

Accepted Manuscript
Accepted Manuscript (Uncorrected Proof)

Title: An AI-Based Early Warning Model for Rapid Intensification of Tropical Cyclones in
Disaster Management Applications

Authors: Kalach Andrew Vladimirovich¹, Shkerin Alexey Leonodivch^{1,*}

1. *Russian Technological University, Moscow, Russia.*

To appear in: **Disaster Prevention and Management Knowledge**

Received date: 2025/12/24

Accepted date: 2026/03/13

First Online Published: 2026/10/04

This is a “Just Accepted” manuscript, which has been examined by the peer-review process and has been accepted for publication. A “Just Accepted” manuscript is published online shortly after its acceptance, which is prior to technical editing and formatting and author proofing. Archives of Rehabilitation provides “Just Accepted” as an optional service which allows authors to make their results available to the research community as soon as possible after acceptance. After a manuscript has been technically edited and formatted, it will be removed from the “Just Accepted” Website and published as a published article. Please note that technical editing may introduce minor changes to the manuscript text and/or graphics which may affect the content, and all legal disclaimers that apply to the journal pertain.

Please cite this article as:

Andrew Vladimirovich, K., & Alexey Leonodivch, S. [An AI-Based Early Warning Model for Rapid Intensification of Tropical Cyclones in Disaster Management Applications]. Disaster Prevention and Management Knowledge. Forthcoming 2026.

نسخه پذیرفته شده پیش از انتشار

نسخه پذیرفته شده پیش از انتشار

عنوان: یک مدل هشدار زودهنگام مبتنی بر هوش مصنوعی برای تشدید سریع طوفان‌های گرمسیری در کاربردهای

مدیریت بحران

نویسندگان: کالاج اندرو ولادیمیرویچ^۱، شکرین الکسی لئونودیوچ^{۱*}

۱. دانشگاه فناوری روسیه، مسکو، روسیه.

نشریه: دانش پیشگیری و مدیریت بحران

تاریخ دریافت: 1404/10/3

تاریخ پذیرش: 1404/12/22

تاریخ انتشار اولیه: 1405/07/12

این نسخه «پذیرفته‌شده پیش از انتشار» مقاله است که پس از طی فرایند داوری، برای چاپ، قابل پذیرش تشخیص داده شده است. این نسخه در مدت کوتاهی پس از اعلام پذیرش به صورت آنلاین و قبل از فرایند ویراستاری منتشر می‌شود. نشریه آرشیو توانبخشی گزینه «پذیرفته‌شده پیش از انتشار» را به عنوان خدمتی به نویسندگان ارائه می‌دهد تا نتایج آن‌ها در سریع‌ترین زمان ممکن پس از پذیرش برای جامعه علمی در دسترس باشد. پس از آنکه مقاله‌ای فرایند آماده سازی و انتشار نهایی را طی می‌کند، از نسخه «پذیرفته‌شده پیش از انتشار» خارج و در یک شماره مشخص در وبسایت نشریه منتشر می‌شود. شایان ذکر است صفحه آرای و ویراستاری فنی باعث ایجاد تغییرات صوری در متن مقاله می‌شود که ممکن است بر محتوای آن تأثیر بگذارد و این امر از حیطه مسئولیت دفتر نشریه خارج است.

لطفا این گونه استناد شود:

Andrew Vladimirovich, K., & Alexey Leonodivch, S. [An AI-Based Early Warning Model for Rapid Intensification of Tropical Cyclones in Disaster Management Applications]. Disaster Prevention and Management Knowledge. Forthcoming 2026.

Abstract

Rapid intensification (RI) of tropical cyclones remains a key source of forecast uncertainty and a major driver of rapid hazard escalation in early warning systems. We propose a lightweight temporal mixing neural model for probabilistic RI prediction from a compact set of physically interpretable, storm center predictors: Sea Surface Temperature (*SST*), Mean Sea Level Pressure (*MSLP*) and vertical wind shear, augmented with 6-24 h lagged values. We also consider storm-wise splits by cyclone identifier and report discrimination with ROC-AUC and average precision, probabilistic accuracy via Brier score and decision-oriented verification with POD and FAR metrics. We apply post-hoc Platt scaling on the validation split and assess calibration with reliability diagrams. On the held-out test split the model achieves strong discrimination with ROC-AUC approximately 0.79 and AP metric approximately 0.20, Brier score approximately 0.054. Based on SHAP analysis, *MSLP*- and *SST*-related features identified as most the impactful features for the forecast. All experiments were carried out on a single graphics processing unit (GPU) NVIDIA RTX 5080 Ti. Therefore, the proposed approach is suitable for environments with limited computational resources.

Keywords: tropical cyclone, rapid intensification, early warning, disaster risk reduction, intensity forecast, TimeMixerDrop, ERA5, IBTrACS.

چکیده

تشدید سریع (RI) طوفان‌های گرمسیری همچنان منبع کلیدی عدم قطعیت پیش‌بینی و عامل اصلی تشدید سریع خطر در سیستم‌های هشدار اولیه است. ما یک مدل عصبی ترکیبی زمانی سبک برای پیش‌بینی احتمالی RI از مجموعه‌ای فشرده از پیش‌بینی‌کننده‌های مرکز طوفان که از نظر فیزیکی قابل تفسیر هستند، پیشنهاد می‌کنیم: دمای سطح دریا (SST)، فشار متوسط سطح دریا (MSLP) و برش باد عمودی، که با مقادیر تأخیری ۶-۲۴ ساعته تقویت شده‌اند. ما همچنین تقسیم‌بندی‌های طوفان را با استفاده از شناسه طوفان در نظر می‌گیریم و تفکیک را با ROC-AUC و دقت متوسط، دقت احتمالی از طریق امتیاز Brier و تأیید تصمیم‌گیری محور با معیارهای POD و FAR گزارش می‌دهیم. ما مقیاس‌بندی Post-hoc Platt را روی تقسیم اعتبارسنجی اعمال می‌کنیم و کالیبراسیون را با نمودارهای قابلیت اطمینان ارزیابی می‌کنیم. در تقسیم‌بندی آزمون hold-out، مدل با ROC-AUC تقریباً ۰.۷۹ و معیار AP تقریباً ۰.۲۰ و امتیاز Brier تقریباً ۰.۵۴ به تفکیک قوی دست می‌یابد. بر اساس تجزیه و تحلیل SHAP، ویژگی‌های مرتبط با MSLP و SST به عنوان تأثیرگذارترین ویژگی‌های پیش‌بینی شناسایی شدند. تمام آزمایش‌ها بر روی یک واحد پردازش گرافیکی (GPU) NVIDIA RTX 5080 Ti انجام شد. بنابراین، رویکرد پیشنهادی برای محیط‌هایی با منابع محاسباتی محدود مناسب است.

کلیدواژه: سیکلون حاره‌ای، شدت‌گیری سریع، پیش‌بینی شدت، IBTrACS, ERA5, TinyTimeMixer

INTRODUCTION

Tropical cyclones are among the most destructive atmospheric phenomena on the planet. Particularly dangerous are cases of rapid intensification, defined as an increase in the maximum wind speed in a cyclone by at least 30 knots in 24 hours (Kim et al., 2024). Such sudden bursts of intensity significantly complicate short-term forecasting and measures to mitigate the consequences. According to recent studies, there has been an increase in the frequency of intense tropical cyclone intensification, especially in coastal regions: between 1980 and 2020, the number of such cases tripled (Li et al., 2023). Potential sources of this trend are general ocean warming and changes in atmospheric circulation. In view of these facts, the task of operational forecasting of the rapid intensification of such a rapidly developing hydrometeorological hazard as a tropical cyclone is a pressing scientific challenge.

From a disaster risk reduction (DRR) perspective, rapid intensification (RI) is a hydrometeorological hazard escalation process that can sharply reduce effective warning lead time. Probabilistic RI forecasts can support early warning action workflows by enabling earlier evacuation planning, pre-positioning of response assets, and risk communication under uncertainty. In the context of the Sendai Framework priorities, specifically risk understanding and preparedness, such forecasts can be used as an input to impact-based warning services that translate hazard likelihood into expected consequences for exposed and vulnerable coastal communities.

In disaster management, RI represents a hazard escalation process that can rapidly increase the likelihood and severity of impacts such as storm surge, destructive winds and extreme precipitation in exposed coastal zones. Probabilistic RI forecasts can therefore serve as an operational trigger within end-to-end early warning systems, supporting “what the weather will do” in impact-based warning practice. Such services are aligned with the Sendai Framework priorities on understanding disaster risk and enhancing preparedness for effective response (UNDRR, 2015), where timely and actionable information is essential for reducing losses.

For the Russian Federation, direct impacts from tropical cyclones are relatively rare because such systems typically form in lower latitudes. Nevertheless, remnants of tropical cyclones and associated storms can affect the Russian Far East through heavy precipitation and gusty winds. Improving forecasting approaches remains important in the broader context of global atmospheric monitoring, risk assessment and management of hydrometeorological hazards (Kalach & Paramonov, 2025; Kalach et al., 2025).

Despite significant progress in understanding and modeling tropical cyclones, predicting intense intensification remains a scientific and technical challenge. Traditional static models, such as the Rapid Intensification Index developed since the early 2000s, still do not show reliable results despite improvements (Griffin et al., 2022). High-resolution dynamic numerical models are capable of simulating such cyclone intensification, but they require significant computational resources and do not always take into account the multiscale processes that influence the formation of such intensification.

Modern machine learning methods offer an alternative approach, learning directly from historical data to identify complex nonlinear dependencies that precede the process of rapid intensification. For example, in the works of foreign colleagues (Griffin et al., 2022), a model based on a convolutional neural network (CNN) was developed to estimate the probability of tropical cyclone intensification based on satellite images. The results of their work showed better accuracy than traditional statistical methods. Another study (Rogers et al., 2022) revealed a link between the configuration of air flows in a hurricane and its ability to intensify, emphasizing the importance of considering the vertical wind shear and the angle of air inflow into the inner regions of hurricanes. In addition to specialized models for cyclones and storms, the general application of artificial intelligence (AI) models in meteorology is also developing. In addition to specialized cyclone-focused models, the general applications of AI in meteorology are rapidly developing. For example, DeepMind introduced the GraphCast global forecasting model based on graph neural networks, reporting skillful medium-range in up to 10-day predictions on a broad verification set (Lam et al., 2023). In parallel, benchmark initiatives such as WeatherBench2 provide a standardized evaluation protocols that stimulate the development of compact and fast next-generation models (Rasp et al., 2023).

Previously, the authors of this work patented a number of programs, reports on which were then tested at conferences. Within the framework of this series of programs, digital twins were studied (Paramonov & Kalach, 2025; Paramonov & Kalach, 2024) digital twins for analyzing and modeling tropical cyclones based on ERA5 reanalysis, the IBTrACS dataset, the use of numerical methods for solving Navier-Stokes equations, and two-dimensional convolutional neural networks. Despite high physical adequacy, the research results showed high computational resource and time costs. To overcome these limitations and ensure operational forecasting, a model based on the TimeMixerDrop architecture was proposed.

This research paper aims to develop a method for predicting the intense intensification of tropical cyclones, combining high physical validity and computational efficiency, allowing the implemented model to be used in conditions of limited resources. (Kim et al., 2024) addressed

a similar problem, and the results of their work will be taken as the baseline model against which the results of the proposed solution will be compared.

1 MATERIALS AND METHODS

1.1 Physically based model

In developing the current model, special attention was paid to physical factors that determine the potential of a tropical cyclone to rapidly intensify. It is known that for a cyclone to develop intensively, it needs a warm ocean as a source of energy and moisture, sufficiently low atmospheric pressure at its center, and a weak vertical wind shear in the troposphere. The latter factor is considered critical: a strong shear disrupts convection and prevents cyclone consolidation. Thus, the minimum set of forecast parameters for intense intensification includes:

- sea surface temperature (SST, °C);
- minimum sea level pressure (MSLP, hPa) at the center of the cyclone;
- vertical wind shear (SHEAR, m/s) between pressure levels of 200 and 850 hPa.

We intentionally restrict predictors to three widely available variables to keep the pipeline minimal and reproducible. However, additional relevant to RI factors such as mid-level humidity, ocean heat content and others are expected to improve the achieved results.

The choice of data characteristics is based on physical considerations and is confirmed by numerous studies of hurricane and cyclone dynamics, for example in equations 1-2:

$$Q_{in} = \rho_w c_w h (SST - T_{ref}) \quad (1)$$

where Q_{in} – is the heat energy reserve in the ocean, J/m²,

ρ_w – is the density of seawater, kg/m³,

c_w – is the specific heat capacity of water, J/(kg·K),

h – is the depth of the mixed layer of the ocean, m,

SST – is the sea surface temperature,

°C, T_{ref} – is the threshold temperature for cyclone attenuation, °C.

$$SHEAR = \sqrt{(u_{200} - u_{850})^2 + (v_{200} - v_{850})^2} \quad (2)$$

where u_{200}, u_{850} – are zonal wind components from west to east at pressure levels of 200 and 850 hPa,

v_{200}, v_{850} – are meridional wind components from north to south at the corresponding pressure levels.

Focusing on these variables allows us to reduce the dimensionality of the problem, improve interpretability, and reduce the computational resources required by the model.

The application of the principle of physical plausibility consists not only in the selection of informative input parameters, but also in the adaptation of the neural network structure to the nature of the target task. In particular, the rapid intensification of a cyclone is a process that occurs over time, so it is advisable to consider a daily forecast based on an analysis not only of the values of these parameters themselves, but also of their dynamics over the previous hours. These conditions are suitable for the use of recurrent or convolutional-recurrent architectures that could extract temporal dependencies.

This work uses prior knowledge about the processes occurring in atmospheric phenomena: the TimeMixerDrop architecture was originally designed for efficient processing of low-dimensional time series with the ability to take external factors into account. Thus, a compact pre-trained model is used, which is then retrained and tuned for a specific task, avoiding overfitting. This strategy is consistent with the concept of combining knowledge and data: a model pre-trained on a broad class of tasks carries generalized knowledge about the entire family of processes, and then physically meaningful choices of input data deepen the model's understanding of a specific topic. As a result, the proposed model is capable of making meaningful predictions even on a relatively small training set, while maintaining high physical interpretability.

1.2 ERA5 and IBTrACS data processing

An extensive set of meteorological data and observations of tropical cyclones for the period from 2000 to 2024 was used to train and test the model. The ERA5 global reanalysis was selected as the source of atmospheric field data, providing historical data with a high resolution of approximately 30 km and hourly steps (Harrigan et al., 2020). ERA5 is the fifth generation of ECMWF reanalysis and is characterized by the use of an improved analysis scheme and a huge amount of satellite observations, which provides a reliable representation of atmospheric conditions. Sea surface temperature, surface pressure, and wind components at 200 and 850 hPa levels with a 6-hour interval were taken from this reanalysis. The choice of a 6-hour interval corresponds to the standard time step in tropical cyclone data. The data were extracted near the position of each cyclone: for each tropical cyclone trajectory record, the reanalysis point closest to the specified coordinates of the cyclone center was selected, i.e., linear interpolation was performed by longitude and latitude. In a similar way, a time sample of the values of the listed parameters was formed for each time step of each cyclone under study.

Information about tropical cyclones themselves, their location, intensity, and wind speed were taken from the IBTrACS dataset, an international archive of data on tropical cyclones (Knapp et al., 2010). IBTrACS aggregates data from meteorological services around the world and provides a unified global catalog of cyclone trajectories and characteristics. All cyclones that began no earlier than 2000 in all ocean basins were selected. For each 6-hour time mark when the cyclone existed, the following parameters were recorded:

- current maximum wind speed;
- minimum pressure;
- geographical coordinates.

The target variable for the model was a binary indicator of rapid intensification over the next 24 hours, which was formed as follows (3):

$$RI = \begin{cases} 1 & v_{t+24} - v_t \geq 30 \\ 0 & v_{t+24} - v_t < 30 \end{cases} \quad (3)$$

Thus, each sample is the state of a cyclone at time t , described by three features: SST, MSLP, and shear. The target label is a binary RI label.

We adopt the widely used binary RI definition of ≥ 30 *kt* intensity increase within 24 hours to ensure comparability with prior literature mentioning RI. At the same time, RI is a continuum of intensification behavior. Future work can generalize this formulation to multi-threshold or to a regression model prediction the full intensification distribution.

RI is less frequent than non-RI at 6-hourly resolution, resulting in class imbalance. To mitigate imbalance during training, we use imbalance aware optimization with class weighted loss function. Threshold free metrics are reported with PR-AUC alongside with ROC-AUC because PR-AUC is more informative under extreme class imbalance and threshold-based results will be discussed separately.

The data was divided into training and test samples in such a way as to exclude information leakage between cyclones. Data on different cyclones are considered independent, but within a single cyclone, consecutive records are strongly correlated, so the sample was divided not by time stamps but by storm groups. This principle provides a realistic assessment of the model's generalizability to new storm seasons. The training sample contains more than 1,500 different storms and approximately 50,000 time stamps. All quantitative features are normalized and scaled for numerical stability of the model.

1.3 TimeMixerDrop architecture

In this study, we configure the TimeMixerDrop architecture as a compact neural network model for forecasting time series in meteorology. It is based on an architecture developed in

2024 at IBM Research. The key idea behind the model is to replace computationally expensive recurrent or transformer-based layers with lightweight ones that sequentially mix information in the temporal and spatial dimensions. This solution has made it possible to create extremely compact models, starting from 1 million parameters, but demonstrating high accuracy on a variety of time series with extremely low computational costs (Ekambaram et al., 2024). The basic architecture includes multiple levels of Mixer layers: each layer first applies a small fully connected neural network to all time steps of a feature, then another fully connected network is applied to all features at a single time step. These steps alternate and allow the model to capture both temporal dependencies and cross-channel relationships in a multivariate time series.

To stabilize training, residual blocks are applied after each block and simple ReLU activation functions are used. Formally, the model output can be represented as:

$$\widehat{y_{t+24}} = \sigma(f_{\theta}(X_{t-L+1} \dots X_t)) \quad (4)$$

where X_{t-i} – is the vector of input features at time $t - i$,
 f_{θ} – is the mapping defined by TimeMixerDrop layers,
 σ – is the sigmoid function that converts the model output into probability,
 L – is the time window value.

In this study, the parameter $L = 4$, thus considering 4 preceding time stamps with a step of 6 hours, in other words, considering the preceding 24 hours.

The TimeMixerDrop-RI implementation uses an open implementation from IBM. This model was pre-trained on sets of public heterogeneous time series, which ensures high transfer learning capability. In this work, the architecture is modified with a fully connected output layer of neurons with a sigmoid activation function, as well as a Dropout normalization layer, after which the improved architecture with frozen initial weights is retrained using the Low Rank Adaption (LoRA) method. The LoRA method allows for the effective retraining of large models by introducing a pair of low-rank adapter matrices into each fully connected transformation instead of a full weight update (5):

$$W' = W_0 + BA \quad (5)$$

$$B \in \mathbb{R}^{n \times r} \quad A \in \mathbb{R}^{r \times m} \quad r \ll \min(n, m)$$

where W' – is the updated weight matrix after LoRA adaptation,

W_0 – is the initial weight matrix,

B, A – are low-rank adapter matrices.

In this work, the rank of the adapter matrices is $r = 8$, which adds no more than 0.1 million parameters to the original model.

The architectural parameters of TimeMixerDrop were configured as follows:

- time window $L = 24$ hours;
- size of the internal layers of the Mixer – 64 neurons per time block and 64 per feature block;
- number of Mixer layers = 4.

The pre-computed weights of TimeMixerDrop were fixed, and the LoRA adapters were trained on the prepared sample. Training was performed using the Adam optimizer for 50 epochs. In the presence of class imbalance, a class frequency boosting factor was introduced to increase the stability of the computational experiment.

1.4 Binary loss function

The described model was trained using a loss function typical for binary classification tasks—binary cross-entropy (BCE), which has proven itself in many scientific and research works. For this model, BCE looks as follows:

$$L = -\frac{1}{N} \sum_{i=1}^N [y_i \ln \hat{p}_i + (1 - y_i) \ln(1 - \hat{p}_i)] \quad (6)$$

where $y_i \in \{0,1\}$ – actual class label for the i -th example,

$\hat{p}_i$ – probability of label 1 predicted by the model for the i -th example.

This loss function measures the discrepancy between the resulting distributions: it takes a minimum value of 0 when the prediction matches the actual labels. During training, the algorithm minimizes this loss function.

In addition to the BCE loss function, the Brier score metric was used for the final quality assessment of the probability forecasts, which is calculated as the mean square error between the predicted probability and the observed outcome:

$$Br = \frac{1}{N} \sum (\hat{p}_i - y_i)^2 \quad (7)$$

This metric is analogous to the loss function relative to the probability of an event: a lower Brier score indicates better calibration and greater accuracy of the probability forecast.

2. RESULTS

The main comparison of the results achieved is shown in Table 1:

Table 1 – Comparison of evaluation metrics for the proposed and baseline models

Model	ROC-AUC	PR-AUC	Brier-score
Logistic regression baseline	0.766	0.171	0.056
TimeMixerDrop	0.793	0.195	0.054

The table shows that the TimeMixerDrop-based model outperforms the baseline model on all comparable metrics.

Another considered indicators are ROC-AUC and ROC-PR metrics, which are shown below in Figure 1:

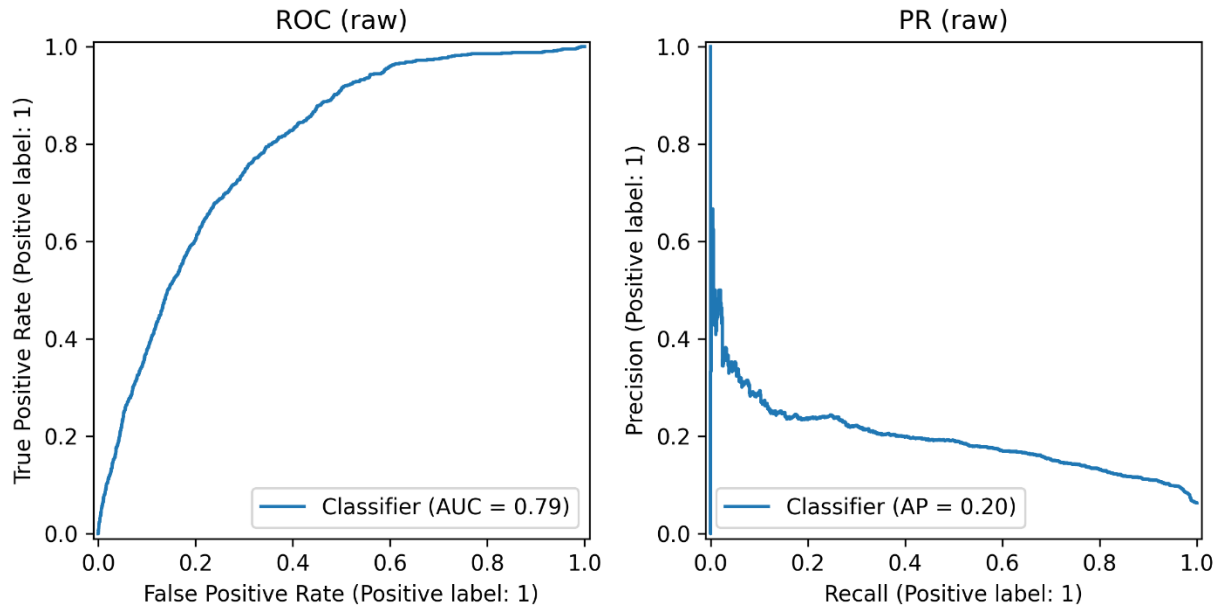


Figure 1 – Results of ROC-AUC and ROC-PR metrics on the test split

The decrease in the Brier score also indicates an increase in the probabilistic reliability of the model, i.e., the proposed model can also produce reliable probabilistic forecasts of events, which is extremely important when testing the model in real conditions.

Distribution of calibrated RI probabilities for RI vs. non-RI samples on log scale is demonstrated within Figure 2.

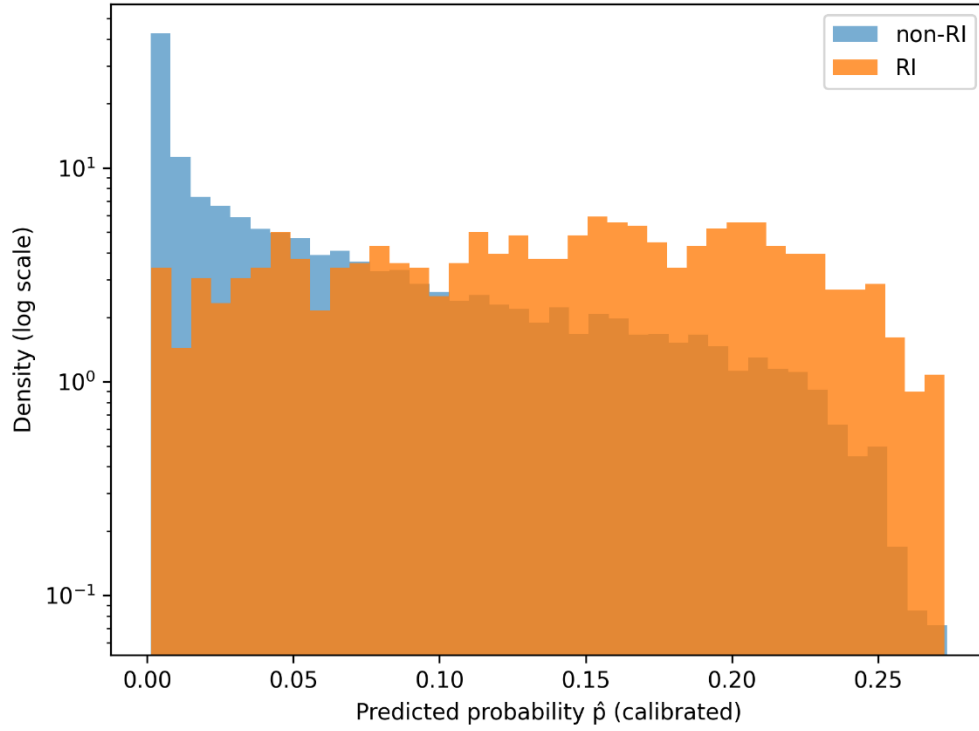


Figure 2 – Calibrated predicted probability distributions for RI (orange) and non-RI (blue) samples on a log-density scale.

Post-hoc calibration keeps the probability estimates well aligned with the empirical event frequency and does not materially change discrimination. On the held-out test split, the calibrated probabilistic accuracy improves relative to climatological baseline. We report the reliability diagram (Figure 3) to provide an auditable summary of calibration across the probability range.

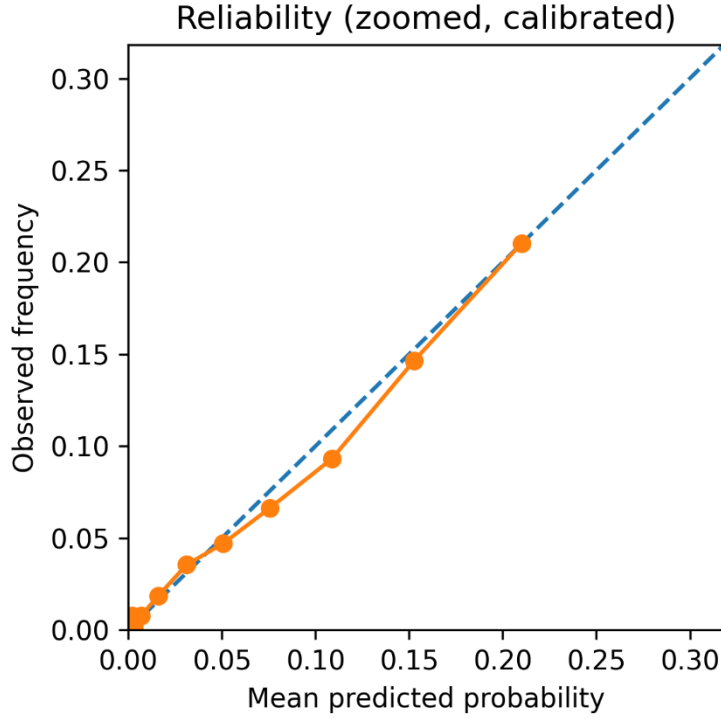


Figure 3 – Reliability diagram for calibrated probabilities.

In addition to metrics free of threshold and probabilistic calibration, early warning applications require threshold-based verification that explicitly accounts for false alarm and missed events. We therefore evaluate the contingency table at selected probability level thresholds τ and report Probability of Detection (POD) and False Alarm Ratio (FAR):

$$POD = \frac{TP}{TP + FN}, FAR = \frac{FP}{TP + FP} \quad (8)$$

where TP – number of true positive predictions, FN – number of false negative predictions, FP – number of false positive predictions. For clarity, we also report false positive rate (FPR) and the alert rate r , which approximates the fraction of time steps that would trigger an escalation action in an early warning workflow:

$$FPR = \frac{FP}{FP + TN}, r = \frac{TP + FP}{N} \quad (9)$$

where TN – number of true negative predictions, N – number of all predictions.

RI remain less frequent than non-RI at the 6-hour track, thus, in the held-out test split we evaluate $N = 12981$ samples with 823 RI positives, which means base rate is about 6.34%. Therefore, probability thresholds in the conventional 0.2-0.27 range are operationally meaningful for this model, and threshold selection can be framed as an operational trade-off between detection capability and alert workload. We provide selected thresholds in Table 2 by

targeting feasible alert budgets in which reflects realistic workload constraints for warning operation.

Table 2 – Threshold based early warning verification on the test split

τ	TP	FP	TN	FN	POD	FAR	Precision	Alerts	alert_rate
0.207	154	504	11654	669	0.187	0.767	0.234	658	0.050
0.242	45	92	12066	778	0.055	0.672	0.328	137	0.010
0.270	2	2	12156	821	0.002	0.5	0.5	4	0.0003

As shown in Table 2, increasing the threshold reduces alert workload but also decreases POD. Notably, at the 1% alert rate operating point the precision reaches 0.328, which is $\sim 5.2\times$ higher than the climatological event frequency, indicating meaningful enrichment for an operational trigger under constrained capacity.

To assess regional variability, we report verification by ocean basin, what is shown in Table 3, including the number of test samples and RI positives per basin. Skill varies across regions and ROC-AUC and PR-AUC are undefined when the basin subset contains single class, which we denote explicitly. These results motivate basin-adaptive thresholding and region-specific feature extensions in future work.

Table 3 – Basin-wise metrics evaluations

basin	Number of 6-h samples	Number of RI in samples	RI positive rate	ROC-AUC	PR-AUC	Brier score
EP	1808	75	0.041	0.893	0.232	0.035
NA	2393	103	0.043	0.759	0.149	0.038
NI	560	47	0.083	0.670	0.125	0.075
SI	2676	156	0.058	0.800	0.155	0.051
SP	1361	93	0.068	0.757	0.165	0.060
WP	4183	349	0.083	0.770	0.248	0.070

Skill varies across basins, reflecting regional differences in climatology and sample composition. These results motivate basin-aware thresholding and targeted feature extensions in future work.

3 DISCUSSIONS

The proposed model demonstrates strong discrimination on the held-out test split, ROC-AUC = 0.793, AP = 0.195 while using only a compact physically interpretable predictor set. Probability calibration yields a small but consistent improvement in Brier score, indicating improved probabilistic reliability. Compared to a logistic regression baseline trained on the same standardized predictors, TimeMixerDrop achieves higher ROC-AUC, supporting the benefit of nonlinear temporal mixing even under a minimal feature design.

The results of the model indicate that it is physically adequate: the warm ocean surface provides the cyclone with a large amount of energy, creating favorable conditions for the formation of a tropical cyclone. The least influential feature was the vertical wind shear. Possible reason for this assessment is the consideration of only the difference between two levels of wind speed; increasing the number of levels may improve the model's accuracy.

To understand which input features, contribute most to forecasting, a post-forecasting analysis was performed using the SHAP (Shapley Additive explanations) method. This method estimates the influence of each feature on the predicted probability by calculating conditional Shapley values (10):

$$\phi_j(x) = \sum_{S \subseteq F \setminus j} \frac{|S|! (|F| - |S| - 1)!}{|F|!} [f_{S \cup j}(x_{S \cup j}) - f_S(x_S)] \quad (10)$$

where $\phi_j(x)$ – SHAP value of feature j ,

F – set of all features,

S – subset of features,

f_S – model using only features from set S .

The calculated mean absolute values of SHAP analysis are shown in Figure 4.

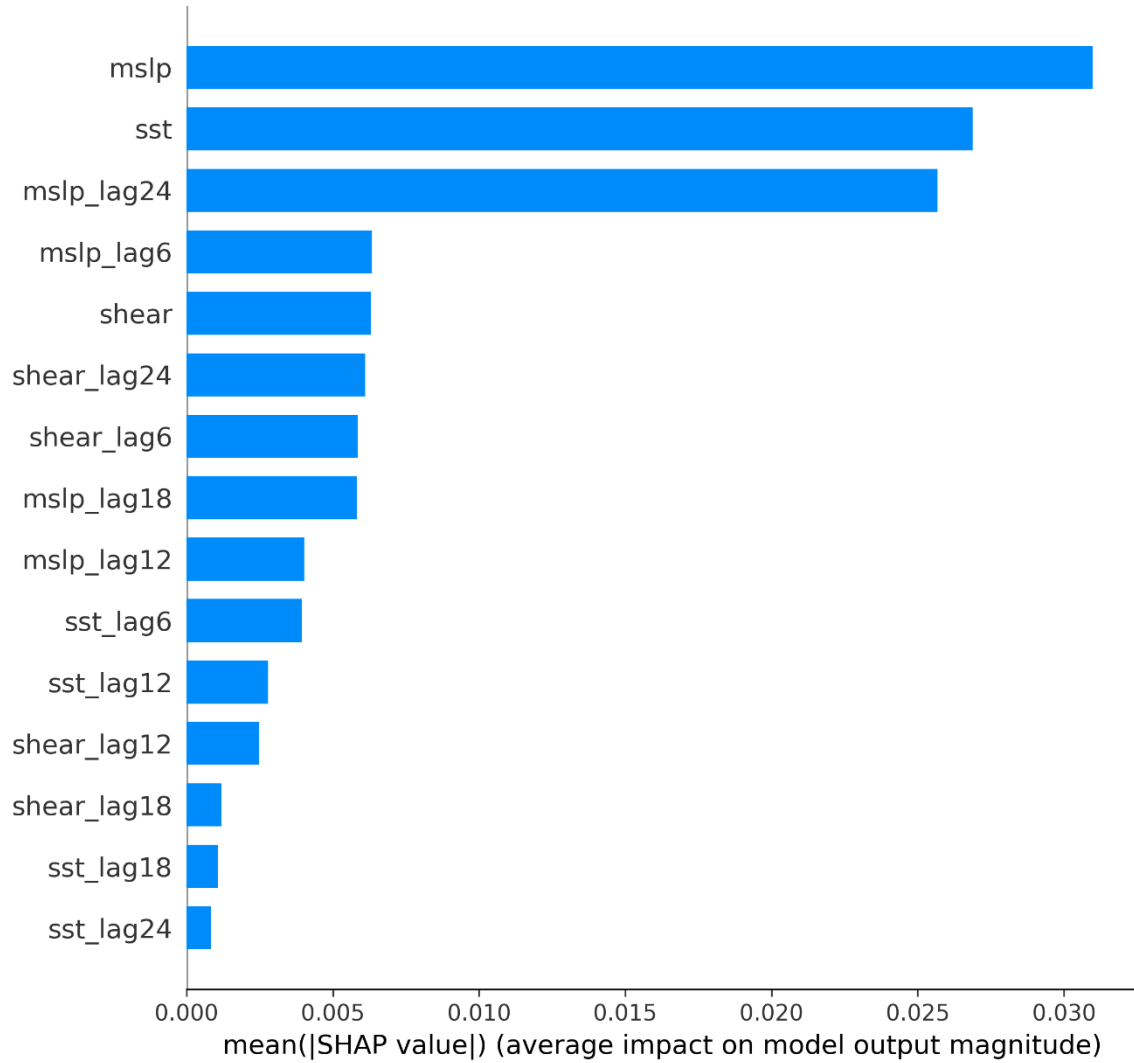


Figure 4 – Feature importance based on SHAP analysis. Lag features represent rolling window values of a different feature set (MSLP, SST, shear)

The average absolute value of SHAP, as a percentage of the contribution to the model forecast. The higher the column, the greater the contribution of the feature to the forecast result. As can be seen in Figure 4, the dominant feature of the model is MSLP and SST, whose average significance is more than four times higher than wind shear feature. A curious fact to consider, MSLP and shear lags have comparatively meaningful impact on the model output, meanwhile SST lag features show the least importance, in compare to SST without lag being second most impactful feature.

The comparatively lower attribution of vertical wind shear should be interpreted cautiously. In the present setup, shear is represented by a simplified two levels, which may not capture deep layer shear structure, directional changes, or its interaction with storm motion and convection. Therefore, the reduced importance may reflect feature under specification rather than a

physically irrelevance of shear. A more complete representation is expected to yield more physically consistent importance patterns.

To demonstrate proposed solutions in we provide a qualitative example of a model behavior, shown in Figure 5.

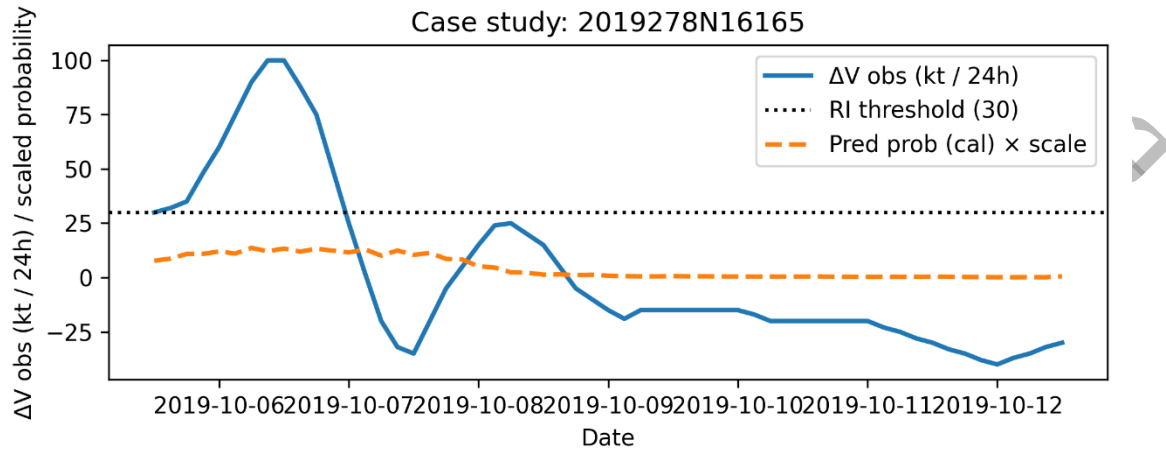


Figure 5 – the model results for the cyclone 2019278N16165 (IBTrACS id)

Figure 5 shows the observed 24-h intensity change together with the calibrated RI probability over time. This single storm visualization supports interpretability signal, while the main conclusions are based on aggregated verification over the full held-out test split. The case study is provided for qualitative illustration and does not replace aggregated verification. Evaluation is performed on the full held-out test split.

SHAP analysis also confirmed the physical consistency of the model: the influence of factors is consistent with their physical significance, which increases the interpretability of the model results.

One of the key advantages of the proposed model is its compactness and speed. The number of its parameters is many times smaller than that of large transformer-based models, which allows the model to be run on computing power comparable to that of home computers.

The model's inference time has been reduced significantly, allowing for the prediction of sudden cyclone intensification several hours before the event occurs, which is critical for emergency response and ensuring the safety of coastal areas.

Taken together, these results show that the proposed architecture and methods have significantly improved both the accuracy and efficiency of the forecast, while the model can be easily integrated into low-power computing devices.

Operational value in early warning and decision support, such as in disaster management workflows, the primary operational product of the model is a calibrated probability of RI within

the next 24 hours. This probability can be used as a decision support signal in early warning systems: to update warning levels and internal readiness states; to support evacuation planning and staging decisions when high-impact escalation becomes likely; to communicate risk under uncertainty to the public and partner agencies. Because the consequences of missed RI events can be severe, emergency managers often require explicit trade-offs between detection capability and false alarms. Therefore, probabilistic forecasts are preferable to hard discrete prediction, such as will be or not will be.

In an impact-based forecasting setting, the RI probability can be combined with exposure and vulnerability information to prioritize area and assets at risk. Practically, this means that the model output can be mapped to action thresholds that reflect local cost-loss ratios and acceptable false-alarm tolerance. The lightweight computational consumptions of TimeMixerDrop-RI enables frequent updates, within 6 or less hours, making it suitable for integration into operational pipelines and crisis response coordination even where high-performance computing is not available.

Although, the proposed TimeMixerDrop-RI model demonstrates promising skill in forecasting intensification, several limitations should be acknowledged.

First, the predictor set is intentionally minimal and includes only three physically motivated variables: sea surface temperature, mean sea-level pressure and vertical wind shear between 200 and 850 hPa. While this choice improves interpretability and reduces the dimensionality of the problem, it ignores other factors that are known to influence rapid intensification, such as upper-ocean heat content, mid-level humidity, low-level vorticity, and convective instability indices. As a result, some physically important scenarios may not be fully captured by the current model.

Second, the temporal context considered by the model is limited to 24-hour window prior to the forecast time. Rapid intensification can be conditioned by longer-term preconditioning of the ocean-atmosphere system, including multi-day accumulation of heat and moisture. Restricting the model to a single-day history simplifies the architecture and reduces computational cost but may limit its ability to capture slowed background processes that precede explosive strengthening.

Third, the model is trained in this study focuses on standard probabilistic metrics and one historical case. These diagnostics are sufficient to demonstrate the potential of the approach, but they do not fully represent complex model performance. In particular, cost-loss consideration, false alarm tolerance and missed-event penalties are not specifically analyzed.

Finally, the current research considers only the probability of rapid intensification as the target quantity. Other aspects that are critical for crisis management are not directly addressed by the model.

Building on the limitations discussed above, several directions of improvements can be considered.

From a methodological perspective, the first natural extension is to enrich the predictor data set, incorporate variables not mentioned in the current work, extend temporal window of forecast.

Global TimeMixerDrop-RI model architecture can be complemented by basin specific or regionally fine-tuned variants. Such regional adaptations would allow the architecture to account for local climatology, typical shear patterns.

Future studies should focus on integrating the rapid intensification forecast into end-to-end impact-based early-warning system. In this way, the model's output can be transformed from a purely meteorological indicator into risk-oriented metrics directly usable by emergency operation centers.

Considering the ongoing changes in climate systems, it is important to investigate the robustness of the model under non-stationary conditions.

CONCLUSION

An author's model based on the TimeMixerDrop architecture has been developed and researched for predicting the intense intensification of tropical cyclones. The model is based on a modern and compact architecture adapted for the task of binary classification based on physical parameters. The key achievements of the work are a number of facts.

The proposed TimeMixerDrop model outperforms a logistic regression baseline on the held-out test split: ROC AUC = 0.79 and PR = 0.20. The model achieves a lower Brier score than the climatological baseline indicating improved probabilistic accuracy. Analysis of the proposed metrics indicates more confident and reliable detection of cases of rapid intensification, which was confirmed by the specific example and aggregated results. The level achieved is comparable to or exceeds the results of individual foreign studies on this topic.

The model has high computational efficiency due to its small number of parameters and short inference time, and can be used on computing power comparable to home computers, which is extremely important for rapid response.

In the context of disaster prevention and crisis management, the proposed TimeMixerDrop-RI model can serve as a building block for impact-oriented early-warning systems. Its ability to generate calibrated probabilistic forecasts of rapid intensification with modest computational requirements makes it suitable for integration into operational workflows, including in regions with limited access to high-performance computing. When combined with surge, rainfall, and exposure models, the approach presented in this study can support more informed and timely decisions by emergency managers, ultimately contributing to the reduction of tropical-cyclone-related risks for coastal communities.

The present implementation uses a compact set of predictors: SST, MSLP, simplified shear and a 24-hour before forecast history. While this choice supports operational simplicity and reproducibility, rapid intensification can be influenced by additional environmental and oceanic factors (humidity, ocean heat content, and others) and by longer-term coupled processes. Future work should extend the feature set and explore longer temporal contexts. Also, the current formulation relies on point-wise representations at the cyclone center and does not explicitly capture spatial predictors from gridded fields is a clear direction to improve POD and FAR metrics. More to it, because RI is a rare event at 6-hour resolution, threshold selection must be framed as an operational decision problem. Future studies should evaluate

basin-specific generalization more extensively, compare against additional baselines and assess post-processing rules to suppress isolated false alarms while preserving lead time.

The results obtained in this way allow us to recommend, under considered limitations, their use in the creation of intelligent systems for forecasting rapidly developing hydrometeorological hazard in real time.

نسخه پذیرفته شده پیش از انتشار

REFERENCES

1. **Ekambaram, V., Jati, A., Dayama, P., et al.** (2024), Tiny Time Mixers (TTMs): Fast pre-trained models for enhanced zero/few-shot forecasting of multivariate time series, *arXiv preprint*, 2401.03955.
2. **Griffin, S. M., Wimmers, A., & Velden, C. S.** (2022), Predicting rapid intensification in North Atlantic and eastern North Pacific tropical cyclones using a convolutional neural network, *Weather and Forecasting*, 37(8), 1333–1355, <https://doi.org/10.1175/WAF-D-21-0194.1>.
3. **Harrigan, S., Zsoter, E., Alfieri, L., et al.** (2020), GloFAS-ERA5 operational global river discharge reanalysis 1979–present, *Earth System Science Data*, 12, 2043–2060, <https://doi.org/10.5194/essd-12-2043-2020>.
4. **Kalach, A. V., & Paramonov, A. A.** (2025), Ontology model of almost periodic data analysis with ordered argument, *Bulletin of Voronezh State University. Series: Systems Analysis and Information Technologies*, (1), 101–114, <https://doi.org/10.17308/sait/1995-5499/2025/1/101-114>. [In Russian].
5. **Kalach, A. V., Paramonov, A. A., & Shkerin, A. L.** (2025), Hazard assessment of tropical cyclone structure zones based on almost periodic analysis, *Bulletin of Voronezh State University. Series: Systems Analysis and Information Technologies*, (2), 16–28, <https://doi.org/10.17308/sait/1995-5499/2025/2/16-28>. [In Russian].
6. **Kim, S.-H., Lee, W., Kang, H.-W., & Kang, S. K.** (2024), Predicting rapid intensification of tropical cyclones in the western North Pacific: A machine learning and net energy gain rate approach, *Frontiers in Marine Science*, 10, 1296274.
7. **Knapp, K. R., Diamond, H. J., Kossin, J. P., et al.** (2010), International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4 [Data set], NOAA National Centers for Environmental Information. Accessed 14 May 2025.
8. **Li, Y., Tang, Y., Wang, S., et al.** (2023), Recent increases in tropical cyclone rapid intensification events in global offshore regions, *Nature Communications*, 14, 5167, <https://doi.org/10.1038/s41467-023-40605-2>.
9. **Paramonov, A. A., & Kalach, A. V.** (2025), Modification of the mathematical apparatus of almost periodic analysis for processing spatio-temporal data with an ordered argument, *Bulletin of the South Ural State University. Series: Mathematics. Mechanics. Physics*, 17(2), 44–49, <https://doi.org/10.14529/mmph250205>. [In Russian].
10. **Paramonov, A. A., & Kalach, A. V.** (2024), Modeling emergencies using near-periodic analysis of typhoon structure images, *Bulletin of the South Ural State University. Series: Mathematics. Mechanics. Physics*, 16(4), 67–74, <https://doi.org/10.14529/mmph240408>. [In Russian].
11. **Pathak, J., Subramanian, S., Kashinath, K., et al.** (2022), FourCastNet: A global data-driven high-resolution weather model using adaptive Fourier neural operators, *arXiv preprint*, 2208.05419.
12. **Rasp, S., Hoyer, S., Dueben, P., et al.** (2023), WeatherBench 2: A benchmark for the next generation of data-driven global weather models, *arXiv preprint*, 2308.15560.
13. **Rogers, R. F., Liu, M., & Zhang, J. A.** (2022), The shear-relative variation of inflow angle and its relationship to tropical cyclone intensification, *Journal of Geophysical Research: Atmospheres*, 127(13), e2022JD037280, <https://doi.org/10.1029/2022JD037280>.