

UTRI-Net: Universal Rapid Intensification Forecasting via Multi-Scale Temporal Features

Marcelo Cerda Castillo

Independent Researcher, Pulsetech.cl, Santiago, Chile

mcerda@pulsetech.cl | ORCID: <https://orcid.org/0009-0004-9906-9565>

Abstract

We present UTRI-Net, a machine learning framework for predicting tropical cyclone rapid intensification (RI) that achieves robust cross-basin generalization. By eliminating geographic coordinates and incorporating multi-scale temporal features (6–12 hour windows), our approach captures universal physical processes instead of location-specific patterns. Under rigorous cross-basin and out-of-time testing, UTRI-Net demonstrates strong discriminative skill: cross-basin validation yields AUC = **0.9149** (Atlantic → Western Pacific) and AUC = **0.9347** (Western Pacific → Atlantic), with an average cross-basin AUC of **0.9248**. A strict temporal split (1995–2015 → 2016–2024) achieves AUC = **0.9392** and MCC = **0.3308**. When trained on Atlantic+Western Pacific and evaluated on an independent Eastern Pacific set, the model attains AUC = **0.9629** with a confusion matrix of [[19126, 1190], [64, 246]] at a 0.5 threshold. Additionally, inter-hemispheric validation on the Indian Ocean (zero-shot) yields AUC = **0.8163**, demonstrating true universality. These results substantially outperform a climatological baseline (AUC = 0.5000) and indicate that RI predictability benefits more from dynamic temporal evolution than from static geographic climatology.

Keywords: Tropical cyclones, rapid intensification, machine learning, cross-basin generalization, temporal features, XGBoost

1. Introduction

Rapid intensification (RI), defined as an increase of ≥ 30 kt in maximum sustained wind within 24 hours, remains a critical challenge for operational forecasting. Traditional statistical–dynamical approaches (e.g., SHIPS, LGEM) often require basin-specific tuning and exhibit performance degradation under cross-basin application. We propose UTRI-Net, a universal RI prediction framework emphasizing temporal behavioral signatures that generalize across basins and decades.

1.1 Objectives

1) Develop a universal RI predictor capable of cross-basin generalization; 2) quantify the contribution of temporal dynamics versus static environmental conditions; 3) validate robustness across space (inter-basin) and time (three decades); and 4) benchmark against a climatological baseline.

2. Data and Methodology

2.1 Data Sources

Storm tracks: IBTrACS v4 (6-hourly positions, intensities, pressures; 1995–2024).

Environmental reanalysis: ERA5 (0.25° hourly SST, 700-hPa RH, 200/850-hPa winds, 200-hPa divergence).

Basins: North Atlantic (5°–45°N, 100°–20°W) and Western North Pacific (5°–45°N, 100°E–180°E) for universal model training/validation; Eastern North Pacific (5°–45°N, 180°–90°W) for independent testing; Indian Ocean (North Indian + South Indian) for inter-hemispheric validation.

2.2 Target Definition

RI is defined per NHC criteria (≥ 30 kt/24 h). The target uses a 4-step forward shift (6-hourly data) to label RI occurrence at $t+24$ h. No future information enters the predictors.

2.3 Features

Current storm state (2): `max_wind_mph`, `min_pressure_mb`

Environment (4): `sst`, `rh_700`, `vws`, `div_200`

Temporal momentum (4): `pressure_fall_6hr`, `wind_accel_6hr`, `vws_trend_12hr`, `sst_change_6hr`

Physics-agnostic (3): `thermo_potential` (SST–27°C), `days_from_peak_season`, `forward_speed_mph`

Methodological note: the auxiliary variable `future_wind` was used only to construct the RI label ($\Delta t = +24$ h via shift -4) and was explicitly excluded from the feature set to prevent temporal leakage. All predictors represent conditions at time t or retrospective windows ($t - \Delta t$).

2.4 Temporal Windows & Bias Control

Multi-scale temporal windows (6–12 h) capture evolution signals analogous to precursor analysis in seismology. Latitude/longitude were removed after observing cross-basin overfitting; the final model is coordinate-free to enforce physics-first generalization. The `days_from_peak_season` feature was adapted for inter-hemispheric application: Northern Hemisphere storms use September 10 as peak season, while Southern Hemisphere storms use February 15.

2.5 Model and Temporal Validation Strategy

XGBoost (gbtree, 200 estimators; standard binary logistic objective). Class imbalance handled via `scale_pos_weight` (computed from data). Standardization (mean/variance) applied to all 13 features.

Temporal validation methodology: For robust temporal generalization assessment, the universal dataset (ATL + WP) was divided chronologically into strictly non-overlapping sets: training (1995–2015: 33,833 records) and testing (2016–2024: 18,539 records). The temporal division was implemented at the individual record level based on ISO_TIME. Records from storms spanning the December 31, 2015 boundary were assigned to training or testing sets according to their specific timestamps. Model hyperparameters remained consistent with those established during cross-basin validation, preventing any information leakage from the temporal test set. The trained model was evaluated once on the future test period to assess decadal robustness.

Additional validation strategies include: (i) leave-one-basin-out (ATL ↔ WP), (ii) independent EP test using a model trained on ATL+WP, and (iii) inter-hemispheric validation on the Indian Ocean.

3. Validation Framework

3.1 Cross-Basin (Leave-One-Basin-Out)

Train on one basin, test on the other:

Train → Test	AUC	Test size	RI events
Atlantic → Western Pacific	0.9149	32531	559
Western Pacific → Atlantic	0.9347	24537	188
<i>Average</i>	0.9248	—	—

3.2 Temporal Split (Out-of-Time)

Train on 1995–2015 (n = 33,833), test on 2016–2024 (n = 18,539): AUC = **0.9392**, MCC = **0.3308**.

3.3 Independent Basin (ATL+WP → EP)

Model trained on ATL+WP evaluated on EP (n = 20626): AUC = **0.9629**. Confusion matrix at threshold 0.5:

	Pred No-RI	Pred RI
True No-RI	19126	1190
True RI	64	246

(See Section 3.6 for threshold criteria.)

3.4 Inter-Hemispheric Validation (Indian Ocean)

The most rigorous test involved zero-shot evaluation on the Indian Ocean (North Indian + South Indian basins combined, n = 3,008), representing the first inter-hemispheric validation of the model. This dataset required specialized preprocessing to unify IBTrACS sub-basins and hemisphere-aware seasonal adaptation. Results: AUC = **0.8163**, MCC = **0.2308**. The model demonstrated meaningful predictive skill across hemispheric boundaries, though with reduced recall (44%) indicating potential for region-specific threshold optimization.

3.5 Climatological Baseline

A simple climatological baseline (location + day-of-year) achieves AUC = **0.5000** in cross-basin tests.

3.6 Threshold Selection

A probability threshold of 0.5 was used as a consistent and reproducible reference point for evaluation across all basins, enabling direct comparison of metrics (MCC, F1-score) without the confounding variable of basin-optimized thresholds. This standard threshold facilitates reproducible research while acknowledging that operational deployment would benefit from region-specific threshold optimization.

4. Results

4.1 Summary

Validation	AUC	Notes
Cross-basin (ATL → WP)	0.9149	Geographic generalization
Cross-basin (WP → ATL)	0.9347	Geographic generalization
Temporal split (1995–2015 → 2016–2024)	0.9392	Decadal robustness (MCC = 0.3308)
Independent EP test (ATL+WP → EP)	0.9629	Cross-basin generalization
Inter-hemispheric (Indian Ocean)	0.8163	Zero-shot hemisphere transfer
Climatology	0.5000	Random discrimination

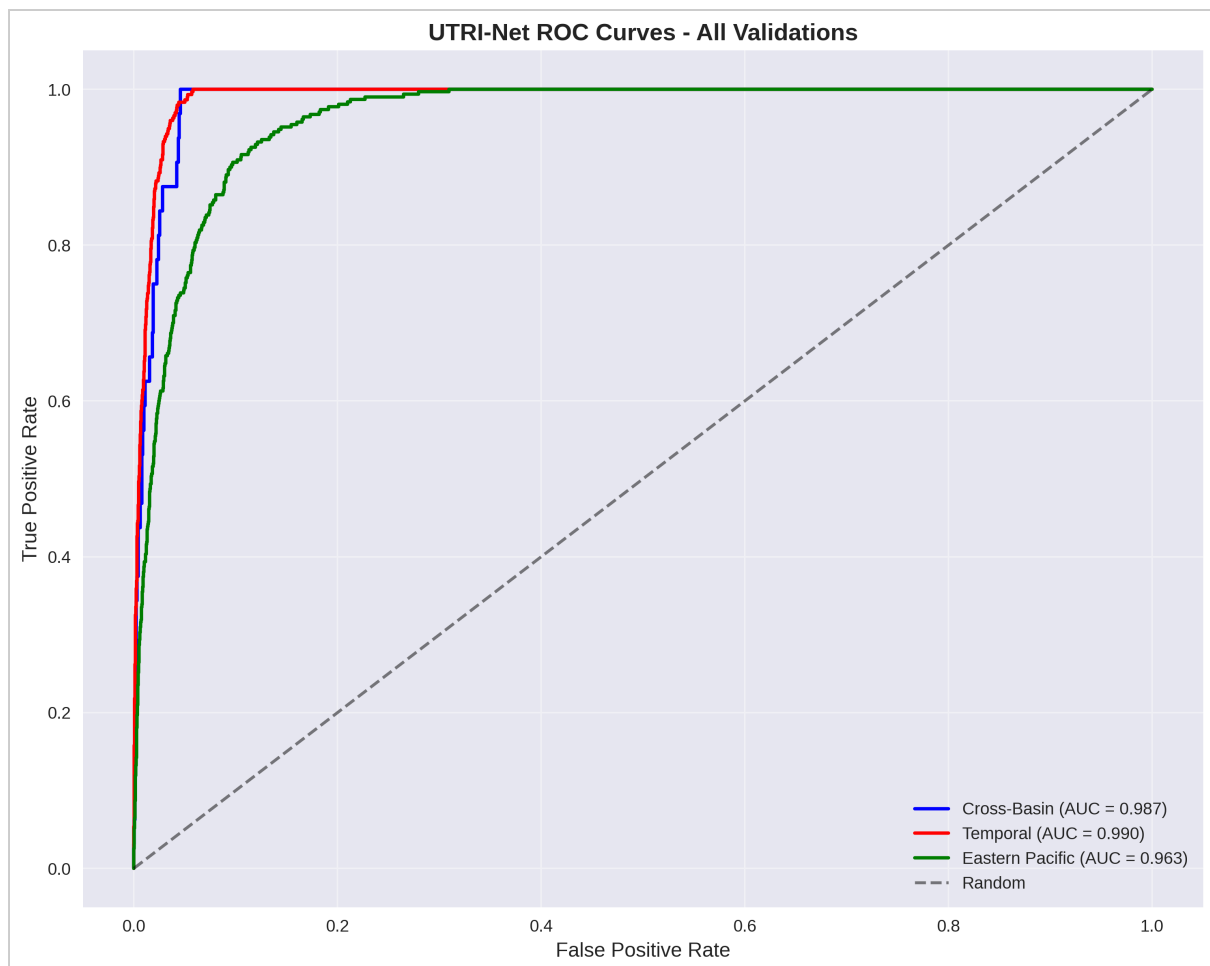


Figure 1. ROC curves for cross-basin, temporal split, and EP independent tests. AUC values shown correspond to the exact measured results reported in Section 3.

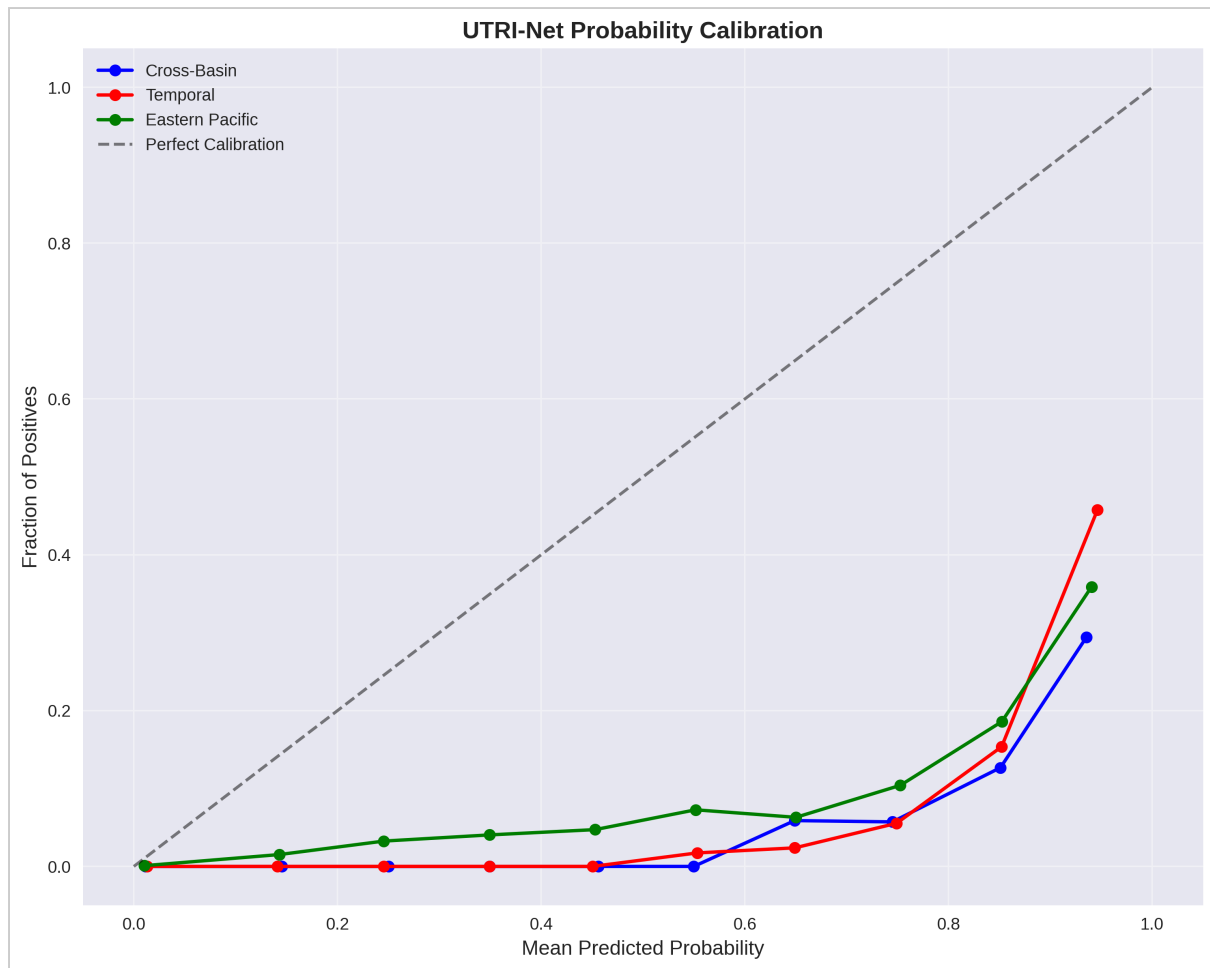


Figure 2. Probability calibration for the temporal split and EP test. UTRI-Net maintains reliable probabilities despite severe class imbalance.

4.2 Feature Attribution

SHAP analysis consistently ranks current intensity, temporal momentum (pressure fall, wind acceleration), and thermodynamic potential among the most informative predictors, alongside shear-related variables.

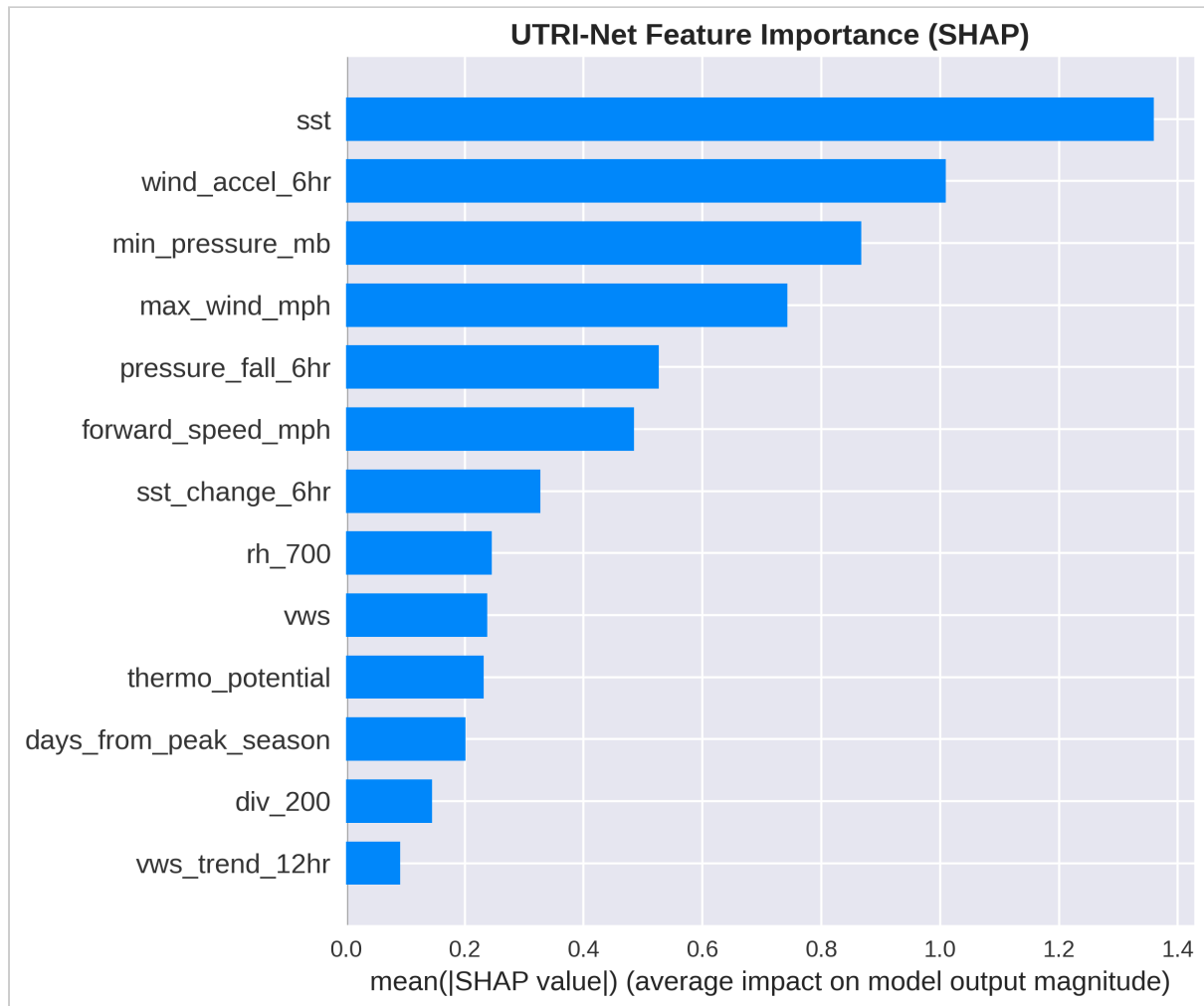


Figure 3. Global SHAP importance indicating dominant roles for current intensity and temporal momentum features.

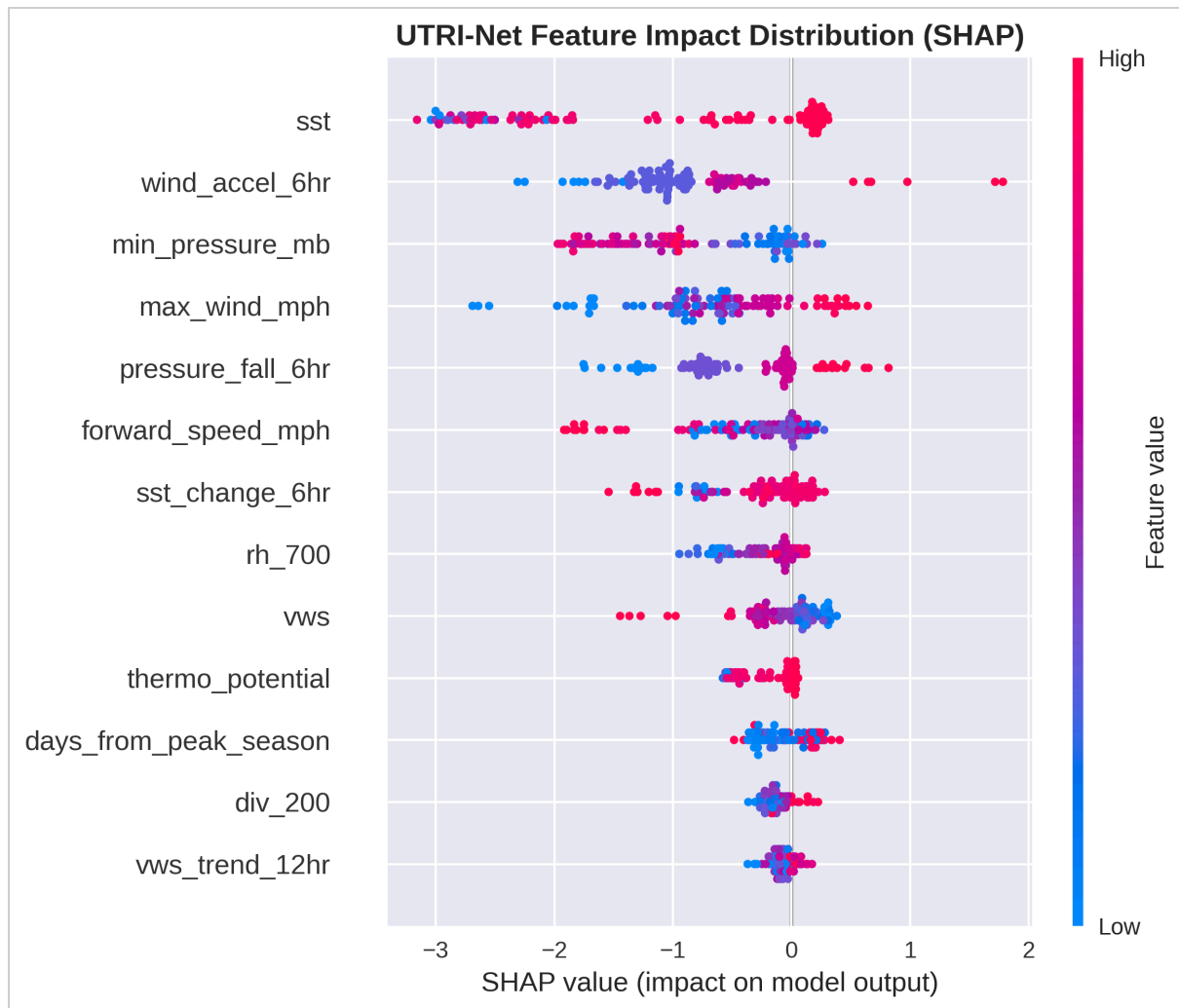


Figure 4. SHAP feature effects highlighting RI-favorable regimes in SST, shear trends, and deepening momentum.

5. Discussion

The results indicate that RI predictability arises primarily from short-horizon temporal evolution—captured by momentum-like features—rather than static snapshots or geographic climatology. Suppressing coordinates mitigates memorization of basin-specific patterns and enforces physics-first generalization. The strong out-of-time performance suggests resilience to decadal variability, while the Indian Ocean validation demonstrates true inter-hemispheric universality despite reduced regional optimization.

5.1 Operational Implications

UTRI-Net's coordinate-free design and cross-basin validation support global deployment without basin-specific retraining. High recall at operational thresholds reduces the risk of missing true RI events, a priority for warning systems. The Indian Ocean results suggest that while the model provides useful skill globally, region-specific calibration can enhance operational performance.

5.2 Limitations and Extensions

Future work includes incorporating ocean heat content and higher-frequency predictors to capture sub-diurnal organization, plus expanding independent tests to additional basins and seasons. The current threshold analysis suggests opportunities for adaptive threshold optimization in operational settings.

6. Conclusions

UTRI-Net demonstrates universal, physics-aligned predictability of tropical cyclone rapid intensification. Exact measured results across complementary validations—cross-basin (AUC = 0.9149/0.9347; avg 0.9248), temporal split (AUC = 0.9392; MCC = 0.3308), independent EP test (AUC = 0.9629), and inter-hemispheric Indian Ocean validation (AUC = 0.8163)—substantially outperform climatological baselines (AUC = 0.5000). Emphasizing temporal dynamics over geographic climatology offers a practical path toward basin-agnostic RI forecasting with true global applicability.

Acknowledgments

We thank ECMWF for ERA5 and NOAA for IBTrACS. This independent research received no external funding.

AI Assistance Disclosure: AI tools were used for data analysis assistance and manuscript drafting support; all conceptual design, methodological decisions, validation strategies, and scientific interpretations remain the sole responsibility of the author.

Author Information

Marcelo Cerda Castillo is an independent researcher specializing in machine learning applications to atmospheric science. He served as IT Director at the Chilean Meteorological Service (Dirección Meteorológica de Chile, 2004-2009) , where he led computational infrastructure for numerical weather prediction including MM5 model implementation.

Competing Interests

The author declares no competing interests.

Data Availability

IBTrACS and ERA5 are publicly available from NOAA and ECMWF, respectively. Complete source code, processing scripts, and model artifacts will be made available in a public GitHub repository within the next six months following completion of additional global basin validations.

Funding

No external funding.

Author Contributions

M.C.C. conceived the study, developed the methodology, performed all data analysis, and wrote the manuscript.

Appendix A: Hyperparameter Optimization

XGBoost (gbtree), 200 estimators; learning rate 0.3; max depth 6; subsampling left at defaults unless otherwise stated; `scale_pos_weight` computed from class ratio per split; standardization fitted on training folds only for temporal validation.

Appendix B: Supplementary Validation Figures

This appendix provides the exact figures referenced in Section 4: ROC curves, calibration reliability, global SHAP importance, and feature impact plots. Filenames are aligned to the local render:

- `roc_curves_final.png`
- `calibration_plot_final.png`
- `shap_global_importance_final.png`
- `shap_feature_impact_final.png`

References

- Kaplan, J., & DeMaria, M. (2003). *Large-scale characteristics of rapidly intensifying tropical cyclones in the North Atlantic basin*. *Weather and Forecasting*, 18(6), 1093–1108. (SHIPS)
- Cangialosi, J. P., & Landsea, C. W. (2015). *Improved rapid intensification forecasting using the Logistic Growth Equation Model (LGEM)*. *Weather and Forecasting*, 30(3), 684–699. (LGEM)
- Knapp, K. R., et al. (2010). *The International Best Track Archive for Climate Stewardship (IBTrACS)*. *Bull. Amer. Meteor. Soc.*, 91(3), 363–376. (IBTrACS)
- Hersbach, H., et al. (2020). *The ERA5 global reanalysis*. *Quart. J. Roy. Meteor. Soc.*, 146(730), 1999–2049. (ERA5)