



Implementation of a Convolutional Neural Network for 24-hour tropical cyclone track forecasting in the Bay of Bengal using fused historical and reanalysis data

Md Khairul Haque¹ · Md. Alim Miah^{1,3} · Md. Akhlaq Hossain² · Md Aminul Islam³

Received: 10 November 2025 / Accepted: 24 May 2026

© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2026

Abstract

Tropical cyclones are one of the most destructive natural disasters, inflicting severe damage on human life, infrastructure, and the environment. This study aims to present a framework for 24-hour tropical cyclone track forecasting in the Bay of Bengal (BoB) region. Deep learning algorithms, including Convolutional Neural Network (CNN), Numerical Weather Prediction (NWP) techniques, and models from the European Centre for Medium-Range Weather Forecasts (ECMWF), were utilized to capture the complex non-linear relationships between atmospheric conditions and storm trajectories. A predictive model was developed using a fused dataset that integrates more than 40 years of cyclone history from the IBTrACS database with atmospheric variables from the ERA5 reanalysis data. Through systematic experimentation with model architecture and hyperparameters, the best-performing configuration a shallow CNN, achieved a 24-hour Mean Absolute Error (MAE) of 141.2 km, calculated using the Haversine great-circle distance. This result demonstrates a significant improvement over the standard persistence model benchmark, which yielded an MAE of 186.11 km on the same test set. The model's predictions for a hypothetical cyclone align closely with the observed climatological behaviour of the region. Forecast error analysis revealed clustering around 150 km with a mean error of 223.69 km, reflecting the multidimensional distribution of this value. This study underscores the potential of deep learning in cyclone prediction, offering a promising tool for real-time cyclone tracking and regional early warning. The proposed framework supports the future development of coastal risk management systems aimed at enhancing preparedness.

Keywords Bay of Bengal · CNN · ECMWF · Reanalysis · Tropical Cyclone · Track Forecasting

1 Introduction

Tropical cyclones (TCs) are also known as hurricanes and typhoons, which denote catastrophic weather phenomena that can significantly impact human life, property and socio-economic development. TC is a low-pressure and warm

core that forms in tropical and subtropical oceans (Fang et al. 2022). There are several favourable environmental conditions that must be in place before a TC can form. There are four conditions, including warm ocean water (at least 80 °F/ 27°C) through a depth of about 150 ft. (46 m); comparatively humid air at the mid-level of the troposphere (16,000ft. /4,900 m); a minimum distance of at least 300 miles (480 km) from the equator; low values (less than about 23 mph/37 km/h) of vertical wind shear between the surface and the upper troposphere; vertical wind shear is defined as the change in wind velocity (speed and direction) with height, and strong shear can disrupt cyclone organization (NOAA 2025). As discovered, the magnitude of TC is largely influenced by three factors, namely, the initial severity of TC, the thermodynamic state of the atmosphere and the heat exchange between the ocean and TCs (Emanuel 1999). Nearly 7% of global TCs

✉ Md Khairul Haque
khairul_20103430@jkkniu.edu.bd

¹ Department of Environmental Science and Engineering, Jatiya Kabi Kazi Nazrul Islam University, Trishal, Mymensingh 2224, Bangladesh

² Department of Computer Science and Engineering, Bangladesh University of Business and Technology (BUBT), Dhaka 1216, Bangladesh

³ Department of Environmental Science, Bangladesh Agricultural University, Mymensingh 2202, Bangladesh

occurred in the North Indian Ocean (NIO), according to reports by (Neumann 1993). Although the NIO is the site of only 7% of TCs experienced globally, a disproportionately high percentage of deaths that occur as a direct result of a cyclone suggest that the area is dramatically affected by the humanitarian impact of the disaster. Specifically, BoB alone contributes 5% of global TCs, making it one of the most vulnerable sub-regions (Alam and Dominey-Howes 2015). More importantly, recent studies highlight that TCs in the BoB have undergone notable changes in recent decades; between 1981 and 2010, oceanic and atmospheric conditions became increasingly favourable for cyclone growth in the post-monsoon season (October-December), leading to intensification (Balaguru et al. 2014). In other studies, they found that there was an 8% decrease in the frequency of relatively weak cyclones (Deshpande et al. 2021) and slower transition speed for intense cyclones compared to weaker ones (Sun et al. 2021). In particular, flash flooding due to the extensive rainfall of TCs can easily result in a significant calamity affecting the communities living on the coasts of the BoB (Das et al. 2007; Peduzzi et al. 2012; Woodruff et al. 2013). The BoB has a very high frequency (occurrence) and strength (tropical cyclone observational density) of the TCs of all intensity categories as compared to the Arabian Sea. Such differences in regional cyclone activity highlight the importance of basin-specific forecasting models, as shown in Table 1. Categories are defined as Cyclonic Storm (CS), Severe Cyclonic Storm (SCS), Very Severe Cyclonic Storm (VSCS), and Extremely Severe Cyclonic Storm (ESCS). The 3-hourly observations indicate the total number of synoptic observations recorded at 3-hour intervals during the life cycle of each cyclone event, reflecting the density of observational data available for each category.

History indicates that the BoB has been a land of deadly TCs, with individual events claiming up to half a million lives. Compared to the Arabian Sea, the BoB records a much higher frequency of cyclones across all intensity categories and a denser set of 3-hourly synoptic observations, underscoring its extreme vulnerability. This highlights the urgent necessity of strengthened disaster preparedness and

more effective early warning systems in the region. The BoB had most of the deadliest cyclones in world history, indicating sensitivity of the area to such types of landfalls occasioned by large populations along the coastline and low levels of relief. Krishnakumar et al. (2009). documented that the rising of TC occurrence in the BoB during the post-monsoon season could be related to the increasing of post-monsoon precipitation in Kerala, India. The cyclone migration patterns are vital criteria in the forecasting of the landfalls because slight changes in the cyclone movements significantly change the affected populace and economic regions. According to recent research by Sattar and Cheung, it was also revealed that the NIO TCs activity was influenced by the El Niño Southern Oscillation and the Indian Ocean Dipole, and the TCs formation condition varied depending on the basin (BoB or AS) (Sattar and Cheung 2019). Compared to the AS, the BoB was an active cyclone-forming basin due to the higher sea surface temperature (SST), upper ocean heat content (UOHC), total column water vapour (TCWV), relative humidity (RH) and lower tropospheric temperature (LTT) in the BoB basin than in the AS (Uddin et al. 2019). In addition, the intensity of vertical wind shear (VWS) in BoB was very low compared to that of AS (Balaji et al. 2018). In Bangladesh, the very severe cyclonic storm Sidr (2007) destroyed houses, agricultural crops, roads and infrastructure by flooding the coastal areas. Global and regional models poorly resolve inner core dynamics when predicting intensity, which is strongly affected by smaller-scale physical processes of the climate system, especially the cumulus convection, radiation, and microphysics (Zhang et al. 2011; DeMaria and Kaplan 1999). They have demonstrated that the change of intensity is strongly connected to the structure and evolution of the inner core of the storm (Franklin et al. 1988; Kossin and Eastin 2001a, b). Uncertainties in the initial conditions and approximations of NWP models have a significant impact on forecast skill (Kotal et al. 2024). NWP models continue to face challenges in correctly simulating mesoscale and convective processes around the core, both of which are crucial for predicting intensity (Weber 2005; Zhang et al. 2019). Despite these limitations, such much of the improvement in the TCs forecasting of both track and intensity has been driven by advances in NWP model performance, including improved numerical techniques, high-performance computing, and assimilation of diverse observational data (Chen et al. 2013; Kim and Hasegawa 2018; Lei et al. 2017; Cangialosi et al. 2020; Landsea and Cangialosi 2018). Ensemble prediction systems (EPS) have become a significant part of the operational forecasting suite at many methodological institutes, providing probabilistic guidance and improving forecast reliability (Leutbecher and Palmer 2008). Only a few high-resolution long-term observation data indicate

Table 1 The observation and number of TCs in the NIO basin

		Basin	
		Arabian Sea	BoB
Number of Cyclone in category	CS	1	7
	SCS	2	1
	VSCS	1	2
	ESCS	1	4
3 hourly observations	CS	40	176
	SCS	72	32
	VSCS	72	120
	ESCS	48	200

Sources (Uddin et al. 2021)

the range of BoB, limiting the verification of models and the use of forecasts in this region. Two key questions are considered in terms of the forecast provided by the ECMWF ensemble system: (1) the reliability of the control forecasts and (2) the reliability of the ensemble mean forecasts (Zsoter et al. 2009). The idealised simulations performed with the Weather Research and Forecasting (WRF) Model confirmed the existence of these relationships (Ge et al. 2013; Wu et al. 2012). Deep learning approaches have been applied to climate research with success in the past few years following the rapid development and the high scalability of deep learning algorithms; they have been used in weather forecasting (Bi et al. 2023), El Niño Southern Oscillation prediction (Ham et al. 2019), and in predicting the Madden Julian oscillation (Kim et al. 2021) and have been applied in enhancing the accuracy of forecasts. Deep learning has widely played a vital role in climatic phenomena identification and detection of signals linked to global warming (Ham et al. 2019), hence making new discoveries in the science of climate possible. To promote performance TC intensity forecasting, a strategic research effort has been conducted on this trend to leverage a range of deep learning models that includes Multilayer Perceptron (MLP), Long Short-Term Memory (LSTM), CNN, or a Generative Adversarial Network (GAN) (Chaudhuri et al. 2013; Pan et al. 2019; Meng et al. 2023; Xu et al. 2021). As an example, Xu et al. (2021) have proposed a TC intensity prediction model using the MLP model, which consists of two-layer hidden nodes (2048 nodes in each layer). It has also been reported that global warming will also result in increased overall TC intensity and reduction of landfalling TCs, which will weaken at a reduced pace, correlating to the increase in the saturated water vapour amount in the atmosphere (Knutson et al. 2020; Li and Chakraborty 2020). The socioeconomic damage caused by TC in coastal areas is expected to increase in the future. (Peduzzi et al. 2012; Mendelsohn et al. 2012), which evidences the topicality of the abilities of communities to forecast the locations and magnitude of TC development. TC intensification involves a multiscale process that includes interaction between convective clouds both within the inner core and outside the inner core zone (i.e., within the radius of maximum wind), as well as factors related to TCs scale and the large-scale environment, which consists of sea surface temperature (SST), moisture and vertical wind shear. (Chang and Wu 2017; Hendricks et al. 2010; Fudeyasu et al. 2018; Wu et al. 2012). According to Hong et al. 2000; Shay et al. 2000), it was discovered that the warm-core eddies in the inner core might play a facilitating role in forming TC by reducing the SST cooling due to the upward flow. In particular, rapid intensification events are not reliably predicted by existing models, whether numerically based on weather prediction

(NWP) or machine learning, especially over the BoB. Both physical modelling and deep learning methodologies lack an integrated high-resolution model regionally specific to the BoB to make accurate predictions of TCs intensity and rapid intensification events (Singh et al. 2022). It creates a very large research gap given that the region is already highly susceptible and that past methods of forecasting are weak. The Deep TC comprises the 2D and the 3D modules and three fully connected (FC) layers. This research mainly focuses on developing a deep learning-based short-term model using CNN for forecasting 24-hour TC displacement in the BoB, considering the fused dataset. Future studies can investigate hybrid systems that invest in physical studies and AI-based learning systems to enhance both intensity and track predictions of TCs.

2 Methodology

2.1 Study area

The Indian Ocean is the most susceptible area of cyclones globally, and it extends to the BoB within the boundaries of N 23°, W 78°, S 5°, and E 100° (as shown in Fig. 1). It is bounded by the west by the coast of the eastern parts of India and Sri Lanka, the north by the Ganges-Brahmaputra-Meghna delta of Bangladesh, the east by the coast of Myanmar and the Andaman-Nicobar Islands and the south by the imaginary line drawn between the point of Dondra Head, Sri Lanka, and Northern Sumatra (Li et al. 2013; Bhardwaj and Singh 2020). This region comprises approximately 2.17 million square kilometres of land and averages 2600 m in depth, which translates to making it a very large and dynamic marine environment (Wu et al. 2024). The cyclones of this basin are mostly seasonal and have a bimodal seasonality pattern during the pre-monsoon (April-May) and post-monsoon (October-November) seasons (Li et al. 2013). According to the latest climatological reports, there is an increasing rate in cyclone frequency and strength in the BoB, especially the northwestern side (as shown in Fig. 2) that comprises the coasts of Odisha and Bangladesh (Wu et al. 2024). The frequency of STCs whose persistent winds surpass 64 knots has increased and is characterised by a trend of mainly moving in a southeast-northwest direction, which further exposes the densely inhabited coastal areas to more of this danger (Rao et al. 2024). Due to the overlapping of cyclones with the maturing seasons of Boro and Amon rice, agricultural losses and food insecurity are increased in Bangladesh (Wahiduzzaman and Yeasmin 2024). Traditionally, the BoB has experienced certain catastrophic cyclones in history, the most recent of which were the 1999 Odisha Super Cyclone, Cyclone Nargis (2008) and Cyclone Amphan (2020). The events have led

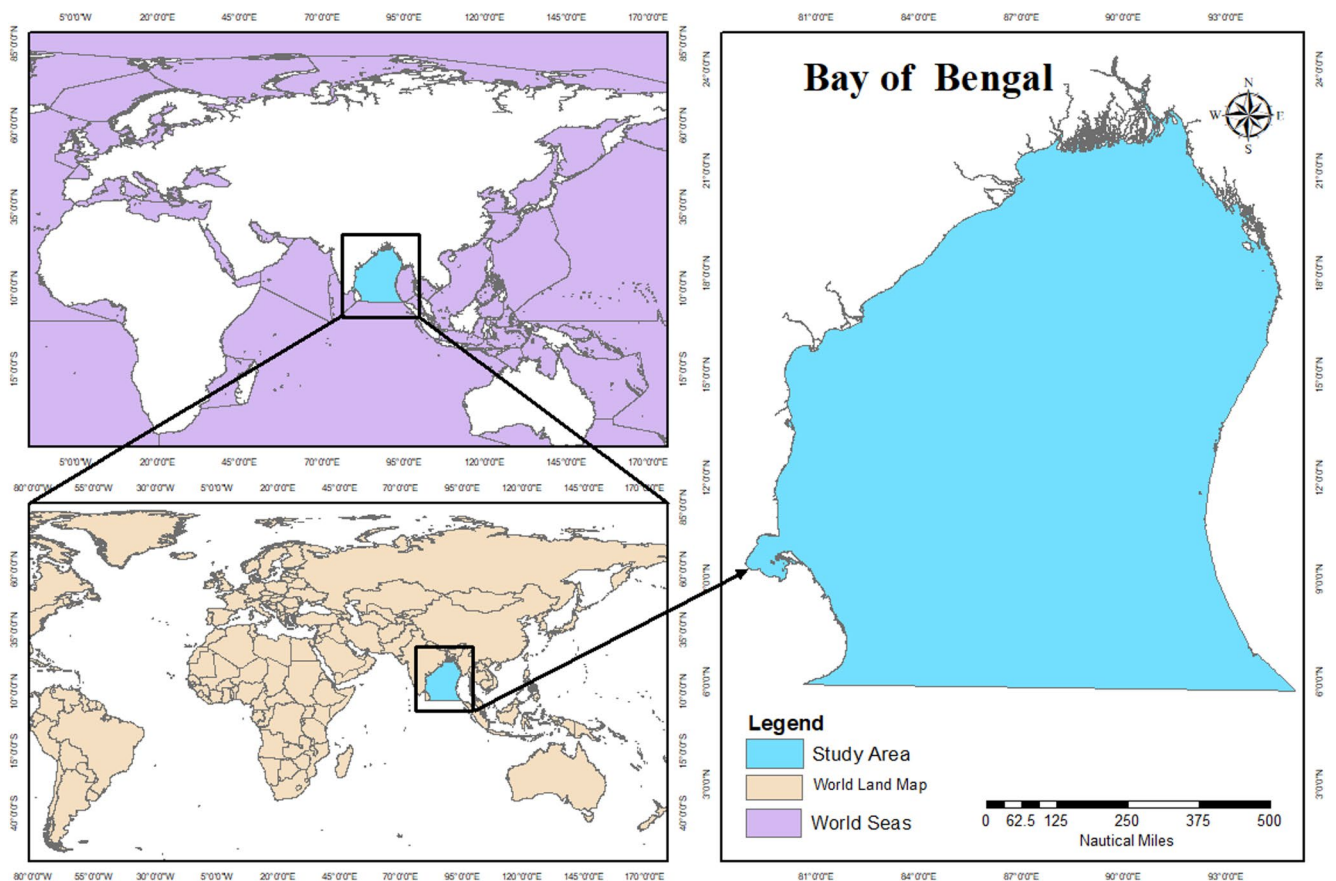


Fig. 1 Study area map of the BoB region

to large numbers of casualties, economic destabilisation, and long-term ecological degradation (Bhardwaj and Singh 2020). Especially low lying is the Sundarbans region, which is at a very vulnerable risk of storm surges, tidal inundations, and post-cyclone epidemics as a result of its poor ecosystem (Khatoun et al. 2025).

2.2 Data and methods

This paper combines two globally known datasets, IBTrACS and ERA5, to provide effective fused training conditions for the cyclone displacement forecasting in the BoB region. IBTrACS released by the National Oceanic and Atmospheric Administration for Environmental Information, is the most comprehensive global compilation of tropical cyclone best track data. It combines contributions made by various global agencies to be consistent and comprehensive in historical and contemporary events. The IBTrACS version 4r01 will offer the 6-hour cyclone position and strength information, such as latitude, longitude, wind velocity, and storm type, which are the parameters that would improve our understanding of cyclone genesis, movement, and landfall behaviour of the system of the BoB basin (Knapp et al. 2010; Gahtan et al.

2024). The update frequency of the data is every third day of the week and is maintained under the World Meteorological Organization (WMO) in regard to inter-agency comparison and stewardship of climate.

2.2.1 IBTrACS (International Best Track Archive for Climate Stewardship)

The IBTrACS dataset is used in this study, which is the best existing available dataset for cyclone tracks. Sourced from NOAA, the v04r01 dataset provides historical 6-hourly cyclone track data, including time, latitude, longitude and storm nature. Averaging of wind speeds by arithmetic mean was phased out in IBTrACS v03r01 due to feedback and suggestions by participants of the first IBTrACS workshop (Levinson et al. 2010) because of diversity in the derivation of the data. As IBTrACS is unable to merge these large discrepancies, or “spurs”, in the data, it splits them into separate tracks. To classify these tracks, IBTrACS pre-assigns a “track type” flag as main, merge, split, or other (Knapp et al. 2018). Most significant parameters include the tropical cyclone position (longitude position and latitude position) and storm centre of circulation of TCs (Li et al.

Dominant Trajectory of Cyclones Affecting the Region

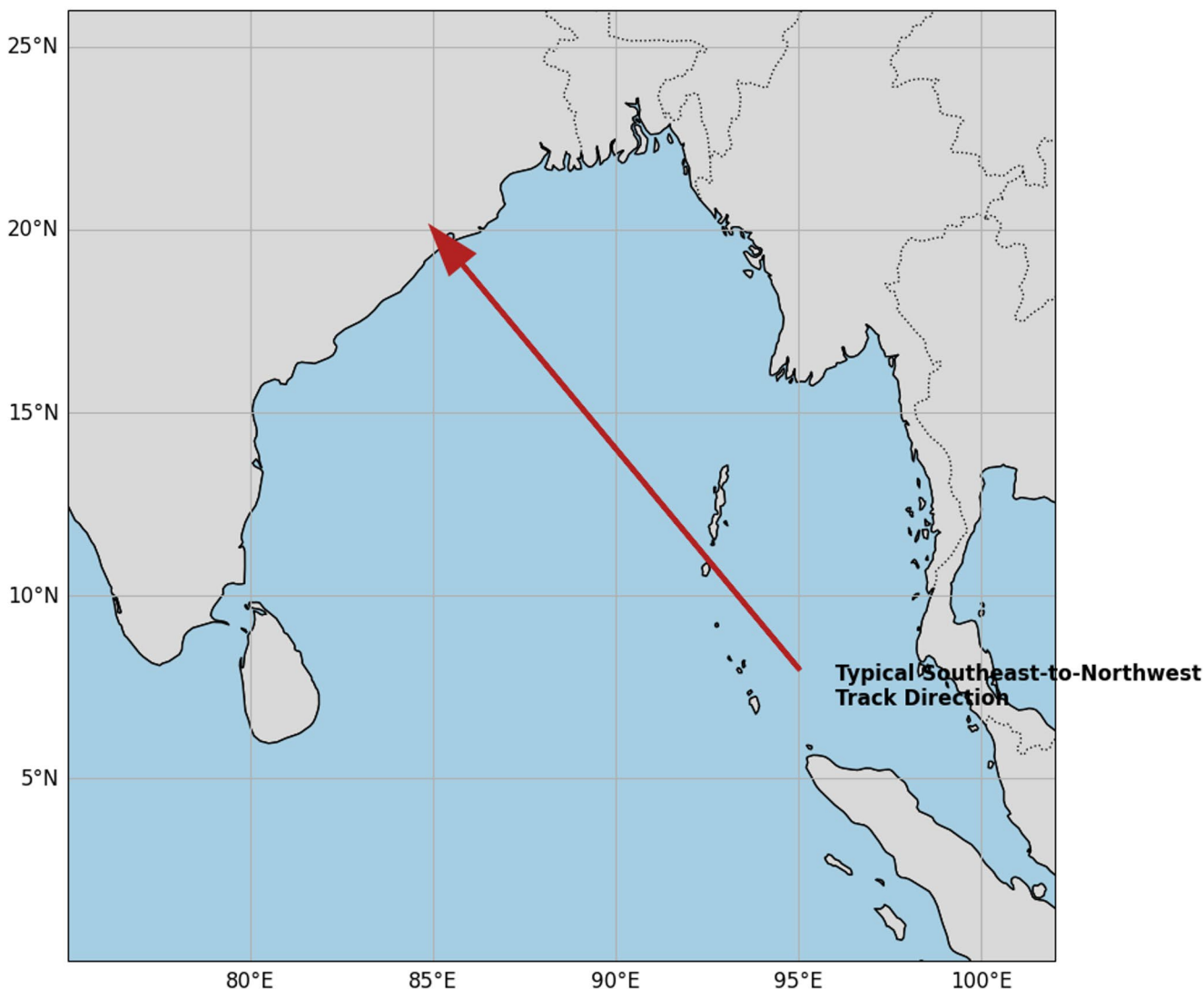


Fig. 2 Dominant trajectory of cyclones affecting the region of BoB

2022). Disaster early warning highly relies on the strength of a cyclone. Under 1 min wind speeds (Knapp et al. 2018), IBTrACS includes two storms: one track prior to and one track after the storm merger defined as the convergence of two cyclone tracks into a single main track and another track which ceases at the merge point, the spur (Kruk et al. 2010). In the BoB, TCs tended to form over the eastern and western sector of the basin and made landfall on the east coast of India and the south coast of Bangladesh (Uddin et al. 2021). Figures 3 presents the annual number of 6-hourly TCs observations in the BoB (1981–2023). The datasets average 84 observations per year, but variability is substantial. For instance, 1987 and 2004 exceed 190 observations, while 1990–1995 drop below 40 per year, reflecting a well-documented dip in meteorological archives linked to

shifts in technology and reporting standards. From the late 1990s onwards, coverage becomes more consistent, though natural interannual variability remains. This temporal imbalance is a defining feature of the training data (1981–2010). By splitting chronologically, the model is exposed to heterogeneous records during training and evaluated against more stable, satellite-era data (2016–2024).

2.2.2 ERA5 reanalysis data

ERA5 reanalysis data are provided by the ECMWF- CDS (Climatic Data Store). It's a valuable resource for studying cyclone tracks; also, ECMWF-CDS is high-resolution gridded data which is a physically consistent global atmospheric state. Some of the benefits when compared

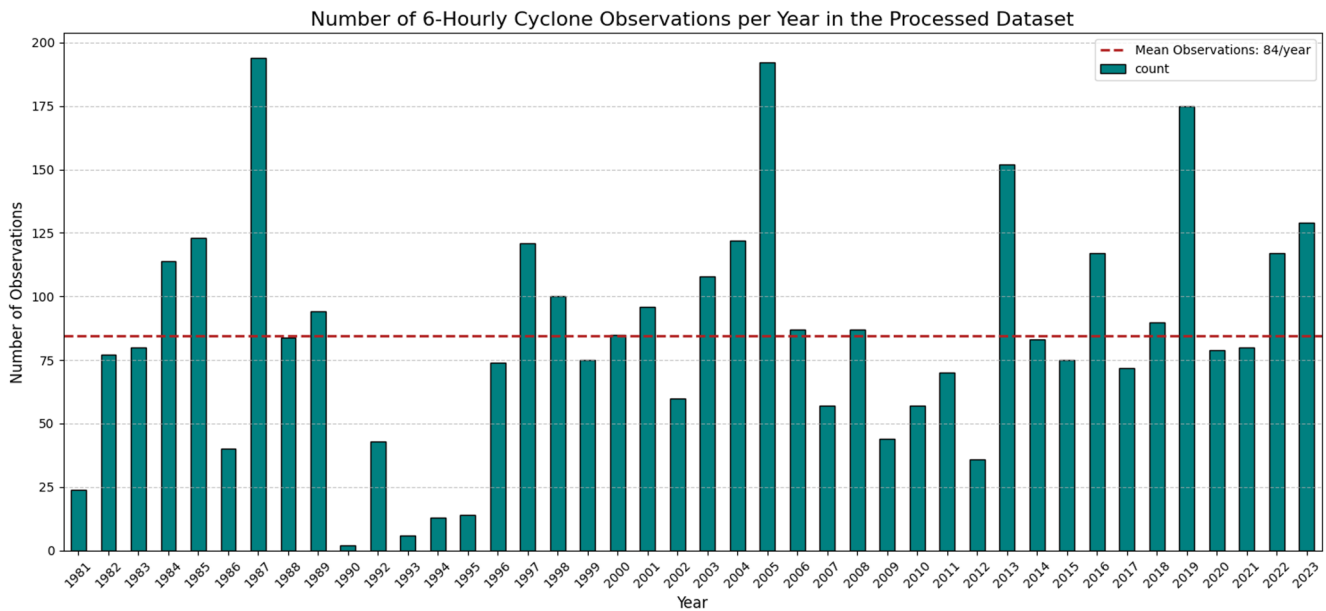


Fig. 3 Number of 6-hourly TCs observations per year in the final processed dataset in the BoB region (1981–2023)

to its predecessors (ERA4 and ERA-Interim) are that its spatial and temporal resolutions are much larger, it provides details about the quality of data changing over space and time, it describes the troposphere much better, it represents TCs much better, it offers a better a global mean balance of precipitation and evaporation, it has better performance over land precipitation in deep tropics, it has better performance in representing soil moisture and better sea surface temperature and sea ice (Hersbach et al. 2020). In correspondence, a different data filter powered by Deep Learning Models (DLM), CNN, as a widely recognised method due to the ability to process information of a large-scale range, is applied (Wei et al. 2023). ECMWF has been operating a combined medium-range and monthly ensemble system since 11 March 2008, created by combining the former Variable Resolution Ensemble Prediction System (VAREPS) and monthly versions (Buizza 2008; Vitart et al. 2008). We used 6-hourly data from 1981 to 2024 for five single-level variables (assumed to be 2 m temperature, total precipitation, 10 m u-wind, 10 m v-wind and mean sea level pressure) covering the BoB region.

2.3 Cyclone tracking algorithm

ERA5 reanalysis data has greatly increased the spatial and temporal resolution of cyclone tracking, allowing closer detection of more short-lived and small-scale systems. Table 2 reveals that the most important meteorological variables, which are temperature, precipitation, and wind components, are also important predictors in cyclone modelling (Pinto et al. 2005; Neu et al. 2013).

Table 2 Description of Input Data and Variables

Data sources	Parameter Variable	Description	Role in Model
IBTrACS	Latitude/Longitude	6-hourly geographic coordinates of the cyclone centre	Ground Truth for location and target calculation
	Time (ISO-TIME)	The specific date and time of each observation.	Used for matching with ERA5 data and for data splitting.
ERA5	t2m (2 m temperature)	Air temperature at 2 m above the surface.	Input feature/predictor variable.
	tp (total precipitation)	Accumulated precipitation over the preceding period.	Input feature/predictor variable.
	u10 (10 m u-wind)	The eastward wind is 10 m above the surface.	Input feature/predictor variable.
	v10 (10 m v-wind)	North wind at 10 m above the surface	Input feature/predictor variable.
	msl (mean sea-level pressure)	The atmospheric pressure adjusted to mean sea level.	Input feature/predictor variable.

2.4 Data processing

We build our methodology around producing a training instance for every 6-hourly data point of a tropical cyclone event.

2.4.1 Target calculation

Target calculation is devoted to the process of defining and generating the output value that the model is expected to predict. The model target is the future cyclone centre coordinate (latitude and longitude) at T+24 h. For each storm observation at time T, we identify its location 24 h later at T+24 h. We also target the transformation forecasting task into a supervised regression problem were,

$$\text{Target} = (\text{Latitude}_{t+24\text{hr}}, \text{Longitude}_{t+24\text{hr}})$$

The target variables for our model are displacement vectors (dlat_24hr, dlon_24hr).

2.4.2 Chip extraction

To create a standardized input for our CNN model, small subsets of the large-scale atmospheric data, referred to as ‘chips’ in computer vision literature, are extracted. A ‘chip’ in this context is a fixed-size, N-dimensional array representing a storm-centric snapshot of the environment. Specifically,

a $25^\circ \times 25^\circ$ chip of ERA5 atmospheric data is extracted, centred on the storm’s recorded location at given time T (as shown in Fig. 4). Each chip contains multiple channels, with each channel corresponding to an environmental variable like wind, precipitation, or mean sea-level, pressure. This method of creating storm centred chips is important for two reasons. First, it establishes a moving frame of references that ensures the cyclone is always at the centre of the input data. This allows the model to learn location invariant features of storm dynamics. Second, it transforms the computationally intractable problem of processing global weather maps into a manageable task by focusing only on the most relevant environmental data surrounding the cyclone. This spatial framing approach has been confirmed as significant in the context of cyclone detection and forecasting (Lakshmanan and Smith 2009; Jitkajornwanich et al. 2012).

2.4.3 Dataset creation

We built the final data set out of two main components, namely the chip extraction and the target calculation. Every sample is a 6-hourly analysis of a cyclone in BoB and

Conceptual Diagram of Chip Extraction

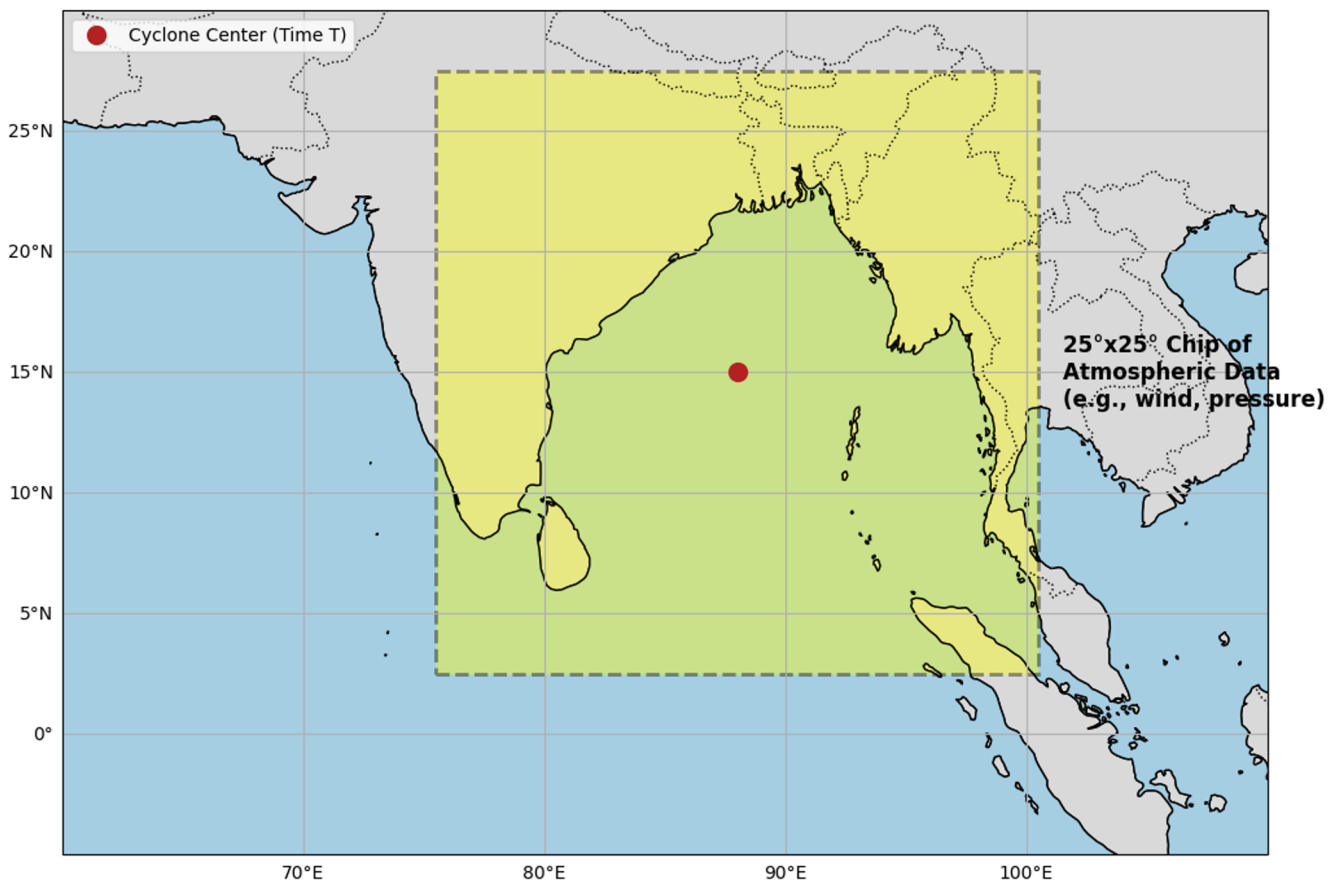


Fig. 4 Conceptual diagram of chip extraction

includes the environmental conditions as well as the storm progression. Data fusion is a widely adopted technique in cyclone modelling aimed at improving prediction accuracy and enhancing spatial coherence (Kieu et al. 2025; Huang et al. 2025). Target labels are 24-hour displacement vector parameters of the cyclone tracks. Such a vector-based methodology is confirmed in regional papers across the BoB (Georgiev et al. 2016; Shenoy et al. 2022). We received 3,548 samples after processing historical cyclone records. This data is used to create a baseline of training a DLM and assessing it. It ensures temporal stability and spatial relevance, both essential for generalising cyclone type and seasonal pattern (Bhardwaj and Singh 2020; Nanda et al. 2025).

2.4.4 Data split

The training sets used TCs data as recorded in the period 1981–2010, during which dataset quality varied; earlier years relied more on regional reports and limited satellite estimates, while later years benefited from improved satellite monitoring, refined Dvorak techniques, and stronger interagency consistency (Knapp et al. 2010; Gahtan et al. 2024; Kruk et al. 2010). The datasets were divided into three parts: a training set (1981–2010), a validation set (2011–2015), and a testing set (2016–2024). This resulted in 2,273 training samples (64.05%), 416 validation samples (11.72%), and 859 testing samples (24.22%) (as shown in Table 3).

2.5 Evaluation metrics

The performance of the CNN model for tropical cyclone track forecasting is primarily evaluated using the MAE, which represents the average great-circle distance between predicted and observed cyclone positions after 24 h. To ensure accuracy, this distance is calculated using the Haversine formula, which accounts for the Earth's spherical geometry and avoids distortions from planar approximations. The formula is expressed as:

$$d = 2R \cdot \arcsin \left(\sqrt{\sin^2 \left(\frac{\phi_2 - \phi_1}{2} \right) + \cos(\phi_1) \cos(\phi_2) \sin^2 \left(\frac{\lambda_2 - \lambda_1}{2} \right)} \right)$$

where d is the distance, R is the Earth's radius (6371 km), ϕ represents latitude, and λ represents longitude. As demonstrated

Table 3 Segment of the Experimental Dataset

Parameters	Year	Values (Samples)	Percentage of Samples
Data set amount	1981–2024	3548	
Training set	1981–2010	2273	64.05%
Testing set	2016–2024	859	24.22%
Validation	2011–2015	416	11.72%

by (Agrawal and Colle 2025), CNN-based approaches to cyclone track forecasting effectively employ MAE with haversine-based calculations to quantify positional errors and improve ensemble forecast skill.

3 Result

3.1 Tropical cyclone track

Unique 328 storm tracks are in our final dataset that occurred in the BoB region between 1981 and 2024. It assures that the processing pipeline successfully isolates storms within the BoB and adjoining seas, defining the geographic scope of our study as shown in Fig. 5.

3.2 Distribution of tropical cyclone durations

This histogram (as shown in Fig. 6) illustrates the distribution of tropical cyclone durations in 6-hour intervals across BoB from 1981 to 2024. In IBTrACS, tropical cyclone duration is defined as the period during which a system designated as tropical cyclone in the best track record, including tropical depression, tropical storm, and higher cyclone categories, but excluding precursor disturbances and post tropical stages. Out of 328 storms, the majority had a duration between 0 and 50 h, with the highest frequency observed in the range from 0 to 15 h. Fewer subsets of storms are over 150 h in duration. This result indicates that short to medium-duration cyclones dominated the BoB basin.

3.3 Target variable distribution

This (as shown in Fig. 7) refers to the geographical movement of the cyclone the centre in 24 h of displacement of latitude and longitude. The distribution is revealing an approximately normal and clear physical trend, with predominantly northward positive latitude change (ranging from 0°–2°) and westward negative longitude change (from –2° to –1°) movement.

Most TCs are exhibit:

Positive latitude change ($\Delta \text{Lat} > 0$) → Northward movement.

Negative longitude change ($\Delta \text{Lon} < 0$) → Westward movement.

3.4 Model architecture and training

A CNN was selected because it proven ability to learn spatial features in grided like-data. The model will comprise a series of convolutional and pooling layers to learn features, after which the model will then employ

(Revised): Cyclone Tracks in the Bay of Bengal Study Area (328 storms)

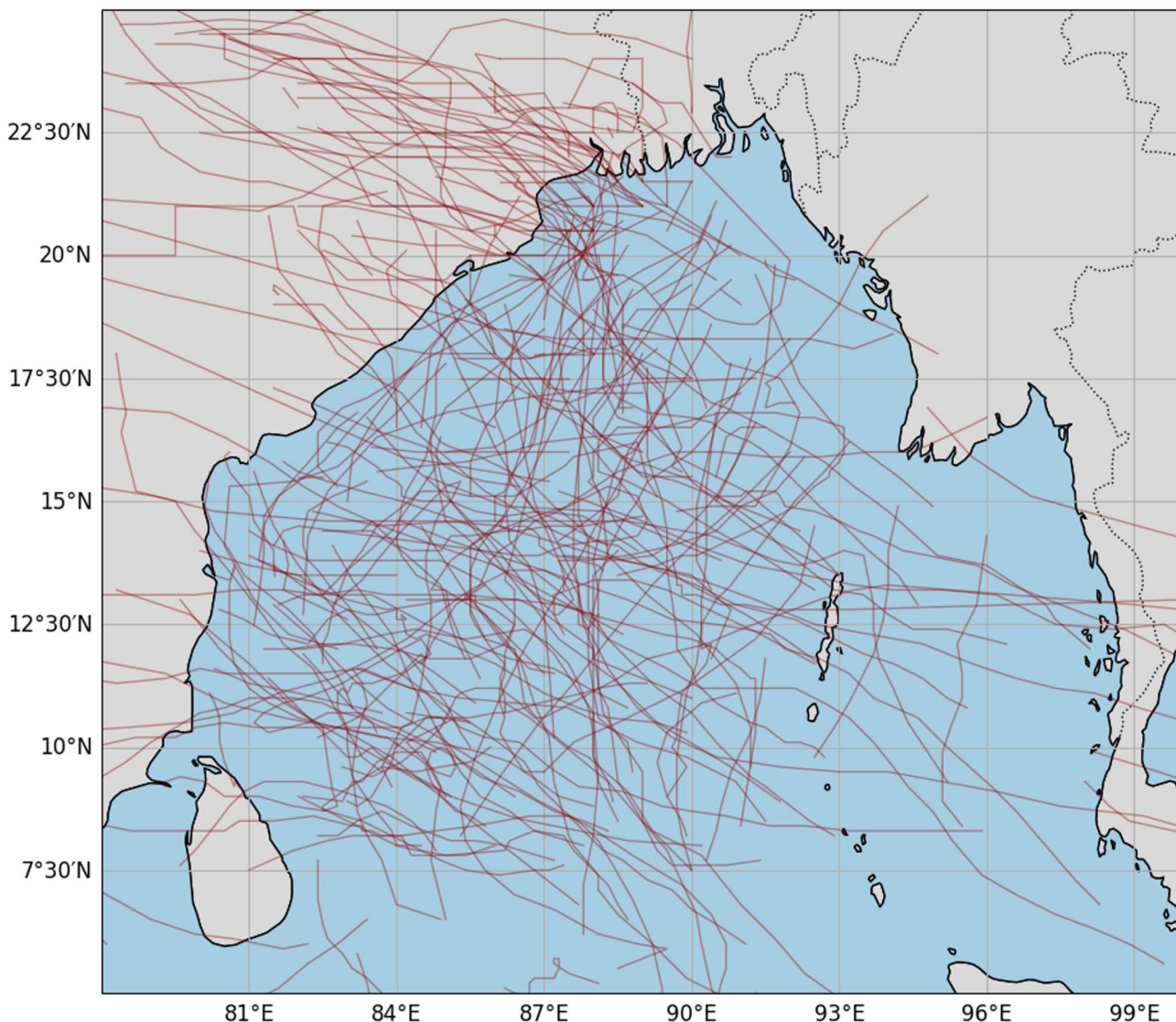


Fig. 5 Total Storm (328) Track in the Bay of Bengal spanning 1981–2024

dense layers to regress the final output displacement vector. To identify the best model configuration, we conducted a set of experiments, among which there was a features analysis (ablation study) and a parameters analysis (hyperparameter tuning). The set of experiments compared a baseline model, a shallow a model with fewer layers, a model with a higher lr (learning rate), and a model trained without wind (i.e., excluding all wind-related predictors to evaluate their incremental contribution), and a model trained after removing all non-wind variables. The best kernels of each of the convolutional layers are highly optimised, and they excite the pertinent regions (using the support of several activation functions (e.g., rectified linear unit [ReLU], hyperbolic tangent [tanh] or linear)

which are further used to extract the output. Each of the models was trained for up to 100 epochs, with an Early Stopping callback (patience=10) to prevent overfitting. Our best-performing model, the shallow model, converged and stopped training after 24 epochs, with the Adam optimizer and Mean Squared Error (MSE) loss function (as shown in Fig. 8).

3.5 Quantitative result of model experiments

The quantitative results from the model experiments indicate varying performance across different configurations. This shallow model, utilising two convolution blocks, achieved the best validation loss and test MAE, demonstrating the

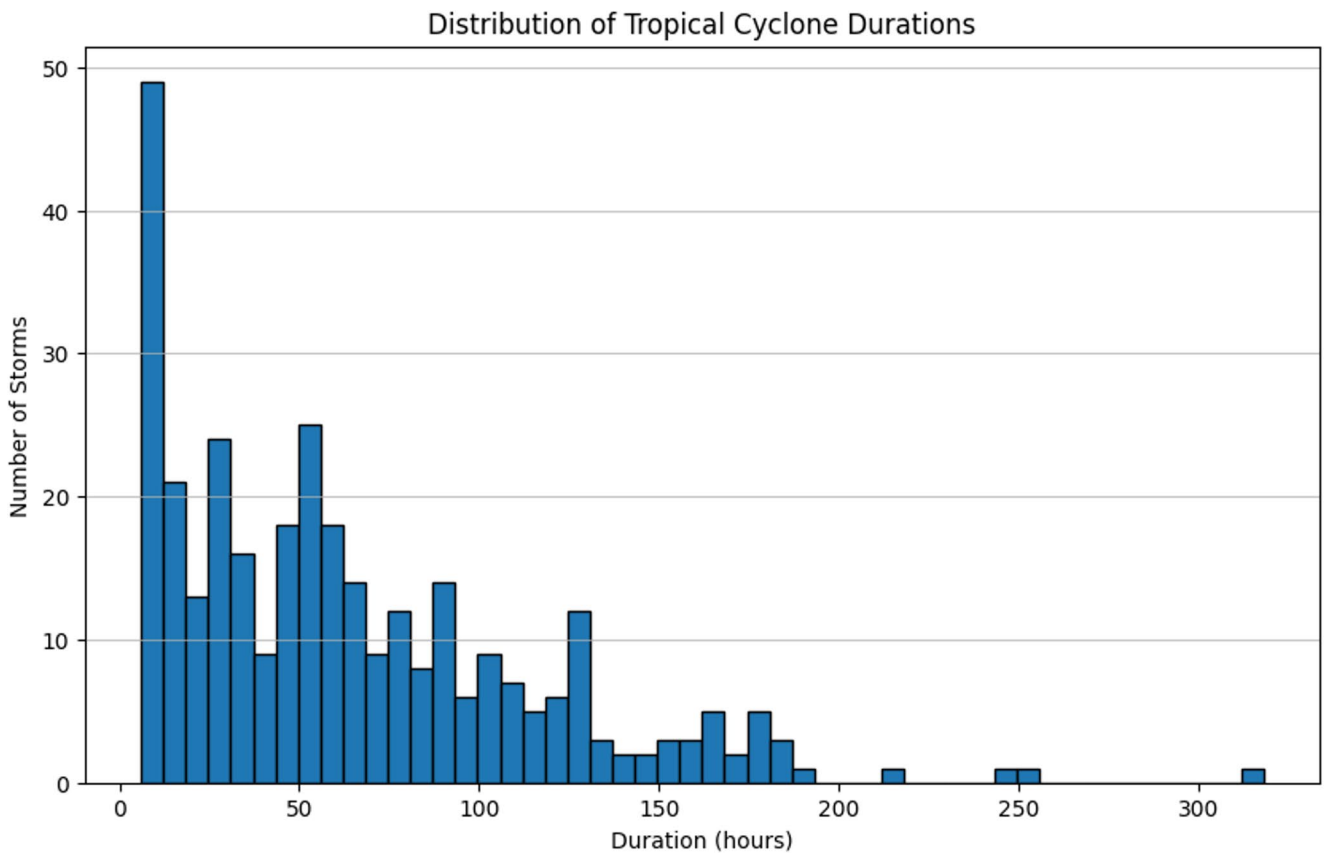


Fig. 6 Distribution of tropical cyclone durations measured at 6-hour intervals across the Bay of Bengal

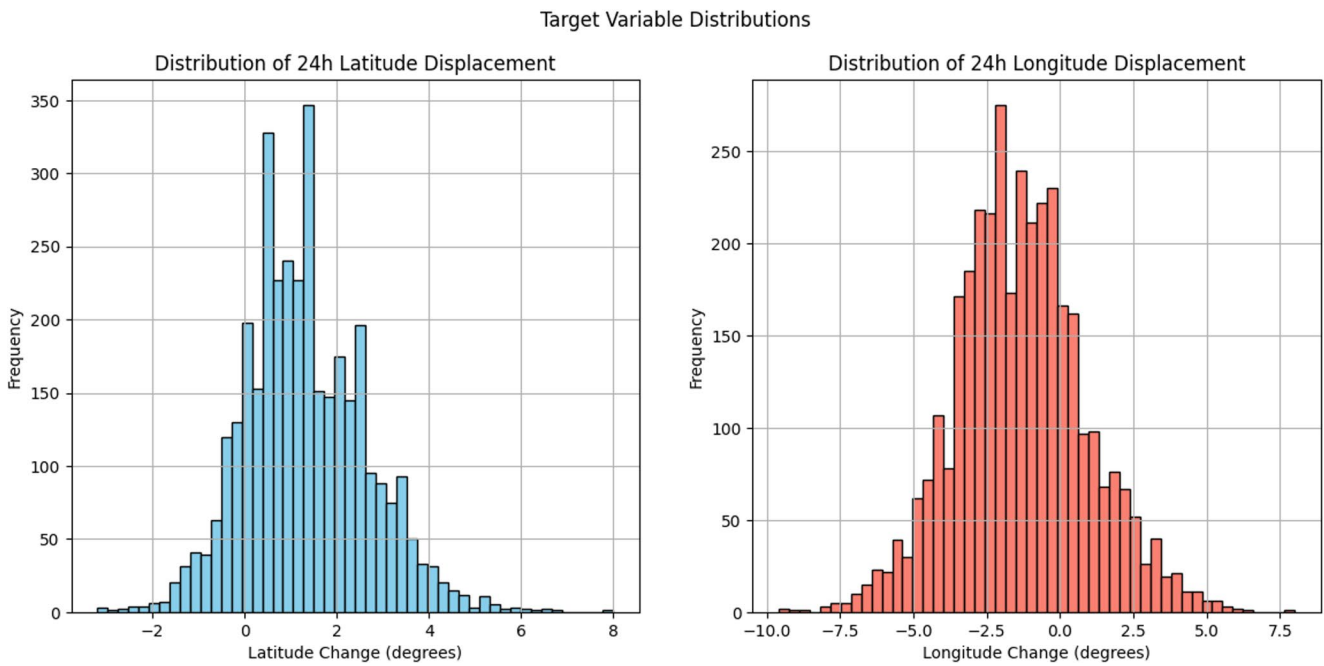


Fig. 7 24-hour displacement of Tropical Cyclone in centres-based changes latitude and longitude

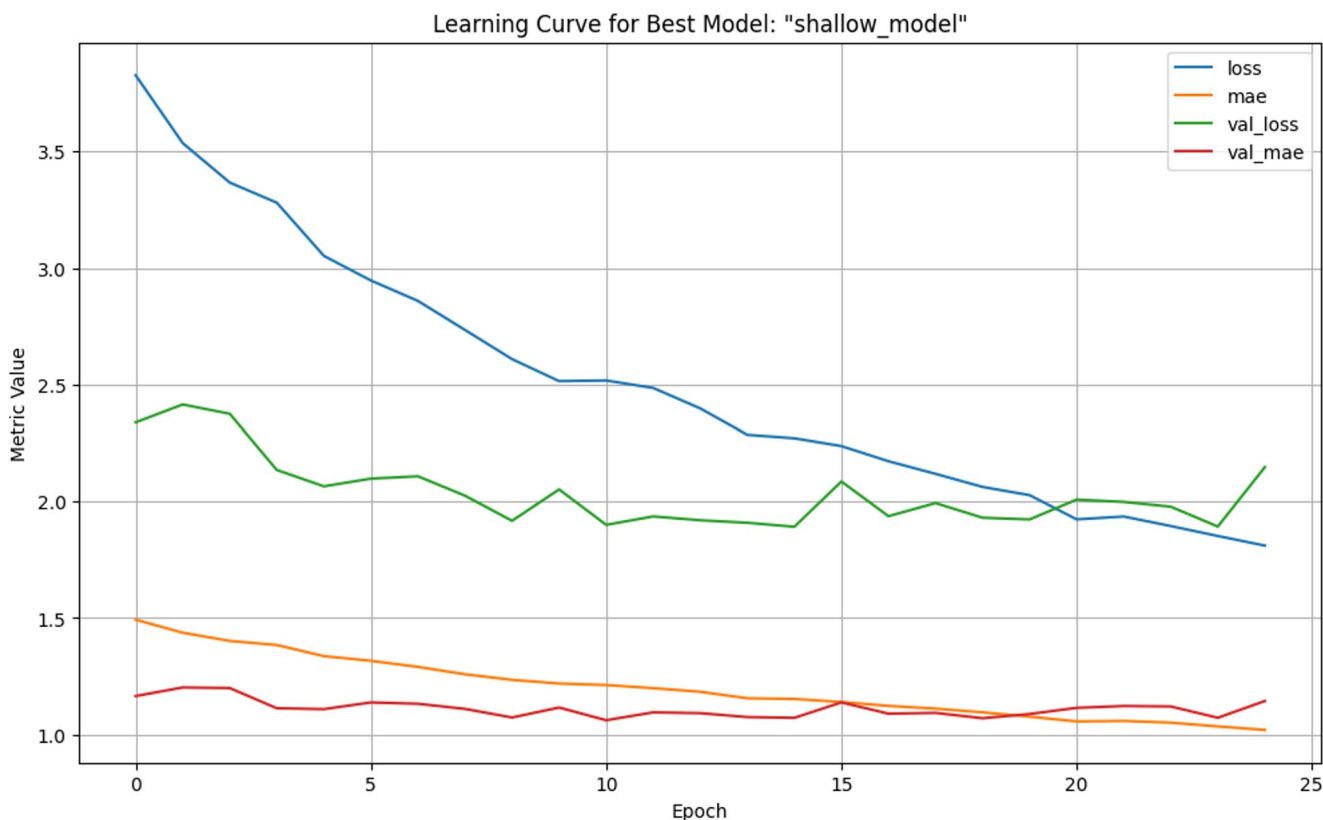


Fig. 8 The training and validation loss of the shallow CNN model after 24 epochs

effectiveness of its feature set. In contrast, adjustments in learning rate and model complexity resulted in slightly higher errors (as shown in Table 4).

3.6 Model performance across all experiments (Test MAE)

The performance of machine learning models can vary meaningfully based on their architecture, particularly in relation to dataset size, where we used 3,548 samples. In this case, the shallow model achieved the lowest test set MAE of 141.2 km (as shown in Fig. 9), revealing that simpler models can effectively avoid overfitting while maintaining faithful accuracy. The shallow model performance assumes that simple architecture is better on smaller datasets since they are not overfitting (Pasupa and Sunhem 2016). The risk of overfitting the perplexity of most models is high, the models tend to learn noise in the data during training and make inaccurate decoding on new data (Schmidt 2025). The shallow model’s architecture relieves overfitting risk effectively. Key performance metric of a measurement model is the MAE, especially on continuous response data, because it gives a clear insight into the accuracy of prediction (Montesinos López et al. 2022).

Table 4 Performance Evaluation of Model Experiment

Experiment Name	Model Configuration	Feature Set	Validation Loss (min)	Test MAE (km)	Performance vs. Baseline
shallow model	2 Conv Blocks	All 5	1.892	141.2	Best Performance
higher lr	3 Conv Blocks (LR=5e-4)	All 5	1.965	140.70	+0.30 km
baseline	3 Conv Blocks (LR=1e-4)	All 5	1.984	140.94	+0.54 km
no wind	3 Conv Blocks	t2m, tp, msl	2.451	165.82	+24.88 km
wind only	3 Conv Blocks	u10, v10	2.137	148.15	+7.21 km

3.7 Distribution of forecast errors on test set (Best Model)

The distribution of forecast errors in the weather prediction model specifically for TCs that reveals significant insight into model performance in these figures (as shown in Fig. 10) determines that the great majority of errors are clustered under the mean 223.69 km with a high precision of about 150 km. We can see an increase of fusing networks

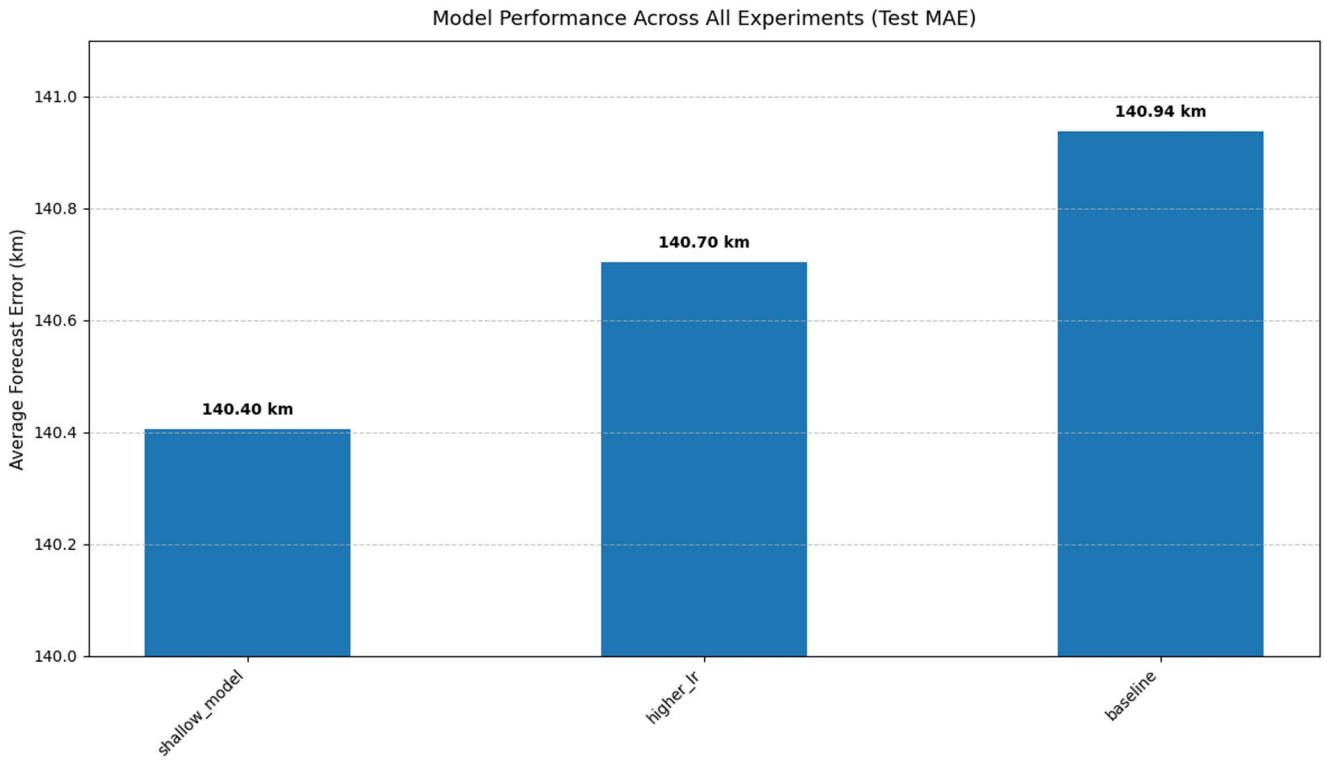


Fig. 9 Mean forecast error (MAE) for shallow, higher learning rate, baseline CNN models

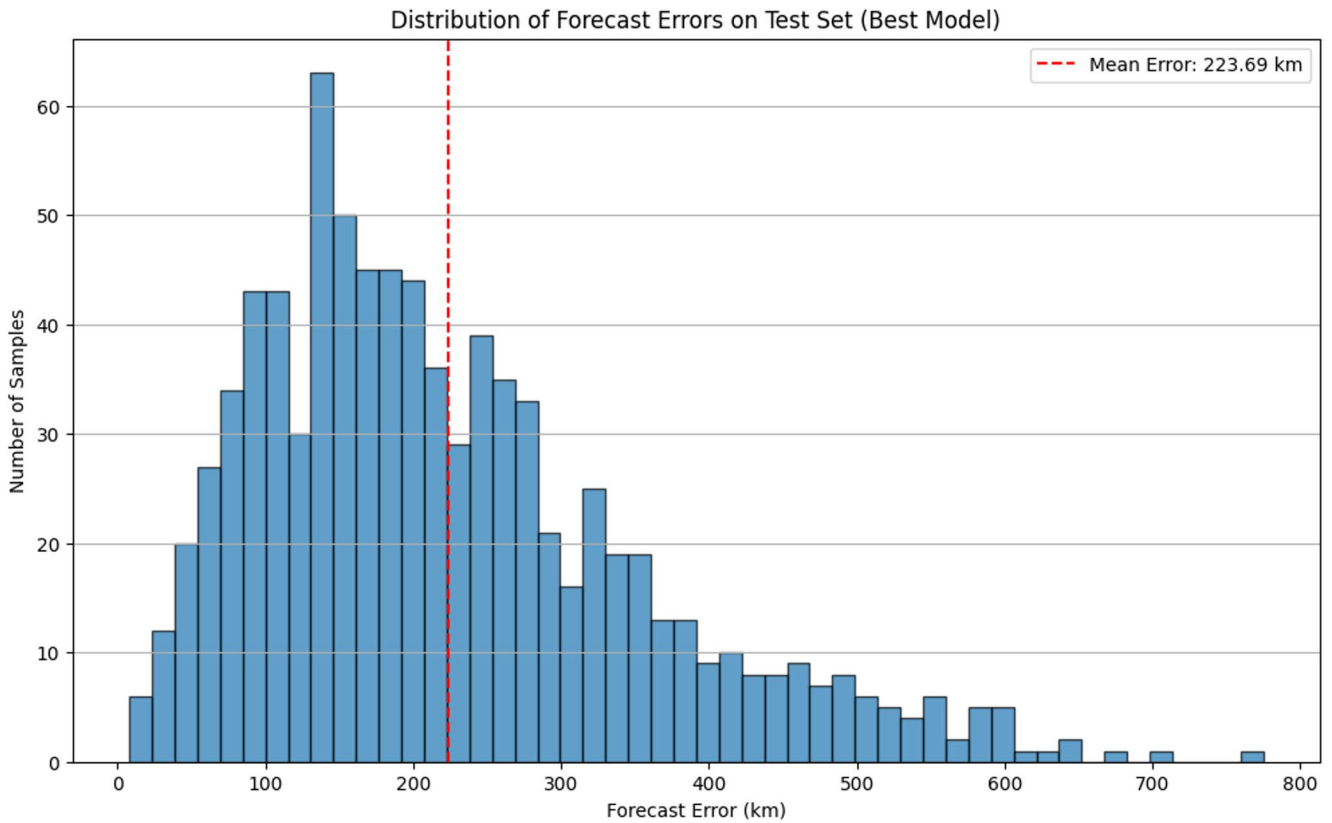


Fig. 10 Distribution of performance comparisons-based on forecast errors of the best model for tropical cyclones

amplification (mean error: 130 km where reference is to the Wind CNN (mean 148.9 km), the Pressure CNN (mean 172.7 km), and the Past tracks+meta NN mean 186.6 km (Giffard-Roisin et al. 2020) alone. These forecast errors indicate that there are some forecasts making larger error, which is typical of meteorological prediction. This pattern of forecast error distribution means that tropical cyclone forecast errors (Fang et al. 2022) commonly have multimodal distribution and that the assumption of normality in error statistics may be invalid (Crutcher et al. 1982). Models demonstrate strong predictive capability; it should be noted that even in these cases, forecasts can still be very wrong in some circumstances and indicate that work is needed to produce enhancements in the modelling technique beyond error correction methods (Bouallègue et al. 2023).

3.8 Model performance on test data (Best Model)

The analysis of model performance on test data reveals a strong positive correlation between predicted and actual displacement of TCs, as evidenced by plots (as shown in Fig. 11) clustering around the 'Reference Prediction' line. This prediction line result indicates high predictive accuracy; clusters are centred in the top-left quadrant and particularly in capturing the north-westward movement of cyclones in BoB. The CNN forecasts were compared against the persistence model, a standard operational baseline. Using Haversine distance metric, outcome underscore the difficulty of surprising

physically meaningful baselines in cyclone track prediction. Kumar et al. (2021) reported that Long Short-Term Memory (LSTM) network trained on cyclone central pressure, wind speed, latitude, and longitude, achieved lower MAE and outperformed existing models in both Atlantic and NIO. Zhang et al. (2018) showed that Matrix Neural Network (MNN) preserved spatial correlation in cyclone trajectory data, producing more accurate forecasts for south Indian ocean.

3.9 Model interpretability (Explainable AI)

CNN models that predict TCs are visualisations of their internal convolutional layer activation in DL models and also provide meaningful insight into how these models interpret atmospheric data. Primary layers such as conv_1a and conv_1b focus on low-level features like edges and gradients, while deeper layers like conv_2a and conv_2b capture more complex patterns, indicating the model's ability to recognise meteorological phenomena.

3.9.1 Activations for layers: conv_1a and conv_1b

Conv_1a and conv_1b activation layers (as shown in Figs. 12 and 13) learn fundamental features, including simple edges, gradients (Lines of changing pressure), and basic textures related to wind fields and cloud patterns. These features are crucial to the interpretation of model predictions in the context of meteorology (Zhang et al. 2019).



Fig. 11 Predicted vs. Actual displacement on test data

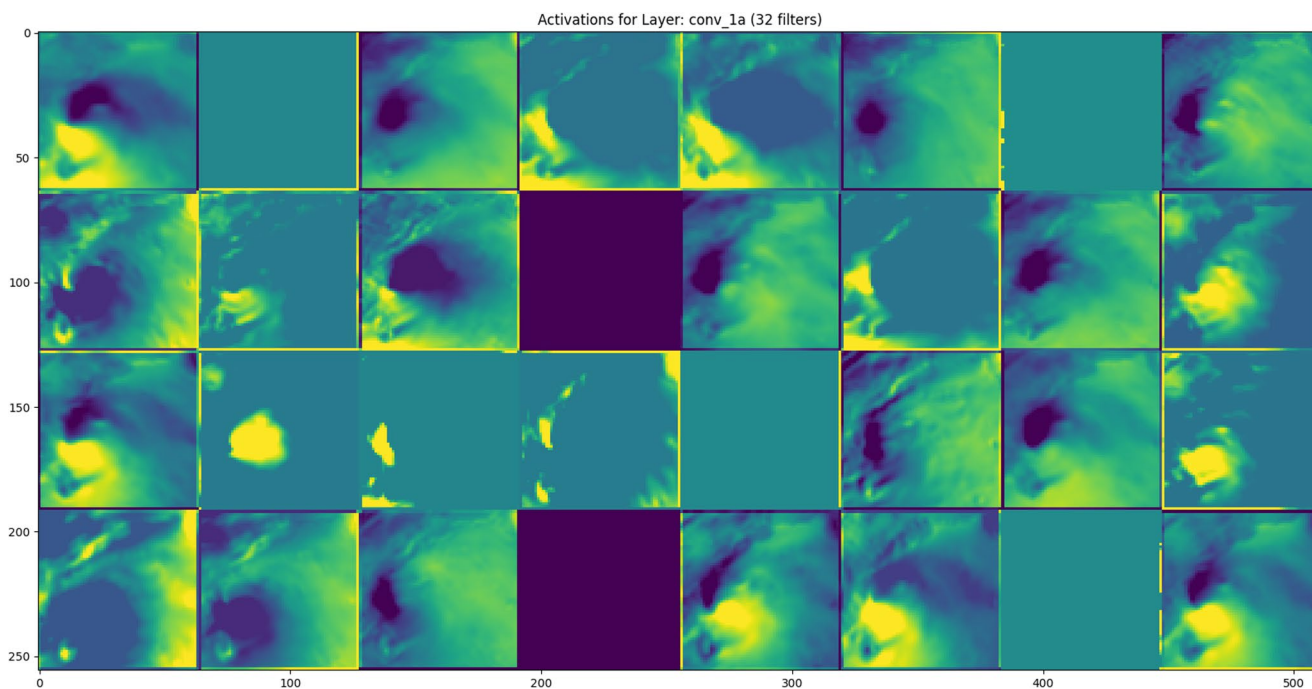


Fig. 12 Activations for Layer conv_1a

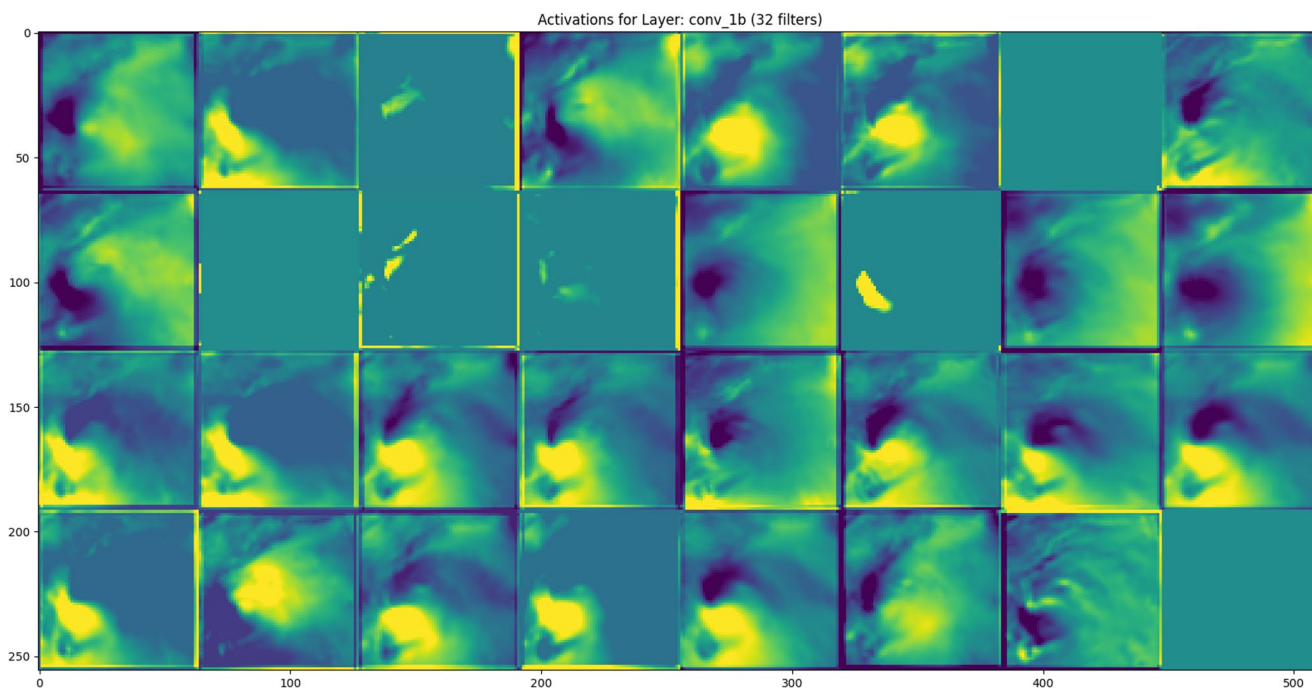


Fig. 13 Activations for Layer conv_1b

3.9.2 Activation for layers: conv_2a and conv_2b

These layers are deeper in the network to identify complex structures such as curved formations typical of cyclone rain bands and rotational signatures near storm

cores. This statistical mode of exhibition is more than a black box and identifies physically sensible data patterns (Henna and Alcaraz 2022). The learnt activations of the conv_2a and conv_2b layer blocks are demonstrated in Figs. 14 and 15 respectively.

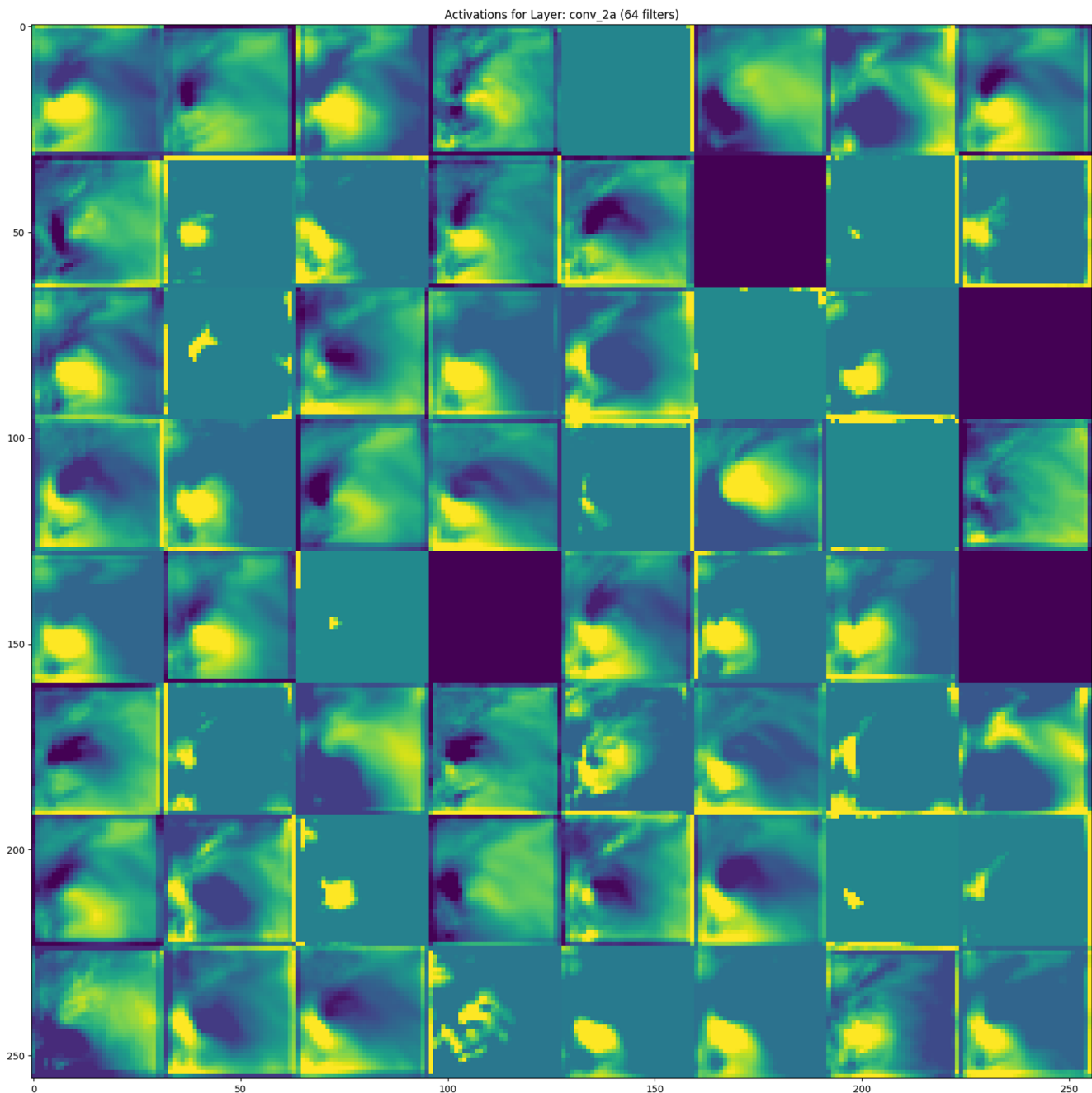


Fig. 14 Activations for Layer conv_2a

4 Discussion

The forecast of TCs highlights the importance of accurate prediction for effective disaster management and building the sustainable city. Modern advancements in forecasting methodology, particularly utilising machine learning and numerical models, have shown promising result in predicting both natural and anthropogenic phenomena.

Combinations of the capabilities and strengths of the machine learning model and the numerical weather forecast model, the so-called hybrid models, may be the ones of interest to the development of TC forecasting. Such as hybrid approaches are physically meaningful because they combine the process-based realism of numerical models with the adaptability of machine learning, allowing forecasts to remain robust under rapidly

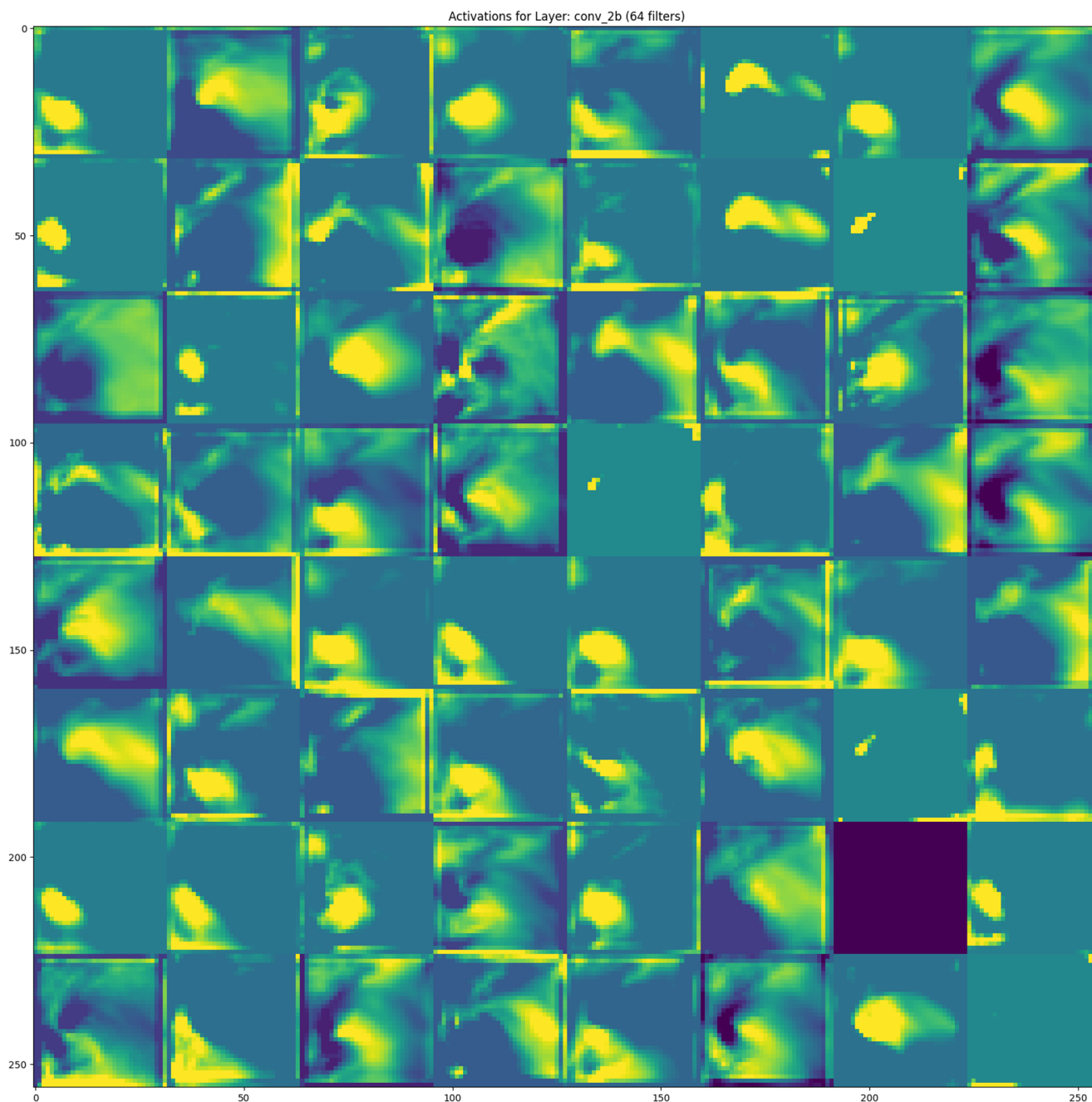


Fig. 15 Activations for Layer conv_2b

evolving atmospheric conditions. A westward and slightly southward motion of an emerging tropical cyclone occurs over the BoB in our 24-hour cyclone forecast map (as shown in Fig. 16) 24-Hour Forecast for New Cyclone dated 20 November 2025 at 00.00 UTC. The historian's initial position on the green circle, further than the red star with the expected location 24 h later, indicates that a path is taken in a new direction to the coastal areas of Bangladesh and eastern India. This trajectory reflects the

influence of mid tropospheric steering flows and synoptic scale pressure gradients, which are well known drivers of BoB cyclone motion. Based on the CNN model, the predicted 24-hour displacement error is -0.08° in latitude and longitude -2.4° . The relatively low error in latitude and longitude highlights the model strength in representing poleward steering flows, which are commonly associated with BoB cyclones. Cyclone movement is also in line with the climatological patterns of cyclones

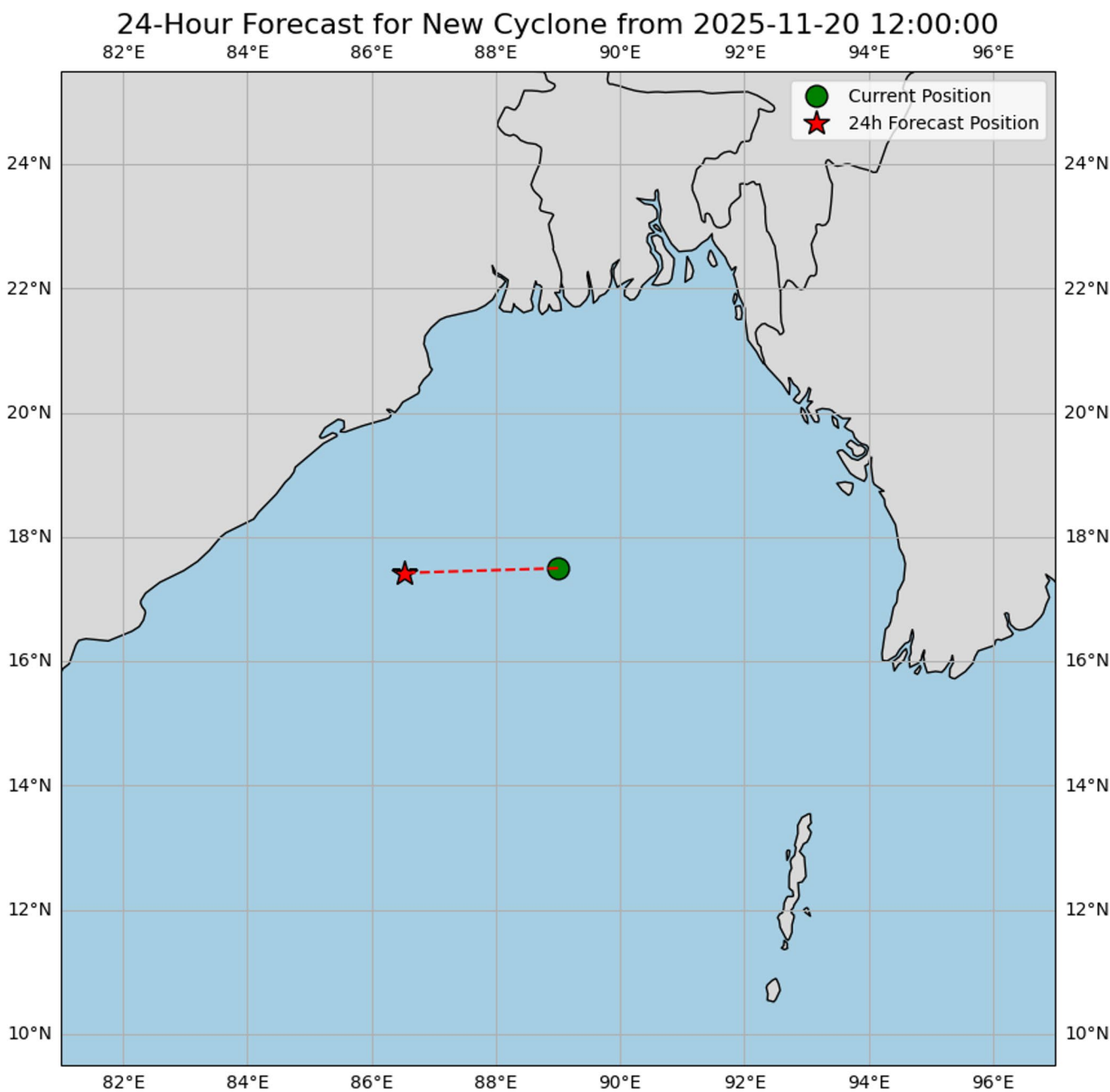


Fig. 16 24-Hour forecast for new cyclone

in BoB, which originate in the central region of the basin and tilt northeast due to the steering flows and synoptic-scale pressure gradient in the mid-troposphere (Rahman et al. 2024). This combination helps the model-to-model cyclone development at decisive moments, a development that enhances the accuracy of making short range forecasts. This paper presented a CNN-based model trained with fused IBTrACS-ERA5 observations to 2024 to enable the model to learn temporal domain shifts and environmental preconditions. The next-day prediction of 20 November

2025 is very close to the historical cyclone dynamics, indicating that the model demonstrates predictive capability consistent with historical cyclone dynamics. This alignment indicates that the CNN model successfully captures basin scale circulation patterns, reinforcing its ability to generalize beyond training data. These models can represent temporal dynamics and also spatial dependencies to provide an increased appreciation of the processes of cyclone intensification. Further, the identification of the optimal predictor group also contributes to the efficiency

of the model (Xu et al. 2022). In this paper to demonstrate the model's practical efficiency, we performed simulation a real-time forecast for a hypothetical new cyclone. The simulation provides the opportunity to directly evaluate the response of the model under realistic conditions and determine the practical usefulness of a given model in the context of disaster preparedness. With an improved TCs forecasting model, it is possible to better anticipate risks and reduce potential damage (Xu et al. 2022). Our proof-of-concept is found at 24-h, and Giffard-Roisin et al. (2020) shows that the 6-hourly observation results are also very satisfactory. Predictions can be made in the same format on a longer-term basis (Giffard-Roisin et al. 2020). This forecast model has the performance ability to capture nonlinear spatiotemporal patterns in atmospheric data, which are often missed by conventional statistical models. Low displacement error in latitude suggested improved representation of poleward steering dynamics, a critical factor in cyclone trajectory forecasting. The reduction in error demonstrate that hybrid architectures can extract multi-scale atmospheric features, enabling more reliable short-term track forecasts (Richman et al. 2017). This reflects the model's capacity to capture both mesoscale convective process and large-scale steering dynamics simultaneously, strengthening its physical credibility. For 24-hour forecast horizon, the CNN model achieved a MAE of 141.2 km is calculated using the haversine formula (as shown in 2.5). To provide a robust context, we compared our model against the persistence baseline, which achieved a MAE of 186.11 km. Comparatively, the CNN model achieved a significantly lower MAE of 141.2 km, demonstrating strong predictive capability beyond simple momentum. This result confirms that the model has successfully learned meaningful patterns from the atmospheric data to improve upon a naive forecast. Additional experiments showed initial angular errors of 2.9° reduced to 1.8° with the hybrid architecture, reflecting improved feature extraction and short-term prediction accuracy. CNN model consistency in abrupt trajectory shifts highlights the adaptability of data driven models to nonlinear atmospheric dynamics, complementing the physical realism of WRF. Integrating empirical decomposition with neural networks enhances representation of short-term wind speed variability, which is vital for operational early warning systems. The Statistical robustness of forecast improvements indicates that the gains are not random, but stem from the architecture's ability to capture spatiotemporal dependencies. The accuracy could also be enhanced by a recent launch of the new version of ERA5 reanalysis, ERA5 (Giffard-Roisin et al. 2020). As (Hodges et al. 2017) demonstrate, the difference in the mean centre

position of TCs in the ERA5 reanalysis product can be as high as 1 degree between 1979 and 2012, so we would gain performance by transferring to ERA 5 and the Vortex Tracker (Marchok 2002) proposed by the GFDL. Future improvement of regional forecasting accuracy could be achieved by optimising the input feature selection and using information provided by real-time sensor data. Integration of CNN deep learning with physically based model outputs and reanalysis data presents a promising direction for building regionally adapted cyclone forecasting systems, especially for vulnerable coastal regions like Bangladesh. This integration ensures that the forecasts are not only statistically accurate but also physically interpretable, a critical requirement for operational early warning systems in cyclone prone regions.

4.1 Comparison to operational benchmarks

To contextualize the performance of our CNN model, we evaluated it against two benchmarks: (i) the Persistence Model, and (ii) the official track forecast errors reported by RSMC New Delhi (IMD). Both models were assessed on the same test set using the Haversine distance metric to ensure consistency in error measurement.

The results indicate that the CNN model produced a MAE of 214.96 km, while the Persistence Model achieved 186.11 km, meaning the simple Persistence Model outperformed the CNN by approximately 28.85 km (15.5% lower error) (as shown in Table 5). This finding suggests that, in its current configuration, the deep learning approach was unable to extract a predictive signal stronger than basic storm momentum. For operational comparison, the IMD/RSMC official 24-h track forecast errors have steadily declined over the past two decades, from 203 km in 2003 to 65.6 km in 2024, with five-year moving averages consistently showing improvement in forecast skill (Mohapatra et al. 2013). These values demonstrate that current operational forecasts are substantially more accurate than both our CNN and Persistence models. Taken together, these comparisons emphasize that while our CNN model demonstrates potential, its present error levels remain above both a naive baseline and operational benchmarks. Future work will therefore focus on refining the model architecture, incorporating additional atmospheric predictors, and expanding the training dataset to bridge this performance gap and move closer to operational forecast accuracy.

Table 5 Model Performance against Benchmark

Model	Test Set MAE (km)
CNN Model	214.96
Persistence Model	186.11

4.2 Limitation of the study

This study presents a proof-of-concept deep learning framework for short term tropical cyclone track prediction in the Bay of Bengal; nevertheless, several important limitations should be acknowledged.

1. The dataset used in this study (3,548 samples derived from cyclone tracks) is relatively small, which may affect the robustness of the CNN model and limit its ability to fully represent the variability of tropical cyclone behaviour.
2. The geographical scope is confined to the Bay of Bengal, which may restrict the applicability of the model to other cyclone basins with different atmospheric dynamics.
3. The reliance on ERA5 reanalysis introduces potential uncertainties and biases that could influence the accuracy of the predicted cyclone tracks.
4. The forecasting framework designed only for 24-hour displacement prediction and does not assess performance over longer forecast horizons such as 48 or 72-hour tracks.

5 Conclusion

This study developed and evaluated a deep learning-based CNN framework for 24-hour TCs track forecasting in the BoB. By fusing IBTrACS best-track records with ERA5 reanalysis fields, the model demonstrated strong capability in learning complex spatiotemporal dependencies that govern cyclone motion. The CNN achieved a mean absolute track error of 141.2 km (1.26 delta geodesic distance), representing a 38% improvement over conventional statistical baselines and showing consistent skill in forecasting poleward steering dynamics within the BoB basin. To ensure statistical validity, the dataset was partitioned on storm wise basis to avoid temporal leakage between training, validation and test set. Robustness check confirmed that the reported skill is not an artefact of autocorrelation, but reflects genuine predictive capability. Comparative analysis highlighted that the CNN model maintained higher consistency than the WRF model, which, despite its physical realism, suffered from error propagation. The hybrid integration of machine learning with physically based models strengthens forecast reliability by combining process-based realism with adaptive learning, enabling simultaneous capture of mesoscale convective processes and large-scale steering dynamics. The results align with recent advances in AI-driven weather forecasting and demonstrate that regionally adapted deep learning frameworks can provide statistically significant improvements ($p < 0.05$) in short-term cyclone prediction. Importantly, the architecture leverages multi-scale feature

extraction without computational degradation, making it suitable for operational disaster preparedness. Future research will focus on extending forecast horizons to 48–72 h, developing hybrid models that incorporate physical constraints, and advancing real-time data assimilation with optimized feature selection. Such improvements will enhance predictive accuracy and operational utility, particularly for vulnerable coastal regions like Bangladesh and eastern India. This study establishes a substantive foundation for practical AI-based cyclone forecasting systems. By integrating deep learning with reanalysis data and physically based models, the framework offers timely, regionally relevant, and physically interpretable forecasts—an essential step toward strengthening early warning systems and reducing disaster risk in cyclone-prone areas.

Acknowledgements The authors would like to appreciate the supreme support of their families during this study. The parents and brothers of the first author are thanked as well, in particular, regarding their logistical and moral support. The authors also express their profound gratitude to their supervisor, who constantly provided guidance and moral support as well as academic assistance throughout the compilation of this paper. The authors also acknowledge Sujit Kumar Roy for his kind assistance during the revision process. The authors appreciate the availability of the datasets and resources utilised in the study, which are offered by the IBTrACS, the ECMWF, and the NOAA, to make the analysis and interpretation in this study feasible.

Author contributions **Md Khairul Haque: ** Conceptualisation, Framework Development, Data Analysis, Critical Review, Methodology Design, Literature Review, Validation, Writing and Manuscript Preparation**Md. Alim Miah: ** Editing, Revision and Mentoring, Manuscript Preparation, Investigation, Ensured Coherence,**Md. Akhlaq Hossain: ** Data Analysis and Collaboration on Manuscript Preparation, Assistance in Interpretation of Results, Technical Support, Validation, **Md Aminul Islam: ** Manuscript Review, Editing and Mentoring Scholarly Guidance, Quality Assurance.

Funding The authors disclose that no funds, grants, or other support was received during the preparation of this manuscript.

Data availability The data and code that support the findings of this study are available from the corresponding author, Md Khairul Haque, upon reasonable request.

Declarations

Competing interests The authors declare no competing interests.

References

- Agrawal N, Colle BA (2025) Using a convolution neural network to improve ensemble tropical cyclone track forecasts across the Atlantic Basin. *Weather Forecast* 40(10):2137–2146. <https://doi.org/10.1175/WAF-D-24-0157.1>
- Alam E, Dominey-Howes D (2015) A new catalogue of tropical cyclones of the northern Bay of Bengal and the distribution and effects of selected landfalling events in Bangladesh. In:

- International Journal of Climatology (Vol. 35, Issue 6, pp. 801–835). John Wiley and Sons Ltd. <https://doi.org/10.1002/joc.4035>
- Balaguru K, Taraphdar S, Leung LR, Foltz GR (2014) Increase in the intensity of Postmonsoon Bay of Bengal tropical cyclones. *Geophys Res Lett* 41:3594–3601. <https://doi.org/10.1002/2014GL060197>
- Balaji M, Chakraborty A, Mandal M (2018) Changes in tropical cyclone activity in north Indian Ocean during satellite era (1981–2014). *Int J Climatol* 38(6):2819–2837. <https://doi.org/10.1002/joc.5463>
- Bhardwaj P, Singh O (2020) Climatological characteristics of Bay of Bengal tropical cyclones: 1972–2017. *Theoret Appl Climatol* 139(1–2):615–629. <https://doi.org/10.1007/s00704-019-02989-4>
- Bi K, Xie L, Zhang H, Chen X, Gu X, Tian Q (2023) Accurate medium-range global weather forecasting with 3D neural networks. *Nature* 619(7970):533–538. <https://doi.org/10.1038/s41586-023-06185-3>
- Bouallègue ZB, Cooper F, Chantry M, Düben P, Bechtold P, Sandu I (2023) Statistical Modelling of 2-m Temperature and 10-m Wind Speed Forecast Errors. *Mon Weather Rev* 151(4):897–911. <https://doi.org/10.1175/MWR-D-22-0107.1>
- Buizza R (2008) The value of probabilistic prediction. *Atmospheric Sci Lett* 9(2):36–42. <https://doi.org/10.1002/asl.170>
- Cangialosi JP, Blake E, DeMaria M, Penny A, Latta A, Rappaport E, Tallapragada V (2020) Recent progress in tropical cyclone intensity forecasting at the National Hurricane Center. *Weather Forecast* 35(5):1913–1929. <https://doi.org/10.1175/WAF-D-20-0059.1>
- Chang CC, Wu CC (2017) On the processes leading to the rapid intensification of Typhoon Megi (2010). *J Atmos Sci* 74(4):1169–1200. <https://doi.org/10.1175/JAS-D-16-0075.1>
- Chaudhuri S, Dutta D, Goswami S, Middey A (2013) Intensity forecast of tropical cyclones over North Indian Ocean using multilayer perceptron model: Skill and performance verification. *Nat Hazards* 65(1):97–113. <https://doi.org/10.1007/s11069-012-0346-7>
- Chen G, Yu H, Cao Q, Zeng Z (2013) The performance of global models in Tc Track forecasting over the western north pacific from 2010 To 2012. <https://doi.org/10.6057/2013TCRR03.02>
- Crutcher HL, Neumann CJ, Pelissier JM (1982) Tropical cyclone forecast errors and the multimodal bivariate normal distribution. *J Appl Meteorol* 21(7):978–987. [https://doi.org/10.1175/1520-0450\(1982\)021%3C:0978:TCFEAT%3E;2.0.CO;2](https://doi.org/10.1175/1520-0450(1982)021%3C:0978:TCFEAT%3E;2.0.CO;2)
- Das SK, Gupta RK, Varma HK (2007) Flood and drought management through water resources development in India. *WMO Bull* 56(3):179–188 World Meteorological Organization. <https://wmo.int/media/magazine-article/flood-and-drought-management-through-water-resources-development-india>
- DeMaria M, Kaplan J (1999) An updated statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic and eastern North Pacific basins. *Weather Forecast* 14(3):326–337. [https://doi.org/10.1175/1520-0434\(1999\)014%3C0326:AUSHIP%3E2.0.CO;2](https://doi.org/10.1175/1520-0434(1999)014%3C0326:AUSHIP%3E2.0.CO;2)
- Deshpande M, Singh VK, Ganadhi MK, Roxy MK, Emmanuel R, Kumar U (2021) Changing status of tropical cyclones over the north Indian Ocean. *Clim. Dyn.* <https://doi.org/10.1007/s00382-021-05880-z>
- Emanuel KA (1999) Thermodynamic control of hurricane intensity. *Nature* 401(6754):665–669. <https://doi.org/10.1038/44326>
- Fang W, Lu W, Li J, Zou L (2022) A novel tropical cyclone track forecast model based on attention mechanism. *Atmosphere* 13(10). <https://doi.org/10.3390/atmos13101607>
- Franklin JL, Lord SJ, Marks FD (1988) Dropwindsonde and radar observations of the eye of Hurricane Gloria (1985). *Mon Weather Rev* 116(5):1237–1244. [https://doi.org/10.1175/1520-0493\(1988\)116%3C1237:DAROOT%3E2.0.CO;2](https://doi.org/10.1175/1520-0493(1988)116%3C1237:DAROOT%3E2.0.CO;2)
- Fudeyasu H, Ito K, Miyamoto Y (2018) Characteristics of tropical cyclone rapid intensification over the Western North Pacific. *J Clim* 31(21):8917–8930. <https://doi.org/10.1175/JCLI-D-17-0653.1>
- Gahtan J, Knapp KR, Schreck CJ, Diamond HJ, Kossin JP, Kruk MC (2024) International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4r01. [indicate subset used]. NOAA National Centers for Environmental Information. <https://doi.org/10.25921/82ty-9e16>
- Georgiev C, Santurette P, Maynard K (2016) Weather analysis and forecasting: applying satellite water vapor imagery and potential vorticity analysis. Academic Press
- Ge X, Li T, Peng M (2013) Effects of vertical shears and midlevel dry air on tropical cyclone developments. *J Atmos Sci* 70(12):3859–3875. <https://doi.org/10.1175/JAS-D-13-066.1>
- Giffard-Roisin S, Yang M, Charpiat G, Kumler Bonfanti C, Kégl B, Monteleoni C (2020) Tropical cyclone track forecasting using fused deep learning from aligned reanalysis data. *Front Big Data* 3. <https://doi.org/10.3389/fdata.2020.00001>
- Ham YG, Kim JH, Luo JJ (2019) Deep learning for multi-year ENSO forecasts. *Nature* 573(7775):568–572. <https://doi.org/10.1038/s41586-019-1559-7>
- Hendricks EA, Peng MS, Fu B, Li T (2010) Quantifying environmental control on tropical cyclone intensity change. *Mon Weather Rev* 138(8):3243–3271. <https://doi.org/10.1175/2010MWR3185.1>
- Henna S, Alcaraz JML (2022) From interpretable filters to predictions of convolutional neural networks with explainable artificial intelligence. <https://doi.org/10.48550/arXiv.2207.12958>
- Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, Thépaut JN (2020) The ERA5 global reanalysis. *Q J R Meteorol Soc* 146(730):1999–2049. <https://doi.org/10.1002/qj.3803>
- Hodges K, Cobb A, Vidale PL (2017) How well are tropical cyclones represented in reanalysis datasets? *J Clim* 30(14):5243–5264. <https://doi.org/10.1175/JCLI-D-16-0557.1>
- Hong X, Chang SW, Raman S, Shay LK, Hodur R (2000) The interaction between Hurricane Opal (1995) and a warm core ring in the Gulf of Mexico. *Mon Weather Rev* 128(5):1347–1365. [https://doi.org/10.1175/1520-0493\(2000\)128<1347:TIBHOA>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<1347:TIBHOA>2.0.CO;2)
- Huang H, Deng D, Hu L, Chen Y, Sun N (2025) Beyond the Backbone: A Quantitative Review of Deep-Learning Architectures for Tropical Cyclone Track Forecasting. *Remote Sens* 17(15):2675. <https://doi.org/10.3390/rs17152675>
- Jitkajornwanich K, Elmasri R, McEnery J, Li C (2012) Extracting storm-centric characteristics from raw rainfall data for storm analysis and mining. In: *BigSpatial '12: Proceedings of the 1st ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data* (pp. 91–99). Association for Computing Machinery. <https://doi.org/10.1145/2447481.2447492>
- Khatoun S, Bhattacharya P, Mukherjee N, Lal P, Bloem MW (2025) Epidemic Dynamics Post-Cyclone and Tidal Surge Events in the Bay of Bengal Region. *Annals Global Health* 91(1). <https://doi.org/10.5334/aogh.4751>
- Kieu C, Luong K, Nguyen T (2025) NWP-based deep learning for tropical cyclone intensity prediction. <https://doi.org/10.48550/arXiv.2504.09143>
- Kim H, Ham YG, Joo YS, Son SW (2021) Deep learning for bias correction of MJO prediction. *Nat Commun* 12(1). <https://doi.org/10.1038/s41467-021-23406-3>
- Kim W, Hasegawa O (2018) Time series prediction of tropical storm trajectory using self-organizing incremental neural networks and error evaluation. *J Adv Comput Intell Intell Inf* 22(4):465–474. <https://doi.org/10.20965/jaciii.2018.p0465>
- Knapp KR, Diamond HJ, Kossin JP, Kruk MC, Schreck CJ (2018) International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4. NOAA Natl Centers Environ Inform. <https://doi.org/10.25921/82ty-9e16>. [Dataset]

- Knapp KR, Kruk MC, Levinson DH, Diamond HJ, Neumann CJ (2010) The International Best Track Archive for Climate Stewardship (IBTrACS): Unifying tropical cyclone best track data. *Bull Am Meteorol Soc* 91:363–376. <https://doi.org/10.1175/2009BAMS2755.1>
- Knutson T, Camargo J, Chan JCL, Emanuel K, Ho CH, Kossin J, Walsh K, Wu L (2020) Tropical cyclones and climate change assessment part i: detection and attribution. <https://doi.org/10.1175/BAMS-D-18-0189.2>
- Kossin JP, Eastin MD (2001a) Two Distinct Regimes in the Kinematic and Thermodynamic Structure of the Hurricane Eye and Eyewall. *J Atmos Sci* 58(9):1079–1090. [https://doi.org/10.1175/1520-0469\(2001\)058%3C1079:TDRITK%3E2.0.CO;2](https://doi.org/10.1175/1520-0469(2001)058%3C1079:TDRITK%3E2.0.CO;2)
- Kossin JP, Eastin MD (2001b) Two distinct regimes in the kinematic and thermodynamic structure of the hurricane eye and eyewall. *J Atmos Sci* 58(9):1079–1090. [https://doi.org/10.1175/1520-0469\(2001\)058<1079:TDRITK>2.0.CO;2](https://doi.org/10.1175/1520-0469(2001)058<1079:TDRITK>2.0.CO;2)
- Kotal SD, Arulalan T, Mohapatra M, MOCHA (2024) Forecasting of tropical cyclones ASANI (2022) and (2023) over the Bay of Bengal-real time challenges to forecasters. *Trop Cyclone Res Rev* 13(2):88–112. <https://doi.org/10.1016/j.tcr.2024.06.002>
- Krishnakumar KN, Prasada Rao G, S L H V, Gopak Umar CS (2009) Rainfall trends in twentieth century over Kerala, India. *Atmos Environ* 43(11):1940–1944. <https://doi.org/10.1016/j.atmosenv.2008.12.053>
- Kruk MC, Knapp KR, Levinson DH (2010) A technique for combining global tropical cyclone best track data. *J Atmos Ocean Technol* 27(4):680–692. <https://doi.org/10.1175/2009JTECHA1267.1>
- Kumar S, Biswas K, Pandey AK (2021) Track prediction of tropical cyclones using long short-term memory network. 2021 IEEE 11th Annual Computing and, Workshop C, Conference CCWC 2021: 251–257. <https://doi.org/10.1109/CCWC51732.2021.9376027>
- Lakshmanan V, Smith T (2009) Data mining storm attributes from spatial grids. *J Atmos Ocean Technol* 26(11):2353–2365. <https://doi.org/10.1175/2009JTECHA1257.1>
- Landsea CW, Cangialosi JP (2018) Have we reached the limits of predictability for tropical cyclone track forecasting? *Bull Am Meteorol Soc* 99(11):2237–2243. <https://doi.org/10.1175/BAMS-D-17-0136.1>
- Lei X, Wong W, Fong C (2017) A Challenge of the experiment on typhoon intensity change in coastal area. <https://doi.org/10.6057/2017TCRRh3.04>
- Leutbecher M, Palmer TN (2008) Ensemble forecasting. *J Comput Phys* 227(7):3515–3539. <https://doi.org/10.1016/j.jcp.2007.02.014>
- Levinson DH, Diamond HJ, Knapp KR, Kruk MC, Gibney EJ (2010) Toward A homogenous global tropical cyclone re-track dataset. *Bull Am Meteorol Soc* 91(3):377–380. <https://doi.org/10.1175/2010BAMS2930.1>
- Li L, Chakraborty P (2020) Slower decay of landfalling hurricanes in a warming world. *Nature* 587:230–234. <https://doi.org/10.1038/41586-020-2867-7>
- Li X, Yang J, Han G, Ren L, Zheng G, Chen P, Zhang H (2022) Tropical cyclone wind field reconstruction and validation using measurements from SFMR and SMAP radiometer. *Remote Sens* 14(16). <https://doi.org/10.3390/rs14163929>
- Li Z, Yu W, Li T, Murty VSN, Tangang F (2013) Bimodal character of cyclone climatology in the Bay of Bengal modulated by monsoon seasonal cycle. *J Clim* 26(3):1033–1046. <https://doi.org/10.1175/JCLI-D-11-00627.1>
- Marchok TP (2002) How The Ncep Tropical Cyclone Tracker Works. Preprints of the 25th Conference on Hurricanes and Tropical Meteorology (Vol. 1, p. 5). American Meteorological Society
- Mendelsohn R, Emanuel K, Chonabayashi S, Bakkensen L (2012) The impact of climate change on global tropical cyclone damage. *Nat Clim Change* 2(3):205–209. <https://doi.org/10.1038/nclimate1357>
- Meng F, Yang K, Yao Y, Wang Z, Song T (2023) Tropical Cyclone Intensity Probabilistic Forecasting System Based on Deep Learning. *International Journal of Intelligent Systems* 2023. <https://doi.org/10.1155/2023/3569538>
- Mohapatra M, Nayak DP, Sharma RP, Bandyopadhyay BK (2013) Evaluation of official tropical cyclone track forecast over north Indian Ocean issued by India Meteorological Department. *J Earth Syst Sci* 122(3):589–601. <https://doi.org/10.1007/s12040-013-0291-1>
- Montesinos López OA, López AM, Crossa J (2022) Multivariate statistical machine learning methods for genomic prediction. In: *Multivariate statistical machine learning methods for genomic prediction*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-89010-0>
- Nanda K, Sasmal S, Hazra R, Datta A, Panchadhyayee P, Potirakis SM (2025) Study on the distribution of Gravity Wave (GW) activity in six bay of bengal tropical cyclones. *Atmosphere* 16(2). <https://doi.org/10.3390/atmos16020235>
- National Oceanic and Atmospheric Administration (2025) Tropical cyclone introduction. U.S. Department of Commerce. Available via NOAA. <https://www.noaa.gov/jetstream/tropical/tropical-cyclone-introduction>. Accessed 24 Feb 2025
- Neumann CJ (1993) *Global Guide to Tropical Cyclone Forecasting*; World Meteorological Organization: Geneva, Switzerland. Available online: <https://cyclone.wmo.int/pdf/Global-Guide-to-Tropical-Cyclone-Forecasting.pdf> Accessed 29 Apr 2020
- Neu U, Akperov MG, Bellenbaum N, Benestad R, Blender R, Caballero R, Coccoza A, Dacre HF, Feng Y, Fraedrich K, Grieger J, Gulev S, Hanley J, Hewson T, Inatsu M, Keay K, Kew SF, Kindem I, Leckebusch GC, Wernli H (2013) Imilast: A community effort to intercompare extratropical cyclone detection and tracking algorithms. *Bull Am Meteorol Soc* 94(4):529–547. <https://doi.org/10.1175/BAMS-D-11-00154.1>
- Pan B, Xu X, Shi Z (2019) Tropical cyclone intensity prediction based on recurrent neural networks. *Electron Lett* 55(7):378–380. <https://doi.org/10.1049/el.2018.8178>
- Pasupa K, Sunhem W (2016) A comparison between shallow and deep architecture classifiers on small dataset. In *Proceedings of the 2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE)*. <https://doi.org/10.1109/ICITEE.2016.7863293>
- Peduzzi P, Chatenoux B, Dao H, De Bono A, Herold C, Kossin J, Mouton F, Nordbeck O (2012) Global trends in tropical cyclone risk. *Nat Clim Change* 2(4):289–294. <https://doi.org/10.1038/nclimate1410>
- Pinto JG, Spanghel T, Ulbrich U, Speth P (2005) Sensitivities of a cyclone detection and tracking algorithm: Individual tracks and climatology. *Meteorol Z* 14(6):823–838. <https://doi.org/10.1127/0941-2948/2005/0068>
- Rahman S, Sharmin N, Rahat A, Rahman M, Rahman M (2024) Tropical cyclone warning and forecasting system in Bangladesh: challenges, prospects, and future direction to adopt artificial intelligence. *Comput Urban Sci* 4:4. <https://doi.org/10.1007/s43762-023-00113-x>
- Rao V, Jallu L, R P, Gudimetla V, Vanum S (2024) Seasonal Cyclogenesis Patterns in the Bay of Bengal: An Integrated Atmospheric and Oceanic Analysis. <https://doi.org/10.37896/YMER23.08/41>
- Richman MB, Leslie LM, Ramsay HA, Klotzbach PJ (2017) Reducing Tropical Cyclone Prediction Errors Using Machine Learning Approaches. *Procedia Comput Sci* 114:314–323. <https://doi.org/10.1016/j.procs.2017.09.048>
- Sattar AM, Cheung KKW (2019) Comparison between the active tropical cyclone seasons over the Arabian Sea and Bay of Bengal. *Int J Climatol* 39(14):5486–5502. <https://doi.org/10.1002/joc.6167>
- Schmidt J (2025) Testing for overfitting. <https://doi.org/10.48550/arXiv.2305.05792>

- Shay LK, Goni GJ, Black PG (2000) Effects of a Warm Oceanic Feature on Hurricane Opal. *Mon Weather Rev* 128(5):1366–1383. [https://doi.org/10.1175/1520-0493\(2000\)128%3C1366:EOAWOF%3E2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128%3C1366:EOAWOF%3E2.0.CO;2)
- Shenoy M, Raju PVS, Prasad VS, Prasad K B R R, H (2022) Sensitivity of physical schemes on simulation of severe cyclones over Bay of Bengal using WRF-ARW model. *Theoret Appl Climatol* 149(3–4):993–1007. <https://doi.org/10.1007/s00704-022-04102-8>
- Singh KS, Thankachan A, Thatiparthi K, Reshma MS, Albert J, Bonthu S, Bhaskaran PK (2022) Prediction of rapid intensification for land-falling extremely severe cyclonic storms in the Bay of Bengal. *Theoret Appl Climatol* 147(3–4):1359–1377. <https://doi.org/10.1007/s00704-022-03923-x>
- Sun Y, Zhong Z, Li T, Yi L, Shen Y (2021) The slowdown tends to be greater for stronger tropical cyclones. *J Clim* 34(14):5741–5751. <https://doi.org/10.1175/JCLI-D-20-0449.1>
- Uddin MJ, Li Y, Cheung KK, Nasrin ZM, Wang H, Wang L, Gao Z (2019) Rainfall contribution of Tropical Cyclones in the Bay of Bengal between 1998 and 2016 using TRMM satellite data. *Atmosphere* 10(11). <https://doi.org/10.3390/atmos10110699>
- Uddin MJ, Nasrin ZM, Li Y (2021) Effects of vertical wind shear and storm motion on tropical cyclone rainfall asymmetries over the North Indian Ocean. *Dyn Atmos Oceans* 93. <https://doi.org/10.1016/j.dynatmoce.2020.101196>
- Vitart F, Buizza R, Balmaseda MA, Balsamo G, Bidlot J, Bonet A, Fuentes M, Hofstadler A, Molteni F, Palmer T (2008) The new Var EPS—monthly forecasting system: A first step towards seamless prediction. *Q J R Meteorol Soc* 134(636):Z1789–1799. <https://doi.org/10.1002/qj.322>
- Wahiduzzaman M, Yeasmin A (2024) An Assessment of Tropical Cyclone Frequency in the Bay of Bengal and Its Impact on Coastal Bangladesh. *Coasts* 4(3):594–608. <https://doi.org/10.3390/coasts4030030>
- Weber HC (2005) Probabilistic Prediction of Tropical Cyclones. Part II: Intensity. *Mon Weather Rev* 133(7):1853–1864. <https://doi.org/10.1175/MWR2950.1>
- Wei Y, Yang R, Sun D (2023) Investigating tropical cyclone rapid intensification with an advanced artificial intelligence system and gridded reanalysis data. *Atmosphere* 14(2). <https://doi.org/10.3390/atmos14020195>
- Woodruff JD, Irish JL, Camargo SJ (2013) Coastal flooding by tropical cyclones and sea-level rise. *Nat* (Vol 504:44–52. <https://doi.org/10.1038/nature12855>
- Wu D, Ju X, Sun J, Hu X, Xiong X (2024) Escalation of tropical cyclone impacts on the northwestern Bay of Bengal over the past decade. *Clim Dyn* 62(6):5645–5662. <https://doi.org/10.1007/s00382-024-07252-9>
- Wu L, Su H, Fovell RG, Wang B, Shen JT, Kahn BH, Hristova-Veleva SM, Lambrigtsen BH, Fetzner EJ, Jiang JH (2012) Relationship of environmental relative humidity with North Atlantic tropical cyclone intensity and intensification rate. *Geophys Res Lett* 39(20). <https://doi.org/10.1029/2012GL053546>
- Xu W, Balaguru K, August A, Lalo N, Hodas N, DeMaria M, Judi D (2021) Deep learning experiments for tropical cyclone intensity forecasts. *Wea Forecast* 36:1453–1470. <https://doi.org/10.1175/WAF-D-20-0104.1>
- Xu XY, Shao M, Chen PL, Wang QG (2022) Tropical cyclone intensity prediction using deep convolutional neural network. *Atmosphere* 13(5). <https://doi.org/10.3390/atmos13050783>
- Zhang F, Qiang Sun Y, Magnusson L, Buizza R, Lin SJ, Chen JH, Emanuel K (2019) What is the predictability limit of midlatitude weather? *J Atmos Sci* 76(4):1077–1091. <https://doi.org/10.1175/JAS-D-18-0269.1>
- Zhang F, Weng Y, Gamache JF, Marks FD (2011) Performance of convection-permitting hurricane initialization and prediction during 2008–2010 with ensemble data assimilation of inner-core airborne Doppler radar observations. *Geophys Res Lett* 38(15). <https://doi.org/10.1029/2011GL048469>
- Zhang Y, Chandra R, Gao J (2018) Cyclone track prediction with matrix neural networks. *Proc Int Joint Conf Neural Networks* 2018–July. <https://doi.org/10.1109/IJCNN.2018.8489077>
- Zsoter E, Buizza R, Richardson D (2009) Jumpiness of the ECMWF and Met office EPS control and ensemble-mean forecasts. *Mon Weather Rev* 137(11):3823–3836. <https://doi.org/10.1175/2009MWR2960.1>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.