Uncertainty-Aware Prediction of Climate Extremes Using Fine-Tuned Time-Series Foundation Models

Imran Nasim[1,3], João Lucas de Sousa Almeida[2]

[1]IBM UK; [2] IBM Research Brazil; [3] Department of Mathematics, University of Surrey

Abstract

AI-driven weather forecasting models, particularly foundation models, have achieved significant advancements in both speed and accuracy [6, 5, 8], however, accurately forecasting rare, high-impact extreme events, such as storms and heatwaves, remains a critical challenge [7]. These models often underestimate event intensity and frequency, limiting their reliability in operational and risk-sensitive contexts. In this study, we investigate uncertainty-aware extreme event forecasting using the recently introduced time-series foundation model, Tiny Time Mixers (TTM) [1]. We develop and compare two uncertainty quantification approaches, hyperparameter ensembling and Monte Carlo (MC) dropout [2], and evaluate their ability to improve classification of extreme events. Our results show that incorporating predictive uncertainty significantly enhances performance compared to zero-shot TTM, and that the choice of uncertainty method and threshold critically affects model behaviour. We find that the hyperparameter ensemble yields more stable and accurate predictions, particularly for rare storm events, highlighting the value of lightweight ensemble models for uncertainty-calibrated forecasting.

Dataset: collection and preprocessing

We obtained weather data from 567 stations across Brazil via the Brazilian National Institute of Meteorology (INMET) [3] web interface [4], covering the period from 2000 to December 2023. Due to varying installation dates and data gaps, not all stations span the full period. After quality checks, we selected data from 2019–2023, which offered the best national coverage with minimal missing values. Stations with more than 1

For training, validation, and testing, we applied a temporal split of 90

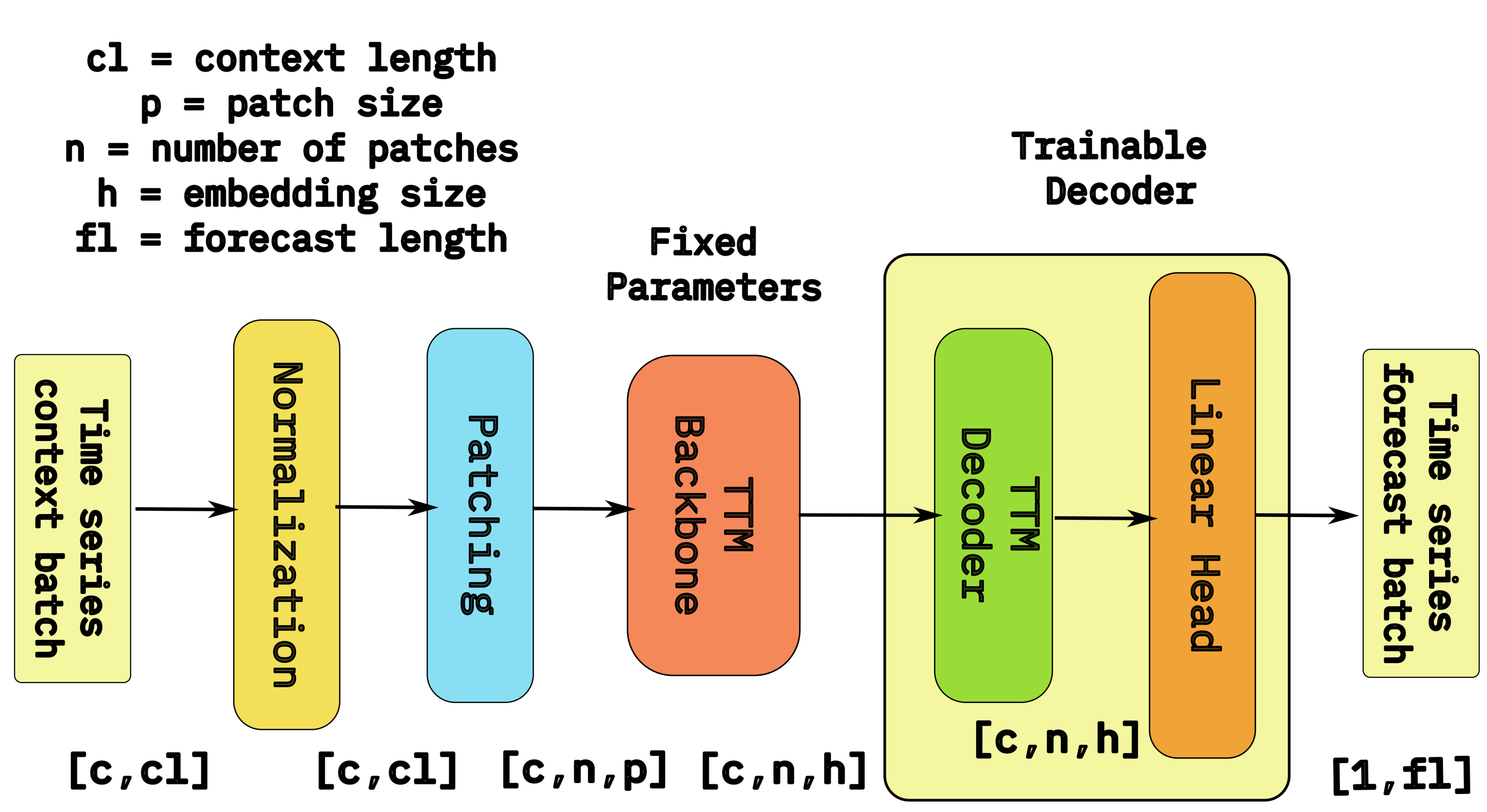

Figure 1: A visual representation of the Tiny Time-Mixer (TTM) architecture.

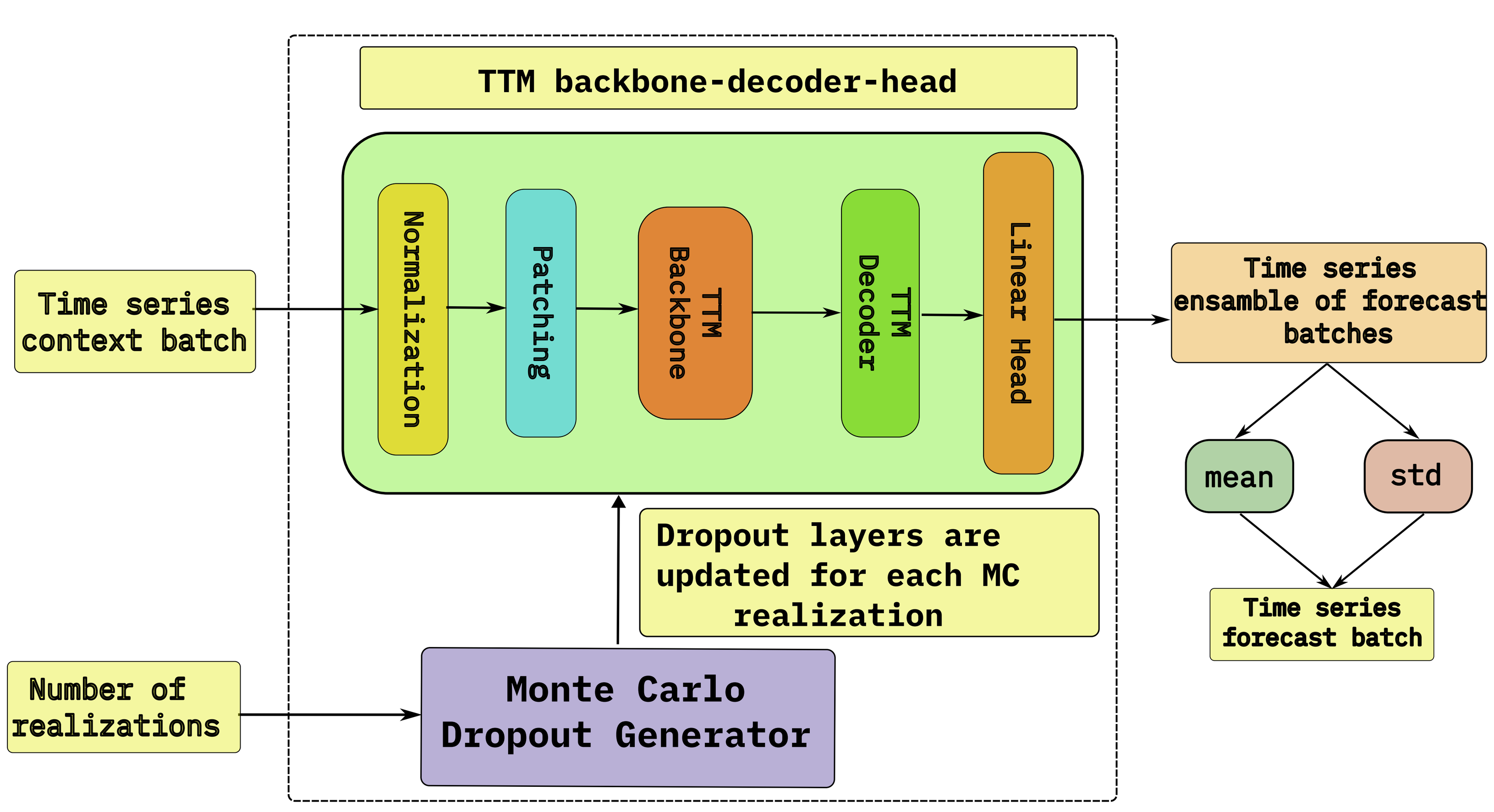

Figure 2: An overview of the ensembling method used to generate the uncertainty quantification metrics

Predictions and Evaluations Metrics

Here we present and example of forecasting and the confusion matrices representing the exactness of the predictions reduced by the finetuned model.

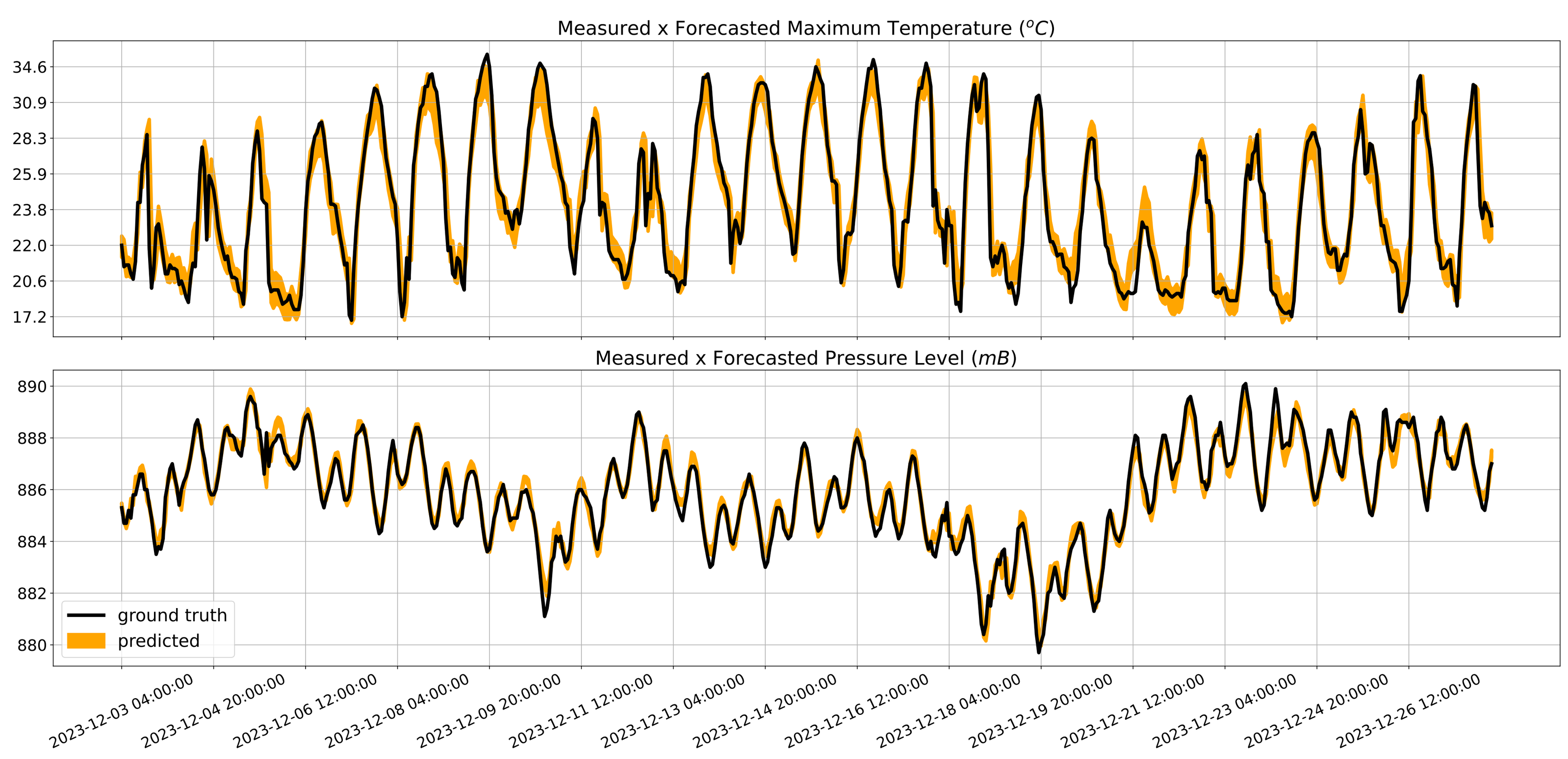

Figure 3: Batchwise extrapolation using PatchTSMixer. The variables were normalized in order to enhance the visualization.

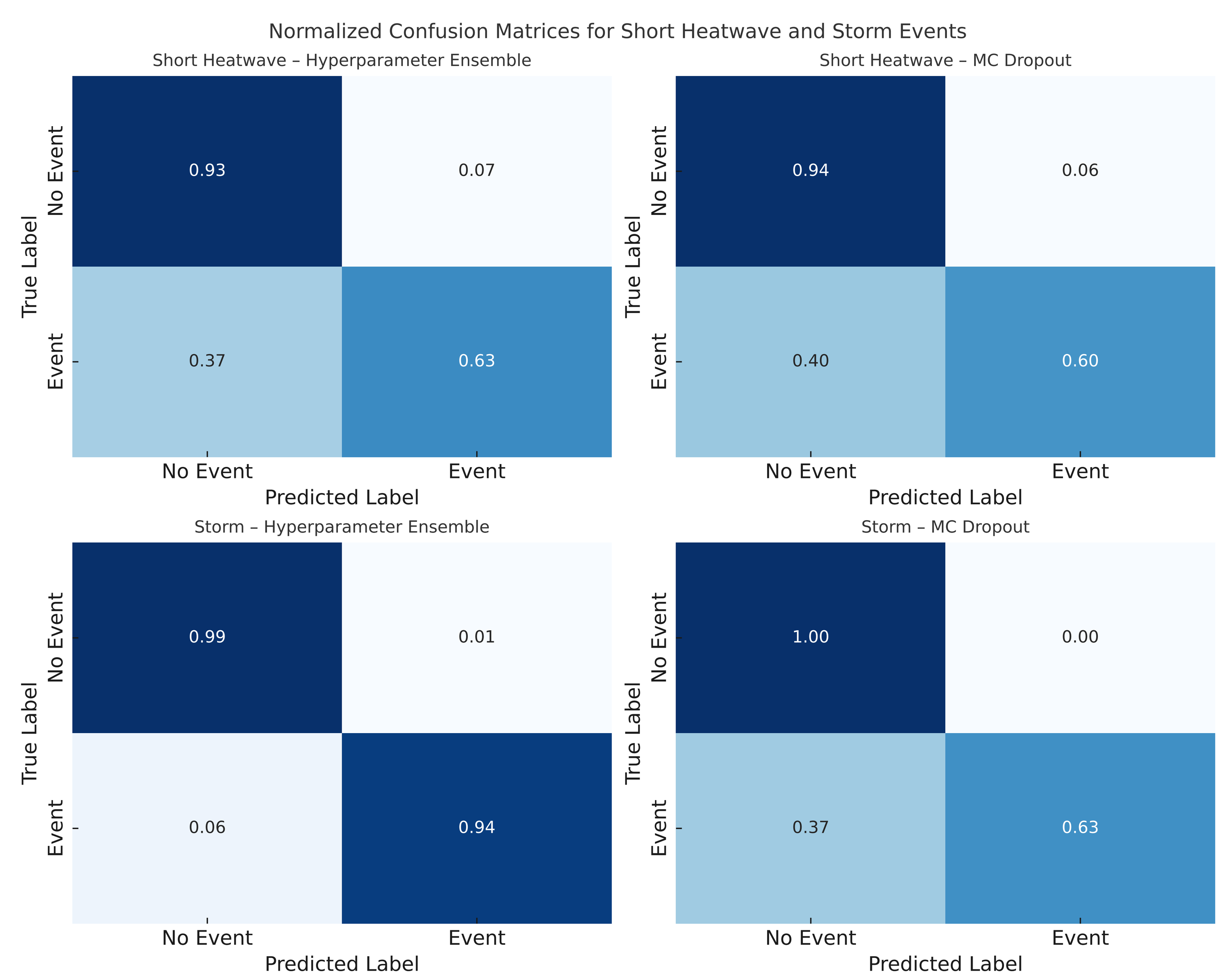

Figure 4: Batchwise extrapolation using PatchTSMixer. The variables were normalized in order to enhance the visualization.

References

[1] V. Ekambaram et al. "TTMs: Fast Multi-level Tiny Time Mixers for Improved Zero-shot and Few-shot Forecasting of Multivariate Time Series". In: *arXiv preprint arXiv:2401.03955* (2024).

[2] Y. Gal and Z. Ghahramani. "Dropout as a bayesian approximation: Representing model uncertainty in deep learning". In: *international conference on machine learning*. PMLR. 2016, pp. 1050–1059.

[3] *Instituto Nacional de Meteorologia, INMET*. https://portal.inmet.gov.br/. Acessed: 2025-01-03.

[4] *Instituto Nacional de Meteorologia, INMET, Dados históricos*. https://portal.inmet.gov.br/dadoshistoricos. Acessed: 2025-01-03.

[5] R. Lam et al. "Learning skillful medium-range global weather forecasting". In: *Science* 382.6677 (2023), pp. 1416–1421.

[6] J. Pathak et al. "Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators". In: *arXiv preprint arXiv:2202.11214* (2022).

[7] Y. Q. Sun et al. "Can AI weather models predict out-of-distribution gray swan tropical cyclones?" In: *arXiv preprint arXiv:2410.14932* (2024).

[8] D. Szwarcman et al. "Prithvi-EO-2.0: A Versatile Multi-Temporal Foundation Model for Earth Observation Applications". In: *arXiv preprint arXiv:2412.02732* (2024).