intensity, time, and location, respectively, using any 21h of data during the course of a cyclone in north Indian ocean basin. Our model outperforms the landfall's location accuracy reported by IMD on its website. The second model achieved the 5-fold cross-validation MAE in the range of $66.18(\pm 2.87)$KM- $158.92(\pm 12.62)$KM and $4.71(\pm 0.54)$h-$8.20(\pm 2.96)$h for location and time prediction respectively across all six ocean basins of the world.

ACKNOWLEDGEMENTS

This research journey has not just fulfilled the requirements for a Ph.D. degree, but it taught me a few important lessons in life - patience, the value of consistency, and balancing life between various roles. It has changed my perspective on life and things. This journey would not have become a reality without a handful of people who supported me in all thick and thin.

I want to begin by thanking my Ph.D. supervisor Dr. Ashish Kumar Pandey, for providing all necessary guidance, support, and time. He diligently helped me in identifying research problems and executing them. I am also immensely indebted to him for funding my research and providing all required resources.

My Ph.D. journey would not have begun without the support provided by Dr. Somitra Kumar Sanadhya, who encouraged me to join IIITD. However, my academic association with him does not last long as he has to move out due to professional commitments.

Next, I would like to acknowledge the contributions of my labmate, Koushik Biswas, who was always with me personally and professionally. He was always my first point of contact whenever I needed any academic input. Also, I want to thank Amit Kumar Chauhan, with whom I started my research journey in cryptography.

Further, I would like to thank my parent organization, Shaheed Bhagat Singh College, which allowed me to pursue a Ph.D. along with my job. There I had many wonderful individuals who deserve mention. Dr. Binay Kumar Sharma and Dr. Kul Anand Sharma always stood by my side and helped fulfill commitments at my workplace.

Last but not the least, I would like to thank the almighty for blessing me with a beautiful family who always took care of me. I am greatly indebted to my parents, who were always with me, supported my studies, and helped me balance professional and family life. Special mention goes to my kids, Abhimanyu and Ruhani, and my wife Richa, who give me motivation for life.

PUBLICATIONS

Publications Related to the Dissertation

1. **Sandeep Kumar**, Koushik Biswas, and Ashish Kumar Pandey, Prediction of landfall intensity, location, and time of a tropical cyclone. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 35. 2021.

2. **Sandeep Kumar**, Koushik Biswas, and Ashish Kumar Pandey, Predicting Landfall's location and time of a tropical cyclone using reanalysis data. In Artificial Neural Networks and Machine Learning – ICANN 2021. Springer International Publishing, 2021. ISBN 978-3-030-86380-7.

3. **Sandeep Kumar**, Koushik Biswas, and Ashish Kumar Pandey, Track prediction of tropical cyclones using long short-term memory network. In 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC) 2021.

4. **Sandeep Kumar**, Koushik Biswas, and Ashish Kumar Pandey, Forecasting formation of a Tropical Cyclone Using Reanalysis , ICMLA 2022 (Accepted).

5. **Sandeep Kumar**, Koushik Biswas, and Ashish Kumar Pandey, Will a Tropical Cyclone make Landfall? Journal of Neural Computing and Applications, 2022 (Accepted).

Other Publications

1. Koushik Biswas, **Sandeep Kumar**, Shilpak Banerjee, and Ashish Kumar Pandey, Smooth Maximum Unit: Smooth Activation Function for Deep Networks using Smoothing Maximum Technique. In Conference on Computer Vision and Pattern Recognition (CVPR), 2022 (Accepted).

2. Koushik Biswas, **Sandeep Kumar**, Shilpak Banerjee, and Ashish Kumar Pandey, ErfAct and Pserf: Non-monotonic smooth trainable Activation Functions. In Proceedings of the AAAI Conference on Artificial Intelligence, 2022 (Accepted).

3. Amit Kumar Chauhan, **Sandeep Kumar**, Donghoon Chang, Somitra Kumar Sanadhya, Revocable identity-based encryption from codes with rank metric. In Topics in Cryptology –CT-RSA 2018. Springer International Publishing, Cham, 2018. ISBN 978-3-319-76953-0.

4. Koushik Biswas, **Sandeep Kumar**, Shilpak Banerjee, and Ashish Kumar Pandey, EIS - Efficient and trainable activation functions for better accuracy and performance. In Artificial Neural Networks and Machine Learning –ICANN 2021. Springer International Publishing, Cham, 2021. ISBN 978-3-030-86340-1.

5. Koushik Biswas, **Sandeep Kumar**, Shilpak Banerjee, and Ashish Kumar Pandey, TanhSoft—Dynamic Trainable Activation Functions for Faster Learning and Better Performance, in IEEE Access, vol. 9, pp. 120613-120623, 2021, doi: 10.1109/ACCESS.2021.3105355.

6. Koushik Biswas, **Sandeep Kumar**, Shilpak Banerjee, and Ashish Kumar Pandey, SAU: Smooth activation function using convolution with approximate identities (Communicated).

7. Koushik Biswas, **Sandeep Kumar**, and Ashish Kumar Pandey (2021). Intensity prediction of tropical cyclones using long short-term memory network. URL: https://arxiv.org/abs/2107.03187.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ANNs	Artificial Neural Networks
BTD	Best Track Data
CNN	Convolutional neural network
ECMWF	European Centre for Medium-Range Weather Forecasts
ECP	Estimated central pressure
EP	East Pacific
ERA5	ECMWF Reanalysis verion 5
IBTrACS	International Best Track Archive for Climate Stewardship
IMD	Indian Meteorological Department
LPS	Low Pressure System
LSTM	Long Short Term Memory Network
MSSWS	Maximum Sustained Surface Wind Speed
NA	North Atlantic
NI	North Indian
NOAA	National Oceanic and Atmospheric Administration
RMC	Regional Meteorological Center
RNN	Recurrent Neural Network
SDGs	Sustainable Development Goals
SI	South Indian

SP	South Pacific
SST	Sea Surface Temperature
TCs	Tropical Cyclones
WMO	World Meteorological Organization
WP	West Pacific

CHAPTER 1

INTRODUCTION

Out of all the disasters that occurred, natural disasters contribute around 77% in terms of economic losses in the 20 years duration 1998-2017 (Wallemacq *et al.*, 2018). During this duration, natural disasters caused a financial loss of $\$2,208$ billion which is up by 68% ($\$895$ billion) compared to damages $\$1,313$ billion reported during the duration 1978-1997. These amounts are underestimated in the sense that the majority of the reported disaster reports (63%) do not include economic loss data. As per the estimate of the World Bank, the actual annual cost of natural disasters is around $\$520$ billion[1] and puts about 26 million people below the poverty line annually. The percentage of climate-related disasters reporting for high income, upper-middle income, lower-middle income, and low-income countries are $52\%, 40\%, 30\%$, and 13% respectively. Whereas climate-related disasters amount to 0.41% of GDP (Gross Domestic Product) for high-income countries, it amounts to 1.77% of GDP for low-income countries, which is way higher than the IMF's threshold of 0.5% for a major economic disaster. The distribution of various natural disasters in terms of their numbers (count), economic losses, number of affected people, and number of deaths are shown in Figures 1.1, and 1.2. From these figures, one can observe that flood and storm are the most frequently occurring disasters with significant contributions in terms of the number of affected people and economic losses caused. In terms of the total number of human lives lost, floods and storms are second only after earthquakes.

During 1998-2017, out of the top 10 climate disasters in terms of economic losses, eight are storms, and two are floods. The hurricanes Katrina, Harvey, and Irma caused financial losses to the extent of $\$156.3, \$95.0$, and $\$80.8$ billion respectively in the USA alone. Recently, the storm Amphan (2021) caused an economic loss of around $\$15.6$ billion in India only[2]. In 2017 alone, 7.5 million people were displaced due to storms and 8.6 million due to floods. As per the Sendai Framework for Disaster Risk Reduction 2015-2030, reducing disaster losses is core to the effective implementation of the 2030

[1] https://www.worldbank.org/en/results/2017/12/01/climate-insurance
[2] https://en.wikipedia.org/wiki/North_Indian_Ocean_tropical_cyclone

Agenda for Sustainable Development Goals (SDGs). In the world's coastal regions, tropical cyclones (TCs) are the primary cause of floods and storms. Early warning systems are key to minimizing economic and human losses due to TCs. Working in this direction, in this dissertation, we try to answer TCs related various prediction problems using deep learning.

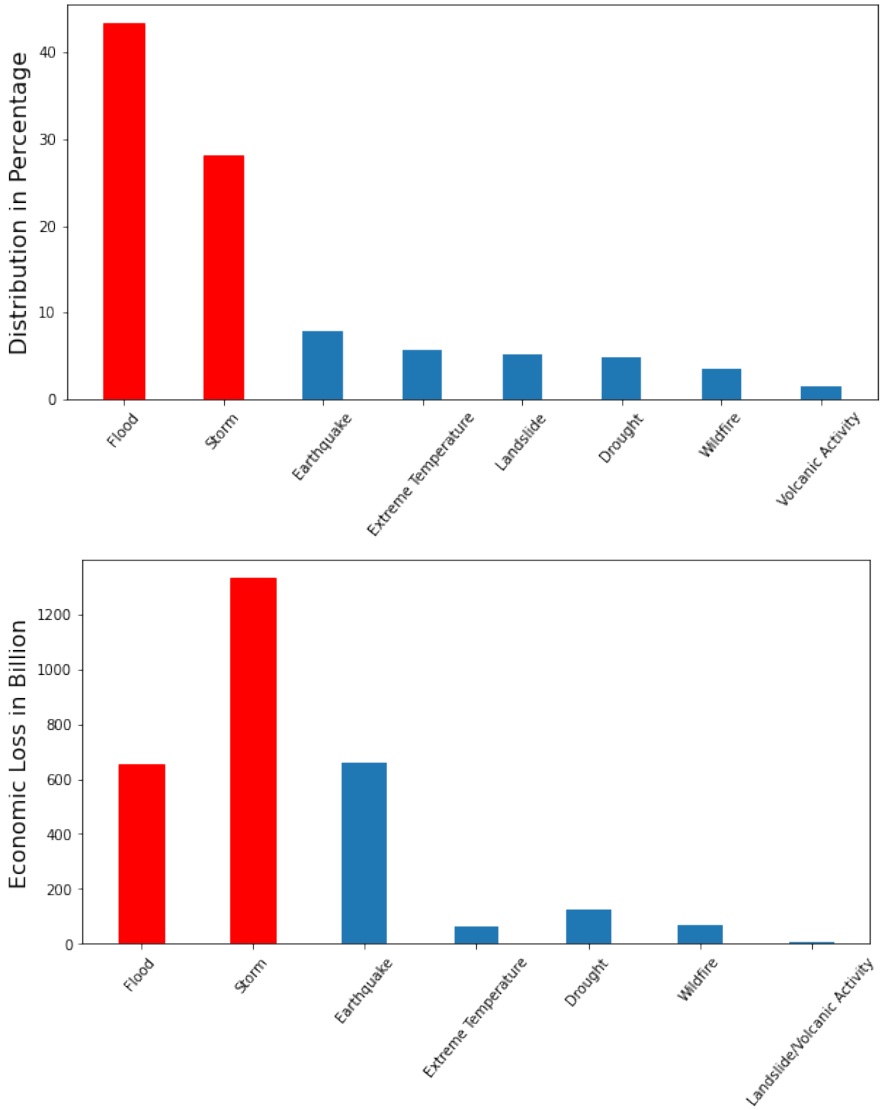


Figure 1.1: Distribution of natural disasters in terms of numbers and economic loss.

1.1 Tropical Cyclones

TCs are the most frequently occurring natural disasters in the tropical and subtropical regions. They are also known as hurricanes, willy-willies, or typhoons in different parts of the world. TCs brings with themselves thunderstorms, strong winds, flash floods,

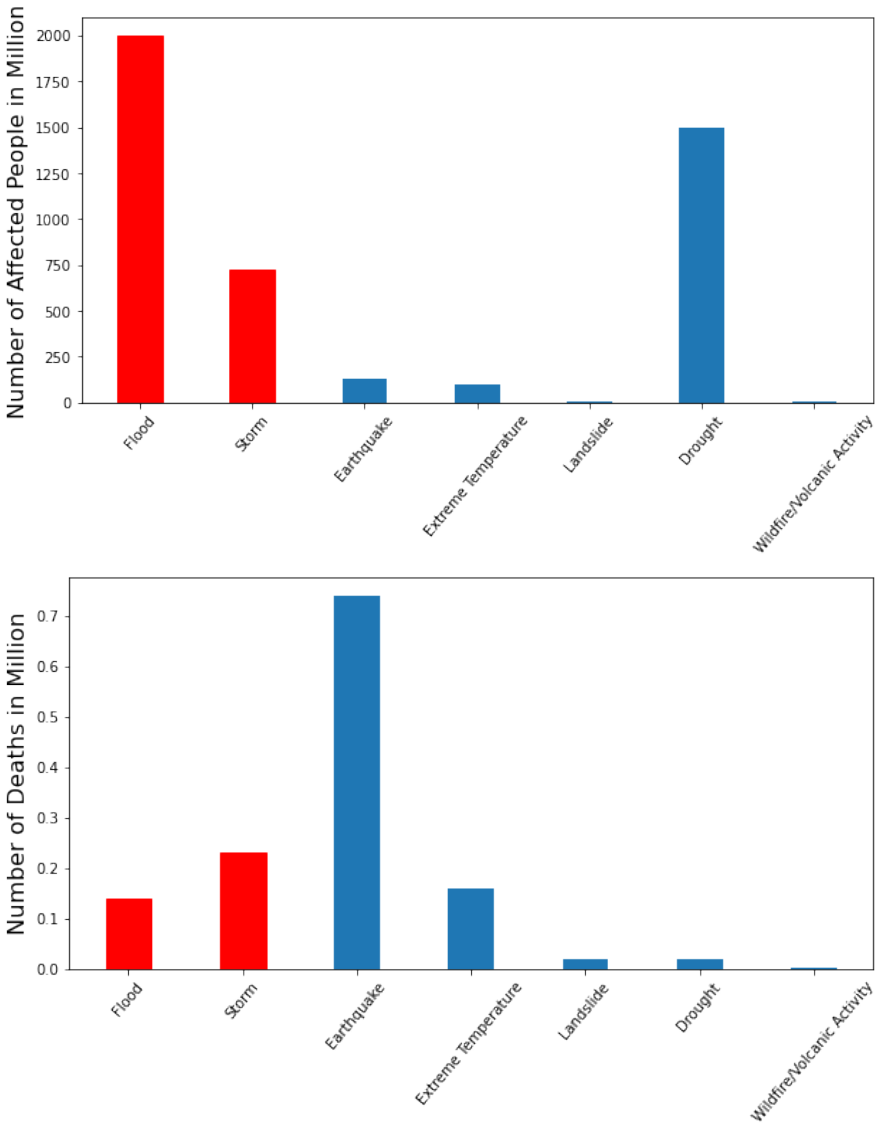


Figure 1.2: Distribution of natural disasters in terms of numbers of affected people and number of death.

lightening, and high tides, thereby causing massive human, economic, and environmental loss (Grinsted *et al.*, 2019; Webersik *et al.*, 2010). With a combined interaction with each other, these hazards further deteriorate to the extent of the disaster. In the last 50 years, tropical cyclones killed around 0.87 million people and caused a monetary loss of $1408 billion[3]. Moreover, there are increasing pieces of evidence that with climate change, the intensity of TCs and associated hazards are going to increase (Knutson *et al.*, 2019). Early information about the formation of a tropical cyclone, its track, and intensification can provide adequate time for disaster managers to take preventive measures to minimize losses. In this work, we have developed deep learning models that can

[3] https://public.wmo.int/en/our-mandate/focus-areas/natural-hazards-and-disaster-risk-reduction/tropical-cyclones

predict the formation (genesis), track, and intensity of a TC in real-time with adequate lead time. Like any other natural disaster, the causal factors that determine the formation of a TC and subsequent track and intensity have a complex non-linear relationship between them. The atmospheric, oceanographic, and geographic causal factors have temporal, spatial, and altitudinal dimensions, making answering any TC-related prediction problem challenging. To understand this, we first describe the formation process of a TC.

1.1.1 Formation of a Tropical Cyclone

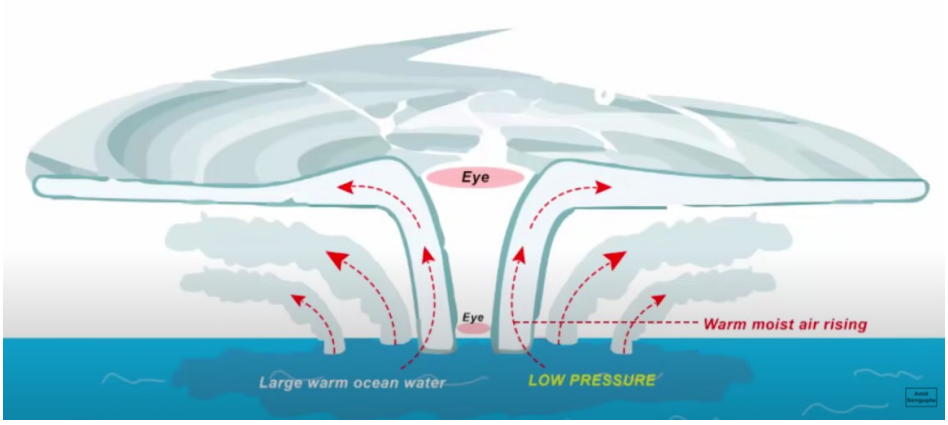


Figure 1.3: Pictorial representation of Tropical Cyclone formation.

Tropical cyclones are like engines fueled by warm ocean waters. Over the tropical and subtropical waters of the world (near the equator), the warmed moist air rises upward over the sea's surface and creates an area of low air pressure below, known as a Low-Pressure System (LPS). This causes the air from the adjoining area of LPS to move towards the LPS, which leads to further warming up of the air, which rises above. When the moist warm air rises above, it cools down and forms clouds. The complete system of cloud formation and wind rotation grows with the help of favorable conditions like- sea surface temperature (SST) greater than 26^oC (Palmen, 1948), low vertical wind shear, high relative humidity, and atmospheric instability (Palmen, 1948; ELSBERRY, 1995; Gray, 1968). The temperature difference between the warm core with rising moist air and the adjoining cool environment leads to rapidly rising buoyant air. As the whole system of wind rotation intensifies, an eye gets formed. The system is very calm, clear with low pressure in the eye. A pictorial representation[4] of the whole system is shown

[4] https://www.youtube.com/watch?v=W2UDbDXXYGE

Table 1.1: The Grade classification of the tropical cyclones by IMD.

MSSWS (knots)	Classification	Grade
<17	Low Pressure Area (LP)	0
17-27	Depression (D)	1
28-33	Deep Depression (DD)	2
34-47	Cyclonic Storm (CS)	3
48-63	Severe Cyclonic Storm (SCS)	4
64-89	Very Severe CS (VSCS)	5
90-119	Extremely Severe CS (ESCS)	6
≥ 120	Super Cyclonic Storm (SS)	7

in Figure 1.3. Due to Coriolis force, the wind rotates anti-clockwise in the northern hemisphere and clockwise in the southern hemisphere. All this makes the TC development process a complicated process that depends on oceanographic, atmospheric, and geographic factors. Moreover, out of these LPSs, only a small number developed in a full-fledged TC under the above favorable conditions (Emanuel, 1989). TCs are categorized in various grades depending on the Maximum Sustained Surface Wind Speed (MSSWS), which is described in Table 1.1, as adopted by Indian Meteorological Department[5] (IMD). Most regional meteorological centers use the definition of MSSWS as adopted by the World Meteorological Organization (WMO), which is the highest average wind speed for ten minutes (one minute in the case of the USA) at the height of 10 meters.

Landfall

Though most of LPSs die over the ocean, few of them intensifies with the help of favorable conditions and travel hundreds of kilometers to hit the sea coast. The event when a TC hits the coast is called landfall. The extent of natural disasters due to landfall is determined by the intensity, location, and time of the landfall. This makes the landfall event the most critical and important during the progress of a TC. Early detection of the landfall event for a TC will give adequate time to the administration and disaster managers to mobilize resources, issue an early warning, and vacate the coastal regions to minimize human, and material loss. **Whereas the existing literature related to TCs focus on predicting the track and intensity of a TC with a specific lead time, we in this work focus on predicting the characteristics (in terms of intensity, location,**

[5] https://rsmcnewdelhi.imd.gov.in/uploads/report/bestrack.pdf

and time) of landfall from the very initiation of TC.

1.2 Data

Each Regional Meteorological Center (RMC) across the world maintain the historical dataset of TCs that is known as Best Track Data (BTD). In this research work, we have used three publicly available dataset: IMD BTD[6], International Best Track Archive for Climate Stewardship (IBTrACS) (Knapp *et al.*, 2010), and ERA5 reanalysis data (Hersbach *et al.*, 2020). The BTD maintained by IMD contains the record of latitude, longitude, MSSWS, grade, estimated central pressure (ECP), and other variables from 1990 onwards. Now, we will briefly describe the two datasets, IBTrACS and ERA5.

IBTrACS

IBTrACS project maintained by World Meteorological Organization (WMO) is a comprehensive dataset of global TCs which collect and merge the TCs data from various RMCs and creates a unified, freely available BTD. The most recent version 4 contains the three hourly BTD across all ocean basins from 1980 onwards and is updated daily. It covers all the TCs originated in the region 70°N to 70°S and 180°W to 180°E. The dataset contains information about ocean basins, reporting regional agency, wind speed, pressure, grade, track (latitude and longitude), distance to land, speed, direction, and many other variables. The genesis location of all the TCs in six ocean basins, North Indian (NI), South Indian (SI), North Atlantic (NA), West Pacific (WP), East Pacific (EP), and South Pacific (SP), is shown in Figure 1.4.

The IBTrACS dataset is available in network Common Data Form (NetCDF) and comma-separated values (CSV) file format. For this work we have used the version 4 of IBTrACS dataset available in CSV file format[7].

[6] https://rsmcnewdelhi.imd.gov.in/report.php?internal_menu=MzM=

[7] https://www.ncei.noaa.gov/data/international-best-track-archive-for-climate-st
v04r00/access/csv/

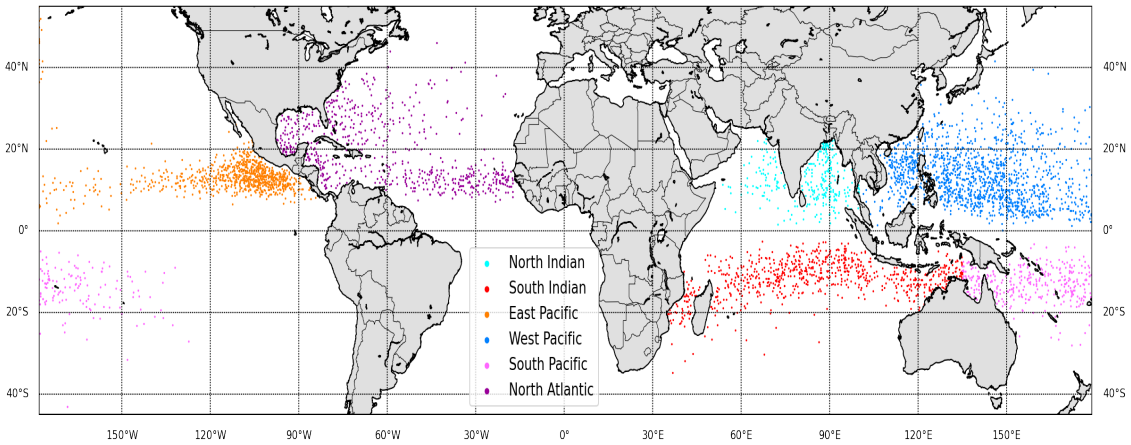


Figure 1.4: Cyclone genesis location in six ocean basins of the world from 1980 onwards.

Reanalysis Data: ERA5

The Copernicus Climate Change Service (C3S)[8] at European Centre for Medium-Range Weather Forecasts (ECMWF) maintain the ECMWF Reanalysis version 5 (ERA5) data. Reanalysis is a scientific way of producing globally complete and consistent data by gathering information from various resources and validating them using the laws of physics. Because of their completeness and correctness, reanalysis data is also known as 'maps without gaps'. Reanalysis data is widely used for understanding climate change, commercial applications, research, education, and policymaking. ERA5 has a record of global reanalysis data from January 1950 to the present. From 1979 onwards, the high-resolution data ERA5 replaces the ERA-interim (Dee *et al.*, 2011). It contains the hourly record of atmospheric, oceanic, climate, and land variables. It covers the whole earth at a resolution of 0.25° (approximately 30 kilometers (KM)) and various pressure levels (heights) from the ocean surface up to 80KM. In (Hodges *et al.*, 2017), the authors show how well a TC is represented in the reanalysis dataset.

The ERA5 dataset can be downloaded in two file formats: General Regularly-Distributed Information in Binary (GRIB) and NetCDF. GRIB is a concise data storage format that is widely used to store meteorological and forecast weather data. It is a standardized file format recommended by WMO. We have downloaded the required reanalysis dataset in GRIB file format using Climate Data Store Application Program Interface (CDS API)[9], which is further processed with eccodes library[10].

[8] https://cds.climate.copernicus.eu/
[9] https://cds.climate.copernicus.eu/api-how-to
[10] https://confluence.ecmwf.int/display/ECC/ecCodes+Home

1.3 Models

As the responsible atmospheric and oceanographic factors behind the development and further evolution of a TC have a complex non-linear relationship, we have used machine learning and deep learning models in our work. In this section, we will describe these models in brief.

Artificial Neural Networks

Artificial Neural Networks (ANNs) (McCulloch and Pitts, 1943) are inspired from the connected neurons in the human mind, where the signals are processed at a neuron and passed on to the connected neurons. Same way, ANNs consist of neurons called nodes, which are partitioned in a sequence of layers. There are three types of layers: input, hidden, and output layers. The intermediate layers between the input and output layers are called hidden layers. Generally, the nodes in one layer are connected with the nodes in preceding and succeeding layers. The incoming information at a node is processed and further sent to the connected nodes. Weights are assigned to the connections (edges) between connected nodes. The flow of information between nodes is determined by the composition of a non-linear function and a weighted linear sum of the incoming input. Generally, the non-linear function in the composition is fixed at a node and called an activation function. The weights assigned to the connecting edges are updated in a way to minimize the suitably chosen loss function, which is done through the Gradient Descent algorithm (Kiefer and Wolfowitz, 1952) by back-propagating the gradients and updating the weights on the way.

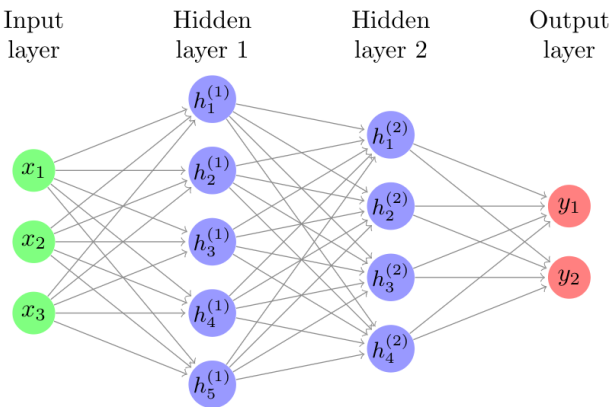


Figure 1.5: A simple fully connected ANN.

A fully connected ANN with two hidden layers is shown in Figure 1.5. Mathematically, the information flow in an ANN can be represented as $h_j^{(i)} = \sigma(W_j^{(i)} X)$, where X is the input matrix containing bias, $h_j^{(i)}$ represents the information at the j^{th} neuron in the i^{th} layer, W_j^i is weight matrix that determines interaction between nodes at $(i-1)^{th}$ and i^{th} layer, and σ is the activation function.

ANN model has successfully been able to capture the complex non-linear relationship between input and output (Ali *et al.*, 2012; Abhishek *et al.*, 2012; Kovordányi and Roy, 2009; Chaudhuri *et al.*, 2015), but it is not the best choice for time-series data as the model is not designed to learn from sequential data. In what follows, we will discuss RNN, which has the ability of information persistence.

Recurrent Neural Network

Recurrent Neural Networks (RNNs) are ANNs with internal loops that help in information persistence (Jordan, 1990; Rumelhart and McClelland, 1987; Cleeremans *et al.*, 1989; Pearlmutter, 1989). RNN can take a data sequence as input and produce a sequence of data as output. The output and intermediate state is determined by weights applied on inputs and a hidden state vector that represents the learned information based on earlier inputs and outputs. A chain-like structure can represent RNN, as shown in Figure 1.6. The upper, middle, and lower chains represent the sequence of outputs, hidden state vector, and sequence of inputs respectively. Mathematically, a simple RNN structure can be represented as:

$$\begin{aligned} h_t &= \sigma(W_h x_t + U_h h_{t-1} + b_h) \\ y_t &= \sigma(W_y h_t + b_y) \end{aligned}$$

where x_t is the input vector at timestamp t, σ is the activation function, h_t and y_t are the hidden state vector and output vector respectively at timestamp t, W_h, W_y and U_h are weight matrices, and b_h, b_y are biases.

One of the drawbacks of RNN is that it fails to remember long history from time series or sequential data (Bengio *et al.*, 1994) because of exploding and vanishing gradient problem in which the gradient vector can increase or decrease exponentially during the training process.

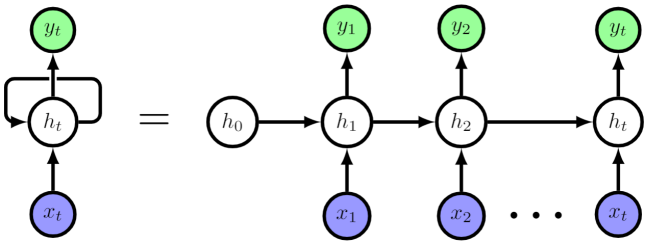


Figure 1.6: Graphical representation of Recurrent Neural Network.

Long Short Term Memory Networks

Long Short Term Memory (LSTM) network overcome the issue of vanishing and gradient problems of RNN. LSTM was originally introduced in 1997 by Hochreiter and Schmidhuber (Hochreiter and Schmidhuber, 1997) and later few improved versions of LSTM were proposed in (Gers *et al.*, 1999; Gers *et al.*, 2003; Gers and Schmidhuber, 2001). LSTM is defined in such a way that it remembers long-term dependencies. It uses three inner cell gates and maintains a memory cell to handle long-term dependencies. LSTM cells can selectively read, selectively write and selectively forget. A general LSTM cell is mainly consists of four gates- an input gate to process newly coming data, a memory cell input gate to process the output of the previous LSTM cell, a forget gate to decide what to be forget and decides the optimal time lag to remember previous states, and an output gate to process all the newly calculated information and generate output. The mathematical representation for the three LSTM gates are:

$$
\begin{aligned}
f_t =& \sigma(w_f.x_t + u_f.h_{t-1} + b_f) \\
i_t =& \sigma(w_i.x_t + u_i.h_{t-1} + b_i) \\
o_t =& \sigma(w_o.x_t + u_o.h_{t-1} + b_o)
\end{aligned}
$$

and the mathematical representation for the cell state, candidate cell state, and output are

$$
\begin{aligned}
\tilde{c_t} =& \sigma(w_c.x_t + u_c.h_{t-1} + b_c) \\
c_t =& f_t.c_{t-1} + i_t.c_t \\
h_t =& o_t.\sigma(c_t)
\end{aligned}
$$

where σ is the activation function, f_t represents forget gate, i_t represents input gate, o_t represents output gate, h_{t-1} is output from previous LSTM block at $(t-1)$ timestamp, x_t is input at current timestamp, $\tilde{c}_t$ is candidate for cell state at timestamp t, c_t represents cell state at timestamp t, and $w_f, w_i, w_c, w_o, u_f, u_i, u_c, u_o$ are weight matrices, b_f, b_i, b_c, b_o are bias vectors, see Figure 1.7.

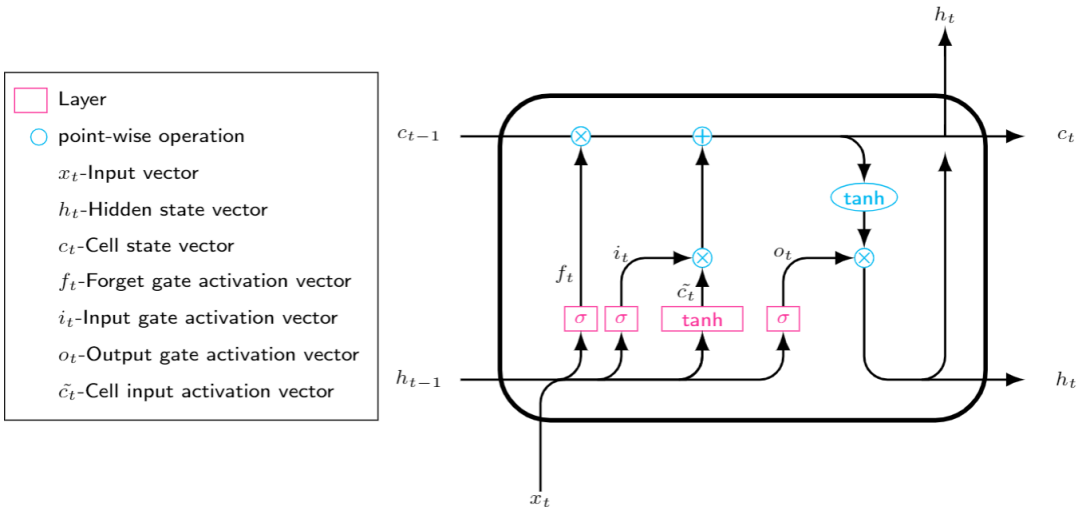


Figure 1.7: Structure of an LSTM Cell.

Stacked LSTM or Deep LSTM (Graves *et al.*, 2013) networks are an updated version of LSTM model that contain various hidden layers and a layer is placed on top of the other layer. Each layer consists of LSTM cells. A LSTM layer provides a sequence output in place of a single output to the below LSTM layer. This structure helps in better learning in sequential and time series data.

Bidirectional LSTM

A Bidirectional LSTM (BiLSTM) (Schuster and Paliwal, 1997) is designed to learn in both forward and backward direction (from past to future and future to past) in comparison of LSTM which learns only in forward direction. It has two different LSTM layers in opposite directions of each other. One layer processes the input in the forward direction and another layer in the backward direction. These layers are connected to the same output layer and simultaneously collect information from the past and future.

Convolutional Neural Network

Convolutional Neural Networks (CNNs) are inspired by the visual cortex system of the eyes where a neuron gathers and processes the information from its neighborhood neuron (LeCun *et al.*, 1989; Krizhevsky *et al.*, 2012; Milletari *et al.*, 2016). They have been successfully utilized in deep learning problems involving image or image-like datasets. A CNN can learn its weights and biases to detect various aspects of an image to differentiate it from the other. CNNs require less pre-processing in comparison to other techniques. For a primitive method, we need to handcraft the filter, whereas a CNN can learn these filters with training. A CNN has multiple layers, generally consisting of convolution, pooling, batch-normalization, Dropout, and fully connected layers that are briefly described below.

Convolution Layer: The convolution operation extracts the high-level abstract features from the input image such as edges, color, orientation, etc. This layer shares the weights across the input and learns the high-level features with small numbers of weights. The weights in the convolution filter are learned during training. A learnable convolution filter has a specific window size and depth equivalent to the input layer's depth. These filters scan through the input and produce an output feature map with depth equals the number of filters. The CNN has multiple such convolution layers that produce high-level abstract output features.

Pooling Layer: Pooling layers help in reducing the dimension of the input maps. It is a down-sampling technique that avoids over-fitting (Krizhevsky *et al.*, 2012). The statistics of a group (in the form of min, max, or average) of features is calculated in place of actual original values to reduce the computational complexity. The pooling layer produces down-sampled maps that are a summarized input version. A slight difference in the feature location in the input detected by the convolution layer will be captured by pooled feature map. This ability of the pooling layer is known as model's in-variance to local translation.

Batch-Normalization Layer: Under normalization, the input data is pre-processed by re-scaling and re-centering it without distorting its shape. It helps model in generalizing input data with faster and stable learning. In batch-normalization (Ioffe and Szegedy, 2015), the normalization is applied on a complete batch of input data.

Dropout: Dropout (Srivastava *et al.*, 2014) is a process under which certain connections in a neural network has been dropped (ignored) while training (backward and forward pass) with a specific probability. Its a popular technique to avoid over-fitting. It makes the network more robust by preventing the network layers to co-adapt with previous layers while minimizing loss.

Fully connected Layer: A fully connected layer is just an ANN, which takes the output from the last final convolution layer or the pooling layer as input after flattening. These form the last few layers of a CNN.

1.4 Research Contributions

Considering the extent of human and economic losses caused by TCs and their adverse effects on people and the environment in coastal regions, it is important to have models that can forecast the different stages of a TC in real-time. If we look at the different phases of TC development, the very first forecast problem is detecting TC formation at a given time and location. An earlier detection of possible TC formation will help marines navigating ships over the high sea. As mentioned earlier, most LPSs die over the sea, but few of these, accompanied by favorable conditions, travel hundreds of kilometers, make landfall, and cause a massive disaster at the landfall location. So, the next significant prediction problem is to know whether a TC will make landfall or not. Early knowledge of possible landfall event will help disaster managers in resource optimization. If they know that the TC will make landfall well advance in time, they will get adequate time for resource mobilization and shifting of people from coastal regions. On the other hand, if they know that a TC will not make landfall, they can ignore it. This will have economic benefits as they do not need to mobilize any resources, and it will not cause unnecessary inconvenience to people residing in coastal areas. Once we know that a TC will make landfall, the next obvious question is the location of landfall, its time, and the intensity at which it is going to hit the coastal region. If the administration knows the exact landfall region, they can focus on a particular area for disaster preparedness. Similarly, knowledge of the precise time of landfall will help in issuing a timely warning to fishermen and coastal population. Knowledge of the intensity of a TC at landfall will help decide the level of disaster preparedness and amount of resource mobilization.

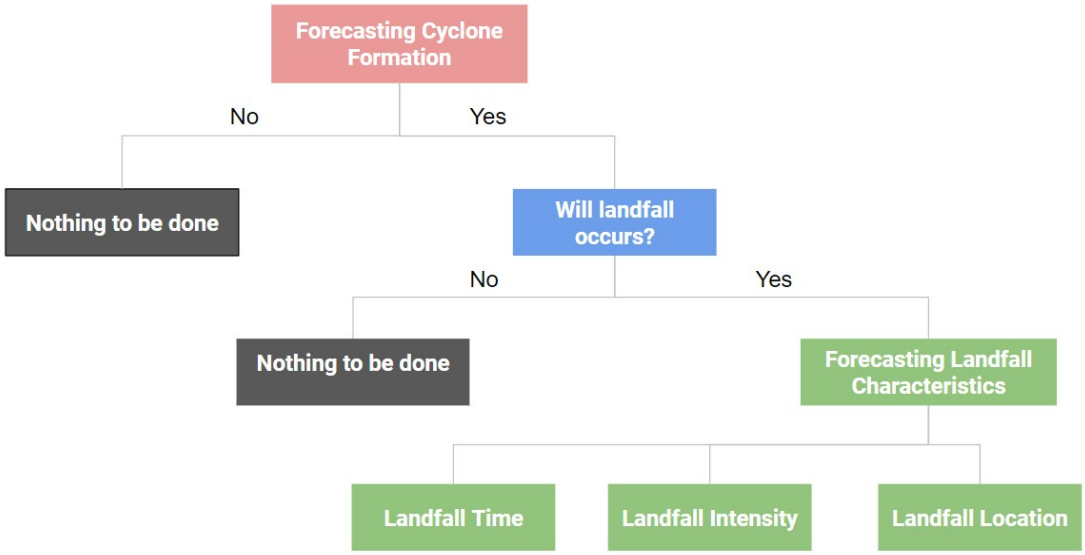


Figure 1.8: Work Summary.

We tried to answer the above TC-related forecast questions in this research work. A summary of the research work is shown in Figure 1.8. The research contributions are as follows:

- **TC Formation Forecast:** Forecasting TC formation is always a challenging problem as the casual factors have complex non-linear relationships and have spatial and temporal dimensions. Also, the theory behind the TC formation process is still not settled, and the research on this problem is still evolving (Halperin *et al.*, 2013). In this work, we have proposed a deep learning classification model using LSTM and CNN to forecast the TC formation (genesis) with high accuracy. An algorithm has been proposed to generate non-TC formation samples. We have identified five variables at three pressure levels from the reanalysis dataset ERA5 that are successfully utilized to detect TC formation in six oceans basins EP, NI, NA, SI, SP, and WP of the world.

- **Landfall Event Forecast:** Predicting whether a landfall will happen or not for a particular TC at an early stage of its development is a challenging problem (Leroux *et al.*, 2018) because the TC dynamics changes rapidly with time. As per our knowledge, there is no other work in the literature that attempts to forecast the landfall event. We identify three variables at three pressure levels from the reanalysis dataset ERA5 and five variables from the IBTrACS dataset that successfully

captures the landfall event in the early phase (first 72 hours of inception) of a TC for the above mentioned six ocean basins. The proposed deep learning classification model consisting of LSTM and CNN has two branches - one branch processes the reanalysis data, and the other processes the IBTrACS data.

- **Landfall's Intensity, Location and Time Forecast:** Forecasting whether a TC will make landfall or not is not sufficient; one needs to know the landfall's intensity, location, and time to have a piece of actionable information. Predicting TC behavior near the landfall is a challenging task as the TC behavior changes abruptly near its landfall (Leroux *et al.*, 2018). Along with earlier mentioned causal factors, the terrain factor also plays a critical role near the landfall. We have proposed two research works in this direction. The first one uses a combination of CNN and LSTM to predict landfall's location and time in six ocean basins of the world. The model uses three variables at three pressure levels and sea surface temperature (SST) from the reanalysis dataset. In the second work, we used an LSTM network that predicts the landfall's intensity, location, and time in the NI ocean basin using BTD data provided by IMD. The model uses five variables along with two derived features from the BTD data.

- **Track Prediction:** Track prediction is another important prediction problem. Most of the track-related work in the literature focus on predicting a TC's track in terms of latitude/longitude with a specific lead time. Recently, in (Alemany *et al.*, 2018) a new approach for track prediction is proposed, where instead of predicting the track of a TC directly in terms of latitude and longitude, the authors predict a GridID, which is obtained by applying a grid function on the ranges of latitudes and longitudes. In this work, we adopted the same approach and predicted the GridID for the Atlantic and NI ocean basins. The proposed model outperforms the existing state of artwork in terms of mean absolute error, time-span of prediction, and training time.

1.5 Thesis Structure

We present various forecast problems in the order they appear in a TC's development process. In Chapter 2, we describe the TC formation (genesis) forecast work. In Chapter 3, the landfall event classification work is presented. Next, we introduce the landfall's intensity, time, and location prediction work in Chapter 4. In Chapter 5, we present the track prediction work in terms of a GridID. Finally, in Chapter 6, we give the conclusion along with future directions.

CHAPTER 2

Forecasting Formation of a Tropical Cyclone[1]

As discussed in Section 1.1, TC formation (also known as genesis of a TC) process is a complicated and dynamic phenomenon. The causal factors have not just temporal and spatial dimensions but also altitudinal dimensions. Despite several years of research, complete understanding of the TC formation is still elusive (Halperin *et al.*, 2013). The tropical depression formation and subsequent intensification mechanism is mainly governed by the convective instability of the second kind (CISK; Charney and Eliassen (1964)) and wind induced surface heat exchange (WISHE; Halperin *et al.* (2013)). These two proposed mechanisms remain questionable even today (Craig and Gray, 1996). As the theories behind the TC formation are still evolving, prediction of TC formation remains a challenging task. In this work, we have proposed a deep learning model that can forecast the formation of Tropical cyclone (first time it developed as a LPS) using reanalysis data. The model takes as less as 12h of data and can forecast TC formation with a lead time up to 60h. The research contributions of this work are as follows:

- Five atmospheric variables from reanalysis dataset are identified that can be used to forecast the TC formation process in six ocean basins of the world.
- We proposed an algorithm to generate negative class samples (non-TC cases) that differentiate from positive class samples (TC cases) in terms of at least 5 days time window or in terms of spatial distance of at least 5°.
- A single deep learning model using LSTM and CNN is proposed that can forecast TC formation with high accuracy across all ocean basins.

[1]This chapter is a slightly modified version of work accepted for publication in conference IEEE 2022 International Conference on Machine Learning and Applications and has been reproduced here with the permission of the copyright holder.

2.1 Related Work

There are mainly two approaches for detecting TCs; one is a model-driven approach based on equations governing the physical phenomenon of TC development (including numerical simulations). The other is a data-driven approach that utilizes historical data relating to TCs (including machine learning methods). The earliest conventional way to detect a TC is early-stage Dvorak analysis (EDA), an extended Dvorak technique (Dvorak, 1975, 1984), which utilizes satellite cloud images. However, this technique includes subjective interpretation of parameters and hence not sufficiently scientific (Dvorak, 1984). EDA is used by the National Hurricane Center (NHC), Central Pacific Hurricane Center (CPHC), and the Japan Meteorological Agency (JMA) to forecast typhoon initiation, up to 48 hours before its formation with an accuracy of up to 57% (Cossuth *et al.*, 2013). In (Yamaguchi and Koide, 2017), authors show that the global ensembles models [European Centre for Medium-Range Weather Forecasts (ECMWF), Japan Meteorological Agency (JMA), National Centers for Environmental Prediction (NCEP), and Met Office in the United Kingdom (UKMO)] along with EDA can be used to improve the accuracy up to 79%. In Halperin *et al.* (2013), authors evaluated the performance of five global Numerical Weather Prediction (NWP) systems [Global Forecast System (GFS), ECMWF, Canada's Global Environment Multiscale Model (CMC), UKMO, and Navy Operational Global Atmospheric Prediction System (NOGAPS)] for TC forecast in the North Atlantic (NA) ocean for the period 2004-2011, and shows the best Precision of around $70\%, 60\%, 50\%$ can be achieved for lead hours 24, 36 and 48 respectively. Over the years, the accuracy of numerical models is improved based on better initialization and improved computing power. But still, these models are not suitable for long lead time forecasts as the numerical methods are prone to error accumulation over iterations.

As the physical mechanisms and causal factors behind the formation of a TC are still not settled, the machine learning techniques that discover information from data and unconstrained by the rule of physics are more suitable. The ability of machine learning techniques to capture the complex non-linear relationship between causal factors makes them a better choice. With the growth in historical data like IBTrACS, satellite data, and reanalysis data in recent times, various machine learning works have appeared that deals with the forecast of TC formation. In Schumacher *et al.* (2009);

Hennon and Hobgood (2003) authors have used statistical method, Linear Discriminant Analysis (LDA), to predict 24h probability of TC formation from derived large scale environmental parameters. In (Wijnands *et al.*, 2016) authors identified the four most essential variables out of 500 variables from the ERA-Interim dataset. They proposed a logistic regression model to forecast TC genesis for the whole globe. It achieves a Precision and Recall of 68.5% and 38.6% respectively for 24 hour lead time. The factors that lead to a TC formation have a complex non-linear relationship, making these linear methods unsuitable for the TC formation prediction task. Recently, machine learning methods and deep learning methods have been successfully applied to answer the TC formation forecast problem.

In (Zhang *et al.*, 2015) Decision trees (DTs) are used to detect developing and non-developing tropical disturbances in the western North Pacific Ocean for the months, June to September of 2004-2013, using five derived parameters from Navy Operational Global Atmospheric Prediction System (NOGAPS). They proposed a genesis potential index (GPI) to represent the potential of a TC formation based on Gray's Index (Gray, 1979). They reported accuracy of 84.6% for a lead time of 24h. The developing and non-developing disturbances are differentiated based on relative vorticity. In Kim *et al.* (2019*a*), DTs, random forest (RF), and support vector machine (SVM) are used to detect the formation of TC in the western North Pacific Ocean for the period 2005-2009 using eight derived predictors from WindSat satellite data. They classify a tropical depression as TC when the MSSWS reaches $13m/s$ (or 25 knots) and the satellite image available with at least 60% coverage in a circle with 4° radius around the center of the tropical disturbance. In (Zhang *et al.*, 2019), the authors use 13 predictors derived from mesoscale convective system (MCS) data and ERA-Interim dataset to predict TC formation using the following machine learning tools - Logistic Regression (LR), Naïve Bayes (NB), DT, K-Nearest neighbors (KNN), Multilayer perceptron (MLP), Quadratic Discriminant Analysis (QDA), SVM, AdaBoost (ADA), and RF. The authors reported the accuracy in terms of F1-score, precision, and recall for lead times 6h, 12h, 24h, and 48h for global (consisting of all ocean basins), NA, and west north pacific (WNP) ocean basins. They took TC that originates in the TC seasons (May to November) into account. The definition of TC genesis coincides with its first appearance in the IBTrACS dataset.

The machine learning techniques are not best suited to deal with satellite and re-

analysis type datasets with both spatial and temporal dimensions. Various deep learning studies have successfully captured the spatial and temporal dimensions of causal factors to answer the prediction problems related to TC's track, intensity (Alemany *et al.*, 2018; Gao *et al.*, 2018; Kumar *et al.*, 2021*c*; Chen *et al.*, 2019*a*; Maskey *et al.*, 2020) and its landfall's characteristics (Kumar *et al.*, 2021*b*,*a*). Recently, a few deep learning studies have been proposed that forecast TC formation. In (Matsuoka *et al.*, 2018), CNN has been used to detect the TC and its precursors in the six ocean basins of the world using 30 years of simulated outgoing longwave radiation (OLR) data generated through a cloud-resolving global non-hydrostatic atmospheric model. The TC and its precursors are categorized as one class and identified based on TC tracking algorithm (Sugi *et al.*, 2002; Yamada *et al.*, 2017) which takes temperature, horizontal components of wind, and sea level pressure (SLP) as inputs. The authors reported the Precision in the range of $79.9\% - 89.1\%$ and FAR in the range of $32.8\% - 53.4\%$ in north Pacific ocean for the period July to November. In (Shakya *et al.*, 2020), the authors have presented a deep learning model to detect an ongoing TC with the help of satellite data of eight TCs in the NI ocean basin. In (Chen *et al.*, 2019*b*) a hybrid CNN-LSTM model is used to predict if an ongoing TC will be intensified to the level of a typhoon (wind speed greater than 64 knots) with a lead time of 24h, using preceding data of 6h, 12h, 18h or 24h. The IBTrACS data and ERA-Interim datasets are used for three ocean basins WP, EP, and NA. Thus we see that very few deep learning studies exists, and each has their own criterion of identifying TC formation or restricted to a particular region or period.

Another line of research is the seasonal prediction of tropical cyclones. In (Richman and Leslie, 2012) the authors proposed the Support Vector Regression (SVR) to study the seasonal prediction of TCs in the north-west portion of Australia. In (Richman *et al.*, 2017) authors use SVR for seasonal prediction of TCs count in the NA ocean. In (Nath *et al.*, 2015) authors presented an ANN model to predict the seasonal number of TCs in the NI ocean basin for the post-monsoon (October, November, December) season using reanalysis data.

2.2 Data

Inspired by the successful usage of reanalysis dataset in the recent works (Chen *et al.*, 2019*b*; Kumar *et al.*, 2021*a*; Boussioux *et al.*, 2020; Giffard-Roisin *et al.*, 2020) to answer TC related track, intensity, and landfall's characteristics problems, we have used ERA5 (Hersbach *et al.*, 2020) high-resolution reanalysis dataset provided by ECMWF[2]. It is hourly weather and climate data for the whole globe. The formation process of a TC is determined by large-scale atmospheric factors at various altitudes around the center of an LPS. For this study we have extracted wind fields u, v, geopotential z, relative humidity r, and temperature t at three pressure levels (altitudes) 225hPa, 500hPa, and 700hPa. These variables largely determine the development process of a TC as follows: u and v fields represent the east-west and north-south movement of air along with its speed, z represents the gravitational potential energy relative to sea level, r represents the water vapor pressure, and t represents the atmospheric temperature. As the atmospheric causal factors behind TC formation may have horizontal extends up to 1000 kilometers (KM), these variables are extracted for a spatial region $10^\circ \times 10^\circ$ with a resolution of 0.25°, centered as TC genesis location, which resulted in a grid of 41×41. As one degree is around 110 km near the equator and decreases as we move pole-wards, this resulted in a spatial extend of around 1000 km with a resolution of around 25 km. A pictorial representation of used reanalysis data is shown in Figure 2.1.

Table 2.1: Number of TC in each ocean basin.

Ocean Basin	**NA**	**NI**	**SI**	**WP**	**SP**	**EP**
Number of TCs	653	360	832	1431	509	983

The IBTrACS dataset (Knapp *et al.*, 2010) maintained by National Oceanic and Atmospheric Administration[3] keeps three hourly global records of all TCs in the form of its time, track (latitude and longitude), intensity, and many more other variables from the very initiation of a TC when it was first detected as a tropical depression or LPS. As the definition of TC genesis time is ambiguous (Horn *et al.*, 2014) we take the time when a TC is recorded first as LPS in IBTrACS as TC formation time. We extracted the record of TC formation time, and corresponding location (latitude, longitude) for all TCs for the earlier mentioned six ocean basins of the world from 1980 to September 2021. All

[2] https://cds.climate.copernicus.eu/
[3] https://www.ncdc.noaa.gov/ibtracs/

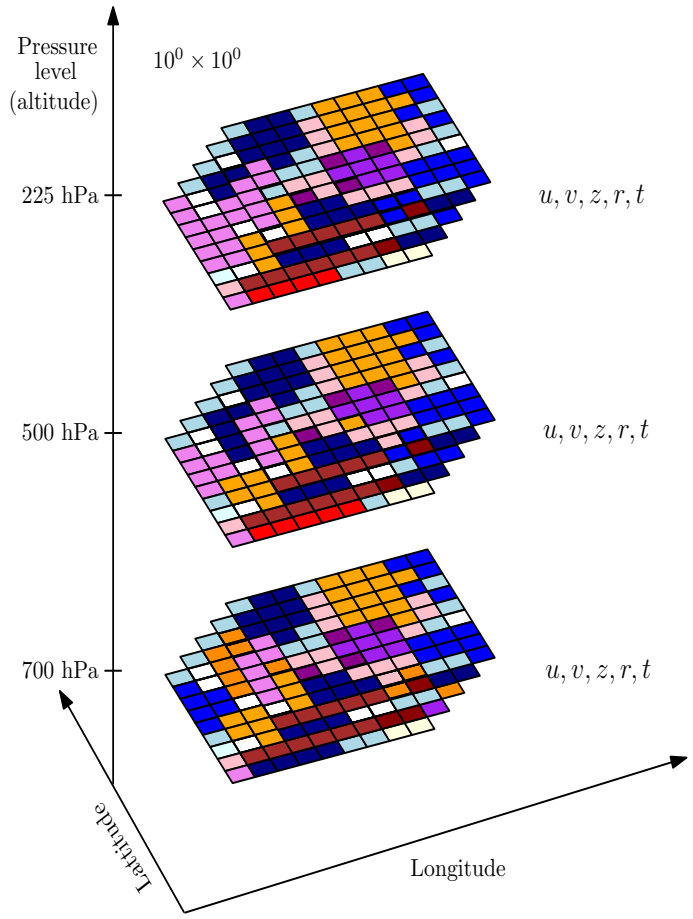


Figure 2.1: Pictorial representation of reanalysis data.

these TCs form the positive class in our classification problem. The total number of TCs (positive cases) are shown in Table 2.1. To generate the negative classes (non-TC formation), for a particular ocean basin, we followed the Algorithm 1. This way, we have an equal number of positive and negative classes in our dataset. A negative class sample represents a time t and location *loc* such that there is no existing TC formation in a time window of 5 days. If t lies within a window of 5 days, then there is no existing TC formation in a spatial window of 5°. This way, we have selected all TC formation samples and non-TC formation samples. Next, for each sample, we downloaded the above-described reanalysis data for grid size 41×41 centered at the location of each sample, for time points $t - 6k$, $12 \geq k \geq 4$, where t is the time of TC formation or randomly selected time of a non-TC formation sample. Thus the reanalysis dataset is extracted for nine time points from $t - 72$h to $t - 24$h at an interval of 6h. The genesis location of all samples (TC formation and non-TC formation) are shown in Figure 2.2 for all six ocean basins.

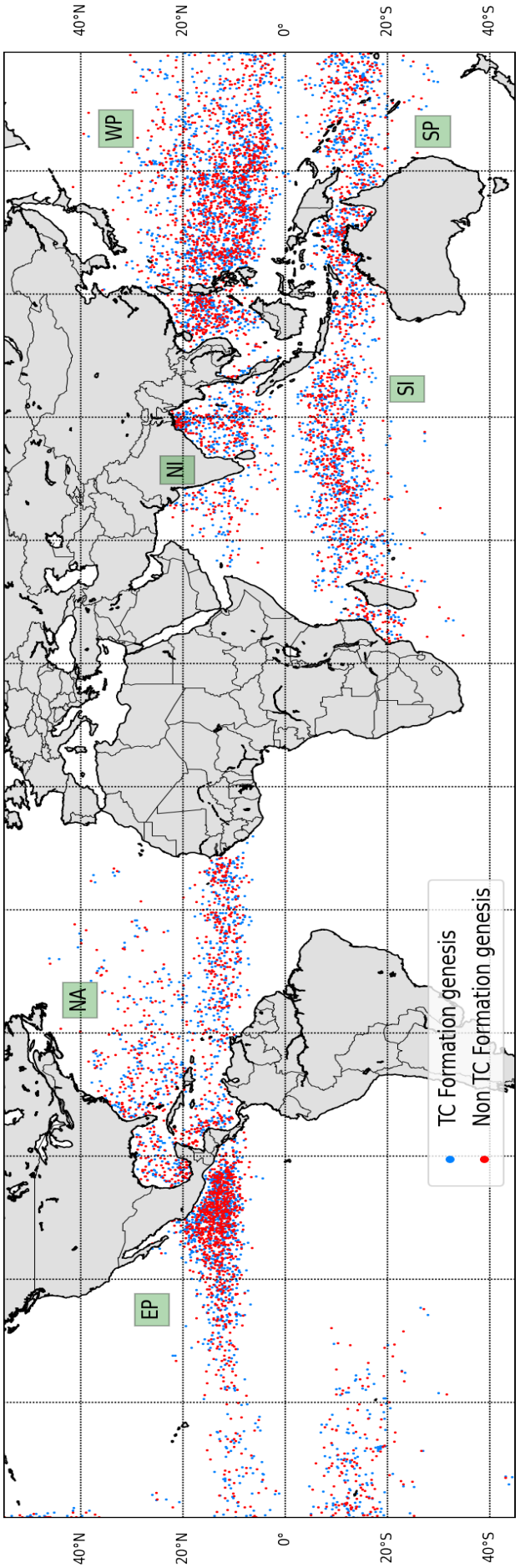


Figure 2.2: Location of TC and non TC formation.

Algorithm 1 Generating negative classes (non TC) in a ocean basin, say *OB*.

Input: Set T and L of time points and location (lat, lon) of all TC formation in ocean basin *OB*.

Parameter: *count* (No. of positive class in *OB*), A Kernel Density Function say *LocGen* fitted on the set L.

Output: Set $T1$ and $L1$ of time points and location *(lat, lon)* of all non TC formation (negative classes) in ocean basin *OB*.

Let $count1 = 0$, $T1 = \{\}$, $L1 = \{\}$.
while $count1 < count$ **do**
  Generate a random time $t1$ (between 01/01/1980 and 01/09/2021) and random location say *(lat1,lon1)* through *LocGen*.
  if *(lat1,lon1)* lies over land **then**
    continue
  end if
  if $Abs(t - t1) >$ 5days $\forall t \in T$ **then**
    Add $t1$ to $T1$ and *(lat1,lon1)* to $L1$, $count1++$.
  else
    if $Abs(lat - lat1) > 5^\circ$ and $Abs(lon - lon1) > 5^\circ$ $\forall (lat, lon) \in L$. **then**
      Add $t1$ to $T1$ and *(lat1, lon1)* to $L1$, $count1++$.
    end if
  end if
end while
return $T1$ and $L1$.

Dataset Preparation

For a TC, 9 data points are available at an interval of 6h as described above. The i^{th} data point ($1 \leq i \leq 9$) is corresponds to the lead time $(72 - 6(i-1))$h. Suppose we want to use T number of data points ((T-1)*6 hours of data) to predict the formation of a TC. For this, we generate $10 - T$ data points, where a single data point is a sequence of T vectors of the form:

$$(\,\mathrm{u225}(t), \mathrm{v225}(t)\,\mathrm{z225}(t), \mathrm{r225}(t), \mathrm{t225}(t), \mathrm{u500}(t), \mathrm{v500}(t)\,\mathrm{z500}(t), \mathrm{r500}(t), \mathrm{t500}(t),$$
$$\mathrm{u700}(t), \mathrm{v700}(t)\,\mathrm{z700}(t), \mathrm{r700}(t), \mathrm{t700}(t))$$

where $k \leq t \leq T+k-1$ and k varies from 1 to $10-T$. For each such training point, the target variable is 1 (positive class) or 0 (negative class). One must note that the above process forms $10 - T$ data points at leads hours $6k$ where $4 \leq k \leq 9 + 4 - T$. All such data points for all the TCs form the dataset.

2.3 Model and its Implementation

As our dataset has both spatial and temporal dimensions, the model utilizes a combination of CNN (LeCun *et al.*, 1989; Krizhevsky *et al.*, 2012; Milletari *et al.*, 2016) and LSTM (Hochreiter and Schmidhuber, 1997; Gers *et al.*, 1999; Gers *et al.*, 2003; Gers and Schmidhuber, 2001) networks to effectively capture the causal factors behind a TC formation. An input training data point is of the dimension $(T, 15, 41, 41)$, where T stands for the length of sequential data points (of $6 * (T - 1)$h), 15 stands for the number of channels (corresponding to u, v, z, r, and t fields at three pressure levels), and $(41, 41)$ is the shape of the grid centered at TC formation location. The model consists of four alternating convolution and max-pooling layers that generate sequential features of length T using $\mathrm{TimeDistributed}$ layer of Keras (Chollet, 2015), which are further fed into a stacked LSTM consisting of three LSTM layers. To avoid over-fitting dropout 0.15 is used between two successive LSTM layers. The model and input-output size of each layer is shown in Figure 2.3 for $T = 3$. This resulted in a lightweight model with just 2,83,073 trainable parameters.

Training and Implementation

We experimented with various configurations of above model by varying number of layers and nodes in it, activation functions, and learning rates. The configuration which works well across all ocean basins is reported. The activation function $\mathrm{ReLU}(x) = \max(0, x)$ (Nair and Hinton, 2010) is used in all layers except the last layer which uses $\mathrm{Sigmoid}(x) = \frac{1}{1+\exp(-x)}$ (Han and Moraga, 1995) activation function. The input variables are scaled in the range $(-1, 1)$ for faster and stable training using $\mathrm{MinMaxScaler}$ of Scikit learn library (Pedregosa *et al.*, 2011). The model uses optimizer Adam (Kingma and Ba, 2014), default learning rate 0.001, binary cross-entropy loss function, 32 batch size, and 30 epochs. The model is implemented in Keras API developed over low-level language $\mathrm{TensorFlow}$ (Abadi *et al.*, 2015) on Nvidia Tesla V100 GPU platform with 16 GB RAM, that takes around 5-15 minutes for 30 epochs depending on ocean basin and T.

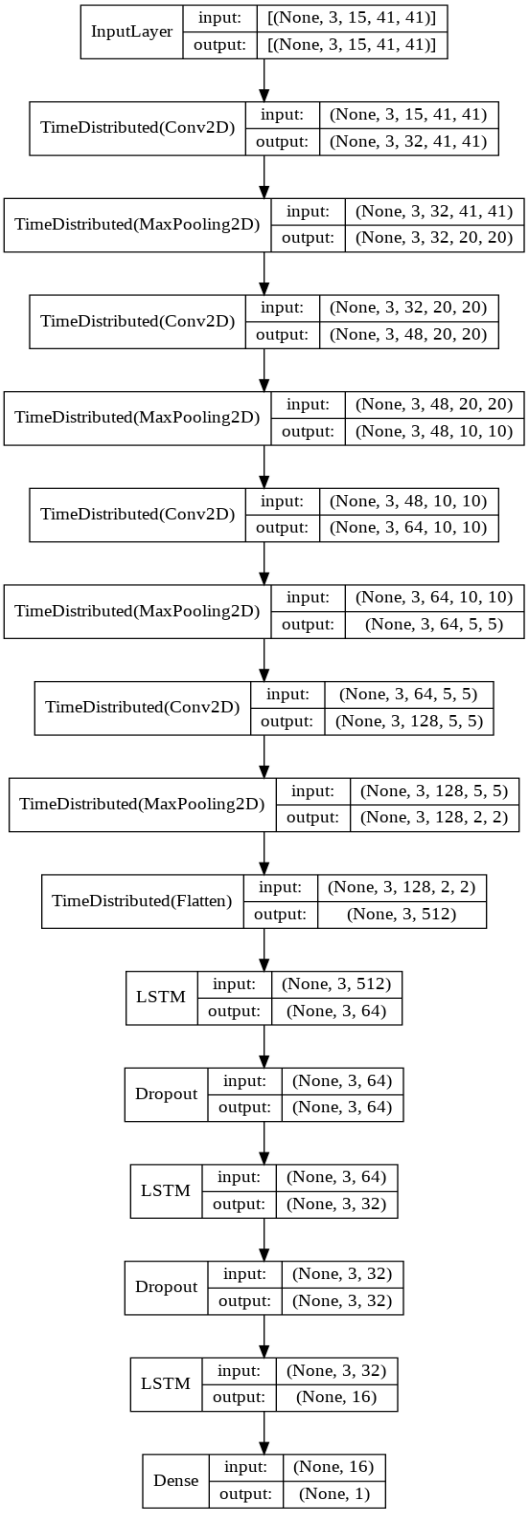


Figure 2.3: Model description for $T = 3$.

2.3.1 Evaluation Metrics

As reported in (Zhang *et al.*, 2019), we have evaluated the performance of proposed model in terms of metrics - Precision, Recall, Accuracy and F1-score (F1) which are defined as:

$$Precision(P) = \frac{TP}{TP + FP}, \quad Recall(R) = \frac{TP}{TP + FN}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad F1 = \frac{2PR}{P + R}$$

Table 2.2: Truth Table for TC formation

	Predicted	
Actual	Non TC	TC
Non TC	TN	FP
TC	FN	TP

where TN, TP, FN, FP are shown in Table 2.2 for our classification problem. A higher precision and recall are desirable. A higher precision indicates that a warning from the model for a possible TC formation can not be ignored, whereas a higher recall suggests that the model can detect a possible TC formation with a high probability. F1 score is a measure of the balance between precision and recall. To report the model's performance in terms of these metrics, we have used the k-fold cross-validation technique described below.

k-Fold Cross Validation

k-fold cross-validation is one of the widely used reliable technique to evaluate a machine learning model. The parameter k represents the number of folds. The technique is described as follows:

1. Shuffle the dataset randomly.
2. Partition the dataet into k equal groups (folds).
3. For each group, do the following iterations:
   - Take this group as test set and remaining (k-1) groups as training set.
   - Fit the model on training set and evaluate on the test set.
   - Retain the evaluation score on the test set and discard the model.
4. Return the average of evaluation scores in above iterations with variation.

In the above process, each sample is given a chance to be in the test set one time and a part of the training set (k-1) times. This technique results in a less biased (or less

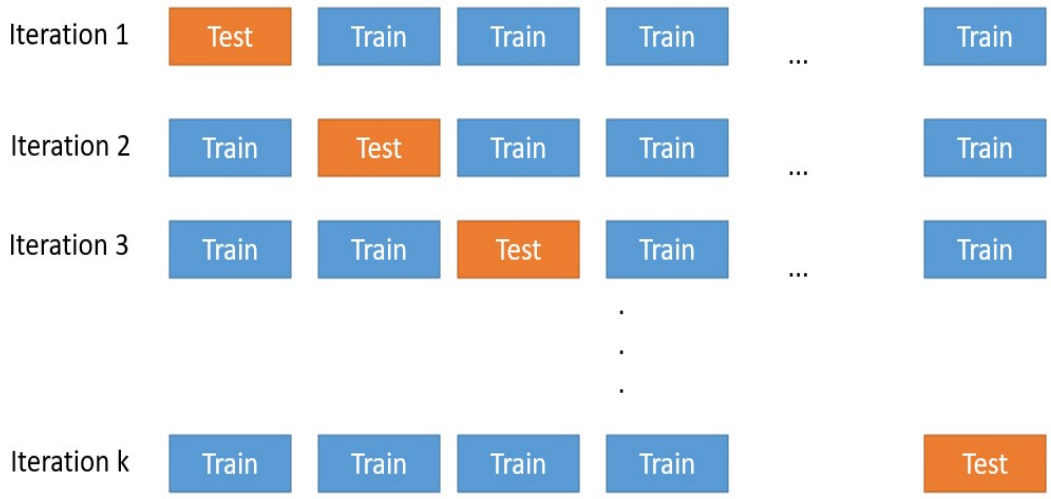


Figure 2.4: Graphical representation of k-fold cross validation process.

optimistic) model performance estimate compared to the usual train/test split. The results are considered to be more reliable for practical purposes. The process is described graphically in the Figure 2.4.

We have reported our model's 5-fold cross-validation score in terms of the metrics mentioned above for various lead times.

2.4 Results and Analysis

The proposed model takes any 12h, or 24h (corresponding to T = 3, or 5) of continuous data at an interval of 6h as input from $t - 72$ to $t - 24$ hours, where t is the time of possible TC formation and predicts whether a TC will be formed or not. This way for a particular T, the model predicts at lead times $6k$, $4 \leq k \leq 13 - T$. Thus the model predicts at a lead time of at least 24h, which is a minimum requirement for practical utility purposes. For various leads time, the model's 5-fold cross-validation accuracy along with the variation (std) is reported, in terms of metrics as mentioned above in Tables 2.3 and 2.4. Increasing or decreasing values of T do not improve the results further.

The accuracy for 24h lead time, vary in a range of $91.7\% - 97.7\%$ and $93.5\% - 97.4\%$ for T equals to 3 and 5 respectively. So for a 24h lead time forecast, 36h (for T = 5) of data gives better results. The Precision remains quite high in the range of $88.4\% - 96.4\%$ for both $T = 3, 5$. This implies that the model has a very low false alarm rate and any

Table 2.3: 5-fold performance (±std) of the model for T = 3 (12h) (Hochreiter and Schmidhuber, 1997; Gers *et al.*, 1999; LeCun *et al.*, 1989; Krizhevsky *et al.*, 2012).

Ocean Basin	Lead Time(h)	Accuracy	Precision	Recall	F1
NA	24	0.943 ±0.01	0.922 ±0.02	0.965 ±0.01	0.943 ±0.01
	36	0.982 ±0.01	0.981 ±0.02	0.983 ±0.01	0.982 ±0.01
	48	0.977 ±0.00	0.977 ±0.01	0.976 ±0.01	0.976 ±0.00
	60	0.912 ±0.02	0.925 ±0.02	0.901 ± 0.03	0.912 ± 0.02
NI	24	0.977 ±0.01	0.964 ±0.02	0.992 ±0.01	0.978 ±0.01
	36	0.990 ±0.01	0.983 ±0.01	0.997 ±0.01	0.990 ±0.01
	48	0.989 ±0.01	0.984 ±0.02	0.994 ±0.01	0.989 ±0.01
	60	0.929 ±0.02	0.939 ±0.03	0.918 ± 0.03	0.928 ±0.02
SI	24	0.955 ±0.01	0.932 ±0.02	0.982 ±0.01	0.956 ±0.01
	36	0.993 ±0.01	0.989 ±0.01	0.996 ±0.00	0.992 ±0.00
	48	0.991 ±0.01	0.995 ±0.00	0.987 ±0.01	0.991 ±0.00
	60	0.913 ±0.01	0.932 ±0.02	0.892 ± 0.02	0.912 ±0.01
WP	24	0.931 ±0.01	0.920 ±0.02	0.944 ±0.02	0.931 ±0.01
	36	0.975 ±0.00	0.982 ±0.01	0.968 ±0.00	0.975 ±0.00
	48	0.972 ±0.01	0.978 ±0.01	0.966 ±0.01	0.972 ±0.01
	60	0.869 ±0.02	0.910 ±0.03	0.818 ± 0.02	0.862 ±0.02
SP	24	0.930 ±0.03	0.898 ±0.06	0.977 ±0.03	0.933 ±0.03
	36	0.964 ±0.03	0.945 ±0.05	0.992 ±0.01	0.967 ±0.02
	48	0.954 ±0.04	0.926 ±0.07	0.991 ±0.01	0.956 ±0.03
	60	0.899 ±0.02	0.872 ±0.05	0.943 ± 0.05	0.903 ±0.02
EP	24	0.917 ±0.01	0.884 ±0.02	0.959 ±0.01	0.920 ±0.02
	36	0.966 ±0.02	0.955 ±0.02	0.979 ±0.01	0.979 ±0.01
	48	0.966 ±0.01	0.963 ±0.01	0.971 ±0.02	0.967 ±0.01
	60	0.904 ±0.02	0.915 ±0.02	0.891 ± 0.02	0.903 ±0.02

warning by model regarding possible TC formation can not be ignored. Also, the Recall remains in the range of $94.4\% - 99.2\%$ for both $T = 3, 5$, which is quite high, implying that model is detecting nearly all TC formation cases, and it can be used reliably for practical purposes. The F1-score varies in the range of $92.0\% - 97.8\%$.

The accuracy for 36h lead time is even better than 24h lead time across all ocean basins, which is in the range $96.4\% - 99.3\%$ and $97.2\% - 99.1\%$ for T equals to 3 and 5 respectively. A possible reason for this is that the reanalysis variables that we have selected in our study are more distinguishable and represent the TC formation well at this lead time. The Precision and Recall remain quite high again in the range of $94.5\% - 99.3\%$ and $96.8\% - 99.7\%$ respectively for both $T = 3, 5$.

The accuracy for lead time 48h is in the range $95.4\% - 99.1\%$, which is greater than

Table 2.4: 5-fold performance (std) of the model for T = 5 (24h) (Hochreiter and Schmidhuber, 1997; Gers *et al.*, 1999; LeCun *et al.*, 1989; Krizhevsky *et al.*, 2012).

Ocean Basin	Lead Time(h)	Accuracy	Precision	Recall	F1
NA	24	0.953 ±0.01	0.951 ±0.02	0.956 ±0.03	0.953 ±0.01
	36	0.986 ±0.01	0.993 ±0.01	0.980 ±0.02	0.987 ±0.01
	48	0.913 ±0.02	0.949 ±0.02	0.873 ±0.03	0.909 ±0.02
NI	24	0.974 ±0.02	0.959 ±0.03	0.992 ±0.01	0.974 ±0.02
	36	0.991 ±0.01	0.988 ±0.01	0.992 ±0.01	0.990 ±0.01
	48	0.925 ±0.01	0.964 ±0.03	0.941 ±0.02	0.952 ±0.01
SI	24	0.960 ±0.02	0.942 ±0.03	0.982 ±0.02	0.961 ±0.01
	36	0.987 ±0.01	0.986 ±0.01	0.988 ±0.01	0.987 ±0.01
	48	0.936 ±0.01	0.941 ±0.03	0.933 ±0.04	0.936 ±0.01
WP	24	0.935 ±0.02	0.912 ±0.04	0.965 ±0.01	0.937 ±0.02
	36	0.972 ±0.02	0.966 ±0.03	0.980 ±0.01	0.973 ±0.02
	48	0.888 ±0.02	0.899 ±0.04	0.877 ±0.02	0.887 ±0.02
SP	24	0.960 ±0.01	0.936 ±0.03	0.988 ±0.01	0.961 ±0.01
	36	0.980 ±0.02	0.978 ±0.02	0.984 ±0.01	0.980 ±0.02
	48	0.930 ±0.01	0.940 ±0.03	0.920 ±0.04	0.928 ±0.01
EP	24	0.949 ±0.01	0.933 ±0.02	0.968 ±0.01	0.950 ±0.01
	36	0.982 ±0.01	0.980 ±0.01	0.984 ±0.01	0.982 ±0.01
	48	0.919 ±0.01	0.936 ±0.01	0.900 ±0.01	0.917 ±0.01

lead time 24h and slightly less than lead time 36h in the case of $T = 3$. The accuracy for 48h lead time decreases by approx $5\% - 9\%$ in comparison of lead time 24h and 36h for $T = 5$. This implies that for 48h lead-time prediction $T = 3$ is a better choice. For $T = 3$, the model can predict with a lead time of 60h, which is quite a large time for early prediction of TC formation. In this case also model achieves an accuracy in the range of $86.9\% - 92.9\%$, which can be considered good because the dynamics of causal factors behind TC formation change rapidly with time. The variation (std) in most of the cases remains in the range of $1\% - 3\%$ except few cases where it goes upto 7%. This means that model can be utilized reliably for practical usage.

In Figure 2.5 the epoch-wise accuracy of the train and test set is shown for all ocean basins. One can observe that the model converges within 30 epochs.

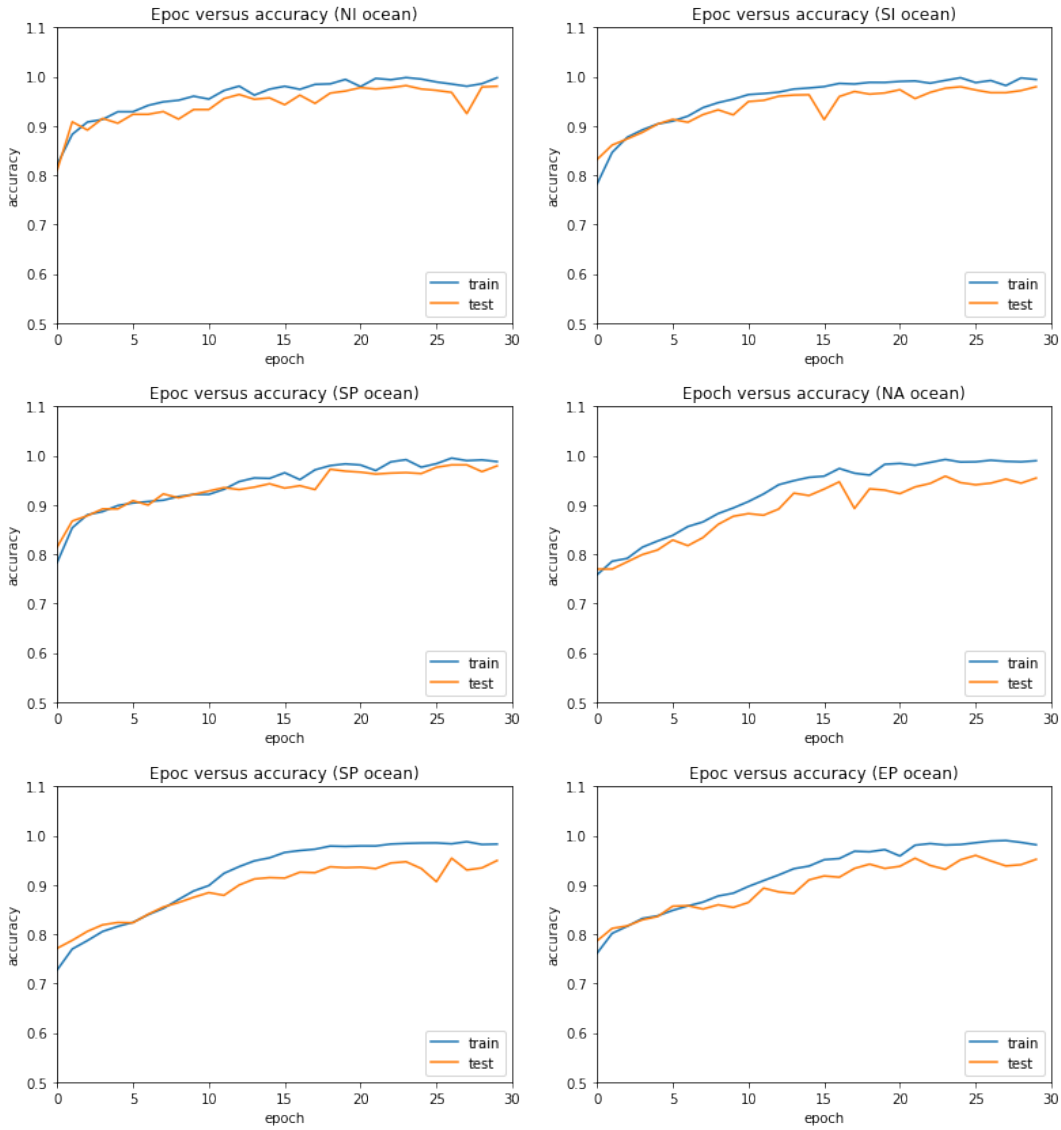


Figure 2.5: Train versus Test accuracy for T = 3.

Table 2.5: Best 5-fold performance reported in Zhang *et al.* (2019).

Ocean Basin	Lead Time(h)	Precision	Recall	F1
NA	24	0.937	0.880	0.908
	48	0.888	0.683	0.757
WNP	24	0.948	0.754	0.817
	48	0.889	0.642	0.701

2.4.1 Comparison

As discussed in section 2.1, the existing deep learning studies are not suitable for a direct comparison, as (Shakya *et al.*, 2020; Chen *et al.*, 2019*b*) deals with detecting an ongoing TC, and in (Matsuoka *et al.*, 2018) the TC formation definition is based on

cloud cover and uses simulated satellite data. We will make a direct comparison with the machine learning work (Zhang *et al.*, 2019), where the TC formation definition coincides with our definition, and authors evaluated nine machine learning models in terms of Precision, Recall, and F1-score for lead times up to 48h in NA, and WNP ocean basins. Out of nine classifiers, overall, AdaBoost works best in all cases. In Table 2.5, we have reported the AdaBoost accuracy for lead times 24h and 48h. From Table 2.5, we observe that for 24h lead time, Precision is more or less the same, but there is a big difference in terms of Recall, whereas we achieve a recall value of 96.5% both in NA and WP ocean basin in comparison of 88% and 75.4%. In the case of lead time 48h, we achieve a better performance both in WP and NA ocean basins with an improvement of more than 9% in Precision and more than 28% in the Recall. Also in (Zhang *et al.*, 2019), TC originated during months May to November has been considered, whereas our model do not have any such restriction.

2.5 Conclusion

In this work, a deep learning model is proposed that can forecast a TC formation using as less than 12h of data and lead time up to 60h with high precision, recall, and F1-score across six ocean basins of the world. Early information regarding potential cyclone formation has enormous social, economic, and environmental benefits. Through this work, the authors establish that the reanalysis dataset has enough information to capture the complex and non-linear natural phenomenon behind a cyclone formation. One can further attempt to use the reanalysis dataset for a longer lead time forecast. The reanalysis dataset provides many other variables like cloud cover, vorticity, sea surface temperature, etc. One can explore these variables to improve the model further.

CHAPTER 3

Will a Tropical Cyclone Make Landfall?[1]

In the different development phases of a tropical cyclone, the most exciting and complex phase is its landfall, which is when a tropical cyclone moves over to the land after crossing the ocean's coast. As discussed in subsection 1.1, the majority of low-pressure systems die over the ocean. But few develop as a TC and travel hundreds of kilometers to move over to coastal regions finally. If a TC does not make landfall, then it may not cause any human and infrastructure loss, but if it makes landfall, then it may bring catastrophic effects. Early detection of the landfall event for a TC will give the administration and disaster managers adequate time to mobilize resources, issue an early warning, and vacate the coastal regions to minimize human and material loss. On the other hand, early detection of no landfall for a TC will save money as preparedness for a possible landfall involves huge resources arrangement, and it will cause no inconvenience to the people residing in the coastal areas. Therefore, predicting if a given tropical cyclone will make landfall or not with high accuracy is very important. In this work, we investigate a fundamental question: will a tropical cyclone make a landfall? A deep learning model has been proposed that can predict the landfall event in the early phase of a tropical cyclone - particularly using any 12 hours or 24 hours of data from the first 72 hours of its inception with very high accuracy. The model can be trained within 05 to 20 minutes depending on the ocean basin and can predict the above-stated problem within seconds, making it suitable for real-time application. The research contributions of this work are as follows:

- Three atmospheric variables from reanalysis dataset and five from IBTrACS are identified that captures the landfall event of a TC.
- A fusion deep learning model has been proposed that gives a 5-fold cross validation accuracy in the range of 97.6% to 99.2% across all ocean basins.

[1]This chapter is a slightly modified version of work accepted for publication in Journal Neural Computing and Applications and has been reproduced here with the permission of the copyright holder.

3.1 Related Work

The traditional and operational TC-related models are mainly numerical, statistical, and ensemble in nature[2]. Whereas the statistical methods (Hall and Jewson, 2007) have inherent limitations in terms of the inability to capture the complex relationship between a huge number of cause factors, the numerical and ensemble methods are computationally involved and time-consuming. Recently deep learning techniques consisting of Convolutional Neural networks (CNNs) and Recurrent Neural networks (RNNs) have been successfully applied to forecast TCs related forecast problems to capture the complex relationship between various causal factors that have both spatial and temporal dimensions (Chen *et al.*, 2019*b*; Moradi Kordmahalleh *et al.*, 2016; Chaudhuri *et al.*, 2017; Giffard-Roisin *et al.*, 2020; Boussioux *et al.*, 2020; Biswas *et al.*, 2021). These earlier works mainly focus on track and intensity prediction, which we will discuss more in the next section. As per our knowledge, this is the first work that successfully attempted to forecast the landfall event for a TC.

CNN and LSTM based deep learning models have been successfully developed to forecast various TC-related prediction problems. In (Moradi Kordmahalleh *et al.*, 2016; Alemany *et al.*, 2018; Gao *et al.*, 2018; Kim *et al.*, 2019*b*; Kumar *et al.*, 2021*c*), CNN or LSTM or a combination of both have been used to predict the trajectories of TCs with a certain lead time. In (Pradhan *et al.*, 2018; Chen *et al.*, 2019*a*; Maskey *et al.*, 2020), CNN has been used to forecast the intensity of a TC. In (Shakya *et al.*, 2020), CNN has been used to predict the TC formation detection using satellite data. Recently, reanalysis and historical TC data have been used to predict TC's track and intensity in a few works. For example, in (Boussioux *et al.*, 2020), authors have used reanalysis data to construct various models based on CNN, GRU, Transformers, and XGBoost, which predict the track and intensity of a TC with a lead time of 24h in two ocean basins NA and EP. In (Giffard-Roisin *et al.*, 2020), a deep learning model using CNN and LSTM is proposed to predict the track of TCs in six ocean basins of the world with a lead time of 24h using ERA-Interm reanalysis data (Dee *et al.*, 2011) and historical cyclone data.

Several researchers have also researched different problems associated with the landfall of a TC. In (Vitart *et al.*, 2003), authors have presented a seasonal analysis of landfall and the effect of various atmospheric and oceanographic factors on landfall of

[2] https://www.nhc.noaa.gov/modelsummary.shtml

TCs for the Mozambique region. In (Wahiduzzaman *et al.*, 2017) a generalized additive model to forecast the seasonal landfall location in terms of affected country in the NI ocean basin has been presented. In (Mohapatra *et al.*, 2015), authors have evaluated the official tropical cyclone landfall forecast issued by Indian Meteorological Department (IMD). In (Powell and Aberson, 2001), authors have presented the official accuracy levels achieved by NOAA of landfall time and location prediction for the Atlantic basin during 1976-2000.

More recently, in (Kumar *et al.*, 2021*b*,*a*), we have proposed a deep learning model using CNN and LSTM networks to predict the landfall's location, time, and intensity of a TC for the six ocean basins of the world using reanalysis and IBTrACS data. These works predict landfall's characteristics from the very initiation of the TC without a restriction of lead time. These studies and earlier studies that focus on landfall's location, time, and intensity prediction, do so without investigating whether a landfall event will happen for a TC or not. As per our knowledge, there is no existing work that deals with the classification problem of whether landfall will occur for a TC or not. To fill this gap, we develop a classification model that predicts whether a landfall event will happen for a TC or not. The current work supplements the earlier works and can be used together to answer the landfall-related forecast questions in completeness.

3.2 Data

We have used publicly available datasets in our study, the IBTrACS dataset Knapp *et al.* (2010) maintained by National Oceanic and Atmospheric Administration[3] (NOAA) and reanalysis data ERA5 (Hersbach *et al.*, 2020) maintained by European Centre for Medium-Range Weather Forecasts[4] (ECMWF). The IBTrACS dataset contains the historical record of a tropical cyclone at an interval of 3h in terms of its ocean basin, time, latitude, longitude, pressure, wind speed, distance to land, category, distance, and direction between the successive location of a TC and many more. From this dataset, we have extracted three variables latitude, longitude, and distance to land at an interval of six hours for this study. Two more features distance and direction between two successive locations of a TC for this six-hourly data have been generated. We have taken

[3] https://www.ncdc.noaa.gov/ibtracs/index.php?name=ib-v4-access
[4] https://cds.climate.copernicus.eu/

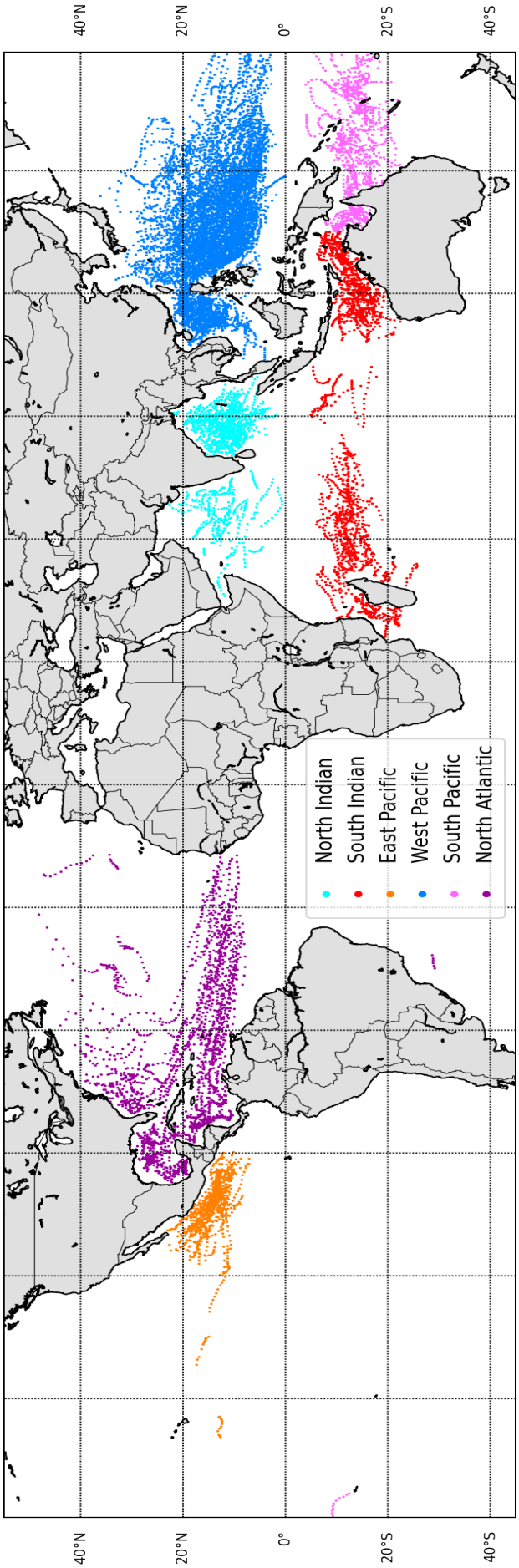


Figure 3.1: Trajectory of first $6 * (S - 1)$ hours of all landfall occurring cyclones.

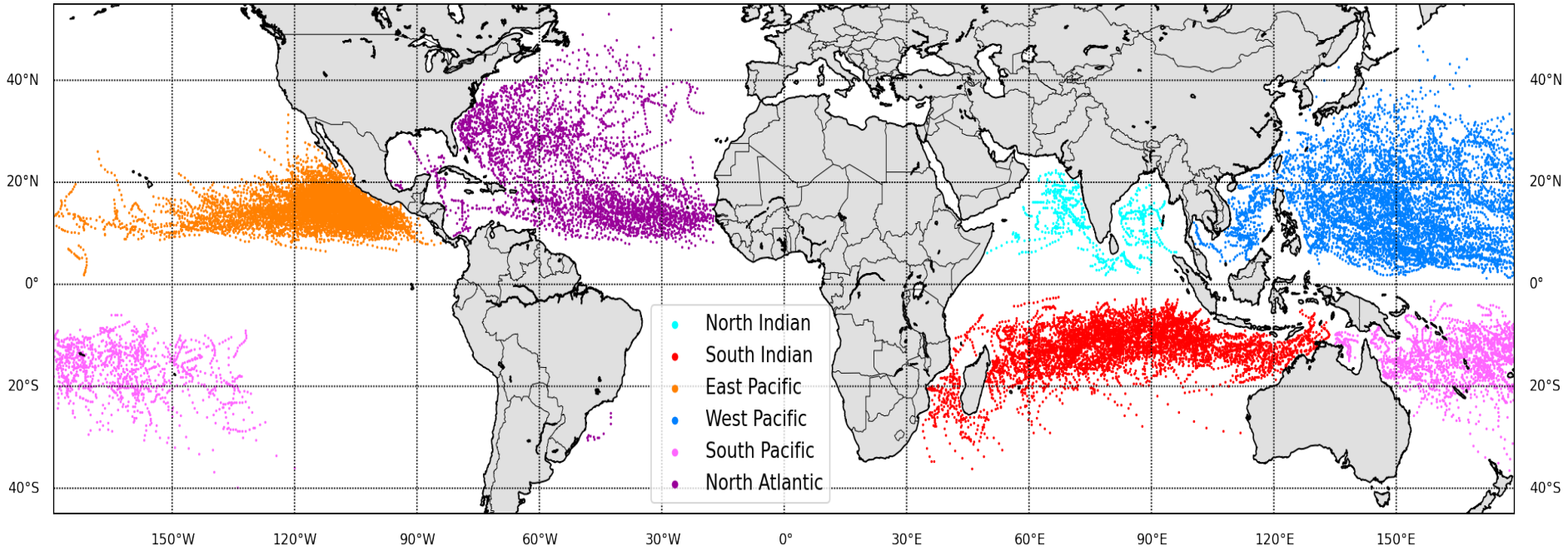


Figure 3.2: Trajectory of first $6 * (S - 1)$ hours of all non landfall occurring cyclones.

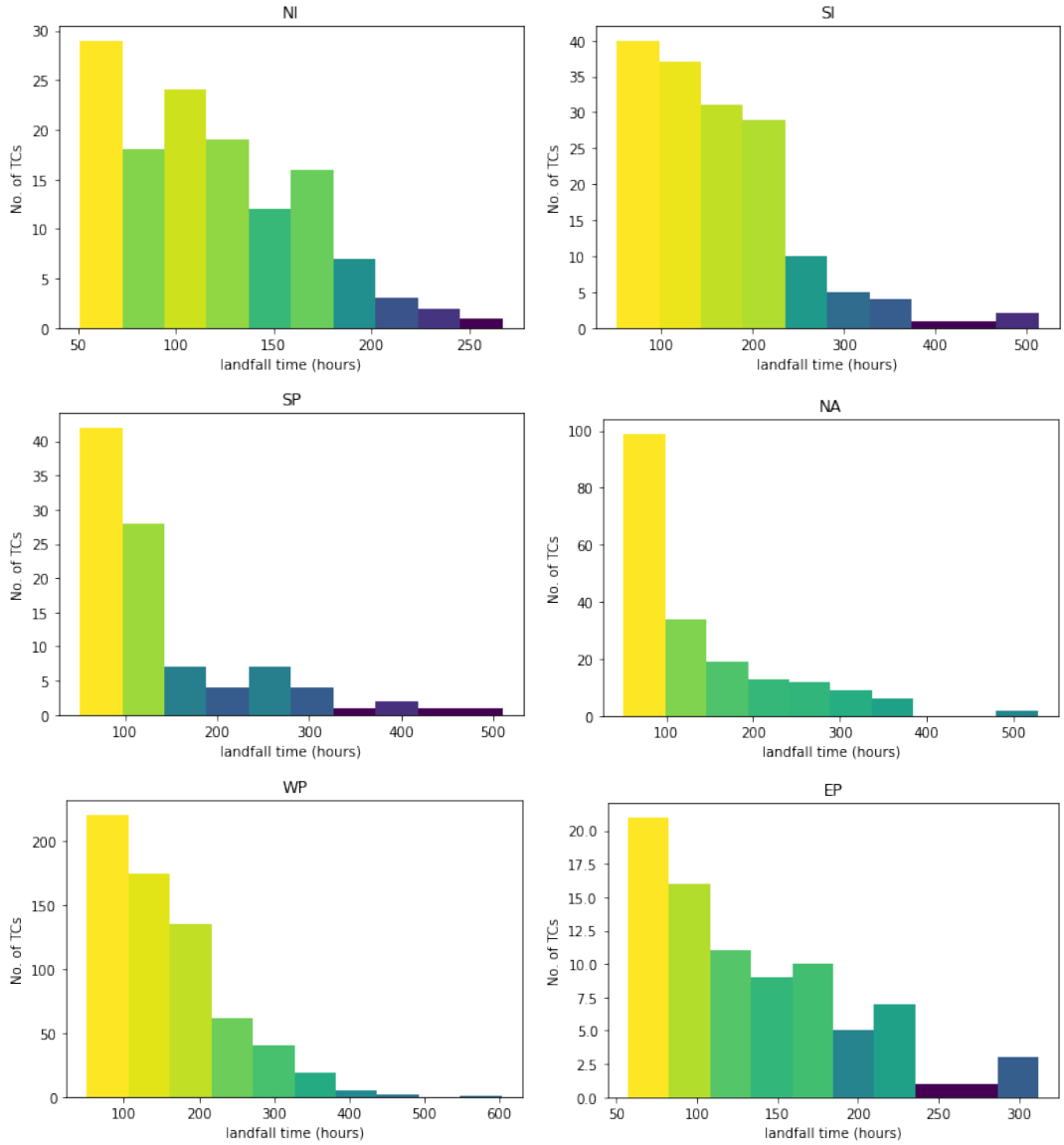


Figure 3.3: Distribution of landfall time for six ocean basins.

only those cyclones into account which remain over the ocean for at least the first 48h of its inception. TCs that remain for fewer than 48h over the ocean will not intensify to a level to cause any destruction and can be ignored. We have extracted $\min(72, t - a)$ hour of data at an interval of 6h, where

$$\begin{cases} a = 0, t \text{ is the life span of TC} & \text{if landfall doesn't occur} \\ a = 24, t \text{ is the time to landfall} & \text{if landfall occurs} \end{cases}$$

This way, we are taking a maximum of 72h of data and keeping away 24h from the landfall time, in case it happens. The study contains cyclones from January 1981 to May 2021. Trajectories of all cyclones used in this study are shown in Figures 3.1 and 3.2 for cyclones with landfall and without landfall, respectively, in six ocean basins.

The number of cyclones in each ocean basin, average, minimum, and maximum time to landfall is shown in Table 3.1. The landfall time distribution for all ocean basins is shown in Figure 3.3.

Table 3.1: Data Description.

Ocean Basin	**NA**	**NI**	**SI**	**WP**	**SP**	**EP**
Number of TCs	494	185	642	1178	383	745
Landfall	194	131	160	660	97	84
Not Landfall	300	54	482	518	285	661
Average Landfall Time	134h	116h	161h	154h	138h	132h
Max Landfall Time	528h	267h	513h	603h	510h	312h
Min Landfall Time	51h	51h	51h	51h	51h	57h

As discussed in Section 1.2, ERA5 is the 5^{th} generation reanalysis data containing climate and weather data for the whole globe. ERA5 replaces the ERA-interim dataset and contains hourly data at a resolution of $0.25°$ from 1979 onwards. The long-term behavior of a TC depends on large atmospheric conditions around its center that has vertical, horizontal, and temporal dimensions. To capture the large-scale atmospheric wind circulations and geopotential at various pressure levels, u, v, and z fields are extracted from ERA5. Geopotential (z field) is the gravitational potential energy per unit mass relative to sea level. Geopotential at different pressure levels affects the weather systems such as TC. The u and v wind fields provide information about east-west and north-south movement, respectively, along with speed. Inspired by the other related works described above, wind fields u, v, and the geopotential field z are extracted at pressure levels of 700hPa, 500hPa, and 225hPa. Since the atmospheric system that determines a TC characteristic can have a horizontal extent up to 1000 kilometers (km), the above variables are extracted for a spatial extent of $16° \times 16°$ and resolution $0.25°$ around the center of the TC for each time point in the dataset. This way, we have extracted these variables' values for a grid of 65×65 centered at the TC location. One degree of latitude/longitude is approximately 110kms near the equator. A pictorial representation of reanalysis data used in this study is shown in Figure 3.4.

3.2.1 Dataset Preparation

For a TC, let $S = \frac{\min(72, t-a)}{6} + 1$ be the number of data points available at an interval of 6h, where t and a are defined as above. Suppose we want to use T number of data

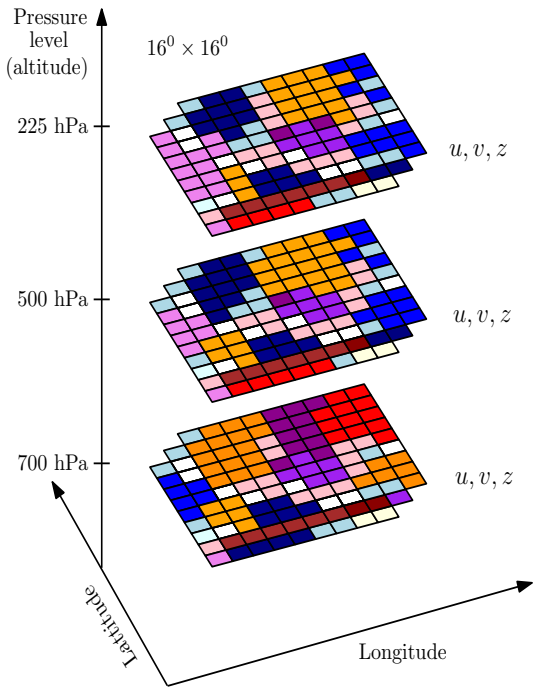


Figure 3.4: Pictorial representation of reanalysis data.

points ((T-1)*6 hours of data) to predict the landfall event. We generate $S - T + 1$ data points, where a single data point is a sequence of T vectors of the form:

$$(\,\mathrm{lat}(t), \mathrm{long}(t), \mathrm{distance}(t), \mathrm{direction}(t), \mathrm{distToLand}(t), \mathrm{u225}(t), \mathrm{v225}(t),$$
$$\mathrm{z225}(t), \mathrm{u500}(t)\ \mathrm{v500}(t), \mathrm{z500}(t), \mathrm{u700}(t), \mathrm{v700}(t), \mathrm{z700}(t))$$

where $k \leq t \leq T + k - 1$ and k varies from 1 to $S - T + 1$. For each such data point, the target variable is 1 (positive class) or 0 (negative class) as per landfall happens or not for this cyclone. The $i^{th}(1 \leq i \leq S - T + 1)$ data point corresponds to the time interval $(t + (i-1)6)$h to $(t + (i-1)6 + (T-1)6)h$, where t is the time corresponds to first data of TC. One must note that the above criterion of choosing S ensures that we are predicting landfall event at least 24h before, in case it happened. The collection of all such data points for all the cyclones form the dataset.

3.3 Model and its Implementation

To capture the spatial and temporal dimensions of our dataset, we have used a combination of CNN (LeCun *et al.*, 1989; Krizhevsky *et al.*, 2012; Milletari *et al.*, 2016) and LSTM (Hochreiter and Schmidhuber, 1997; Gers *et al.*, 1999; Gers *et al.*, 2003;

Gers and Schmidhuber, 2001) networks. As our dataset has two types of data, reanalysis data which is 2-dimensional in shape, and cyclone's track-related data, which is 1-dimensional in shape, our proposed model has two branches, the Reanalysis branch, and the Track branch, as shown in Figure 3.5. The Reanalysis branch takes reanalysis data as input, where one training data is of the shape $(T, 9, 65, 65)$, where T represents the number of sequential data points (corresponding to $6 * (T - 1)$ hours of data), 9 represents the number of channels (corresponding to u, v and z fields for three pressure levels), and $(65, 65)$ is the shape of the grid centered at TC center. The reanalysis branch consists of three convolution layers, where each layer is succeeded by a batch-normalization (Ioffe and Szegedy, 2015) layer, and a max-pooling (Ciresan *et al.*, 2011) layer. These layers generate sequential features of length T using Keras (Chollet, 2015) $\mathrm{TimeDistributed}$ layer, which is further fed into a stacked LSTM with two layers. The last layer of the Reanalysis branch is a LSTM layer with an output size of 16. The Track branch process the track-related data latitude, longitude, distance to land, distance, and direction between the successive position of TC. As we have sequential data of length T, this branch consists of stacked LSTM with three layers. The input shape of Track branch is (T, 5), where 5 corresponds to features: latitude, longitude, distance to land, distance, and direction. The last layer of the Track branch is an LSTM layer with an output size of 8. Further, the last two layers of both branches are concatenated and fed into a dense layer of size 8, which is further connected to the last output layer of size one. The input and output shapes of all layers are described in Figure 3.5, which is generated using Keras. All the layers uses the activation function $\mathrm{ReLU}(x) = \max(0, x)$ (Nair and Hinton, 2010), except the last layer which uses activation function $\mathrm{Sigmoid}(x) = \frac{1}{1+\exp(-x)}$ (Han and Moraga, 1995). The resulted model is a lightweight model with $3,90,061$ trainable parameters.

3.3.1 Training and Implementation

We tried a few configurations of the above-described model by experimenting with the number of layers, number of nodes in a layer, activation function, and learning rate and reported the one which gives the best results across all the ocean basins. Keras API developed over low-level language $\mathrm{TensorFlow}$ (Abadi *et al.*, 2015) is used to implement the model. For faster training the input features are scaled in the range $(-1, 1)$

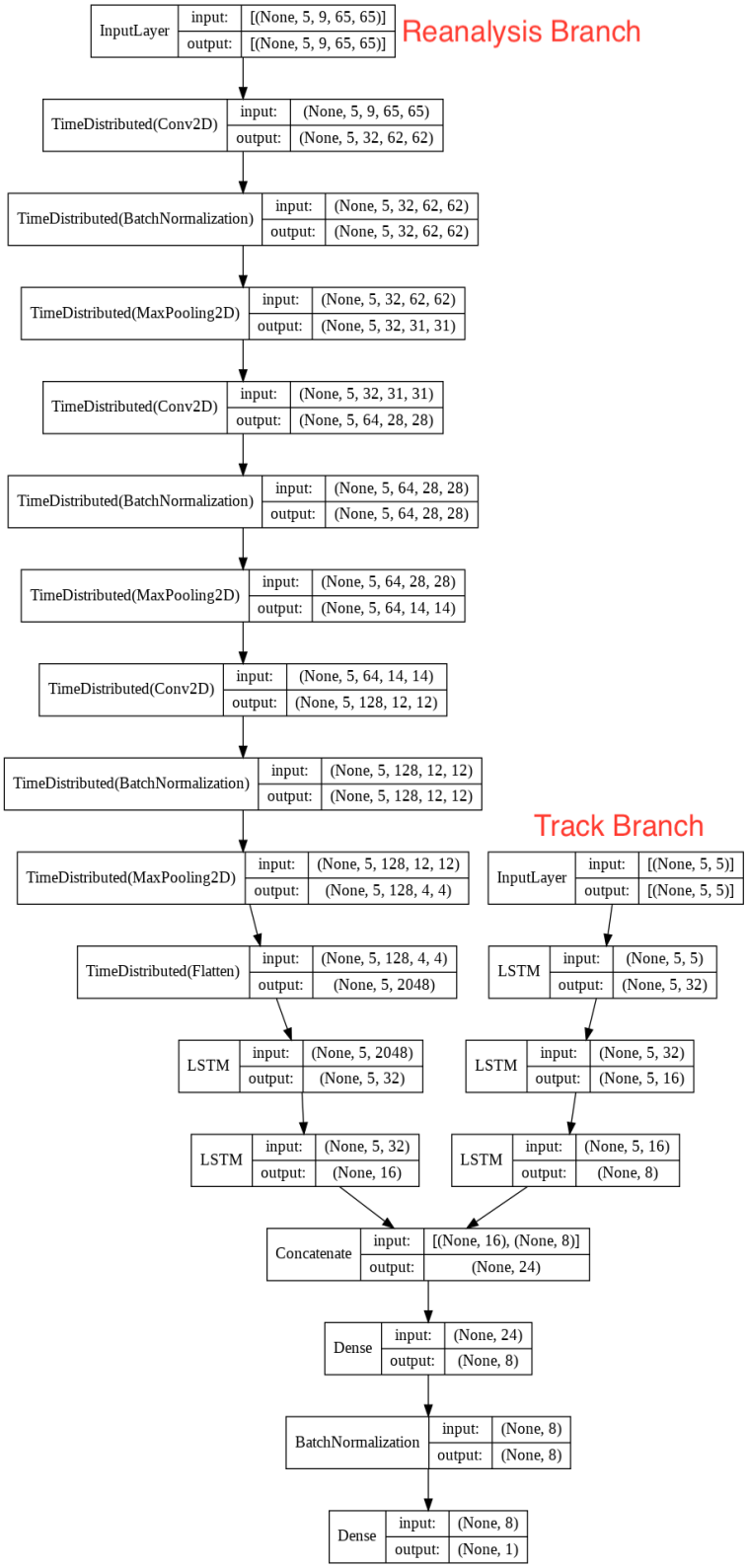


Figure 3.5: Model description for $T = 5$.

using the MinMaxScaler of Scikit learn library (Pedregosa *et al.*, 2011). The model uses binary cross-entropy as loss function, Adam (Kingma and Ba, 2014) optimizer, batch size 128, and default learning rate 0.001. The model converges within 40 epochs, and experiments are performed on Nvidia Tesla V100 GPU with 16 GB RAM, which

takes around 05-20 minutes (depending on the ocean) to execute the 40 epochs.

3.3.2 Evaluation Metrics

In a classification problem in the area of weather forecasting, probability of detection (POD) and false alarm rate (FAR) are two important evaluation criteria (Barnes *et al.*, 2007), which are defined as:

$$FAR = \frac{FP}{FP + TP}, \; POD = \frac{TP + TN}{TP + TN + FP + FN}$$

Table 3.2: Truth Table for TC Landfall

	Predicted	
Actual	Not Landfall	Landfall
Not Landfall	TN	FP
Landfall	FN	TP

TN, FP, FN, TP are shown in Table 3.2 for our binary classification problem. The other important metrics that are important for critical landfall prediction problems are False Positive Rate (FPR) and False Negative rate (FNR), which are defined as:

$$FPR = \frac{FP}{FP + TN}, \; FNR = \frac{FN}{FN + TP}$$

A smaller FAR is desirable as a false landfall alarm will result in a waste of resources in disaster preparedness. A higher POD (also known as accuracy) is expected as it gives an overall performance of the model in terms of the proportion of correctly predicted samples. FPR denotes the proportion of non-landfall cases predicted as landfall, and FNR denotes the proportion of landfall cases predicted as non-landfall. Both FPR and FNR desired to be smaller. Another important metric for a binary classifier is the AUC (area under the curve), which is defined as $AUC = \int ROC$, where ROC denotes the Receiver Operating Characteristics curve that plots FPR versus TPR (1-FNR). A higher AUC means that the model can separate the two classes well.We have reported the performace of our model in terms of 5-fold cross validation score as described in Section 2.3.1.

3.4 Results and Analysis

The proposed model takes any 12h or 24h of continuous data (for T = 3 or 5) from the maximum first 72 hours of cyclone data and predicts whether landfall event will happen or not. Also, in case of landfall occurring cyclone, the model predicts at least 24 hours before. The 5-fold accuracy of the model along with variation (std) in terms of POD, FAR, FPR, FNR, and AUC are shown in Table 3.3, for $T = 3, 5$. Increasing or decreasing T further does not improve the model's performance. The best results for a particular ocean basin corresponding to $T = 3$ or 5 in terms of POD are highlighted in Table 3.3. For $T = 3$, we get the best results in the case of NA, SI, SP, and EP ocean basins and for $T = 5$ in the case of NI and WP ocean basins. The model performs well across all ocean basins with the best POD of 0.987, 0.976, 0.987, 0.983, 0.992, and 0.991 for NA, NI, SI, WP, SP, and EP ocean basins, respectively. One can see that the variation (std) remains low in the range of 0.002 (0.2%) to 0.018 (1.8%), which makes the model reliable to use. Considering the complexity of the landfall event forecasting problem that depends on multiple atmospheric, oceanographic, and geographic factors and the unbalanced number of classes (especially in EP, SI, and SP ocean basins), the model performance can be considered suitable for all practical purposes.

From the Table 3.3, we can see that for the best choice of T, the other criterion FAR, is in the range of 0.007 to 0.027 (0.7% to 2.7%), which is relatively low; therefore, a warning for possible landfall by the model needs to be acted upon without fail.

For the best choice of T, the FPR is negligible for NA, SI, SP, and EP oceans. The FPR for NI and WP oceans are 0.070 (7%) and 0.018 (1.8%), which are pretty low for all practical purposes. This means that with a very high probability, the model will detect all non-landfall events correctly.

For the best choice of T, the FNR for NI, WP, and SP ocean basins is almost negligible and varies in a range of 0.025 to 0.083 for the other three ocean basins. For EP and SI ocean basins, FNR is at a higher side 0.083 (8.3%) and 0.041 (4.1%) respectively, which is probably because of the unbalanced number of classes (landfall and non-landfall) in these two ocean basins as shown in Table 3.1.

The AUC is greater than or equal to 0.975 across all ocean basins for both choices of T, which means the model is effectively able to learn the characteristics of negative

Table 3.3: Five fold performance (std) of the model (Hochreiter and Schmidhuber, 1997; Gers *et al.*, 1999; LeCun *et al.*, 1989; Krizhevsky *et al.*, 2012).

T (h)	Ocean Basin	NA	NI	SI	WP	SP	EP
3 (12h)	Dataset Size	3872	1926	6964	11973	4087	7937
	POD	**0.987** ±0.002	0.964 ±0.019	**0.987** ±0.005	0.964 ±0.033	**0.992** ±0.003	**0.991** ±0.002
	FAR	**0.013** ±0.009	0.014 ±0.009	**0.012** ±0.004	0.021 ±0.005	**0.012** ±0.006	**0.007** ±0.009
	FPR	**0.006** ±0.005	0.033 ±0.018	**0.003** ±0.001	0.023 ±0.005	**0.003** ±0.001	**0.001** ±0.001
	FNR	**0.025** ±0.013	0.036 ±0.023	**0.041** ±0.028	0.044 ±0.063	**0.018** ±0.013	**0.083** ±0.028
	AUC	**0.998** ±0.001	0.991 ±0.006	**0.998** ±0.001	0.995 ±0.004	**0.999** ±0.001	**0.994** ±0.005
5 (24h)	Dataset Size	3380	1560	5028	9629	2947	6451
	POD	0.930 ±0.049	**0.976** ±0.018	0.969 ±0.033	**0.983** ±0.010	0.986 ±0.015	0.987 ±0.009
	FAR	0.078 ±0.047	**0.027** ±0.024	0.040 ±0.048	**0.016** ±0.020	0.010 ±0.004	0.007 ±0.009
	FPR	0.036 ±0.024	**0.070** ±0.070	0.011 ±0.013	**0.018** ±0.023	0.003 ±0.001	0.001 ±0.001
	FNR	0.139 ±0.154	**0.003** ±0.007	0.090 ±0.097	**0.014** ±0.010	0.044 ±0.058	0.128 ±0.097
	AUC	0.975 ±0.033	**0.998** ±0.001	0.989 ±0.016	**0.998** ±0.001	0.999 ±0.001	0.995 ±0.007

Table 3.4: POD for random cyclones for best T.

	Ocean Basin	SID (Storm Identifier)	POD	Forecast
Landfall (1)	NA	2007244N12303	1.0	11111111111
	NI	1988327N06100	1.0	111111111
	SI	1988076S14056	0.727	01111100111
	WP	2009270N10148	1.0	111111111
	SP	1985033S12157	1.0	11111111111
	EP	1983285N09260	0.545	00001111110
No Landfall (0)	NA	1991250N10334	0.818	00000110000
	NI	1993316N11070	0.777	011000000
	SI	2016277S30044	1.0	00000000000
	WP	1982286N17145	1.0	000000000
	SP	2009025S12177	1.0	00000000000
	EP	2013180N11256	1.0	00000000000

and positive classes. The ROC is shown in Figure 3.7 for NA ocean basin for $T = 5$ and for WP ocean basin for $T = 3$. The ROC for other ocean basins looks similar.

The epochs-wise POD (accuracy) graph for train and test set is shown in Figure 3.6 for all six ocean basins for the best values of T. One can see that in all the cases, the model starts converging after 25 epochs. In all cases, after 25 epochs, the train and test accuracy almost overlaps, which means that the model generalizes well on the unseen test data.

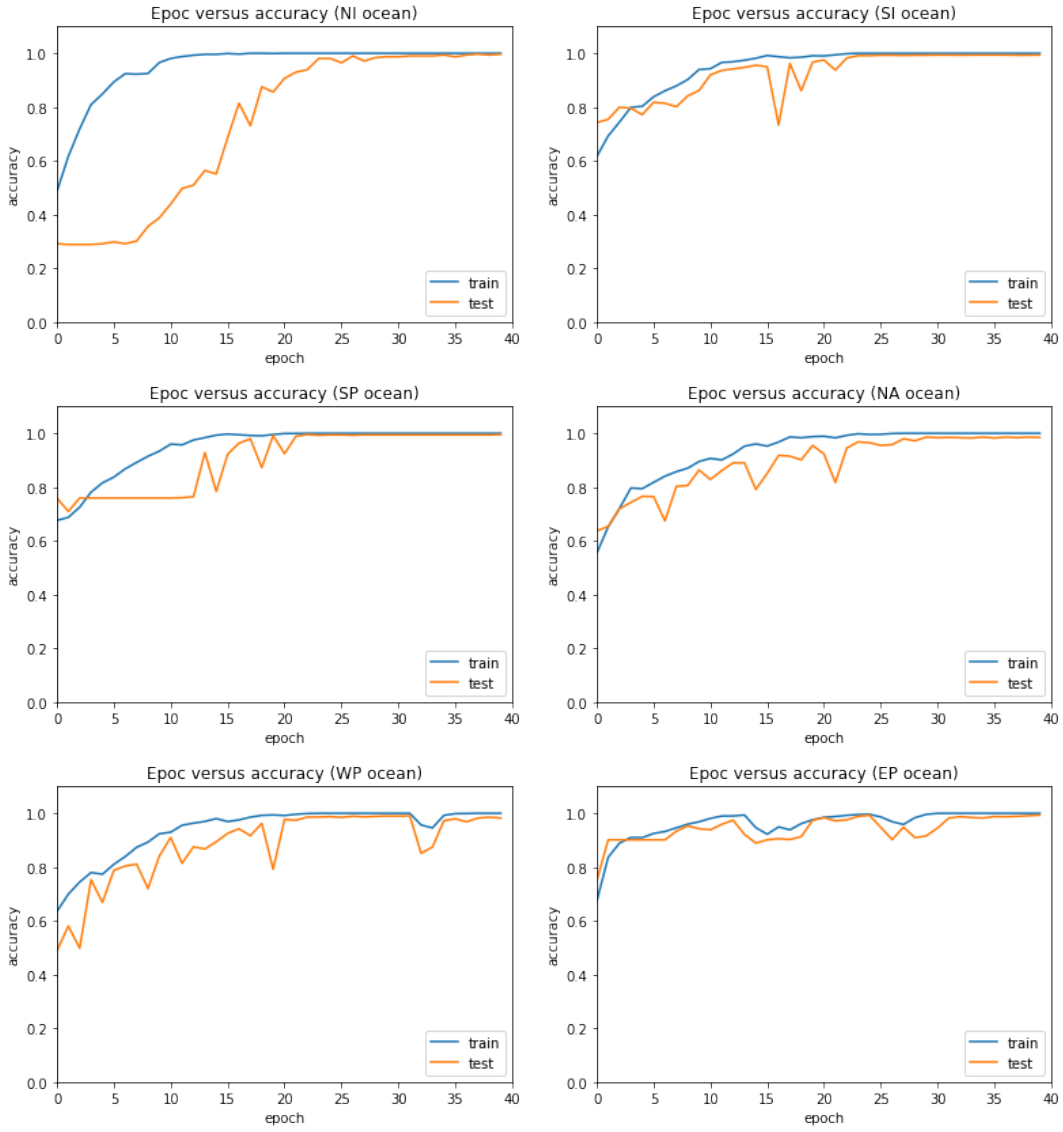


Figure 3.6: Epoch wise train and test accuracy for best T.

To further demonstrate the working and reliability of our model, we have shown the POD for two random TCs (one with landfall and one without landfall) with 72h of data ($13 - T + 1$ training points) from each ocean basin with best T, see Table 3.4. These TCs were never a part of training datasets, and these are also not used during the data

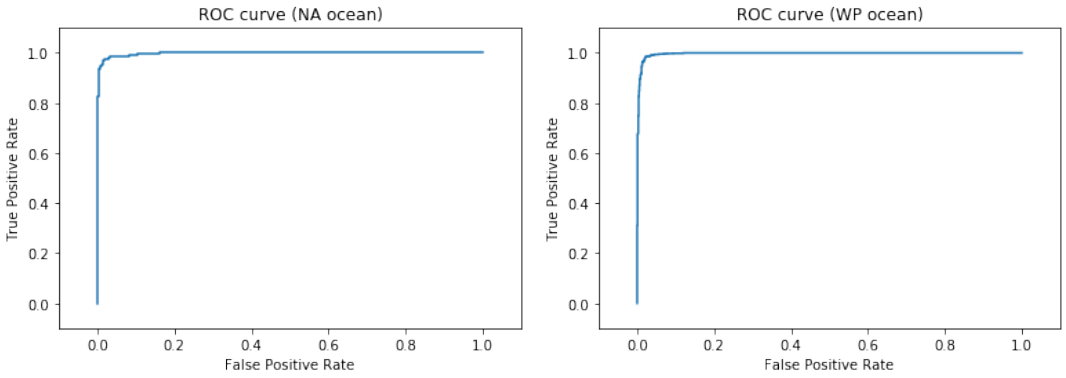


Figure 3.7: ROC curve for NA ocean (T = 5) and WP ocean (T = 3).

preparation process. The data for these cyclones are scaled using the scalers that are fitted on the rest of the data. These cyclones are treated as future cyclones while getting the forecast for these. We have used the best T corresponding to each ocean basin while reporting these results. For example, for the land-falling cyclone with storm identifier (SID) equals 1988076S14056 in the SI ocean basin, the POD achieved is 0.727. The last column (viz. forecast) represents the forecast pattern for eleven timestamps (from 24th hour to 72nd hour at six hours interval using last 12h of data). For this particular SID, the actual forecast pattern is 11111111111, and the predicted one is 01111100111. For the non-land-falling cyclone with SID 2016277S30044 in SI ocean, the POD is 1.0, and the observed and predicted forecast pattern is 00000000000. For land-falling cyclones, we get POD in the range of 0.545 to 1.0, and for non-land-falling cyclones in the range of 0.777 to 1.0.

3.5 Conclusion

We presented a deep learning model based on CNN and LSTM that can forecast the landfall event in the early phase of a tropical cyclone with high accuracy and precision. As per our knowledge, this is the first such work that targeted this prediction problem. We think that this work will bring out a novel and significant problem for the researchers in this area. We can see different approaches to this problem using statistical, numerical, and deep learning methods in the future. The model performs well on all criteria like POD, FAR, FPR, FNR, and AUC across all six ocean basins of the world. This makes the model a practical one because it helps disaster managers avoid unwarranted preparation for a potential landfall in case the model predicts it negatively

and timely preparation, resource mobilization in case the model predicts it positively. Once the model is trained, it can provide a real-time prediction, making it suitable for practical usage. The model also demonstrates that 12h or 24h of reanalysis data and track data of a cyclone in the early phase of a TC have enough information to predict its landfall behavior accurately. Given that reanalysis data has been successfully utilized for track, intensity, and landfall predictions of a TC, one can further explore it to predict TC's genesis and its formation. Reanalysis data also provides many more weather features like humidity, sea surface temperature, vorticity, cloud cover, etc.; one can further explore these features to see their effect on the above-mentioned prediction problems.

CHAPTER 4

Forecasting Landfall's Intensity, Location and Time of a Tropical Cyclone [1]

In the last two research works (Chapter 2 and Chapter 3), we proposed deep learning models that forecast whether a TC will be formed or not and will it make landfall. Once you have an affirmative answer to the last two questions, an important and obvious question to ask is where, at what time, and at what intensity it will make landfall. The development of such a model will help answer the forecast of various critical phases of a TC from its formation to its landfall. The economic and human losses caused by a TC center around a few kilometers of its landfall location. Knowing the location, time, and intensity of landfall of a TC with high accuracy will help the administration in the precision deployment of the resources and optimizing the usage. The knowledge of landfall's location will help focus on the prospective affected area and minimize displacing of people from coastal areas. The intensity at landfall will help decide the extent of resource mobilization, and the knowledge landfall's time will help in timely preparation.

Predicting landfall characteristics (in terms of intensity, location, and time) of a TC is the most challenging task as the TC behaviour near the landfall is not only determined by atmospheric and oceanographic factors but also topographical factors. The TC behavior near the landfall changes abruptly (Leroux *et al.*, 2018). All this makes the development of a model that can predict the landfall characteristic's from the very initiation of a TC important from a disaster mitigation point of view and challenging at the same time. The research contributions in this direction are as follows:

- A LSTM based model is proposed to predict the landfall's characteristics in the NI ocean basin, using BTD data from IMD. Five variables from BTD, along with two derived features, are identified as input to the model.

[1] This chapter is a slightly modified version of work published in Proceedings of the AAAI Conference on Artificial Intelligence, volume 35 2021 and Artificial Neural Networks and Machine Learning – ICANN 2021 and has been reproduced here with the permission of the copyright holder.

- A LSTM and CNN based deep learning model is proposed to predict the landfall's location and time in six ocean basins NI, SI, WP, SP, NA, and EP using reanalysis dataset ERA5. Four variables from the reanalysis dataset and location (latitude and longitude) from IBTrACS are used as input to the model.

4.1 Related Work

There are numerous existing TC track and intensity prediction models that can be classified into numerical (or dynamical) models, statistical models, and ensemble models[2]. The numerical models rely on physical equations governing atmospheric circulations to capture the evolution of the atmospheric fields. These methods are computationally involved and need large supercomputers. The statistical techniques (Hall and Jewson, 2007) learn relationships between TC behavior and its specific details in terms of location, pressure level, wind speed, etc. These are computationally fast but have limitations in capturing complex non-linear relationships between a large number of causal factors. The ensemble models combine a collection of other models and generally give comparatively better results (Krishnamurti *et al.*, 2000). The primary operational numerical models for track and intensity prediction are - European Center for Medium-Range Weather Forecasts (ECMWF), Global Forecast System (GFS), Hurricane Weather and Research Forecasting model (HWRF), and Hurricane Multi-scale Ocean-coupled Non-hydrostatic model (HMON). The primary operational, statistical models for track and intensity prediction are - Climatology and Persistence model (CLIPER5), Trajectory-CLIPER, Statistical Hurricane Intensity Prediction Scheme (SHIP), and Logistic Growth Equation Model (LGEM). ECMWF, Wind Radii Consensus (RVCN), and Florida State Super Ensemble (FSSE) are primary operational ensemble models.

Recently, with the increase in the data related to TCs, various studies have appeared that have successfully applied machine learning-based models to predict various characteristics of a TC. Initial studies regarding tropical cyclones track and intensity forecasts used Artificial Neural Networks (ANNs) (Chaudhuri *et al.*, 2017; Kovordányi and Roy,

[2] https://www.nhc.noaa.gov/modelsummary.shtml

2009; Chaudhuri *et al.*, 2015). Since the prediction problems related to atmospheric conditions involve both spatial and temporal components, deep learning models like RNNS, LSTM networks, CNNs, and combinations of these have been successfully deployed to capture the complex non-linear interplay between various atmospheric components of spatial and temporal nature. In (Moradi Kordmahalleh *et al.*, 2016), a sparse RNN model with flexible topology is presented for the prediction of hurricane trajectory with a lead time of 6h or 12h in the Atlantic ocean. In (Kim *et al.*, 2019*b*), ConvLSTM is used to predict the hurricane trajectory from the past density maps of hurricane paths. In (Alemany *et al.*, 2018), authors have presented a fully connected RNN model to predict cyclone trajectories from historical cyclone data in the Atlantic ocean in terms of grid identifier. In (Gao *et al.*, 2018), a nowcasting model is presented based on an LSTM network to predict typhoon trajectory. Recently, few studies have dealt with TC formation, track, and intensity prediction problems using reanalysis data (Giffard-Roisin *et al.*, 2020; Boussioux *et al.*, 2020; Chen *et al.*, 2019*b*). In (Chen *et al.*, 2019*b*), reanalysis dataset has been used to forecast typhoon formation forecasting in NA, EP, and WP oceans. In (Giffard-Roisin *et al.*, 2020), authors have used historical data of a TC along with reanalysis data from ERA-Interim (Dee *et al.*, 2011) to predict the track of TC with a lead time of 24h in six ocean basins. They propose a fusion network in which the output of CNN trained on wind fields and pressure fields from reanalysis data and output of an ANN trained on historical TC data are fed into another ANN network. Their model does not take the temporal aspect into account as they have stacked input from two-time steps t and $t-6$ to feed into a CNN. In (Boussioux *et al.*, 2020), TC intensity and track prediction task is achieved with a lead time of 24h, using reanalysis data ERA5 (Hersbach *et al.*, 2020), historical TC data, and output from operational forecast models for NA and EP ocean basins. They have proposed framework Hurricast (HURR) consisting of seven different models that used various combinations of CNN, GRU, Transformers, and XGBoost models. They used data between t to $t-21$ hours and captured the spatial and temporal aspects of data, thereby addressing the shortcoming of (Giffard-Roisin *et al.*, 2020).

From the above discussion, we see that the existing works mainly focus on predicting the track and intensity of a TC with a specific lead time. In this work, we have presented deep learning models that can predict the characteristics of the most crucial phase: the landfall of a TC from its very initiation. As there is no existing work with

which we can make a direct comparison, we will try to present a reasonable comparison of our models with the models presented in (Giffard-Roisin *et al.*, 2020; Boussioux *et al.*, 2020) and the landfall's location prediction accuracy reported by IMD on its website. The rest of the chapter is divided into two sections. Section 4.2 presents the model that predicts the landfall's intensity, location and time in NI ocean basin using BTD data and section 4.2 presents the model that predicts landfall's location and time in all six ocean basins NI, SI, NA, WP, SP, and EP using reanalysis data ERA5.

4.2 Forecasting landfall's intensity, location, and time in NI ocean basin

4.2.1 Data

In this study, we have used the BTD data maintained by the Regional Specialised Meteorological Centre (RSMC) of IMD in New Delhi as discussed in section 1.2. BTD contains many features associated with a cyclone, but we selected latitude, longitude, MSSWS, and estimated central pressure (ECP) as inputs in our model. Two more derived features distance and direction of change between two successive recordings of a cyclone is calculated. Another critical factor that affects the course of a tropical cyclone is sea surface temperature (SST), which is obtained from NOAA dataset available at[3].

The dataset contains a few manual errors, which have been corrected carefully, after which a total of 6474 recordings, recorded at an interval of three hours, of 353 cyclones have been extracted. If the difference between two available time points for a cyclone is more than 3 hours, we have filled up missing time points to make the data a continuous time-series data recorded at an interval of 3 hours. If a time series data, $d(t)$ is available for $t = t_0$ and t_{3n} but missing for $t = t_3, t_6, \ldots, t_{3(n-1)}$ then we evaluate $D = (d(t_{3n}) - d(t_0))/n$ and fill the missing data with $d(t_{3k}) = d(t_0) + kD$ for $1 \leq k \leq n-1$. After completing this process and deleting any possible error in the dataset, we get 9088 recordings of 352 cyclones. As we are interested in predicting the landfall and therefore, we retain recordings only until the landfall time for every cyclone. This

[3] http://apdrc.soest.hawaii.edu/erddap/griddap/hawaii_soest_afc8_9785_907e.html

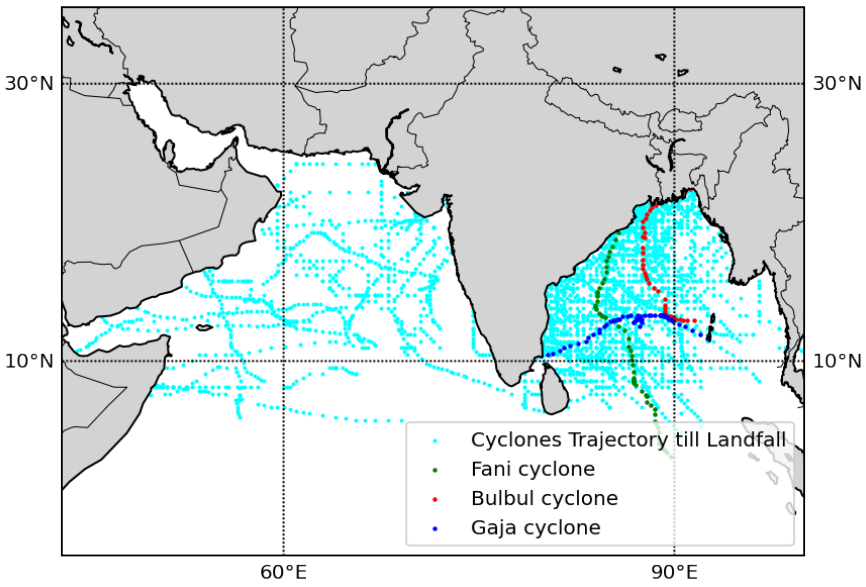


Figure 4.1: Cyclones trajectory till landfall in NI ocean.

further reduces our dataset to 3988 recordings of 206 cyclones (the rest of the cyclones do not hit the coast and die out in the sea). We retained only those cyclones with at least 24 hours of data, and the average time to landfall for such cyclones is around 80 hours. The trajectory of all cyclones till the landfall is shown in Figure 4.1. We have highlighted the trajectory of three TCs, Bulbul, Fani, and Gaja, for which we will show our model's performance along with a 5-fold cross-validation score. These TCs will not be a part of the training dataset.

Generation of Dataset

For a fixed cyclone, let T_L be the number of data points (that is $((T_L - 1)3h)$ recorded after which the landfall occurs. If we want our model to provide predictions after taking T number of data points as input, we need to make sure that our model trains on inputs of size T. To achieve that for each cyclone, we create $T_L - T + 1$ data points. A single data point is a sequence of T vectors of the form:

$$(\text{MSSWS}(t), \text{ECP}(t), \text{SST}(t), \text{distance}(t), \text{direction}(t), \text{latitude}(t), \text{longitude}(t))$$

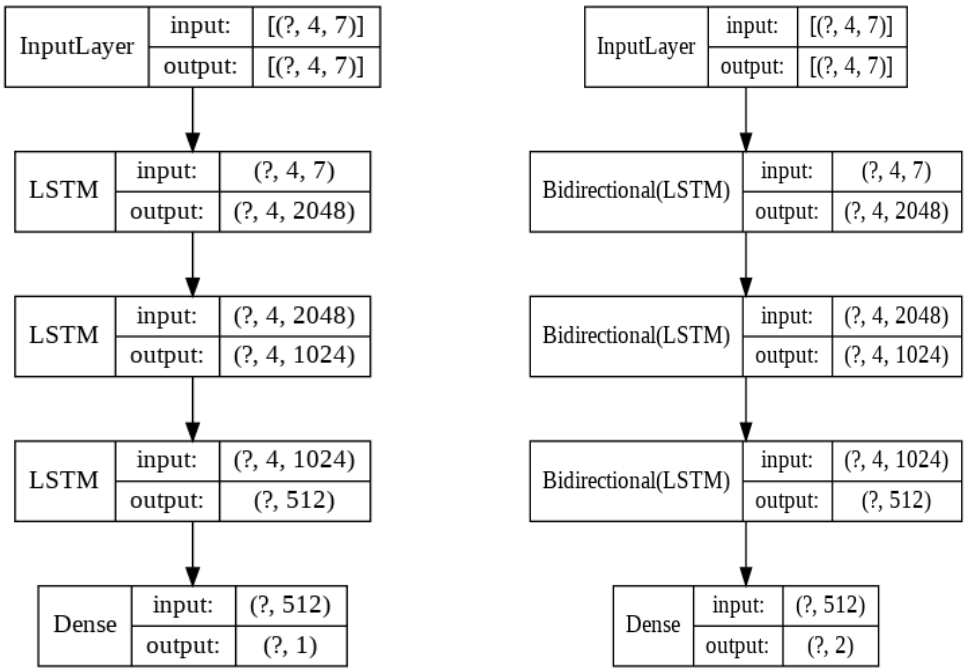
where $k \leq t \leq T + k - 1$. As k varies from 1 to $T_L - T + 1$, we get all such data points for a given cyclone. The target variables for each input are MSSWS (in knots) at landfall or location (latitude and longitude) at landfall, or time (in hours (h)) remaining

to landfall of the cyclone to which the data point corresponds to. For example, the data recording for cyclone Amphan started 00 hours on 16 May 2020, and the landfall occurred at 12 hours on 20 May 2020. Therefore, recordings of 108 hours are available for the Amphan cyclone, which amounts to $T_L = 108/3 = 36$. Suppose, we want to create a model for $T = 4$, then Amphan cyclone will provide $36 - 4 + 1 = 33$ data points. For the models that predict landfall's location and intensity, the target variable for each $i^{th}(1 \leq i \leq 33)$ data point is the location (latitude and longitude) and intensity of landfall. For the model that predicts time remaining to landfall, the target variable for each $i^{th}(1 \leq i \leq 33)$ data point is $(108 - 3 * T - (i - 1) * 3)$h. For a given T, the dataset is a collection of all such data points across all the cyclones. Notice that each data point is a time-series sequence of length T.

4.2.2 Model and its Implementation

We have used two different RNN models based on LSTM, Model-1 for landfall's intensity and time prediction and Model-2 for landfall's location (latitude and longitude) prediction. Model-1 has three stacked LSTM layers and one dense output layer. Model-2 has three stacked BiLSTM layers and one dense output layer. For a faster and better training of our models, we have scaled the features of data using Standard Scaler of Scikit learn library (Pedregosa *et al.*, 2011). The scaling is given by the function, $f(x) = (x - \mu)/\sigma$, where μ is the mean and σ is the standard deviation. Model-1 scales the input variables using the Standard Scaler except the target variables, intensity, and time. Model-2 scales all input variables including target variables, latitude and longitude for training. We have implemented these models in Keras API (Chollet, 2015) which runs on top of low-level language TensorFlow (Abadi *et al.*, 2015), developed by Google. Both models use the default learning rate of 0.001 and Adaptive moment estimation (Adam) (Kingma and Ba, 2014) optimizer to minimize the loss. The model trains the network using the MSE loss function, and the accuracy is measured in terms of MAE and RMSE. The definition of these error measures are as follows:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y}_i)^2, \quad \text{RMSE} = \sqrt{MSE},$$



(a) Model-1 (b) Model-2

Figure 4.2: RNN based on LSTM and BiLSTM for T = 4.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \bar{y}_i| \,,$$

where y_i is the actual value and $\bar{y}_i$ is the predicted value. We have tried various standard activation functions and selected $\text{Swish}(\beta x) = x.\,\text{Sigmoid}(\beta x)$ (Ramachandran *et al.*, 2017) with $\beta = 2$ for Model-1 and $\text{ReLU}(x) = \max(0, x)$ (Nair and Hinton, 2010) for the Model-2 to optimize the accuracy. Both models use a batch size of 512 and a total of 150 epochs. The structures of Model-1 and Model-2 are shown in Figure 4.2 which have been generated using Keras API.

We have used the GPU available on Google Colab to run the experiments, which provides one of the GPU Nvidia K80s, T4s, P4s, or P100s, depending on availability. On average, Model-1 takes 80 seconds and Model-2 takes 90 seconds to complete 150 epochs.

4.2.3 Results and Analysis

The models (Model-1 and Model-2) take a certain number of, say T, continuous data points of a TC, anytime during the course of the TC and predict the intensity, latitude, longitude, and time to its landfall with high accuracy. We also report the distance error

in kilometers (KM) from the predicted landfall location to the actual landfall location. We have reported our model's 5-fold validation accuracy as discussed in section 2.3.1, both in terms of RMSE and MAE for T = 4, 6, 8, and 12 (9h, 15h, 21h, or 33h of data at an interval of 3h) along with standard deviation (std). To further validate the performance of our model, we have also reported the model performance for three recent devastating cyclones Bulbul, Fani, and Gaja. These three cyclones are not the part of the training dataset. The RMSE and MAE values reported for these three cyclones are averages of RMSE and MAE over a sliding window of size T starting from 1st data point till the landfall.

Table 4.1: RMSE and MAE for landfall's intensity (knots) prediction for different values of T (Hochreiter and Schmidhuber, 1997; Gers *et al.*, 1999).

T(Hours)		4(9)	6(15)	8(21)	12(33)
Size of dataset		3189	2843	2544	2039
RMSE (± std)	5-fold Validation	9.35 ±0.63	7.84 ±1.09	7.31 ±0.71	6.19 ±0.69
	Fani	2.66	1.72	3.43	5.53
	Gaja	4.34	3.37	4.78	4.29
	Bulbul	4.35	3.63	3.86	3.40
MAE ± (std)	5-fold Validation	5.17 ±0.51	4.01 ±0.30	4.24 ±0.40	3.87 ±0.36
	Fani	2.03	1.37	2.64	4.10
	Gaja	2.85	2.15	3.51	3.47
	Bulbul	2.30	1.68	2.27	2.35

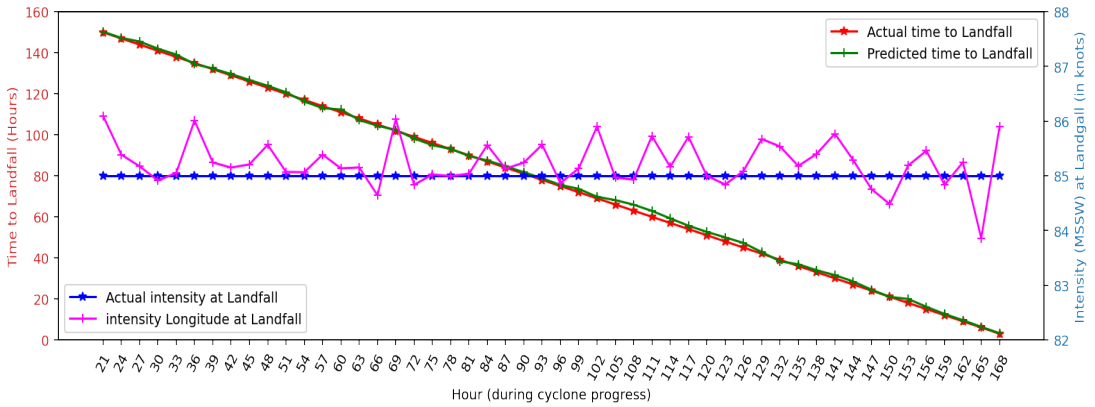


Figure 4.3: Predicted and actual intensity and time of landfall of Fani for $T = 8$.

In Tables 4.1 and 4.2, the RMSE and MAE of prediction of intensity at landfall and time remaining to landfall are reported for different values of T respectively, along with the size of the training dataset. For $T = 8$, that is, if 21 hours data of cyclone is used, then the intensity and time can be predicted within an MAE of 4.24 knots

Table 4.2: RMSE and MAE of time to landfall's for different values of T (Hochreiter and Schmidhuber, 1997; Gers *et al.*, 1999).

T(Hours)		4(9)	6(15)	8(21)	12(33)
Size of dataset		3189	2843	2544	2039
RMSE ± (std)	5-fold Validation	11.21 ±1.03	10.14 ±1.78	8.08 ±0.95	8.52 ±1.51
	Fani	3.6	2.03	2.8	5.34
	Gaja	6.0	5.65	4.0	6.17
	Bulbul	4.0	3.4	3.25	6.21
MAE ± (std)	5-fold Validation	6.25 ±0.2)	5.26 ±0.16	4.5 ±0.58	5.42 ±1.17
	Fani	2.6	1.37	1.8	4.70
	Gaja	3.2	2.74	3.0	4.90
	Bulbul	2.3	1.9	1.77	5.08

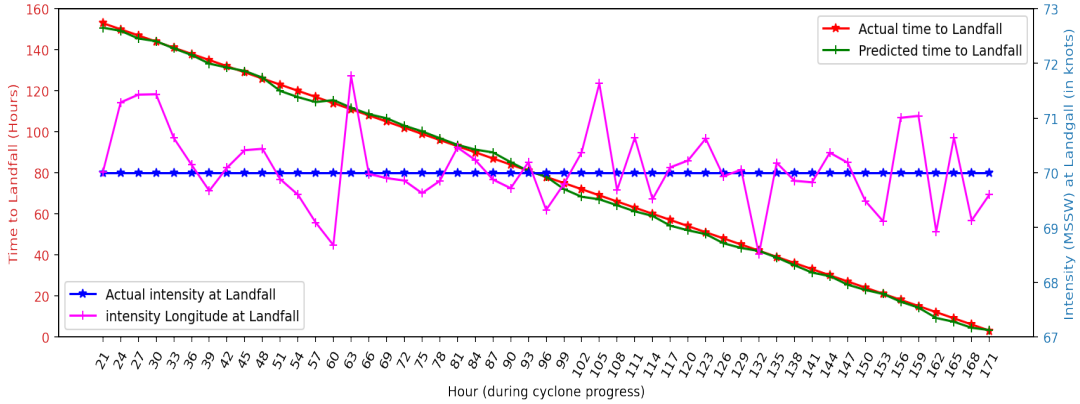


Figure 4.4: Predicted and actual intensity and time of landfall of Gaja for $T = 8$.

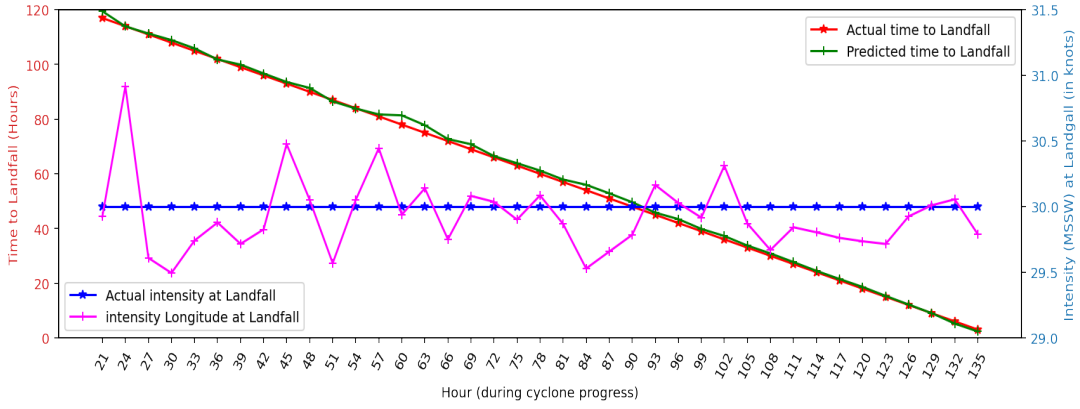


Figure 4.5: Predicted and actual intensity and time of landfall of Bulbul for $T = 8$.

and 4.5h, respectively. From Table 1.1, we can see that the range of MSSWS for all Grades, except Grade 2, is at least 10, which implies that with a very high probability, the model will predict the correct intensity grade at landfall of a TC. Moreover, since we obtain such good accuracy with only 21 hours of observation and the landfall occurs

Table 4.3: RMSE, MAE and Distance Error (kilometers) for Landfall Location Prediction for different values of T (Hochreiter and Schmidhuber, 1997; Gers *et al.*, 1999).

T(Hours)			4(9)	6(15)	8(21)	12(33)
Size of dataset			3189	2843	2544	2039
RMSE ± (std)	5-fold Validation	Lati	0.95 ±0.04	0.67 ±0.06	0.52 ±0.12	0.39 ±0.15
		Long	1.30 ±0.14	0.85 ±0.04	0.72 ±0.08	0.50 ±0.10
	Fani	Lati	0.33	0.16	0.11	0.13
		Long	0.60	0.36	0.19	0.23
	Gaja	Lati	0.85	0.53	0.22	0.13
		Long	0.45	0.22	0.07	0.09
	Bulbul	Lati	0.19	0.15	0.10	0.09
		Long	0.37	0.26	0.17	0.19
MAE ± (std)	5-fold Validation	Lati	0.52 ±0.02	0.33 ±0.01	0.24 ±0.02	0.19 ±0.01
		Long	0.75 ±0.05	0.46 ±0.02	0.37 ±0.02	0.29 ±0.02
	Fani	Lati	0.27	0.11	0.08	0.10
		Long	0.41	0.26	0.14	0.19
	Gaja	Lati	0.28	0.16	0.10	0.1
		Long	0.15	0.09	0.05	0.07
	Bulbul	Lati	0.15	0.09	0.07	0.07
		Long	0.29	0.19	0.14	0.16
Distance ± (std)	5-fold Validation		106.3 ±5.79	67.0 ±2.51	51.7 ±1.20	41.2 ±3.12
	Fani		56.1	32.4	18.7	24.6
	Gaja		38.5	22.7	15.1	15.1
	Bulbul		37.9	24.7	18.2	19.9

on average at the 80th hour in NI ocean, the model can help authorities to prepare well advance in time to take any action. The performance of the model is even better than 5-fold validation accuracy for cyclones Bulbul, Fani, and Gaja as evident from Tables 4.1 and 4.2.

In Figures 4.3, 4.4, and 4.5, the predicted intensity and actual intensity along with predicted time to landfall and actual time to landfall are shown for $T = 8$ (21 hours) for cyclones Bulbul, Fani, and Gaja, respectively. To obtain these figures, we choose a sliding window of 21 hours and get the prediction from the model. For example, the values at 21th hour and 75th hour are the predictions using the data between 0th

and 21th hours and 54th and 75th hours, respectively. It is evident that the model has consistently performed well irrespective of whether the prediction point is close to the landfall or far from the landfall. One should note that these three cyclones took a long time (>135 hours) to hit the coastal region; despite this, the model's predictions are consistently good even at the beginning of the cyclone.

In Table 4.3, the RMSE and MAE of latitude and longitude prediction (in degrees) at landfall are reported for different values of T. The minimum MAE of the model is achieved for $T = 12(33h)$, which is 0.24 and 0.37 for latitude and longitude respectively. A slight error in latitude and longitude may lead to an error of several kilometers in the location. Therefore, we also report the corresponding distance error in kilometers. The distance error is calculated using the distance between actual and predicted landfall location. For example, for $T = 12$, the model can predict the landfall location with an error of 41.2 KM.

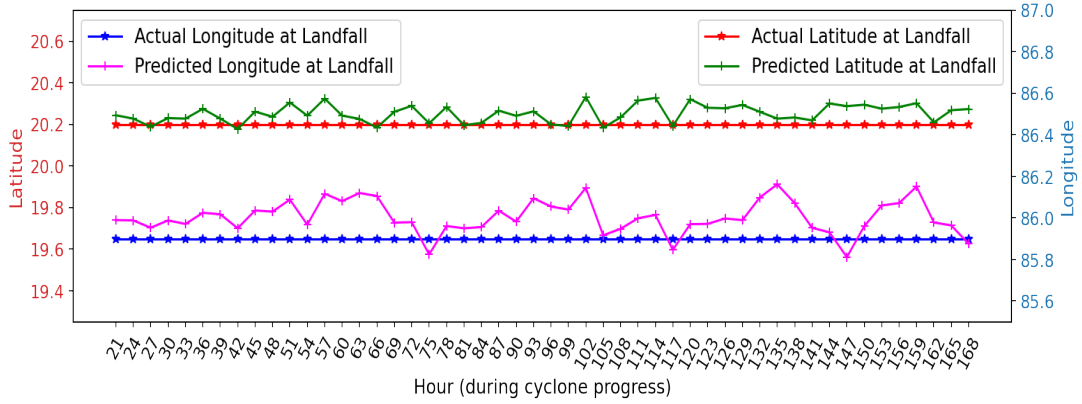


Figure 4.6: Predicted and actual landfall latitude/longitude of Fani cyclone for T = 8

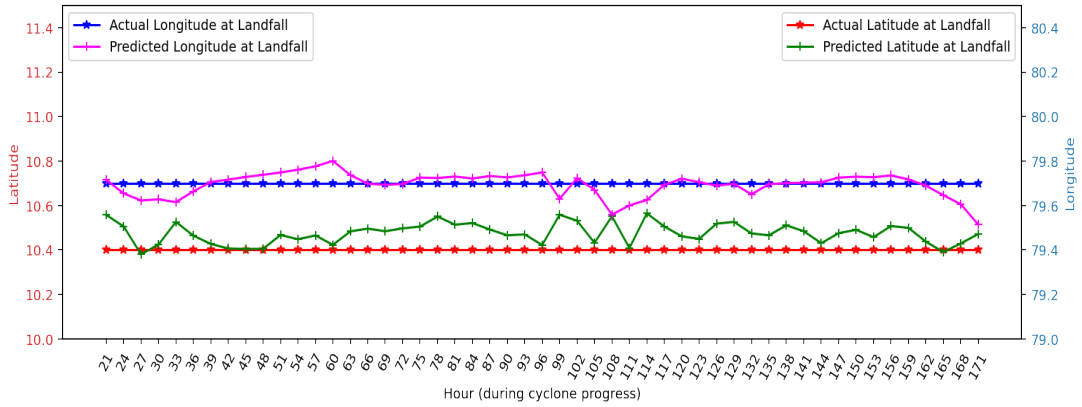


Figure 4.7: Predicted and actual landfall latitude/longitude of Gaja cyclone for T = 8

In the Figures 4.6, 4.7, 4.8, the predicted latitude, longitude, and actual latitude, longitude at landfall are shown for $T = 8$ for cyclones Bulbul, Fani, and Gaja, respectively.

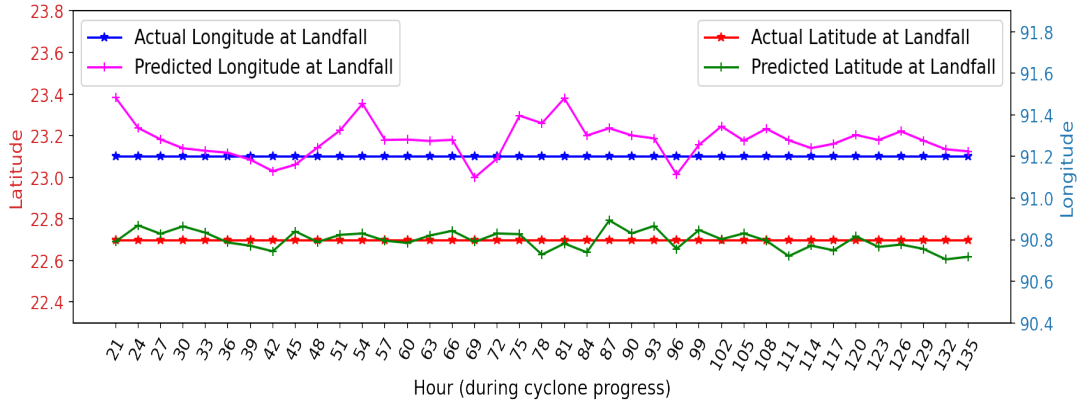


Figure 4.8: Predicted and actual landfall latitude/longitude of Bulbul cyclone for T = 8

It is once again evident that the model has consistently performed well irrespective of whether the prediction point is close to landfall or far from landfall.

Table 4.4: Comparison with IMD's 3/4/5 year (2015-2019) average reported error (KM) for landfall Location.

Lead Time (hours)	36	48	60	72	84	96
IMD	42.84	78.08	92.6	112.5	197	210
Proposed Model (5-fold Accuracy)	33.3	32.19	32.65	30.79	27.88	34.06

Table 4.5: Comparison with IMD's 3/4/5 year (2015-2019) average reported time(hours) to landfall.

Lead Time (hours)	36	48	60	72	84	96
IMD	4.96	5.53	6.8	9.6	5.4	8.83
Proposed Model (5-fold Accuracy)	7.79	6.22	5.51	6.50	8.32	5.58

The standard deviation among 5-fold for all three prediction problems is relatively low, which means the model performs consistently and can be reliably deployed for the above-stated prediction problems. We compare our results with the landfall forecasts reported by IMD on its website[4]. IMD provides average forecasting error of location and time of landfall for each year starting from 2003. Error-values reported from earlier years are large and unsuitable for a fair comparison. Therefore, we have calculated the last 3/4/5 years (as per data availability) average MAE scores reported at the IMD website and provided in Table 4.4 and Table 4.5. The lead time of X hours in Table 4.4 and 4.5 means that forecast was done X hours before the landfall. We have also obtained the 5-fold validation performance of our proposed model at various lead times.

[4] https://rsmcnewdelhi.imd.gov.in/landfall-forecast.php

From Table 4.4, we see that the reported distance error by IMD increases sharply with lead time, but our model performance remains consistent for various lead times. The accuracy of our model for prediction of time to landfall is quite comparable with that of IMD, as shown in Table 4.5.

4.3 Forecasting landfall's location and time in six ocean basins across the world

Next, we tried to replicate landfall's characteristics prediction work of NI ocean from the last section to all other ocean basins of the world. The obvious approach is to use the BTD provided by regional centers, but there is no consistency in the available variables and their recording methodology. The global IBTrACS data that records all regional BTDs has lots of missing values. So for a worldwide study of the problem mentioned above, we explore the reanalysis dataset ERA5.

4.3.1 Data

In this study, we have used two open-source datasets - IBTrACS version 4 Knapp *et al.* (2010) and reanalysis dataset ERA5. From the IBTrACS dataset, we took the information about the cyclone path in terms of latitude and longitudes. As we predict the landfall's location and time of a TC, our dataset consists of all those cyclones that hit the coastal region. Only data corresponding to the time points when a TC was moving over the ocean is taken into account for all such cyclones. If a cyclone moves from ocean to land and then from land to ocean and continues like this during its course, then the data corresponding to its presence over land is not considered. Also, if a cyclone moves from ocean to land and land to ocean multiple times, then each such movement is treated as a separate cyclone while preparing the dataset. We are not using features like MSSWS and ECP, which have many missing values. We extracted data from 1981 to 2020 provided at an interval of three hours. The trajectory of all TCs till their landfall considered in this study is shown in Figure 4.9.

The large-scale atmospheric circulation of wind at different pressure levels plays a crucial role in determining the track of a TC. To capture this information, we have used

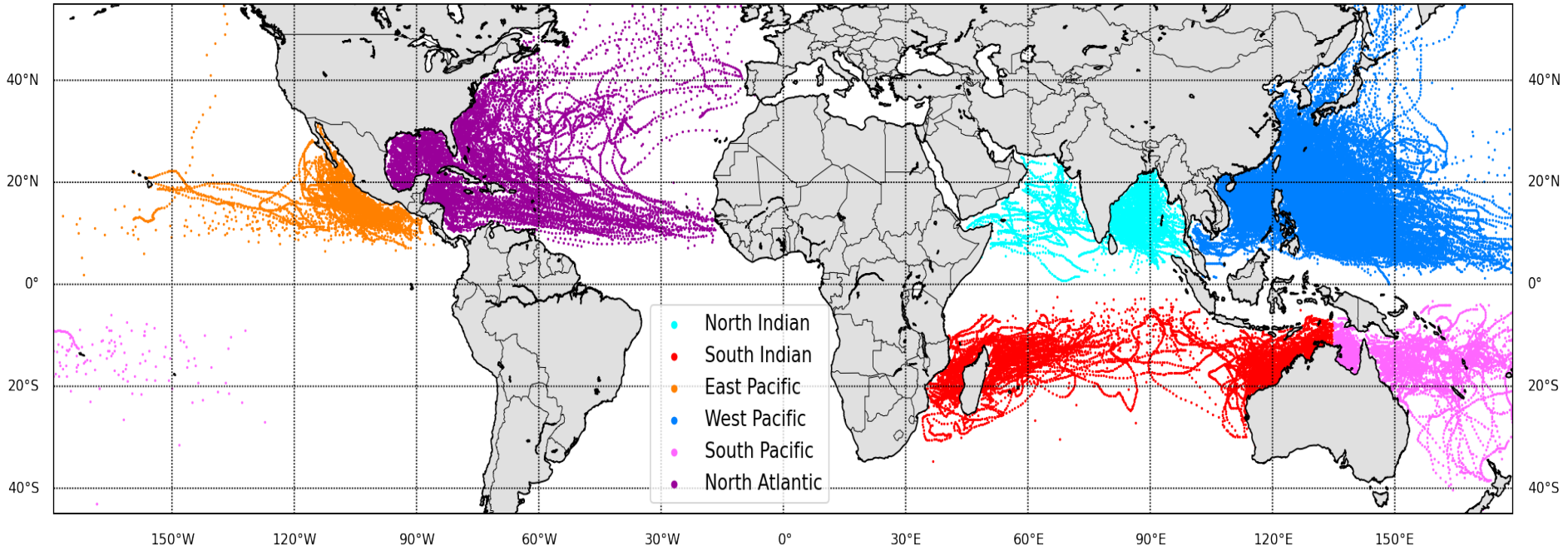


Figure 4.9: Trajectory of all cyclones till landfall in six ocean basins.

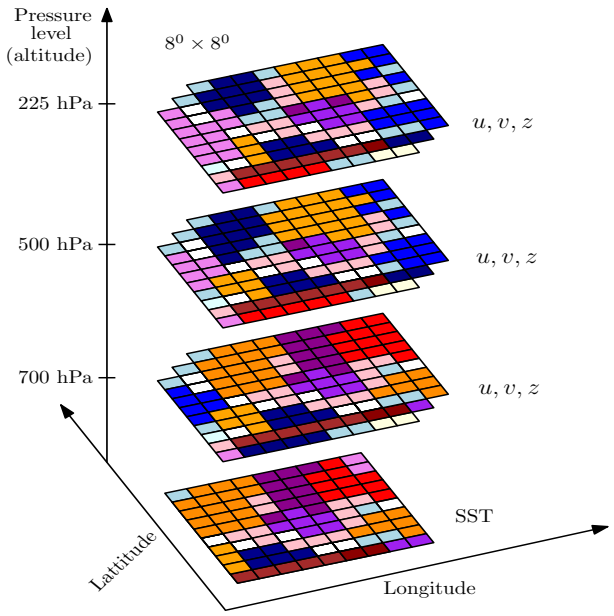


Figure 4.10: A pictorial depiction of u, v, z wind fields and SST

ERA5 (Hersbach *et al.*, 2020) reanalysis data produced by ECMWF. ERA5 is a fifth-generation reanalysis data covering global climate and weather since 1950 as discussed in section 1.2. We extracted the u, v components of wind and z geopotential fields at three atmospheric pressure levels (225hPa, 500hPa, 700hPa) for a spatial extension of 8×8 degree and spatial resolution of 0.25×0.25 degrees (resulting in a grid of size 33×33), centered at the current TC location. The u and v component of wind represents its eastward-westward and northward-southward movement, respectively. The z geopotential represents the gravitational potential energy of a unit mass relative to sea level. The choice of these variables is inspired by (Giffard-Roisin *et al.*, 2020; Boussioux *et al.*, 2020) where authors have used values of these three variables at mentioned three pressure levels for a spatial extension of 25×25 degrees and spatial resolution of 1×1 degree (resulting in a grid of size 25×25) centered at TC location. One degree equals 110 kilometers (KM) on the world map. This way, we are utilizing the mentioned variables values for a spatial spread of around 440 KM (in comparison of 1320 KM by earlier two studies) with a spatial resolution of about 27.5 KM (in comparison of 110 KM by earlier two studies) around the TC center. Apart from these three variables, we extracted the sea surface temperature (SST) for the 33×33 grid centered at the TC location for each time point. A pictorial depiction of reanalysis data is shown in Figure 4.10.

As we are using CNN, to feed the current location information, we have created two

Table 4.6: Dataset size and Landfall time (hours)

Ocean Basin	No. of TCs	Size of Dataset	Average Time	Min Time	Max Time
NI	205	5920	95.61	21	270
SI	282	10600	121.74	21	516
EP	116	4000	112.44	21	315
SP	189	5674	99.06	21	513
WP	1064	39166	119.43	21	606
NA	401	11386	94.18	21	531

more 33×33 matrices lats and longs for each time point of a TC. The each row of lats is equal to vector $(\mathrm{lat} + 0.25 * k| - 16 \leq k \leq 16)$ and each column of longs is equal to vector $(\mathrm{long} + 0.25 * k| - 16 \leq k \leq 16)$ where $(\mathrm{lat}, \mathrm{long})$ denotes the latitude and longitude of TC's current location. Feeding this information in CNN will enable it to generate distance and direction like features between two successive time points of a TC. In Table 4.6, the dataset size along with average landfall time after the initiation of a TC in six ocean basins are shown for all TCs with the minimum time difference between TC formation and its landfall as 21h.

Dataset Preparation

Let T be the number of continuous data points (that is $3(T-1)$ hours of data) taken in the model to predict the target. For a fixed cyclone, let T_L be the number of data points between cyclone formation and its landfall at an interval of three hours. The i^{th} data point is corresponds to data at $(i-1)$3h, in particular T_Lth data point corresponds to landfall that occurs at $(T_L - 1)$3h after cyclone formation. For this TC, we created $T_L - (T-1) - 4 = T_L - T - 3$ inputs, where a single input is a sequence of T vectors of the form:

$$(\,\mathrm{lats}(t), \mathrm{longs}(t), \mathrm{u225}(t), \mathrm{v225}(t), \mathrm{z225}(t), \mathrm{u500}(t),$$
$$\mathrm{v500}(t), \mathrm{z500}(t), \mathrm{u700}(t), \mathrm{v700}(t), \mathrm{z700}(t), \mathrm{SST}(t))$$

where $k \leq t \leq T + k - 1$ and k varies from 1 to $T_L - T - 3$. The target variables for each input are latitude and longitude at landfall or time (in hours) remaining to landfall of the cyclone from the current time t. One must note that by following the above process, we are predicting our target at least 12h before the landfall. For example, BELNA cyclone

formed at 00 hours on 05 December 2019 in EP ocean, and the landfall happened at 15 hours on 09 December 2019, that is $T_L = 38$. Suppose $T = 8$, then this TC will generate $38 - 8 - 3 = 27$ data points. The collection of all such inputs across all TCs for a particular ocean basin will form the dataset.

4.3.2 Model Implementation and Training

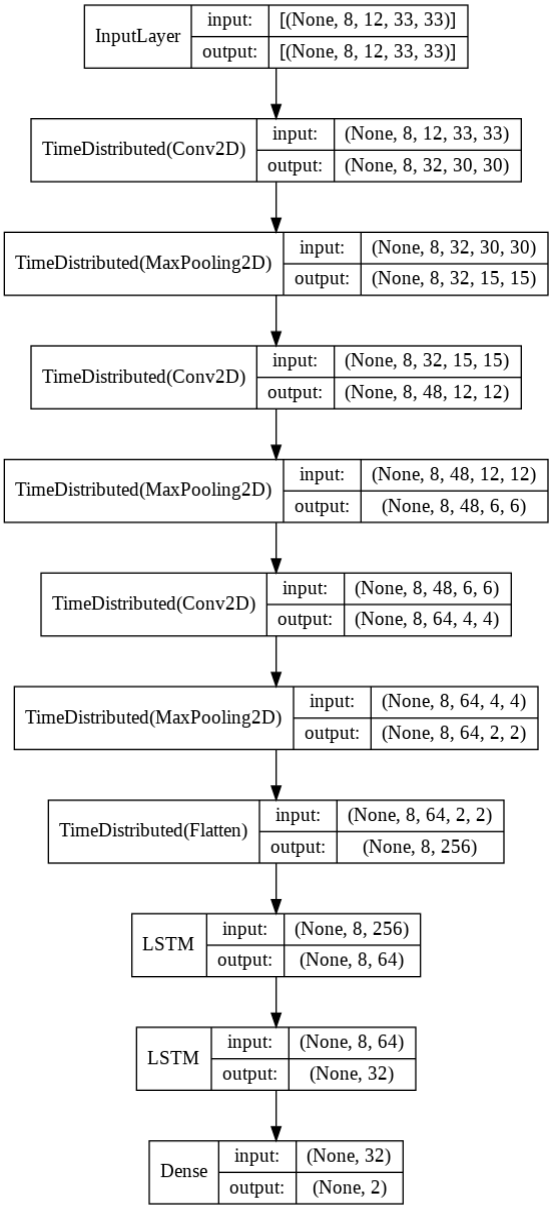


Figure 4.11: Model representation for latitude/longitude prediction for T = 8

We have used a combination of CNN and LSTM to capture our prediction problem's spatial and temporal aspects. We come up with a model that works well for each ocean basin. The structure of the model for location prediction in terms of latitude and longitude is described in the Figure 4.11 for $T = 8$. The model for time prediction

is the same except that the last dense layer's output size is one. It consists of three alternating convolution and max-pool layers that are enclosed within $\mathrm{TimeDistributed}$ layer to generate sequential features of length T, which are fed into stacked LSTM of two layers followed by a Dense layer. The input-output shape of each layer is shown in Figure 4.11. The model is implemented in Keras API (Chollet, 2015) which uses underlying low-level language TensorFlow (Abadi *et al.*, 2015). The input shape of a training data is $(T, 12, 33, 33)$, where T represents the number of sequential data points (corresponding to $3 * (T-1)$ hours of data), 12 represents the number of channels (nine channels of wind and geopotential fields corresponding to u, v and z at three pressure levels, one of SST, two of latitudes and longitudes), and $(33, 33)$ is the shape of the grid centered at TC center.

Training

We tried a few configurations of the above-described model by experimenting with the number of layers, number of nodes in a layer, activation function, and learning rate and reported the one which gives the best results across all the ocean basins. Finally, we have reported the 5-fold cross validation accuracy of our model as described in Section 2.3.1. We trained the model for T equals to 4, 6, and 8; increasing T further does not lead to improved accuracy. The input features are scaled using Standard Scaler of Scikit learn library (Pedregosa *et al.*, 2011), which is given by $f(x) = \frac{x-\mu}{\sigma}$, where μ is the mean and σ is the standard deviation. The target variables are also scaled in case of landfall's latitude and longitude prediction but not for time prediction. The model uses the mean square error (MSE) as the loss function. We have reported the model performance in terms of root mean square error (RMSE) and mean absolute error (MAE), which are defined as follows:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y_i})^2, \quad \mathrm{RMSE} = \sqrt{MSE}, \quad \mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \bar{y_i}|,$$

where y_i is the actual value and $\bar{y_i}$ is the predicted value. The model uses the Adam (Kingma and Ba, 2014) optimizer with a default learning rate 0.001 to minimize the loss function. Convolution layers and LSTM layers uses the activation function $\mathrm{ReLU}(x) = \max(0, x)$ (Nair and Hinton, 2010). The model uses a batch size of 256 and total of 100 epochs. We run our experiments on Nvidia Tesla V100 GPU with 16 GB RAM.

The model takes approximately 30 to 45 minutes (depending on the ocean basin) to complete 100 epochs.

4.3.3 Results and Comparison

Our model, at any time t during the progress of a TC, takes T number ((T-1)*3 h) of data and predicts the landfall's latitude, longitude, and time. For example, if $T = 8$ and a particular TC is at the $t = 52h$ during its progression, then using the data between the time $52 - 21 = 31h$ and t. the model predicts the landfall's characteristics. To avoid any bias when time t is very close to landfall's time, the model predicts only for $t \leq L - 12h$, where L is the landfall's time, which means we are remaining at least 12 hours away from landfall while predicting. For each ocean basin and for different values of $T = 4, 6, 8$ (9h, 15h, 21h). We have reported the size of the dataset and the 5-fold cross validation score as discussed in section 2.3.1 in terms of RMSE and MAE along with standard deviation in Table 4.7. From the predicted latitude and longitude, the distance error between actual landfall location and predicted location is also reported in Table 4.7.

From Table 4.7, we can see that we can predict the landfall's location in ocean basins NI, SI, EP, SP, WP, and NA with a distance error of 66.18KM, 119.96KM, 110.48KM, 144.80KM, 108.0KM, and 158.92KM, respectively for $T = 8$ (21h) with a low standard deviation (std). If we look at the landfall's time prediction results, they are pretty impressive. The model predicts the landfall's time in six ocean basins NI, SI, EP, SP, WP, and NA with an MAE of 4.71h, 6.04h, 8.20h, 6.74h, 5.89h, and 7.42h, respectively, with low standard deviation.

To further demonstrate the working of our model, in Figures 4.12 and 4.13, we have shown landfall's latitude-longitude prediction results for two cyclones BUD (2018) and BELNA (2019) in ocean basins EP and SI, respectively. In Figures 4.14 and 4.15, the landfall's time prediction results for TCs FANI (2019) and NAKRI (2019) in ocean basins NI and WP, respectively, are shown. One can note that the model start predicting at time $t = 21$ and uses data of 21 hours between $t - 21$ and t to predict at time t. One can observe that the model works consistently well irrespective of whether we are close to landfall or away from landfall. All these named cyclones are not part of training data.

Table 4.7: 5-fold validation score of landfall's location and time prediction for different T (Hochreiter and Schmidhuber, 1997; Gers *et al.*, 1999; LeCun *et al.*, 1989; Krizhevsky *et al.*, 2012).

Ocean Basin	**T (hrs)**	**Dataset Size**	**RMSE (std)**			**MAE (std)**			
			Lati degree	**Long** degree	**Time** hours	**Lati** degree	**Long** degree	**Time** hours	**Distance** KM
North Indian	4 (9)	5060	0.58 (0.05)	1.03 (0.09)	10.46 (1.40)	0.38 (0.02)	0.70 (0.05)	7.33 (1.50)	93.78 (6.71)
	6 (15)	4660	0.53 (0.13)	0.90 (0.13)	8.98 (3.47)	0.36 (0.12)	0.59 (0.10)	6.05 (2.38)	81.92 (18.62)
	8 (21)	4284	0.40 (0.04)	0.76 (0.06)	6.72 (0.58)	0.26 (0.02)	0.50 (0.03)	4.71 (0.54)	66.18 (2.87)
South Indian	4 (9)	9441	0.53 (0.03)	1.78 (0.22)	13.17 (0.59)	0.36 (0.02)	1.31 (0.15)	8.32 (0.31)	150.78 (15.55)
	6 (15)	8886	0.46 (0.03)	1.44 (0.06)	11.34 (0.84)	0.30 (0.03)	1.06 (0.04)	7.24 (0.42)	123.32 (5.24)
	8 (21)	8353	0.42 (0.02)	1.42 (0.18)	9.63 (0.96)	0.27 (0.03)	1.05 (0.10)	6.04 (0.45)	119.96 (12.22)
East Pacific	4 (9)	3505	0.70 (0.23)	1.49 (0.33)	10.40 (1.03)	0.52 (0.18)	1.08 (0.24)	7.17 (0.80)	133.43 (33.42)
	6 (15)	3276	0.62 (0.04)	1.33 (0.24)	9.89 (2.17)	0.46 (0.03)	0.95 (0.10)	6.99 (1.74)	117.8 (9.76)
	8 (21)	3056	0.52 (0.03)	1.26 (0.16)	11.28 (3.69)	0.37 (0.02)	0.93 (0.03)	8.20 (2.96)	110.48 (2.86)
South Pacific	4 (9)	4885	0.89 (0.09)	2.12 (0.14)	17.44 (2.70)	0.55 (0.05)	1.50 (0.14)	11.30 (2.19)	179.03 (17.26)
	6 (15)	4520	0.89 (0.19)	2.30 (0.57)	13.24 (1.60)	0.57 (0.14)	1.57 (0.33)	7.69 (0.65)	188.81 (37.51)
	8 (21)	4182	0.72 (0.10)	1.67 (0.23)	10.07 (1.77)	0.44 (0.07)	1.23 (0.20)	6.74 (1.04)	144.80 (22.44)
West Pacific	4 (9)	34874	1.34 (0.34)	1.72 (0.37)	10.84 (0.30)	0.88 (0.24)	1.15 (0.26)	7.44 (0.41)	164.17 (40.08)
	6 (15)	32777	1.04 (0.16)	1.40 (0.22)	10.10 (2.14)	0.72 (0.13)	0.97 (0.16)	7.22 (1.62)	137.86 (24.28)
	8 (21)	30791	0.79 (0.10)	1.09 (0.10)	8.12 (0.93)	0.56 (0.08)	0.76 (0.06)	5.89 (0.72)	108.0 (11.61)
North Atlantic	4 (9)	9782	1.28 (0.12)	2.42 (0.24)	15.10 (4.20)	0.84 (0.04)	1.47 (0.11)	10.47 (3.40)	174.51 (10.68)
	6 (15)	8999	1.17 (0.15)	2.13 (0.32)	10.30 (1.26)	0.79 (0.08)	1.32 (0.14)	6.58 (0.96)	161.74 (15.36)
	8 (21)	8276	1.05 (0.04)	2.10 (0.34)	10.69 (1.14)	0.71 (0.03)	1.38 (0.13)	7.42 (0.98)	158.92 (12.62)

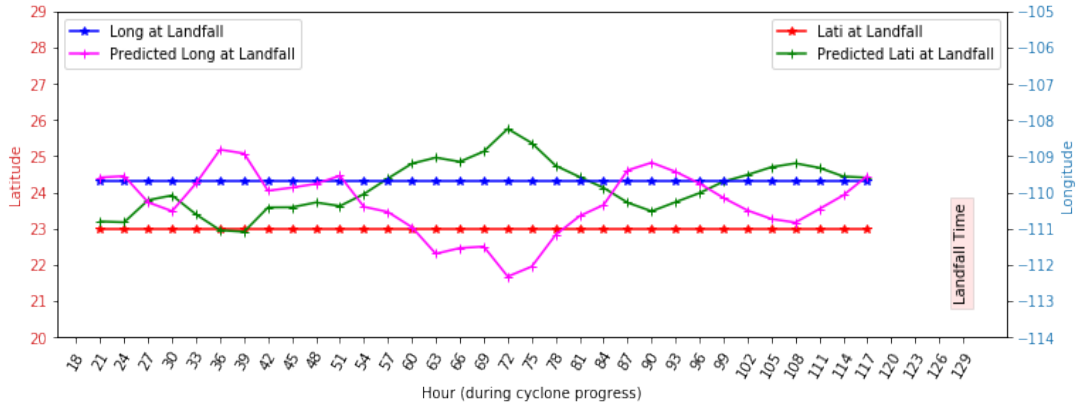


Figure 4.12: Latitude and Longitude prediction for hurricane BUD (2018) in East Pacific ocean for T = 8

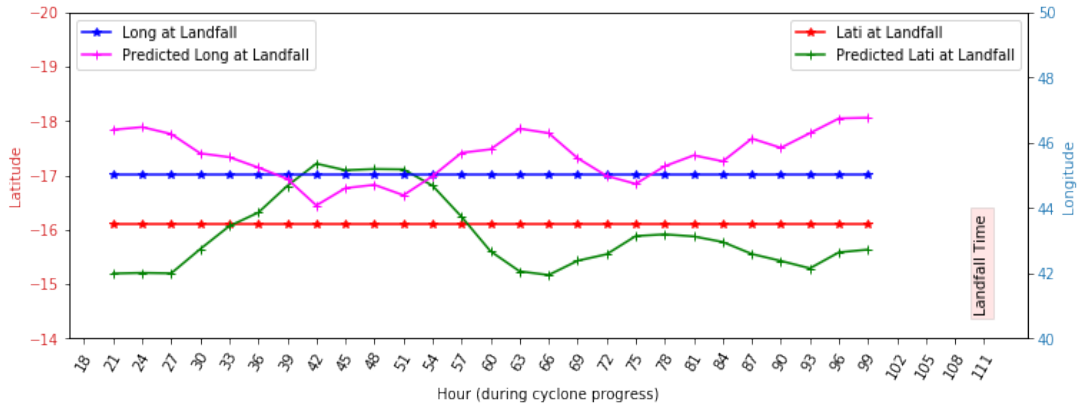


Figure 4.13: Latitude and Longitude prediction for cyclone BELNA (2019) in South Indian ocean for T = 8

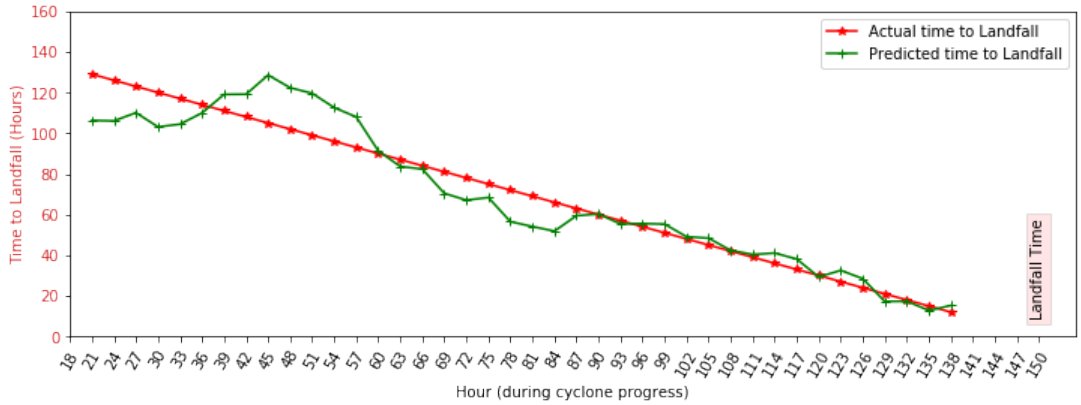


Figure 4.14: Time to Landfall prediction for Fani (2019) cyclone in North Indian ocean for T = 8

4.3.4 Comparison

As per our knowledge, there does not exist any earlier work that predicts the landfall's location and time using reanalysis data. In the absence of directly related work, we will compare our results with closely related works (Giffard-Roisin *et al.*, 2020; Boussioux

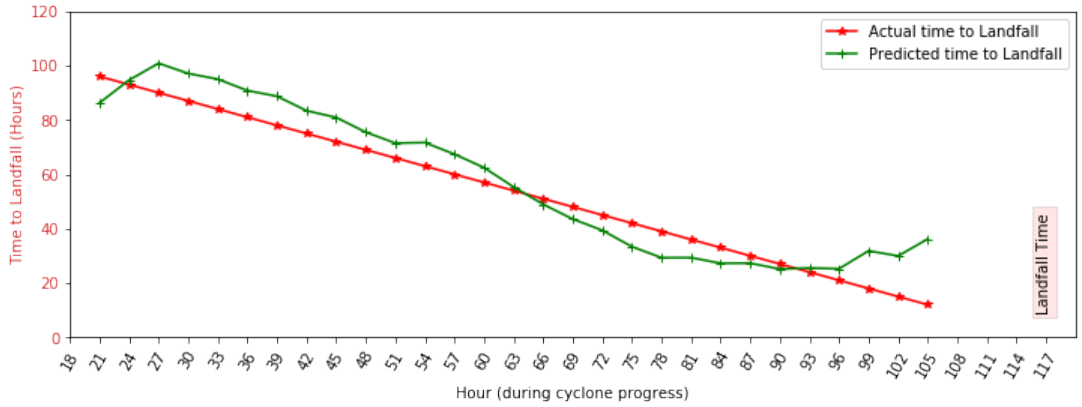


Figure 4.15: Time to Landfall prediction for NAKRI (2019) cyclone in West Pacific ocean for T = 8

et al., 2020). In Table 4.8, the compassion between Fusion model (Giffard-Roisin *et al.*, 2020) and our proposed model is shown. In (Giffard-Roisin *et al.*, 2020) authors used TC reanalysis data of 12 hours at a temporal resolution of 6 hours for track prediction with a lead time of 24 hours. A comparison with our model for $T = 8$ is provided in Table 4.8. We have achieved better results for a much harder problem with arbitrary lead time ($\geq 21h$) depending on the time between TC's initiation and its landfall. The authors of (Giffard-Roisin *et al.*, 2020) have also reported the error range (std) 71KM for a subset of cyclones in the Atlantic ocean, while our model achieves an error range (std) 2.87KM - 22.44KM across all ocean basins. This makes our model more robust and reliable for practical purposes.

In Boussioux *et al.* (2020) authors use a combination of historical TC data, reanalysis data, and output from operational models and propose eight models for track prediction of TCs with 24h lead time in NA and EP ocean basins for years 2017 to 2019. The model which closes to our proposed model is $\mathrm{HURR}\text{-}(\mathrm{viz}, \mathrm{CNN}/\mathrm{GRU})$ which uses CNN-encoder and GRU-decoder (Chung *et al.*, 2014). The authors consider only those hurricanes in their study for which MSSWS reaches 34 knots at some time t_0 and contains at least 60h of data after t_0. Our model does not have any such restrictions. In Table 4.9, we have shown the results for $\mathrm{HURR}\text{-}(\mathrm{viz}, \mathrm{CNN}/\mathrm{GRU})$ and the operational CLP5 model for EP and NA ocean basin. One can see that our results are not as good as that of $\mathrm{HURR}\text{-}(\mathrm{viz}, \mathrm{CNN}/\mathrm{GRU})$ but quite comparable with the operational model CLP5. One can notice that our model's standard deviation is relatively low compared to these two models. Here, we would again point out that we are not making a direct comparison here as our target prediction problem is different from this work but, at the

Table 4.8: Comparison in terms of distance MAE (KM)

Model /Ocean Basin	NI (KM)	SI (KM)	EP (KM)	SP (KM)	WP (KM)	NA (KM)	No. of parameters ($*10^6$)	Training Time	Lead Time	Spacial Extent (Resolution)
Fusion (Track) Giffard-Roisin *et al.* (2020)	138.9	136.1	106.9	161.7	136.1	130.2	≥ 2.27	8h	24h	$25° \times 25°$ ($1° \times 1°$)
Proposed Model (Landfall)	66.1	119.9	110.4	144.8	108.0	158.9	0.157	$\leq$ 0.75h	TC dependent ($>= 12h$)	$4° \times 4°$($0.25° \times 0.25°$)

Table 4.9: Track prediction MAE (std) in KM for HURR model and operational CLP5 model for 24 hour lead time

	Model/Year	2017 (10 TCs)	2018 (15 TCs)	2019 (12 TCs)
EP	HURR-(viz, cnn/gru)	74 (40)	69 (42)	73 (45)
	CLP5	114 (59)	109 (61)	133 (74)
NA	HURR-(viz, cnn/gru)	94 (61)	113 (77)	123 (89)
	CLP5	189 (135)	199 (118)	207 (171)

same time, much more challenging and important.

We do not find any meteorological department across the world which reported the landfall prediction accuracy except the Indian Meteorological Department (IMD) on its website[5]. IMD has reported landfall's location error and time for a certain number of lead hours. In Table 4.4 and 4.5, we have reported the last 3/4/5 years (as per data availability) MAE achieved by IMD for landfall's location and time prediction. From Table 4.6, we can see that in the NI ocean basin, the landfall occurs on average at 95.61 hours. Therefore, it is reasonable to compare our results with that of IMD for 72 lead hours. Clearly, our model performs better than that of models used by IMD for both landfall's location and time. We have not included the results reported by IMD for earlier years, as errors are much higher.

4.4 Conclusion

In section 4.1, we presented a model which used LSTM network based on RNN to predict the intensity, location, and time of landfall of a tropical cyclone in the North Indian Ocean. The model predicts the landfall characteristics with high accuracy and beats the model used by India Meteorological Department on recent cyclones. Using our model, the landfall characteristics of a cyclone can be predicted with high accuracy post a few hours after the origin of the cyclone which will provide ample time to disaster managers to decide if they need to evacuate certain areas (depending on intensity prediction) and if they need to then precisely when (depending on time prediction) and which area (depending on location prediction). In section 4.2, we proposed a model that can predict the landfall's location and time of a TC with high accuracy by observing a TC for 9h, 15h, or 21h at any time of its progression in the world's six ocean basins. The model took only 30 to 45 minutes for training and can predict the landfall characteristics within a few seconds, making our model suitable for practical usage. Our model supports the case that deep learning models like CNN and LSTM can be utilized to predict challenging and complex prediction problems like the landfall of a TC.

One can further work to utilize CNN with Attention and Transformers models for further improvement. One can also develop Consensus models to solve the proposed

[5] https://rsmcnewdelhi.imd.gov.in/landfall-forecast.php

prediction problem. As we can solve a complex landfall prediction problem using data over a small spatial extent with high resolution, it will be interesting to see to use the same kind of data for track or intensity prediction problems. In future work, one can include the characteristics of terrain from which the cyclone is passing as an input feature in the model.

CHAPTER 5

Track Prediction of Tropical Cyclones[1]

One of the most significant prediction problem related to TCs is track prediction, where the track (in terms of latitude and longitude) of a TC for the next few hours is predicted using the TC's data of the last few hours. In section 4.1, we have discussed the existing TC's track prediction-related work. Recently, in (Alemany *et al.*, 2018) a new approach for track prediction is suggested for the Atlantic ocean, where instead of predicting the track of a TC directly in terms of latitude and longitude, the authors predicts a GridID, which is obtained by applying a grid function on the ranges of latitudes and longitudes. In this work, we adopted the same approach and predicted the GridID for the Atlantic and NI ocean basins. It is an incremental work built on the work of (Alemany *et al.*, 2018) in the sense that we proposed a much-simplified model with larger prediction time and better accuracy. The proposed model outperforms the existing model in terms of mean absolute error, time-span of prediction, and training time. Next, we will describe the dataset used, the proposed model, results, and future directions.

5.1 Data

For this study, we pick two BTDs, Atlantic hurricane data from the NOAA database (available on the GitHub page[2] of one of the authors of (Alemany *et al.*, 2018)) and BTD of tropical cyclones in the NI ocean from the Regional Specialized Meteorological Centre, New Delhi[3].

The Atlantic ocean data contains the record of hurricanes from 1920 to 2012, and recordings are provided at 6 hours time intervals. The data consists of a total of 16394 recordings. Each recording contains information about the hurricane's year, time, name

[1]This chapter is a slightly modified version of work published in IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC) 2021 and has been reproduced here with the permission of the copyright holder.

[2]https://github.com/sheilaalemany/hurricane-rnn

[3]https://rsmcnewdelhi.imd.gov.in/report.php?internal_menu=MzM=

(if any), latitude, longitude, ECP, and MSSWS. The trajectories of all the hurricanes are shown in the Figure 5.1. After removing missing and invalid data, there are a total of 789 hurricanes with an average number of 21 recordings. The largest cyclone has 96 such recordings. To capture the relative change of hurricane track progression, two more features, distance, and direction of change[4], are generated between two consecutive recordings of a hurricane. The features ECP, MSSWS, distance, and direction are used for the study.

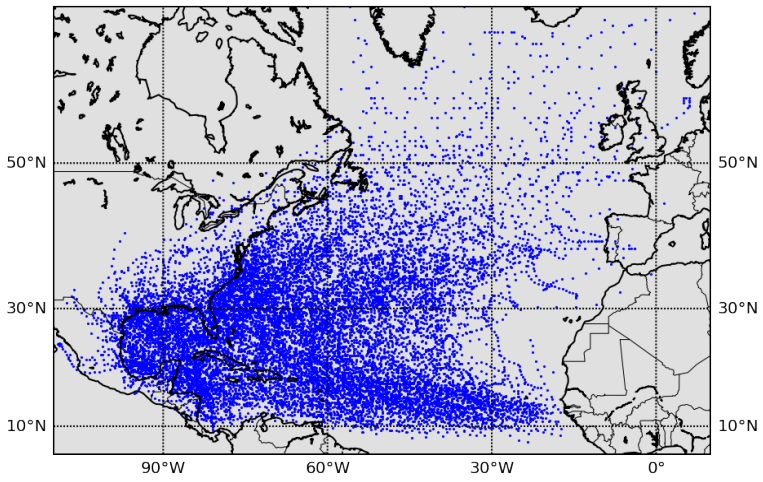


Figure 5.1: Atlantic hurricanes' track.

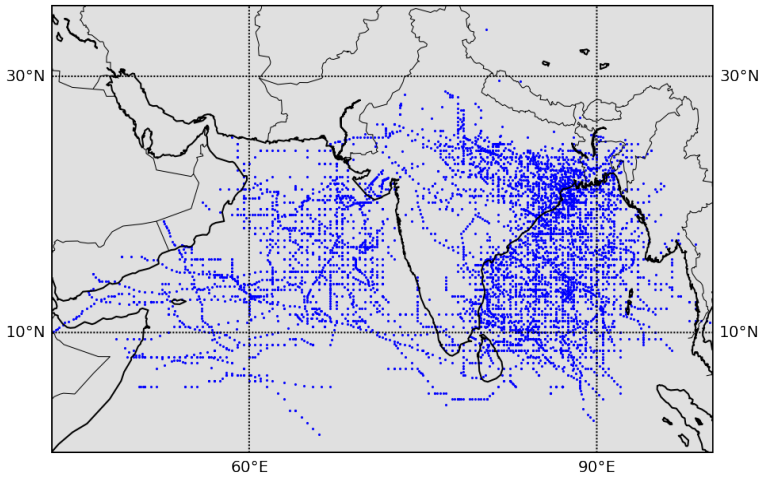


Figure 5.2: North Indian ocean cyclones' track.

The BTD data for NI ocean contains the record of 341 cyclones from 1990 to 2017 in the north Indian ocean, and recordings are provided at 3 hours time interval. The data has a total of 7662 recordings. Each such recording contains information about the cyclone's basin of origin, latitude, longitude, ECP, MSSWS, date and time of occurrence.

[4] https://www.movable-type.co.uk/scripts/latlong.html

The trajectories of all the cyclones are shown in the Figure 5.2. We removed recordings with manual errors, that resulted in a dataset with 341 cyclones, with an average number of 27 recordings. The longest cyclone has 90 such recordings. Imputing techniques are applied to handle the missing data. We generate two new features, distance, and directions, similar to the Atlantic hurricane data and use MSSWS, ECP, distance, and direction as input features in the model.

5.1.1 GridID

To predict the track of a cyclone, we need to estimate latitude and longitude at future times. Instead of estimating latitude and longitude directly, we take the grid approach as done in (Alemany *et al.*, 2018). We divide the region of latitude and longitude into square grid blocks, that is, $x^\circ \times x^\circ$ for some fixed x° along latitude and longitude. Let $[a_{\min}, a_{\max}]$ and $[b_{\min}, b_{\max}]$ be the ranges of latitude and longitude in the data respectively. Then, there will be a total of

$$\left\lfloor \frac{(a_{\max} - a_{\min})(b_{\max} - b_{\min})}{x^2} \right\rfloor$$

grid blocks. We assign a GridID to each grid block, see Figure 5.3 for an example. For the block in ith row (starting from below) and jth column, we assign GridID

$$\left\lfloor \frac{b_{\max} - b_{\min}}{x} \right\rfloor (j-1) + i.$$

Each vertex in the grid has co-ordinates of the form $(a_{\min} + ix, b_{\min} + jx)$. Now, given any general latitude and longitude coordinate (a, b) in the region of the grid, we will have $a = a_{\min} + ix + \alpha, \quad b = b_{\min} + jx + \beta$ for some $i, j \in \mathbb{N}$ and $0 \leq \alpha, \beta < x$. Then,

$$i = \left\lceil \frac{a - a_{\min}}{x} \right\rceil, \qquad j = \left\lceil \frac{b - b_{\min}}{x} \right\rceil.$$

Therefore, we can define a GridID function for given latitude a and longitude b as

$$\text{GridID}(a, b) = \left\lfloor \frac{b_{\max} - b_{\min}}{x} \right\rfloor \left(\left\lceil \frac{b - b_{\min}}{x} \right\rceil - 1 \right) + \left\lceil \frac{a - a_{\min}}{x} \right\rceil,$$

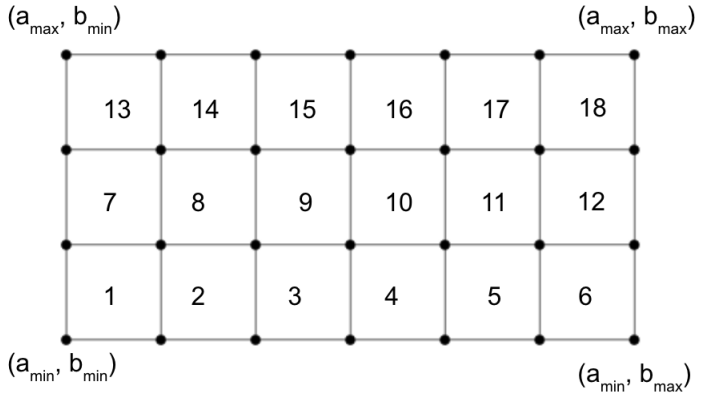


Figure 5.3: Example of a 6×3 grid.

which will assign (a, b) the GridID for the grid block in which it lies. We use the scaled GridID as our output variable.

For the Atlantic hurricane data, we have $(a_{\min}, a_{\max}) = (7, 66)$ and $(b_{\min}, b_{\max}) = (-110, 14)$ and for the Indian ocean cyclone data, we have $(a_{\min}, a_{\max}) = (1, 41)$ and $(b_{\min}, b_{\max}) = (42, 103)$ which generate a total of 7316 and 2440 grid blocks of size $1^\circ \times 1^\circ$ (for $x = 1$) respectively. The grid blocks have almost same area as we are studying tropical cyclones. On average, any point in a $1^\circ \times 1^\circ$ grid block is within 50 kilometers from the center of the block. Therefore, by predicting GridID instead of latitude and longitude, we create a margin error of 50 kilometers on average which is acceptable for practical purposes.

5.1.2 Dataset Preparation

For a TC, let S be the number of data points available at an interval of D hours. Suppose we want to use T_1 number of data points ($(T_1 - 1) * D$ hours of data) to predict the GridID for next T_2 number of time points (at an interval of D hours). We generate $S - T_1 - T_2 + 1$ data points, where a single data point is a sequence of T_1 vectors of the form:

$$(\text{MSSWS}(t), \text{ECP}(t), \text{distance}(t), \text{direction}(t), \text{GridID}(t)), k \leq t \leq T_1 + k - 1$$

and output is a sequence of GridIDs (scaled between 0 and 1) for the next T_2 time points, that is:

$$(\text{GridID}(T_1 + k), \text{GridID}(T_1 + k + 1), \text{GridID}(T_1 + k + 2), \cdots, \text{GridID}(T_1 + k + T_2 - 1))$$

where k varies from 1 to $S - T_1 - T_2 + 1$. The collection of all such data points for all the cyclones form the dataset.

5.2 Model

As the dataset is time-series data, we have used a stacked LSTM model. Here we describe the best model that gives the optimal results after trying various configurations. The input size to the model is (T, 5), where T corresponds to the length of sequential data and 5 is the number of variables. The proposed model consists of four layers: an input layer, two hidden layers, and an output layer. The hidden layers are of size 200 each. To compare, in (Alemany *et al.*, 2018), three hidden LSTM layers have been used. We have used Mean Square Error (MSE) as the loss function, Adam (Kingma and Ba, 2014) optimizer function with a default learning rate of 0.01, and activation function ReLU (Nair and Hinton, 2010). The MSE and Mean Absolute Error (MAE) is defined as:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y_i})^2, \text{RMSE} = \sqrt{MSE}, \text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \bar{y_i}|$$

where y_i is the actual value and $\bar{y_i}$ is the model predicted value. Though we are using MSE as a loss function that model minimizes, we have reported model performance in terms of MAE for comparison with related work. We used Keras API (Chollet, 2015) to implement our RNN architecture on Google Colab (Bisong, 2019). Google Colab provides one of the GPU Nvidia K80s, T4s, P4s or P100s as per availability. The model we used takes only around 50 seconds to fit on the training data. Moreover, the model used in (Alemany *et al.*, 2018) took around 200 seconds for training. Therefore, our model outperforms both in terms of MAE and training time. The model is described in Figure 5.4, for $T_1 = 6$ and $T_2 = 10$, which is generated using Keras API.

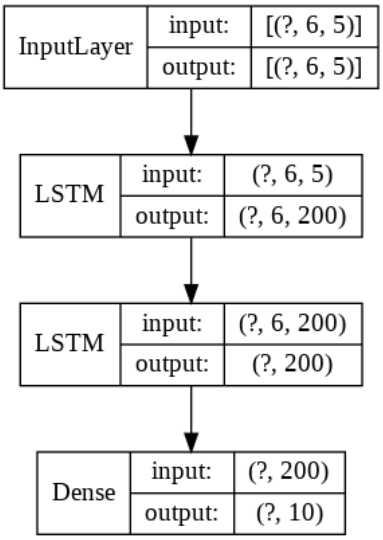


Figure 5.4: Model structure for T_1 = 6, T_2 = 10.

5.3 Results and Analysis

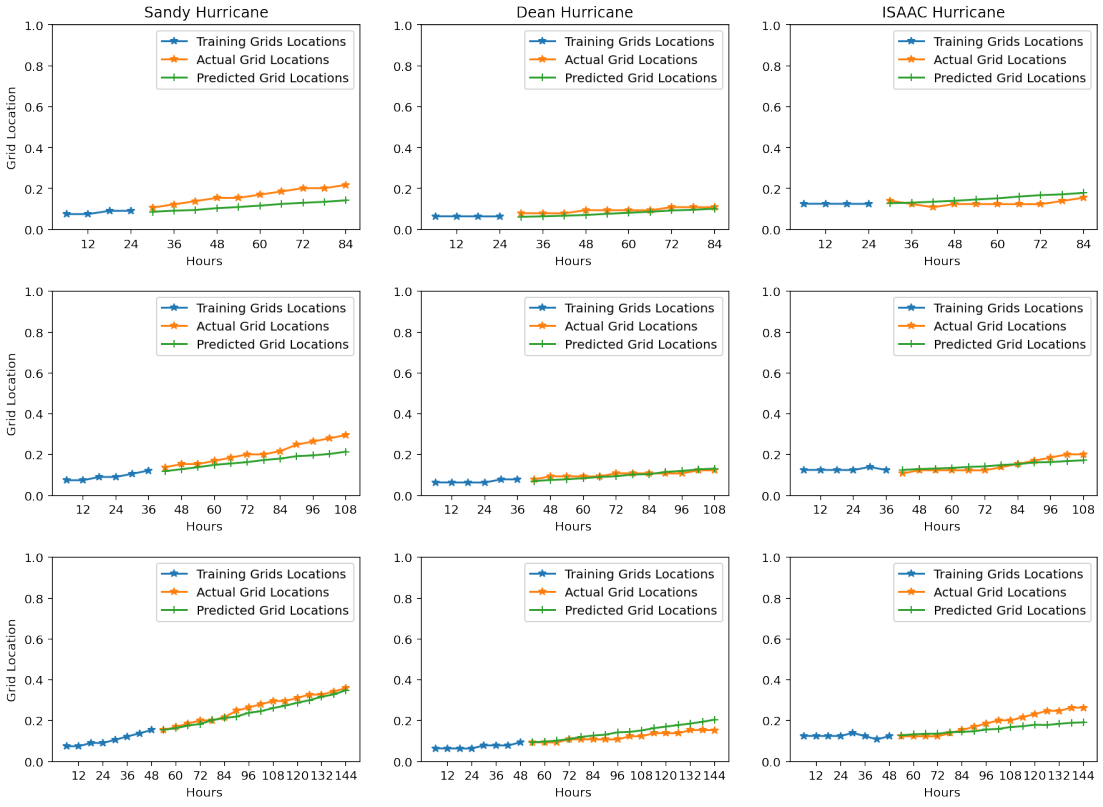


Figure 5.5: Actual and predicted grid predictions for Sandy, Dean and ISAAC for T_1 = 4, 6, 8 and T_2 = 10, 12, 16.

Let T_1 be the length of the training data point and T_2 be the number of predicted time points of a cyclone's GridID. For example, if T_1 is four and T_2 is eight, then we need a model that uses four consecutive time points cyclone data and predicts the GridID for the next eight successive time points. The proposed model is tested for different

Table 5.1: Model performance on Atlantic hurricane data (Hochreiter and Schmidhuber, 1997; Gers *et al.*, 1999).

Data Used			Mean Absolute Error			
Training Points T_1 (hours)	Predict Points T_2 (hours)	Training data size (no. of cyclones)	5-fold MAE	Sandy	Dean	ISAAC
4 (24)	1 (6)	13449 (611)	0.0093	0.0092	0.0060	0.0101
	2 (12)	12838 (592)	0.0127	0.0111	0.0083	0.0087
	4 (24)	11671 (549)	0.0179	0.0162	0.0084	0.0125
	6 (36)	10592 (503)	0.0274	0.0226	0.0176	0.0176
	8 (48)	9606 (464)	0.0287	0.0277	0.0221	0.0212
	10 (60)	8698 (419)	0.0337	0.0290	0.0288	0.0290
	12 (72)	7877 (386)	0.0380	0.0341	0.0219	0.0236
	16 (96)	6409 (334)	0.0468	0.0374	0.0343	0.0283
	20 (120)	5170 (269)	0.0556	0.0400	0.0501	0.0344
6 (36)	1 (6)	12246 (575)	0.0086	0.0145	0.0087	0.0129
	4 (24)	10592 (503)	0.0186	0.0181	0.0142	0.0142
	8 (48)	8698 (419)	0.0291	0.0270	0.0160	0.0220
	12 (72)	7119 (363)	0.0385	0.0290	0.0304	0.0299
	16 (96)	5756 (301)	0.0472	0.0292	0.0319	0.0311
	20 (120)	4642 (250)	0.0550	0.0516	0.0304	0.0280
	24 (144)	3708 (209)	0.0615	0.0674	0.0400	0.0280
8 (48)	1 (6)	11122 (530)	0.0099	0.0123	0.0131	0.0118
	8 (48)	7877 (386)	0.0304	0.0297	0.0212	0.0251
	16 (96)	5170 (269)	0.0500	0.0296	0.0361	0.0356
	24 (144)	3299 (188)	0.0633	0.0635	0.0452	0.0247
12 (72)	1 (6)	9142 (444)	0.0107	**0.0104**	**0.0120**	**0.0084**
	8 (48)	6409 (334)	0.0316	0.0223	0.0221	0.0208
	16 (96)	4158 (231)	0.0550	0.0284	0.0321	0.0380
	24 (144)	2611 (153)	0.0698	0.0370	0.0453	0.0299

combinations of T_1 and T_2. The resulting MAE for various combinations of T_1 and T_2 are reported in Table 5.1 (for Atlantic hurricane data) and Table 5.2 (for Indian ocean cyclone data). For each combination of T_1 and T_2, the size of the training set along with the number of cyclones with at least $T_1 + T_2$ number of data points are reported. We have reported the MAE of 5-fold cross-validation. To further establish the efficacy of our model, we have reported the MAE for the three most significant hurricanes in the Atlantic ocean - Sandy, Dean, and ISAAC, and the two most devastating cyclones - Vayu and Fani in the Indian ocean in recent times. All these named cyclones are not part of the training data.

In the case of Atlantic hurricane data, our results outperform the grid prediction

Table 5.2: Model performance on Indian ocean cyclone data (Hochreiter and Schmidhuber, 1997; Gers *et al.*, 1999).

Data Used			Mean Absolute Error		
Training Points T_1 (hours)	Predict Points T_2 (hours)	Training data size (no of cyclones)	5-fold MAE	Vayu	Fani
4 (12)	1 (3)	6428 (295)	0.0104	0.0092	0.0114
	2 (6)	6133 (290)	0.0127	0.0152	0.0112
	4 (12)	5560 (279)	0.0156	0.0118	0.0266
	6 (18)	5011 (247)	0.0198	0.0173	0.0224
	8 (24)	4524 (236)	0.0233	0.0211	0.0395
	10 (30)	4063 (217)	0.0264	0.0225	0.0428
	12 (36)	3637 (204)	0.0301	0.0281	0.0274
	16 (48)	2865 (174)	0.0334	0.0369	0.0348
	20 (60)	2214 (149)	0.0364	0.0412	0.0296
6 (18)	1 (3)	5843 (283)	0.0107	0.0102	0.0095
	4 (12)	5011 (247)	0.0159	0.0102	0.0185
	8 (24)	4063 (217)	0.0234	0.0304	0.0170
	12 (36)	3233 (187)	0.0316	0.0322	0.0271
	16 (48)	2527 (159)	0.0359	0.0336	0.0318
	20 (60)	1918 (136)	0.0386	0.0320	0.0412
	24 (72)	1421 (104)	0.0452	0.0362	0.0431
8 (24)	1 (3)	5281 (270)	0.0122	0.0068	0.0111
	8 (24)	3637 (204)	0.0260	0.0230	0.0184
	16 (48)	2214 (149)	0.0426	0.0282	0.0434
	24 (72)	1219 (92)	0.0469	0.0368	0.0493
12 (36)	1 (3)	4288 (225)	0.0137	0.0136	0.0163
	8 (24)	2865 (174)	0.0258	0.0174	0.0173
	16 (48)	1652 (122)	0.0441	0.0557	0.0321
	24 (72)	881 (69)	0.0527	0.0455	0.0600
	32 (96)	427 (40)	0.0540	0.0512	0.0558

results obtained in (Alemany *et al.*, 2018). In (Alemany *et al.*, 2018), grid prediction results have been reported only for $T_1 = 12$ and $T_2 = 1$ and the MAE reported for Sandy, Dean and ISAAC are 0.0800, 0.0842 and 0.0592 respectively. For the same combination of T_1 and T_2 for these hurricanes, our results are eight times better (see the values in bold in Table 5.1). Figure 5.5 further supports our claim, which provides the plots of training, actual and predicted grid locations for Sandy, Dean, and ISAAC for different combinations of T_1 and T_2. Moreover, the MAE for 5-fold cross-validation consistently remains low for different combinations of T_1 and T_2 and increase gradually with T_2, for both the datasets. A possible reason for the improved accuracy over the

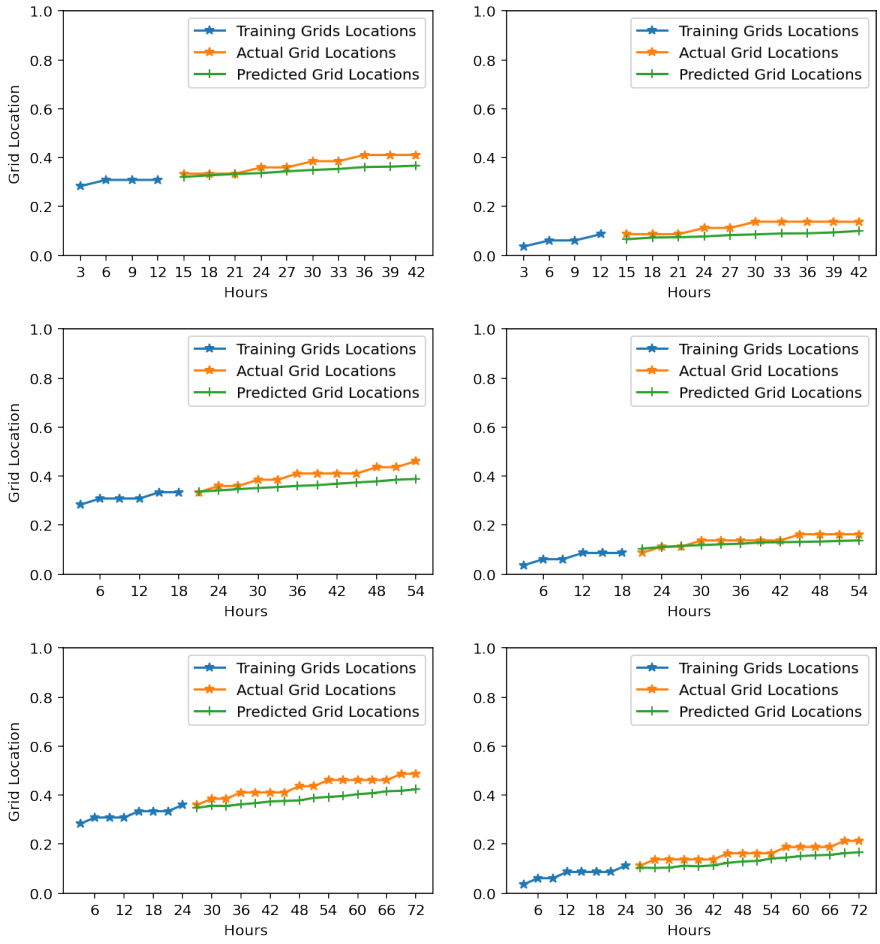


Figure 5.6: Actual and predicted grid predictions for Vayu and Fani for T_1 = 4, 6, 8 and T_2 = 10, 12, 16.

results of (Alemany *et al.*, 2018) in the case of Atlantic hurricane data is the process of generation of training data. From our understanding, in (Alemany *et al.*, 2018), only hurricanes with at most 60 recordings have been used in the model. If a hurricane has less than 60 recordings, then to make it 60 recordings, padding with zero has been used. This padding technique generates invalid training data points if the hurricane length is less than 60. Also, while preparing the training data set, mixing of data of different hurricanes has resulted in invalid training data points. This may result in high MAE as the model will encounter sudden zero values or consecutive values with high jumps. In our case, we avoid these errors, which result in improved MAEs.

MAEs for the NI ocean cyclone data also turn out to be small for all combinations of T_1 and T_2, see Table 5.2. To the best of the authors' knowledge, it is the first such result for the Indian ocean, where the LSTM network is used to predict GridID. The model performs well for the recent cyclones Vayu and Fani. The actual and predicted grid locations for Vayu and Fani are shown in Figure 5.6.

5.4 Conclusion

We proposed an LSTM based RNN model that predicts track of a tropical cyclone using GridIDs. The proposed model is able to capture the complex nonlinear behavior of cyclones and performs pretty well in terms of MAEs outperforming previous grid prediction results. Unlike the previous works that focus on predicting tracks one time-step ahead, our work focus on predicting the cyclone trajectory multiple time steps ahead while maintaining low MAE. Predicting cyclone tracks well in advance helps in taking preventive measures much earlier to minimize human and property loss.

A possible extension of this work could be to train an RNN that can predict the actual location parameters, latitude, and longitude, from GridIDs with low conversion error. One can also experiment with different sizes of grids by changing x value in grid definition and thereby reducing the error margin. A possible challenge in reducing x would be to achieve the low MAE of GridIDs' prediction.

CHAPTER 6

Conclusion and Future Directions

Disaster Risk Reduction (DPR) is one of the critical components of the 2030 agenda of Sustainable Development Goals (SDG). DPR is an integral requirement to social and economic development that has been recognized by various international policy documents like The Yokohama Strategy and Plan of Action for a Safer World (1994), Hyogo Framework for Action (2005-2015), Future We Want" (Rio, June 2012), Sendai Framework for DPR (Sendai, March 2016), and 2030 Agenda for Sustainable Development (New York, September 2015). As discussed in Section 1.1, floods and storms are among the most devastating natural disasters in terms of their count, the number of affected people, and economic loss. TCs are the primary reason behind storms and floods in tropical and subtropical coastal regions. TCs bring various hazards in the form of flooding, storm surge, thunderstorms, lightening, and tornadoes. As per a report in the last 50 years, around 1942 disasters happened due to TCs that killed about 7,79,324 people and caused a financial loss of $US\$1407.6$ billion[1]. Equivalently, the disasters due to TCs cause, on average, 43 deaths and $US\$78$ million economic loss daily. Moreover, natural disasters affect the lower-middle group income countries the most, where it accounts for around 1.77% of their GDP, which is way higher than the IMF's threshold of 0.5% for a major economic disaster. All this make the efforts towards mitigating the disaster caused by tropical cyclones central to sustainable and inclusive development.

An obvious way to mitigate and reduce the adverse effects of TCs is to have models that can forecast various development phases of TCs with high accuracy. Deep learning models have been successfully deployed to predict different weather-related phenomena, including TCs. Motivated with this, in this research work, we have developed various deep learning models using reanalysis and IBSTrACS data to forecast various essential phases of a TC. If we look at the different phases of a tropical cyclone, the first significant problem is to detect its formation well in advance. Therefore, our first proposed model forecasts the cyclone formation with a lead time of up to 60 hours. If

[1] https://public.wmo.int/en/our-mandate/focus-areas/natural-hazards-and-disaster-risk-reduction/tropical-cyclones

a cyclone dies over the sea, then one need not worry, and it can be simply ignored. The next most critical and important phase of a cyclone is its landfall, which is the event when it hits the coastline and moves over to land. Therefore next, we propose a deep learning model that solves the classification problem of whether a cyclone will make landfall or not. The location, intensity, and time of landfall of a cyclone determine the extent of disaster caused by it. An early information about whether a cyclone will make landfall or not, and if it is going to make landfall; where, at what speed and at what time will help the administration in mitigating the adverse effects of disaster. Therefore, next, we proposed a model that forecasts the intensity, location, and time of landfall of a cyclone from the early phase of its development.

The existing literature models (including numerical, statistical, and deep learning) focus primarily on forecasting cyclone formation, its intensity, and track with a specific lead time. As per our knowledge, there is no existing work in the literature that forecasts the landfall event and its characteristics (in the form of intensity, location, and time). The key contributions of this desertion are as follows:

- In Chapter 2, a deep learning model based on LSTM and CNN has been proposed that predicts the cyclone formation at lead times of 24 hours to 60 hours in six ocean basins NI, SI, NA, WP, EP, and SP of the world. Five atmospheric variables at three pressure levels from the reanalysis dataset ERA5 are identified as input to the model. An algorithm has been proposed to generate non-TC samples. The model achieves a 5-fold cross-validation accuracy in the range of $91.7\% - 97.7\%$, $96.4\% - 99.3\%$, $95.4\% - 99.1\%$, and $86.9.7\% - 92.9\%$ at lead times of 24h, 36h, 48h, and 60h respectively using only 12h of data. The model can be trained within 05 to 15 minutes depending on the ocean basin and can make predictions within seconds, making it deployable for practical purposes.

- In Chapter 3, we try to answer the fundamental question of whether a TC will make landfall or not. The proposed deep learning model can do so with a 5-fold cross-validation accuracy in the range of $96.4\% - 99.2\%$ and $93.0\% - 98.7\%$ using 12h or 24h of data (during initial 72h of cyclones progress) respectively across all six ocean basins of the world. It takes three atmospheric variables at three pressure levels from the ERA5 dataset and five variables from the IBTrACS

dataset. The model takes 05 to 20 minutes of training time depending on the ocean basin and can predict within seconds, making it practical for real-life usage.

- In Chapter 4, we proposed two deep learning models that can forecast the landfall's characteristics (its intensity, location, and time) from the very initiation of a cyclone. The first model predicts landfall's characteristics in the NI ocean basin using BTD from IMD. The second model predicts the landfall's location and time for six ocean basins of the world using reanalysis data. The first model achieves a 5-fold cross-validation MAE of $4.24(\pm 0.40)$knots, $4.5(\pm 0.58)$h, and $51.7(\pm 1.20)$KM for landfall's intensity, time, and location, respectively, using any 21h of data during the course of a cyclone. Our model outperforms the landfall's location accuracy reported by IMD on its website.

  The second model achieved the 5-fold cross-validation MAE in the range of $66.18(\pm 2.87)$KM- $158.92(\pm 12.62)$KM and $4.71(\pm 0.54)$h-$8.20(\pm 2.96)$h for location and time prediction respectively across all six ocean basins of the world. The model takes three atmospheric variables (at three pressure levels) and SST as input. It can be trained within 30-45 minutes and predicts within seconds.

- In Chapter 5, an LSTM based model is proposed for track prediction of a cyclone in terms of a GridID, which is obtained by applying the grid function on the ranges of latitudes and longitudes. Recently in (Alemany *et al.*, 2018), a new track prediction approach is proposed in terms of GridID, in place of directly predicting latitudes and longitudes. In the comparison of (Alemany *et al.*, 2018) where the model predicts GridID just one time-step ahead, we proposed a comparatively better model that predicts GridID several time steps ahead with better accuracy. The model performance is reported for Atlantic and NI ocean basins.

Future Directions

The main contribution of this work is to answer the TC's related three prediction problems - TC formation, landfall event, and landfall's characteristics using deep learning models based on LSTM and CNN. A possible and exciting extension of the work will be in the following directions:

- One can explore using the reanalysis dataset to answer other TC-related forecast problems like storm surge, rainfall, continuous tracking of the path, and intensity.
- Reanalysis dataset ERA5 contains an extensive record of various atmospheric and oceanographic variables. In this study, we have mainly used the wind fields - u and v, geopotential z, and SST. It will be interesting to explore the usage of other variables like humidity, temperature, cloud cover, vorticity, and many others in answering above stated prediction problems.
- As per our knowledge, the proposed deep learning model that predicts landfall events and its characteristics is the first attempt to answer these forecast problems. One can explore numerical or statistical methods to target these problems in the future.
- A promising research direction could be to propose ensemble models for these prediction tasks by combining numerical or statistical methods with deep learning techniques.

REFERENCES

1. **Abadi, M.**, **A. Agarwal**, **P. Barham**, **E. Brevdo**, **Z. Chen**, **C. Citro**, **G. S. Corrado**, **A. Davis**, **J. Dean**, **M. Devin**, **S. Ghemawat**, **I. Goodfellow**, **A. Harp**, **G. Irving**, **M. Isard**, **Y. Jia**, **R. Jozefowicz**, **L. Kaiser**, **M. Kudlur**, **J. Levenberg**, **D. Mané**, **R. Monga**, **S. Moore**, **D. Murray**, **C. Olah**, **M. Schuster**, **J. Shlens**, **B. Steiner**, **I. Sutskever**, **K. Talwar**, **P. Tucker**, **V. Vanhoucke**, **V. Vasudevan**, **F. Viégas**, **O. Vinyals**, **P. Warden**, **M. Wattenberg**, **M. Wicke**, **Y. Yu**, and **X. Zheng** (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. URL https://www.tensorflow.org/. Software available from tensorflow.org. 25, 41, 54, 66

2. **Abhishek, K.**, **M. Singh**, **S. Ghosh**, and **A. Anand** (2012). Weather forecasting model using artificial neural network. *Procedia Technology*, **4**, 311 – 318. ISSN 2212-0173. URL http://www.sciencedirect.com/science/article/pii/S221201731200326X. 2nd International Conference on Computer, Communication, Control and Information Technology(C3IT-2012) on February 25 - 26, 2012. 9

3. **Alemany, S.**, **J. Beltran**, **A. Perez**, and **S. Ganzfried** (2018). Predicting hurricane trajectories using a recurrent neural network. *Proceedings of the AAAI Conference on Artificial Intelligence*, **33**. URL https://ojs.aaai.org/index.php/AAAI/article/view/3819. 15, 20, 34, 51, 74, 76, 78, 81, 82, 86

4. **Ali, M.**, **P. Jagadeesh**, **I.-I. Lin**, and **J.-Y. Hsu** (2012). A neural network approach to estimate tropical cyclone heat potential in the indian ocean. *IEEE Geoscience and Remote Sensing Letters*, **9**, 1114–1117. 9

5. **Barnes, L. R.**, **E. C. Gruntfest**, **M. H. Hayden**, **D. M. Schultz**, and **C. Benight** (2007). False alarms and close calls: A conceptual model of warning accuracy. *Weather and Forecasting*, **22**(5), 1140 – 1147. URL https://journals.ametsoc.org/view/journals/wefo/22/5/waf1031_1.xml. 43

6. **Bengio, Y.**, **P. Simard**, and **P. Frasconi** (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, **5**, 157–66. 9

7. **Bisong, E.**, *Google Colaboratory*. Apress, Berkeley, CA, 2019. ISBN 978-1-4842-4470-8, 59–64. URL https://doi.org/10.1007/978-1-4842-4470-8_7. 78

8. **Biswas, K.**, **S. Kumar**, and **A. K. Pandey** (2021). Intensity prediction of tropical cyclones using long short-term memory network. URL https://arxiv.org/abs/2107.03187. 34

9. **Boussioux, L.**, **C. Zeng**, **T. Guénais**, and **D. Bertsimas** (2020). Hurricane forecasting: A novel multimodal machine learning framework. *arXiv preprint arXiv:2011.06125*. 21, 34, 51, 52, 63, 69, 70

10. **Charney, J. G.** and **A. Eliassen** (1964). On the growth of the hurricane depression. *Journal of Atmospheric Sciences*, **21**(1), 68 – 75. URL https://journals.ametsoc.org/view/journals/atsc/21/1/1520-0469_1964_021_0068_otgoth_2_0_co_2.xml. 17

11. **Chaudhuri, S.**, **D. Basu**, **D. Das**, **S. Goswami**, and **S. Varshney** (2017). Swarm intelligence and neural nets in forecasting the maximum sustained wind speed along the track of tropical cyclones over bay of bengal. *Natural Hazards*, **87**. 34, 50

12. **Chaudhuri, S.**, **D. Dutta**, **S. Goswami**, and **A. Middey** (2015). Track and intensity forecast of tropical cyclones over the north indian ocean with multilayer feed forward neural nets. *Meteorological Applications*, **22**(3), 563–575. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/met.1488. 9, 51

13. **Chen, B.-F.**, **B. Chen**, **H.-T. Lin**, and **R. L. Elsberry** (2019*a*). Estimating tropical cyclone intensity by satellite imagery utilizing convolutional neural networks. *Weather and Forecasting*, **34**(2), 447 – 465. URL https://journals.ametsoc.org/view/journals/wefo/34/2/waf-d-18-0136_1.xml. 20, 34

14. **Chen, R.**, **X. Wang**, **W. Zhang**, **X. Zhu**, **A. Li**, and **C. Yang** (2019*b*). A hybrid cnn-lstm model for typhoon formation forecasting. *Geoinformatica*, **23**(3), 375–396. 20, 21, 31, 34, 51

15. **Chollet, F.** (2015). Keras. https://github.com/fchollet/keras. 25, 41, 54, 66, 78

16. **Chung, J.**, **Ç. Gülçehre**, **K. Cho**, and **Y. Bengio** (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *CoRR*, **abs/1412.3555**. URL http://arxiv.org/abs/1412.3555. 70

17. **Ciresan, D. C.**, **U. Meier**, **J. Masci**, **L. M. Gambardella**, and **J. Schmidhuber**, Flexible, high performance convolutional neural networks for image classification. *In Twenty-second international joint conference on artificial intelligence*. 2011. 41

18. **Cleeremans, A.**, **D. Servan-Schreiber**, and **J. L. McClelland** (1989). Finite state automata and simple recurrent networks. *Neural Comput.*, **1**(3), 372–381. ISSN 0899-7667. URL https://doi.org/10.1162/neco.1989.1.3.372. 9

19. **Cossuth, J. H.**, **R. D. Knabb**, **D. P. Brown**, and **R. E. Hart** (2013). Tropical cyclone formation guidance using pregenesis dvorak climatology. part i: Operational forecasting and predictive potential. *Weather and Forecasting*, **28**(1), 100 – 118. URL https://journals.ametsoc.org/view/journals/wefo/28/1/waf-d-12-00073_1.xml. 18

20. **Craig, G. C.** and **S. L. Gray** (1996). Cisk or wishe as the mechanism for tropical cyclone intensification. *Journal of Atmospheric Sciences*, **53**(23), 3528 – 3540. URL https://journals.ametsoc.org/view/journals/atsc/53/23/1520-0469_1996_053_3528_cowatm_2_0_co_2.xml. 17

21. **Dee, D. P.**, **S. M. Uppala**, **A. J. Simmons**, **P. Berrisford**, **P. Poli**, **S. Kobayashi**, **U. Andrae**, **M. Balmaseda**, **G. Balsamo**, **d. P. Bauer**, *et al.* (2011). The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, **137**(656), 553–597. 7, 34, 51

22. **Dvorak, V. F.** (1975). Tropical cyclone intensity analysis and forecasting from satellite imagery. *Monthly Weather Review*, **103**(5), 420 – 430. URL https://journals.ametsoc.org/view/journals/mwre/103/5/1520-0493_1975_103_0420_tciaaf_2_0_co_2.xml. 18

23. **Dvorak, V. F.**, *Tropical cyclone intensity analysis using satellite data*, volume 11. US Department of Commerce, National Oceanic and Atmospheric Administration . . . , 1984. 18

24. **ELSBERRY, R. L.** (1995). Tropical cyclone motion. global perspectives on tropical cyclones. *WMO/TD-No. 693*, 106–197. 4

25. **Emanuel, K. A.** (1989). The finite-amplitude nature of tropical cyclogenesis. *Journal of Atmospheric Sciences*, **46**(22), 3431–3456. 5

26. **Gao, S.**, **P. Zhao**, **B. Pan**, **Y. Li**, **M. Zhou**, **J. Xu**, **S. Zhong**, and **Z. Shi** (2018). A nowcasting model for the prediction of typhoon tracks based on a long short term memory neural network. *Acta Oceanologica Sinica*, **37**, 8–12. 20, 34, 51

27. **Gers, F. A.** and **E. Schmidhuber** (2001). Lstm recurrent networks learn simple context-free and context-sensitive languages. *IEEE Transactions on Neural Networks*, **12**(6), 1333–1340. 10, 25, 41

28. **Gers, F. A.**, **J. Schmidhuber**, and **F. Cummins**, Learning to forget: continual prediction with lstm. *In 1999 Ninth International Conference on Artificial Neural Networks ICANN 99. (Conf. Publ. No. 470)*, volume 2. 1999. viii, ix, 10, 25, 29, 30, 40, 45, 56, 57, 58, 68, 80, 81

29. **Gers, F. A.**, **N. N. Schraudolph**, and **J. Schmidhuber** (2003). Learning precise timing with lstm recurrent networks. *J. Mach. Learn. Res.*, **3**(null), 115–143. ISSN 1532-4435. URL https://doi.org/10.1162/153244303768966139. 10, 25, 40

30. **Giffard-Roisin, S.**, **M. Yang**, **G. Charpiat**, **C. Kumler Bonfanti**, **B. Kégl**, and **C. Monteleoni** (2020). Tropical cyclone track forecasting using fused deep learning from aligned reanalysis data. *Frontiers in Big Data*, **3**, 1. ISSN 2624-909X. URL https://www.frontiersin.org/article/10.3389/fdata.2020.00001. 21, 34, 51, 52, 63, 69, 70, 71

31. **Graves, A.**, **A. Mohamed**, and **G. Hinton**, Speech recognition with deep recurrent neural networks. *In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. 2013. 11

32. **Gray, W. M.** (1968). Global view of the origin of tropical disturbances and storms. *Monthly Weather Review*, **96**(10), 669–700. 4

33. **Gray, W. M.** (1979). Hurricanes: Their formation, structure and likely role in the tropical circulation. *Meteorology over the tropical oceans*, **155**, 218. 19

34. **Grinsted, A.**, **P. Ditlevsen**, and **J. H. Christensen** (2019). Normalized us hurricane damage estimates using area of total destruction, 1900-2018. *Proceedings of the National Academy of Sciences*, **116**(48), 23942–23946. ISSN 0027-8424. URL https://www.pnas.org/content/116/48/23942. 3

35. **Hall, T. M.** and **S. Jewson** (2007). Statistical modeling of north atlantic tropical cyclone tracks. *Tellus*, **59A**, 486–498. 34, 50

36. **Halperin, D. J.**, **H. E. Fuelberg**, **R. E. Hart**, **J. H. Cossuth**, **P. Sura**, and **R. J. Pasch** (2013). An evaluation of tropical cyclone genesis forecasts from global numerical models. *Weather and Forecasting*, **28**(6), 1423–1445. 14, 17, 18

37. **Han, J.** and **C. Moraga**, The influence of the sigmoid function parameters on the speed of backpropagation learning. *In* **J. Mira** and **F. Sandoval** (eds.), *From Natural to Artificial Neural Computation*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1995. ISBN 978-3-540-49288-7. 25, 41

38. **Hennon, C. C.** and **J. S. Hobgood** (2003). Forecasting tropical cyclogenesis over the atlantic basin using large-scale data. *Monthly weather review*, **131**(12), 2927–2940. 19

39. **Hersbach, H.**, **B. Bell**, **P. Berrisford**, **S. Hirahara**, **A. Horányi**, **J. Muñoz-Sabater**, **J. Nicolas**, **C. Peubey**, **R. Radu**, **D. Schepers**, **A. Simmons**, **C. Soci**, **S. Abdalla**, **X. Abellan**, **G. Balsamo**, **P. Bechtold**, **G. Biavati**, **J. Bidlot**, **M. Bonavita**, **G. De Chiara**, **P. Dahlgren**, **D. Dee**, **M. Diamantakis**, **R. Dragani**, **J. Flemming**, **R. Forbes**, **M. Fuentes**, **A. Geer**, **L. Haimberger**, **S. Healy**, **R. J. Hogan**, **E. Hólm**, **M. Janisková**, **S. Keeley**, **P. Laloyaux**, **P. Lopez**, **C. Lupu**, **G. Radnoti**, **P. de Rosnay**, **I. Rozum**, **F. Vamborg**, **S. Villaume**, and **J.-N. Thépaut** (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146**(730), 1999–2049. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803. 6, 21, 35, 51, 63

40. **Hochreiter, S.** and **J. Schmidhuber** (1997). Long short-term memory. *Neural Comput.*, **9**(8), 1735–1780. ISSN 0899-7667. URL https://doi.org/10.1162/neco.1997.9.8.1735. viii, ix, 10, 25, 29, 30, 40, 45, 56, 57, 58, 68, 80, 81

41. **Hodges, K.**, **A. Cobb**, and **P. L. Vidale** (2017). How well are tropical cyclones represented in reanalysis datasets? *Journal of Climate*, **30**(14), 5243–5264. 7

42. **Horn, M.**, **K. Walsh**, **M. Zhao**, **S. J. Camargo**, **E. Scoccimarro**, **H. Murakami**, **H. Wang**, **A. Ballinger**, **A. Kumar**, **D. A. Shaevitz**, **J. A. Jonas**, and **K. Oouchi** (2014). Tracking scheme dependence of simulated tropical cyclone response to idealized climate simulations. *Journal of Climate*, **27**(24), 9197 – 9213. URL https://journals.ametsoc.org/view/journals/clim/27/24/jcli-d-14-00200.1.xml. 21

43. **Ioffe, S.** and **C. Szegedy**, Batch normalization: Accelerating deep network training by reducing internal covariate shift. *In International conference on machine learning*. PMLR, 2015. 12, 41

44. **Jordan, M. I.**, *Attractor Dynamics and Parallelism in a Connectionist Sequential Machine*. IEEE Press, 1990. ISBN 0818620153, 112–127. 9

45. **Kiefer, J.** and **J. Wolfowitz** (1952). Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.*, **23**(3), 462–466. URL https://doi.org/10.1214/aoms/1177729392. 8

46. **Kim, M.**, **M.-S. Park**, **J. Im**, **S. Park**, and **M.-I. Lee** (2019*a*). Machine learning approaches for detecting tropical cyclone formation using satellite data. *Remote Sensing*, **11**(10). ISSN 2072-4292. URL https://www.mdpi.com/2072-4292/11/10/1195. 19

47. **Kim, S.**, **H. Kim**, **J. Lee**, **S. Yoon**, **S. E. Kahou**, **K. Kashinath**, and **M. Prabhat**, Deep-hurricane-tracker: Tracking and forecasting extreme climate events. 2019*b*. 34, 51

48. **Kingma, D.** and **J. Ba** (2014). Adam: A method for stochastic optimization. *International Conference on Learning Representations*. 25, 42, 54, 66, 78

49. **Knapp, K. R.**, **M. C. Kruk**, **D. H. Levinson**, **H. J. Diamond**, and **C. J. Neumann** (2010). The international best track archive for climate stewardship (ibtracs): Unifying tropical cyclone data. *Bulletin of the American Meteorological Society*, **91**(3), 363 – 376. URL https://journals.ametsoc.org/view/journals/bams/91/3/2009bams2755_1.xml. 6, 21, 35, 61

50. **Knutson, T.**, **S. J. Camargo**, **J. C. L. Chan**, **K. Emanuel**, **C.-H. Ho**, **J. Kossin**, **M. Mohapatra**, **M. Satoh**, **M. Sugi**, **K. Walsh**, and **L. Wu** (2019). Tropical cyclones and climate change assessment: Part i: Detection and attribution. *Bulletin of the American Meteorological Society*, **100**(10), 1987 – 2007. URL https://journals.ametsoc.org/view/journals/bams/100/10/bams-d-18-0189.1.xml. 3

51. **Kovordányi, R.** and **C. Roy** (2009). Cyclone track forecasting based on satellite images using artificial neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, **64**(6), 513 – 521. ISSN 0924-2716. URL http://www.sciencedirect.com/science/article/pii/S0924271609000434. 9, 50

52. **Krishnamurti, T. N.**, **C. M. Kishtawal**, **Z. Zhang**, **T. LaRow**, **D. Bachiochi**, **E. Williford**, **S. Gadgil**, and **S. Surendran** (2000). Multimodel ensemble forecasts for weather and seasonal climate. *Journal of Climate*, **13**(23), 4196 – 4216. URL https://journals.ametsoc.org/view/journals/clim/13/23/1520-0442_2000_013_4196_meffwa_2.0.co_2.xml. 50

53. **Krizhevsky, A.**, **I. Sutskever**, and **G. Hinton** (2012). Imagenet classification with deep convolutional neural networks. *Neural Information Processing Systems*, **25**. viii, 12, 25, 29, 30, 40, 45, 68

54. **Kumar, S.**, **K. Biswas**, and **A. K. Pandey**, Predicting landfall's location and time of a tropical cyclone using reanalysis data. *In Artificial Neural Networks and Machine Learning – ICANN 2021*. Springer International Publishing, 2021*a*. ISBN 978-3-030-86380-7. 20, 21, 35

55. **Kumar, S.**, **K. Biswas**, and **A. K. Pandey**, Prediction of landfall intensity, location, and time of a tropical cyclone. *In Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35. 2021*b*. 20, 35

56. **Kumar, S.**, **K. Biswas**, and **A. K. Pandey**, Track prediction of tropical cyclones using long short-term memory network. *In 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*. 2021*c*. 20, 34

57. **LeCun, Y.**, **B. Boser**, **J. S. Denker**, **D. Henderson**, **R. E. Howard**, **W. Hubbard**, and **L. D. Jackel** (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, **1**(4), 541–551. viii, 12, 25, 29, 30, 40, 45, 68

58. **Leroux, M.-D.**, **K. Wood**, **R. L. Elsberry**, **E. O. Cayanan**, **E. Hendricks**, **M. Kucas**, **P. Otto**, **R. Rogers**, **B. Sampson**, and **Z. Yu** (2018). Recent advances in research and forecasting of tropical cyclone track, intensity, and structure at landfall. *Tropical Cyclone Research and Review*, **7**(2), 85 – 105. ISSN 2225-6032. URL http://www.sciencedirect.com/science/article/pii/S2225603219300189. 14, 15, 49

59. **Maskey, M.**, **R. Ramachandran**, **M. Ramasubramanian**, **I. Gurung**, **B. Freitag**, **A. Kaulfus**, **D. Bollinger**, **D. J. Cecil**, and **J. Miller** (2020). Deepti: Deep-learning-based tropical cyclone intensity estimation system. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **13**, 4271–4281. 20, 34

60. **Matsuoka, D.**, **M. Nakano**, **D. Sugiyama**, and **S. Uchida** (2018). Deep learning approach for detecting tropical cyclones and their precursors in the simulation by a cloud-resolving global nonhydrostatic atmospheric model. *Progress in Earth and Planetary Science*, **5**(1), 1–16. 20, 31

61. **McCulloch, W. S.** and **W. Pitts** (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, **5**(4), 115–133. 8

62. **Milletari, F.**, **N. Navab**, and **S. Ahmadi**, V-net: Fully convolutional neural networks for volumetric medical image segmentation. *In 2016 Fourth International Conference on 3D Vision (3DV)*. 2016. 12, 25, 40

63. **Mohapatra, M.**, **D. Nayak**, **M. Sharma**, **R. Sharma**, and **B. Bandyopadhyay** (2015). Evaluation of official tropical cyclone landfall forecast issued by india meteorological department. *Journal of Earth System Science*, **124**. 35

64. **Moradi Kordmahalleh, M.**, **M. Gorji Sefidmazgi**, and **A. Homaifar**, A sparse recurrent neural network for trajectory prediction of atlantic hurricanes. GECCO '16. Association for Computing Machinery, New York, NY, USA, 2016. ISBN 9781450342063. URL https://doi.org/10.1145/2908812.2908834. 34, 51

65. **Nair, V.** and **G. E. Hinton**, Rectified linear units improve restricted boltzmann machines. *In* **J. Fürnkranz** and **T. Joachims** (eds.), *Proceedings of the 27th International Conference on Machine Learning (ICML-10), June 21-24, 2010, Haifa, Israel*. Omnipress, 2010. URL https://icml.cc/Conferences/2010/papers/432.pdf. 25, 41, 55, 66, 78

66. **Nath, S.**, **S. Kotal**, and **P. kundu** (2015). Seasonal prediction of tropical cyclone activity over the north indian ocean using the neural network model. *Atmósfera*, **28**(4), 271–281. ISSN 0187-6236. URL https://www.sciencedirect.com/science/article/pii/S0187623617300115. 20

67. **Palmen, E.** (1948). On the formation and structure of tropical hurricanes. *Geophysica*, **3**(1), 26–38. 4

68. **Pearlmutter**, Learning state space trajectories in recurrent neural networks. *In International 1989 Joint Conference on Neural Networks*. 1989. 9

69. **Pedregosa, F.**, **G. Varoquaux**, **A. Gramfort**, **V. Michel**, **B. Thirion**, **O. Grisel**, **M. Blondel**, **P. Prettenhofer**, **R. Weiss**, **V. Dubourg**, **J. Vanderplas**, **A. Passos**, **D. Cournapeau**, **M. Brucher**, **M. Perrot**, and **E. Duchesnay** (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830. 25, 42, 54, 66

70. **Powell, M. D.** and **S. D. Aberson** (2001). Accuracy of united states tropical cyclone landfall forecasts in the atlantic basin (1976–2000). *Bulletin of the American Meteorological Society*, **82**(12), 2749 – 2768. URL https://journals.ametsoc.org/view/journals/bams/82/12/1520-0477_2001_082_2749_aoustc_2_3_co_2.xml. 35

71. **Pradhan, R.**, **R. S. Aygun**, **M. Maskey**, **R. Ramachandran**, and **D. J. Cecil** (2018). Tropical cyclone intensity estimation using a deep convolutional neural network. *IEEE Transactions on Image Processing*, **27**(2), 692–702. 34

72. **Ramachandran, P.**, **B. Zoph**, and **Q. V. Le** (2017). Searching for activation functions. *arXiv preprint arXiv:1710.05941*. URL https://arxiv.org/abs/1710.05941. 55

73. **Richman, M. B.** and **L. M. Leslie** (2012). Adaptive machine learning approaches to seasonal prediction of tropical cyclones. *Procedia Computer Science*, **12**, 276–281. 20

74. **Richman, M. B.**, **L. M. Leslie**, **H. A. Ramsay**, and **P. J. Klotzbach** (2017). Reducing tropical cyclone prediction errors using machine learning approaches. *Procedia Computer Science*, **114**, 314–323. ISSN 1877-0509. URL https://www.sciencedirect.com/science/article/pii/S1877050917318422. Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems, CAS October 30 – November 1, 2017, Chicago, Illinois, USA. 20

75. **Rumelhart, D. E.** and **J. L. McClelland**, *Learning Internal Representations by Error Propagation*. 1987, 318–362. 9

76. **Schumacher, A. B.**, **M. DeMaria**, and **J. A. Knaff** (2009). Objective estimation of the 24-h probability of tropical cyclone formation. *Weather and Forecasting*, **24**(2), 456 – 471. URL https://journals.ametsoc.org/view/journals/wefo/24/2/2008waf2007109_1.xml. 18

77. **Schuster, M.** and **K. K. Paliwal** (1997). Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, **45**(11), 2673–2681. 11

78. **Shakya, S.**, **S. Kumar**, and **M. Goswami** (2020). Deep learning algorithm for satellite imaging based cyclone detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **13**, 827–839. 20, 31, 34

79. **Srivastava, N.**, **G. Hinton**, **A. Krizhevsky**, **I. Sutskever**, and **R. Salakhutdinov** (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, **15**, 1929–1958. 13

80. **Sugi, M.**, **A. Noda**, and **N. Sato** (2002). Influence of the global warming on tropical cyclone climatology: An experiment with the jma global model. *Journal of the Meteorological Society of Japan. Ser. II*, **80**(2), 249–272. 20

81. **Vitart, F.**, **D. Anderson**, and **T. Stockdale** (2003). Seasonal forecasting of tropical cyclone landfall over mozambique. *Journal of Climate*, **16**(23), 3932–3945. 34

82. **Wahiduzzaman, M.**, **E. C. Oliver**, **S. J. Wotherspoon**, and **N. J. Holbrook** (2017). A climatological model of north indian ocean tropical cyclone genesis, tracks and landfall. *Climate Dynamics*, **49**(7), 2585–2603. 35

83. **Wallemacq, P.**, **UNISDR**, and **CRED** (2018). Economic losses, poverty and disasters 1998-2017. URL https://www.undrr.org/publication/economic-losses-poverty-disasters-1998-2017. 1

84. **Webersik, C.**, **M. Esteban**, and **T. Shibayama** (2010). The economic impact of future increase in tropical cyclones in japan. *Natural Hazards*, **55**, 233–250. 3

85. **Wijnands, J. S.**, **G. Qian**, and **Y. Kuleshov** (2016). Variable selection for tropical cyclogenesis predictive modeling. *Monthly Weather Review*, **144**(12), 4605 – 4619. URL https://journals.ametsoc.org/view/journals/mwre/144/12/mwr-d-16-0166.1.xml. 19

86. **Yamada, Y.**, **M. Satoh**, **M. Sugi**, **C. Kodama**, **A. T. Noda**, **M. Nakano**, and **T. Nasuno** (2017). Response of tropical cyclone activity and structure to global warming in a high-resolution global nonhydrostatic model. *Journal of Climate*, **30**(23), 9703–9724. 20

87. **Yamaguchi, M.** and **N. Koide** (2017). Tropical cyclone genesis guidance using the early stage dvorak analysis and global ensembles. *Weather and Forecasting*, **32**(6), 2133 – 2141. URL https://journals.ametsoc.org/view/journals/wefo/32/6/waf-d-17-0056_1.xml. 18

88. **Zhang, T.**, **W. Lin**, **Y. Lin**, **M. Zhang**, **H. Yu**, **K. Cao**, and **W. Xue** (2019). Prediction of tropical cyclone genesis from mesoscale convective systems using machine learning. *Weather and Forecasting*, **34**(4), 1035 – 1049. URL https://journals.ametsoc.org/view/journals/wefo/34/4/waf-d-18-0201_1.xml. viii, 19, 26, 31, 32

89. **Zhang, W.**, **B. Fu**, **M. S. Peng**, and **T. Li** (2015). Discriminating developing versus nondeveloping tropical disturbances in the western north pacific through decision tree analysis. *Weather and Forecasting*, **30**(2), 446 – 454. URL https://journals.ametsoc.org/view/journals/wefo/30/2/waf-d-14-00023_1.xml. 19