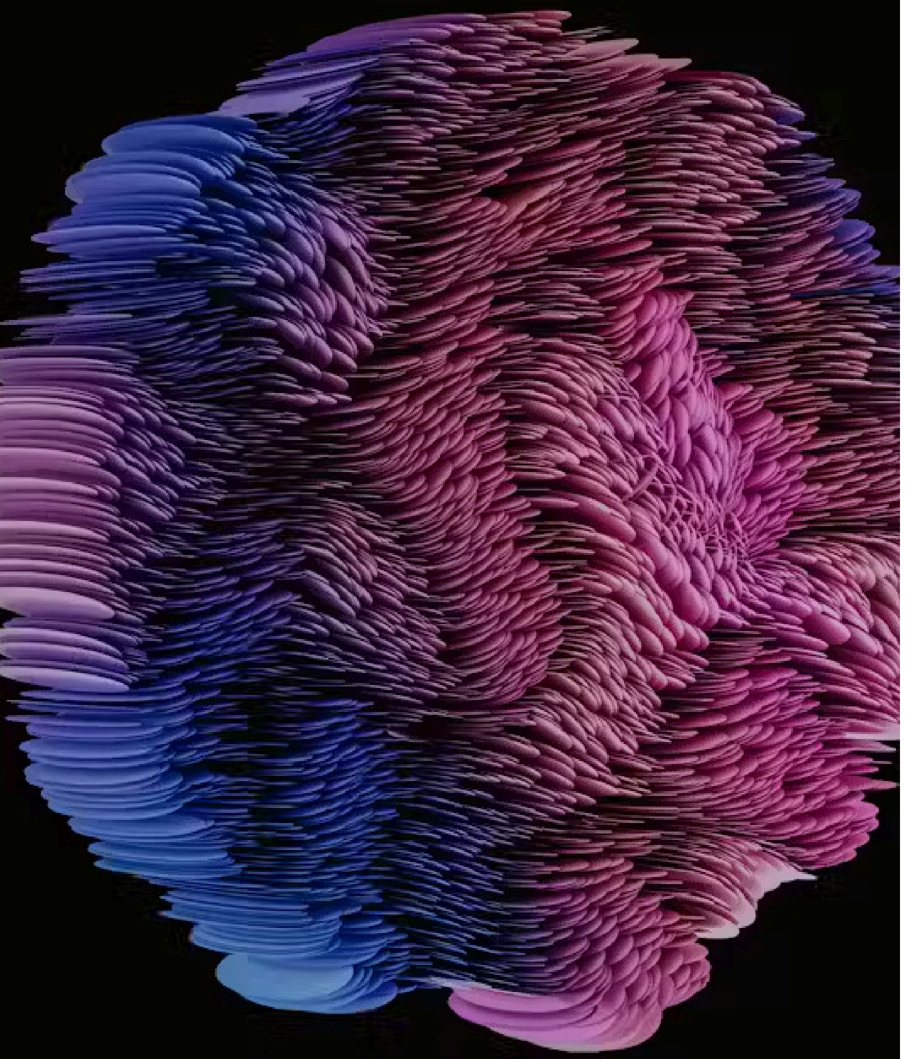


PJDSC 2025

Project Heat Resilience PH

A Reinforcement Learning Approach towards Heat Index Forecast's Adaptive Model Selection (HI'FAMS)

Team Name: PROXIMA LABS

Ducay, Harvey Lemuel Ogao

Empeño, Jhon Kheil Cymon Tomarong

Monfero, John Benedict Angeles

Executive Summary: Improving Heat Index Forecasts

The Philippines, especially Manila, is highly vulnerable to extreme heat events. Project Heat Resilience PH advances a **data-driven forecasting approach** to strengthen heat-risk preparedness.

Our focus is on **improving next-day Heat Index (HI) forecasts**, as accurate predictions are foundational to timely decisions by education and public-health authorities. The effectiveness of static heat-index thresholds depends on the **quality of the forecasts** that trigger them.

Reframing the Problem

We use **Adaptive Model Selection (AMS)**: the system selects the **one best predictive model** from a portfolio (the "model zoo") each day based on context (meteorological features, season, recent errors).

Contextual Bandit Implementation

We use lightweight, data-efficient policies like LinUCB and Contextual Thompson Sampling, which learn from daily outcomes without requiring a full simulation of real-world interventions.

The Differential Reward Mechanism

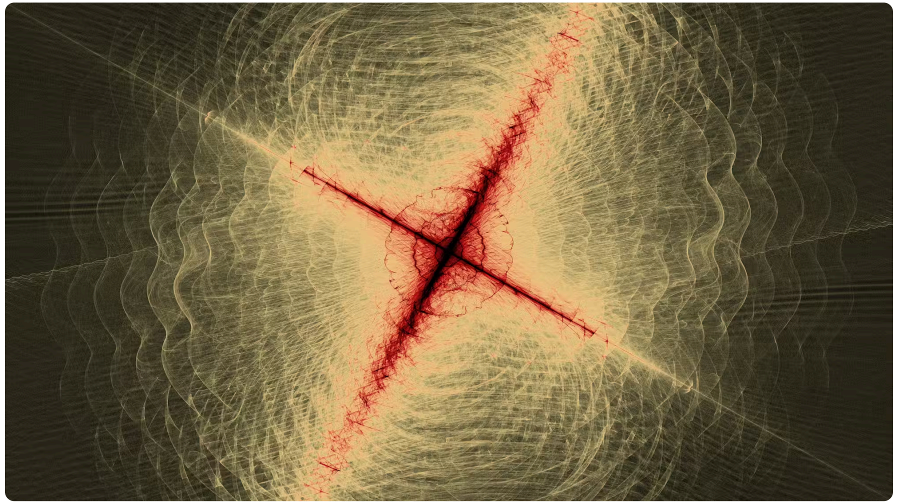
The learning signal for our Adaptive Model Selection (AMS) is a **differential reward** that compares the chosen model's error against a **Best-in-Class Benchmark**.

$$R\_t + 1 = (y\_t + 1 - \hat{y}^{a^*}_{t+1})^2 - \min_{m \in M}(y\_t + 1 - \hat{y}\_t + 1^m)^2$$

This formulation stabilizes learning, remains computationally tractable, and directly targets **predictive utility**. Positive rewards indicate the chosen model improved over the benchmark.

Pipeline Overview

- Built on daily Open-Meteo inputs and historical observations for Metro Manila.
- Runs all candidate models, publishes only the selected forecast.
- Updates the policy when the actual HI truth arrives the next day.



The system aims to reduce forecast errors precisely when decisions matter most, improving information for decision-makers (schools, LGUs, health agencies).

Background: The Challenge of Extreme Heat

Manila, one of the world's densest cities, faces amplified heatwaves due to the **urban heat island** effect, driving the Heat Index (HI) to dangerous levels that threaten public health, especially for students.

> Authorities rely on class suspensions to mitigate heat-related risks, but the decision framework is often **reactive**.

Static Thresholds

Policies hinge on static HI thresholds (e.g., "Danger"), which lack dynamism and do not account for expected duration or school-day context.

Forecast Uncertainty

Suspensions may be announced **too late** or enacted **unnecessarily**, disrupting instruction without proportional health benefit. Accuracy is key.

Project Heat Resilience PH: A Shift in Approach

Project Heat Resilience PH shifts from control-oriented reinforcement learning to a **forecast-improvement approach** grounded in Adaptive Model Selection (AMS).

Forecasting Model

System learns **which forecasting model to trust each day** for next-day HI.

Contextual Bandit

Draws on daily meteorological context to choose a single forecaster from a curated **model portfolio**.

Differential Reward

Measures improvement over a Best-in-Class benchmark, enabling stable, data-efficient updates.

By reducing forecast error precisely where it matters most, AMS strengthens the informational foundation for decision-makers, supporting more **timely, proportionate, and defensible** responses to extreme heat.

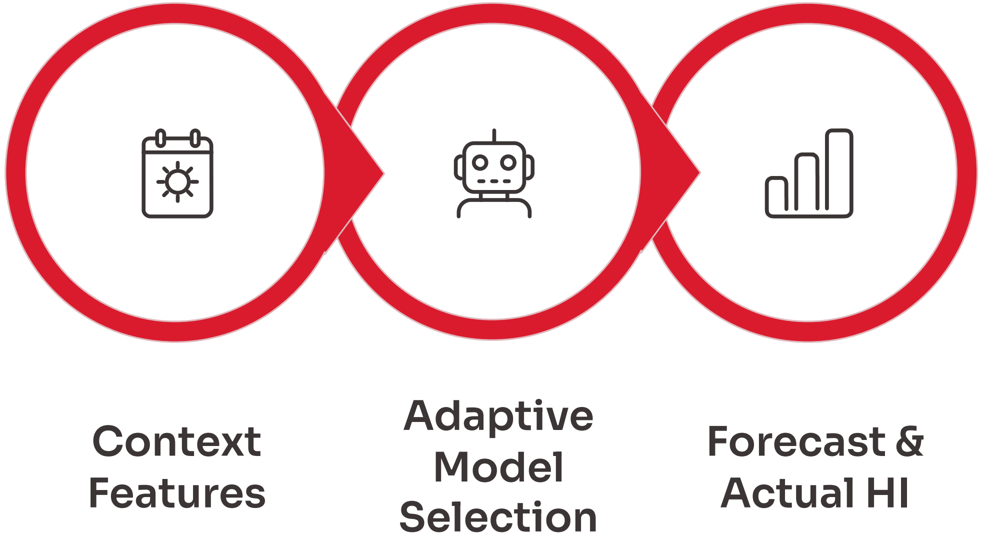
The Core Problem: Minimizing Forecast Loss

Philippine education authorities face a dynamic trade-off: minimizing student health risk versus minimizing educational disruption. Our research addresses the information bottleneck: improving next-day Heat Index (HI) forecasts via Adaptive Model Selection (AMS).

> Bad forecasts—especially near critical thresholds (e.g., 42 °C, 46 °C)—cause costly miscalls (either late/insufficient action or unnecessary suspensions).

Our objective is to:

Pick one model each day from a fixed portfolio to minimize expected, threshold-aware forecast loss (or, equivalently, maximize differential improvement over a Best-in-Class benchmark).



Project Objectives: Building the AMS Pipeline

The primary objectives reflect the shift to Adaptive Model Selection (AMS) for next-day Heat Index (HI) forecasting:

1 Design the AMS Pipeline

Formalize HI prediction as a **contextual bandit** that selects one forecaster from a fixed **model portfolio** using day t-1 context and recent error signals.

2 Engineer the Differential Reward

Implement the benchmark-referenced reward to directly incentivize improvement over a Best-in-Class model, with an optional threshold-aware extension.

3 Develop and Train Selection Policies

Implement LinUCB and Contextual Thompson Sampling (CTS) for daily updating, including a safe fallback (baseline guard) and deterministic tie-breakers.

4 Build a Reproducible Data Stack

Integrate **Open-Meteo** feature ingestion, model-zoo inference, and policy selection using **Supabase** (data/logging) and **Vercel** (scheduled jobs).

Scope and Limitations

The project maintains a clear focus and defined boundaries to ensure a robust and testable foundation for heat-risk response.

Scope

Geographical & Forecasting Focus

Manila City, focusing on next-day **Heat Index (HI)** point forecasting using Open-Meteo inputs.

Decision Space

Adaptive Model Selection (AMS)—selecting **one** forecaster per day from the model zoo; no direct control of class suspensions.

Data & Operations

Daily batch pipeline (ingest → predict → select → publish; verify → learn) with logging and versioning; forecast exposed via public API.

Limitations

Portfolio & Benchmark Dependence

Results are bounded by the quality of the available models; performance relies on the chosen Best-in-Class benchmark.

Data Dependency

Performance depends on Open-Meteo inputs and observation quality; outages can degrade accuracy.

Governance Boundary

The system **does not** automate class suspensions; final policy decisions remain with DepEd/LGUs.

Methodology: The Contextual Bandit MDP

The methodology utilizes a model-selection Reinforcement Learning approach, casting the forecasting task as a daily, single-step decision process.

State Space (S)

Comprehensive view of day-t conditions: Calendar/Seasonality (Day of Year), and Meteorological Summaries (e.g., Temperature, Humidity, Wind Speed) from Open-Meteo.

Action Space (A)

The agent chooses from **all candidate forecasting models** available in the portfolio (the "model zoo").

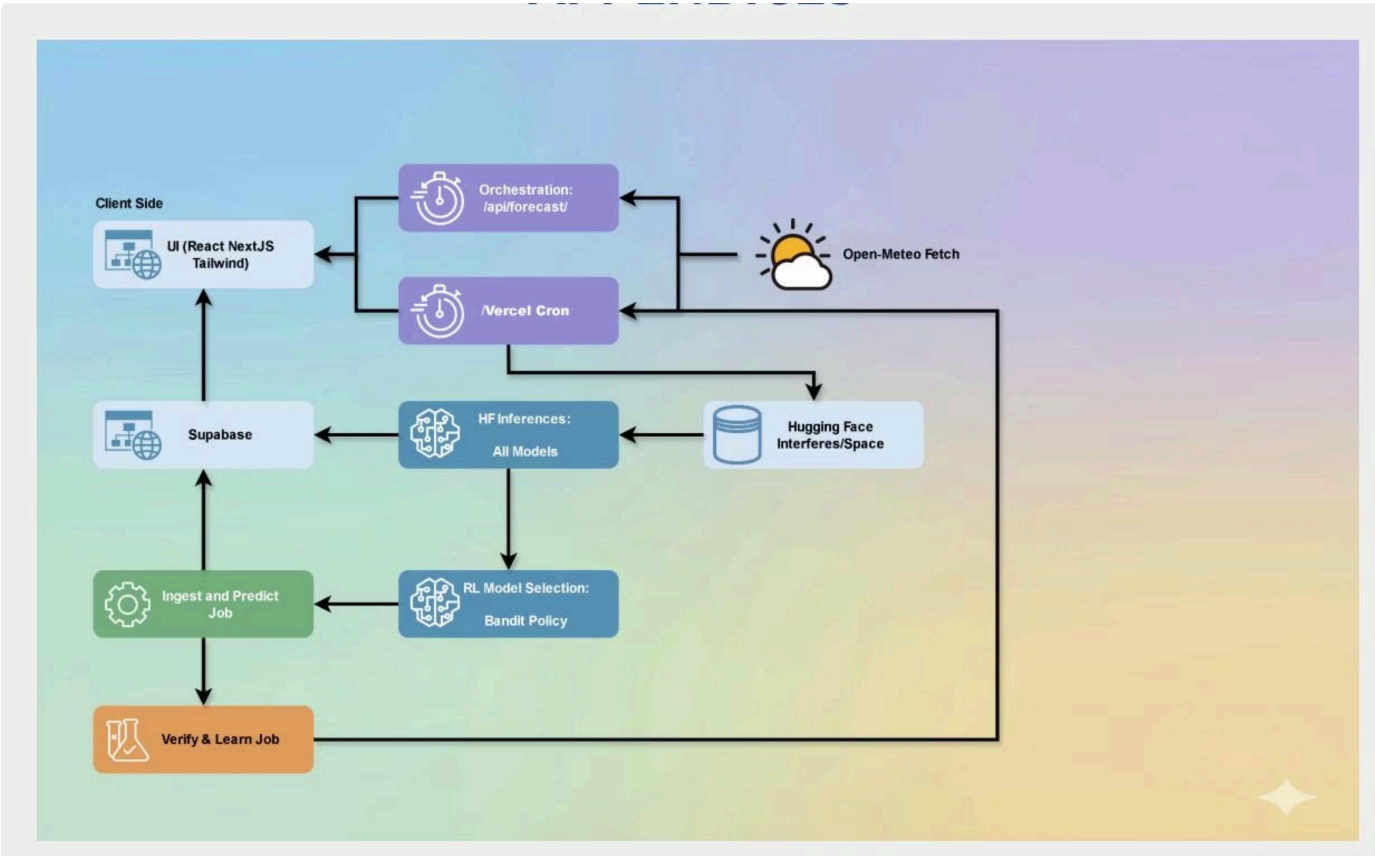
Reward Function (R)

Differential reward computed after the next day's observation is available, comparing the chosen model's MSE against the Best-in-Class Benchmark's MSE.

The environment is implemented with two daily jobs: **Ingest & Predict** (select and publish forecast) and **Verify & Learn** (compute reward and update policy).

Operational Architecture and Transparency

The prototype implementation uses a modern, serverless stack to ensure reliability and transparency for stakeholders.



Compute & Data

Vercel Cron orchestrates jobs, **Hugging Face** handles model inference, and **Supabase Postgres** manages data and logs.

Analytical Dashboard

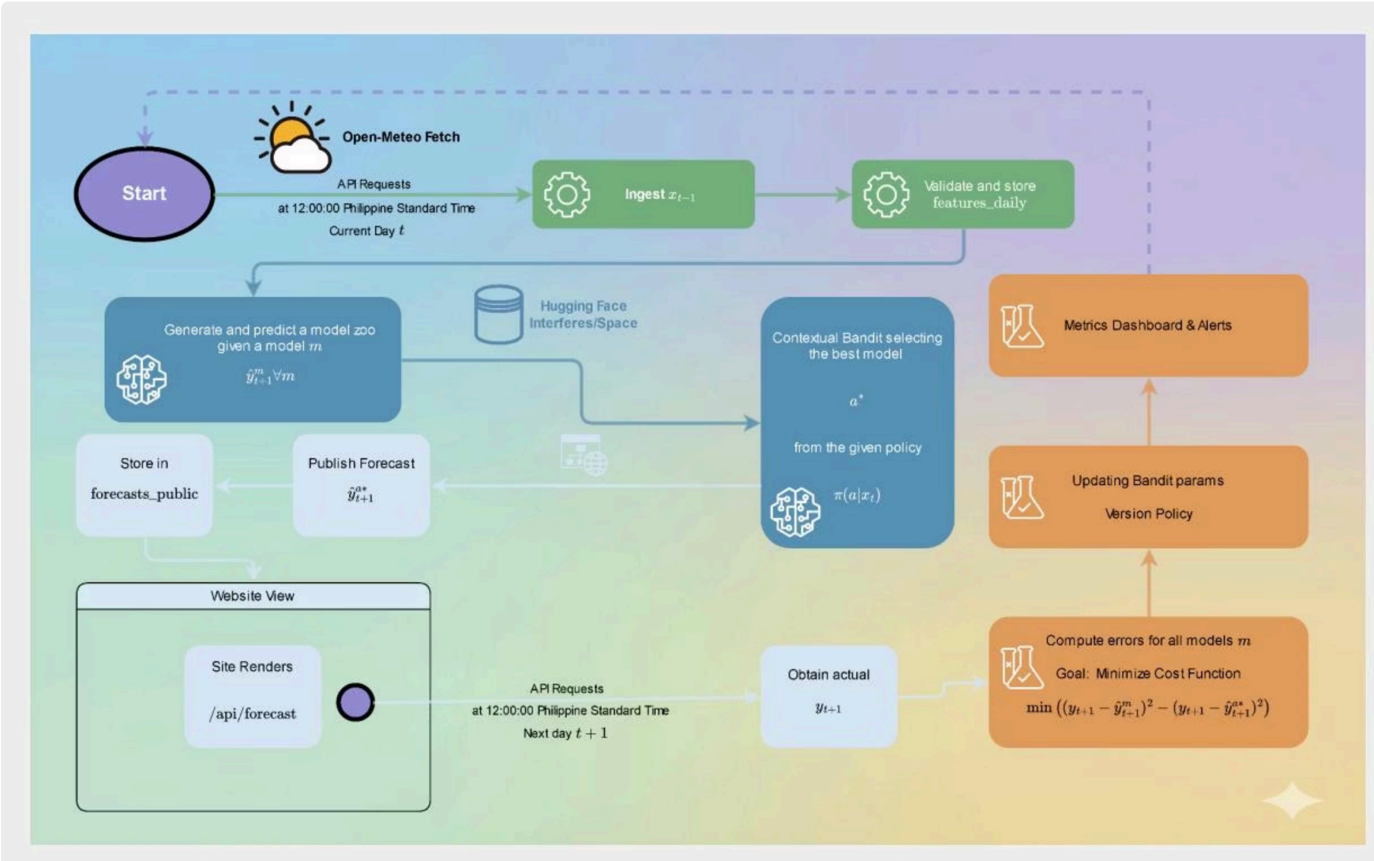
Reports **selected forecasts**, per-model error leaderboards, **regret trends**, and job health to support stakeholder interpretation and adoption.

Guardrail

If the system fails, a guardrail publishes a safe baseline forecast to ensure the site never goes dark, maintaining operational stability.

Daily Dynamics of the Website

The prototype implementation uses a modern, serverless stack to ensure reliability and transparency for stakeholders.



Compute & Data

Vercel Cron orchestrates jobs, **Hugging Face** handles model inference, and **Supabase Postgres** manages data and logs.

Analytical Dashboard

Reports **selected forecasts**, per-model error leaderboards, **regret trends**, and job health to support stakeholder interpretation and adoption.

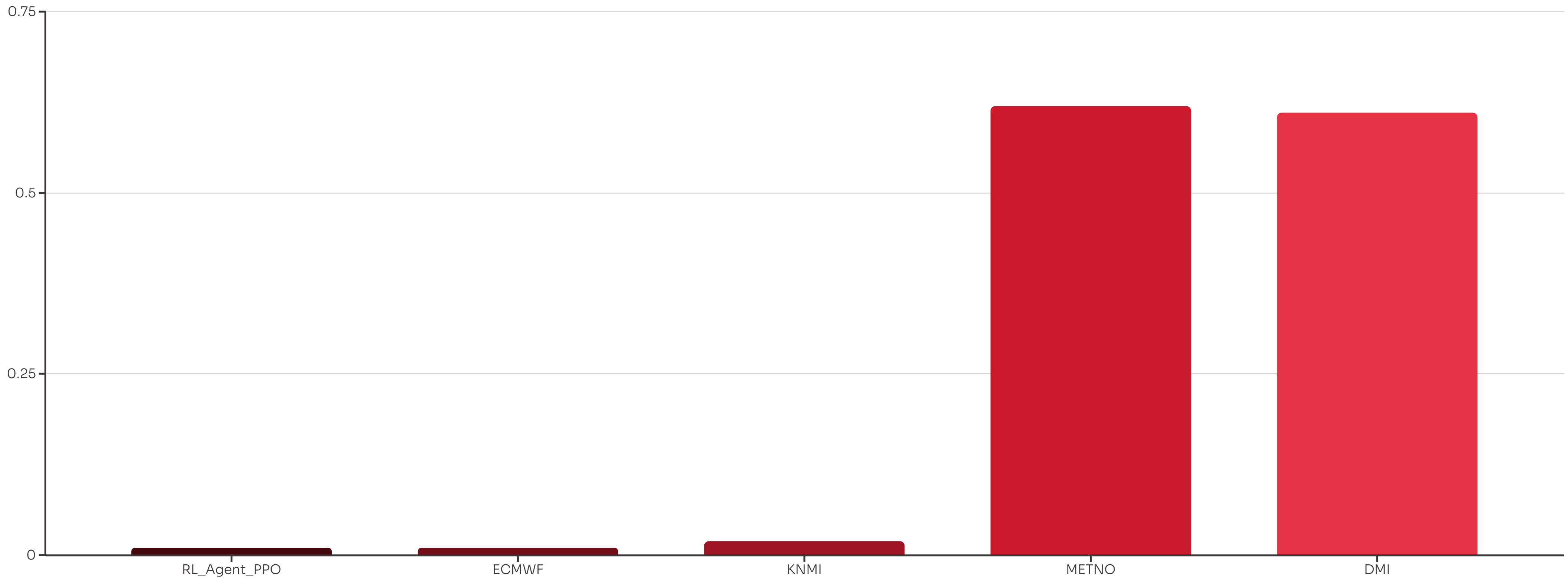
Guardrail

If the system fails, a guardrail publishes a safe baseline forecast to ensure the site never goes dark, maintaining operational stability.

Models Results and Discussion

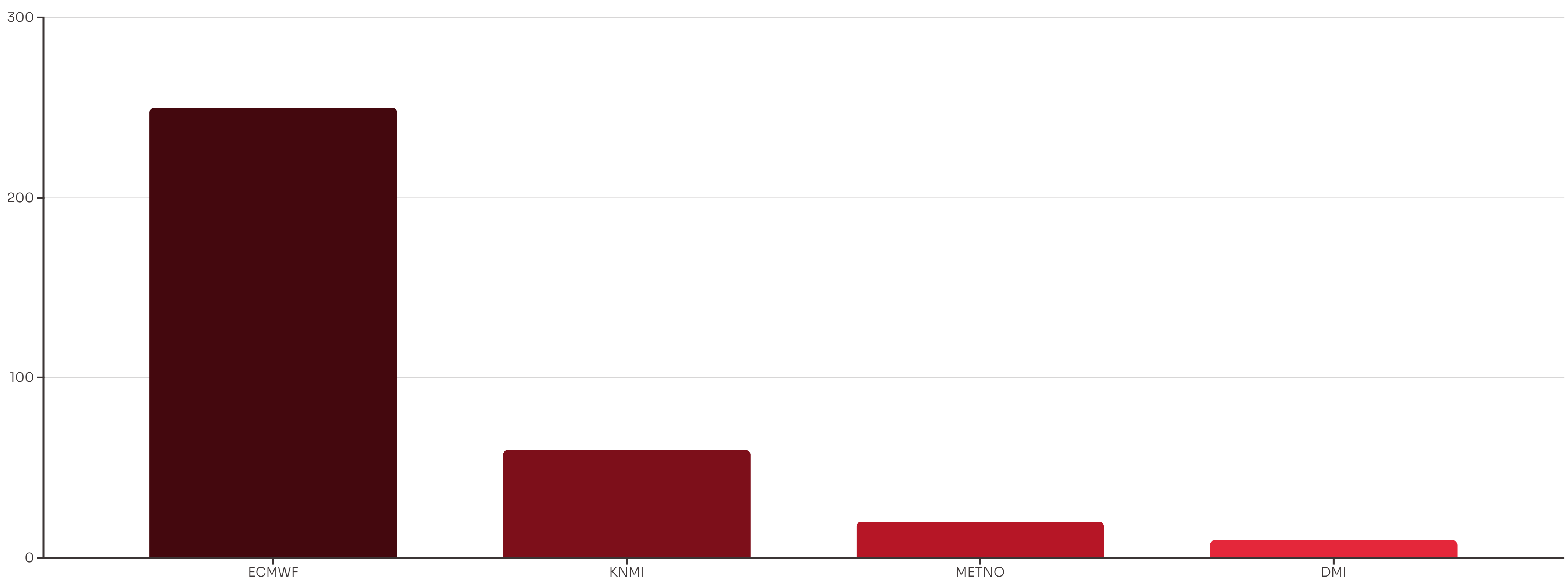
Our analysis of next-day Heat Index (HI) forecasts for Manila City in 2024 reveals the effectiveness of the Reinforcement Learning agent (RL_Agent_PPO) in selecting optimal models from our portfolio.

Forecast Performance: Mean Squared Error (MSE)



The RL_Agent_PPO demonstrates an MSE of approximately 0.01, performing on par with the best individual forecasting models and significantly outperforming the weaker models (METNO/DMI), which showed MSEs around 0.62.

Model Selection Frequency



The agent predominantly selected the ECMWF model, using it on approximately 250 days. KNMI was chosen on about 60 days, while METNO and DMI were rarely selected, indicating their lower reliability within the 2024 dataset.

Error Distributions and Seasonal Behavior

Analysis of daily absolute-error distributions showed that the RL_Agent_PPO and the better models (ECMWF, KNMI) exhibit tight, low-variance error ranges. In contrast, METNO and DMI frequently produced significant outliers, with absolute errors extending up to 1.8. Seasonally, ECMWF was a consistent choice throughout the year, whereas KNMI demonstrated a niche, being preferred during the mid-to-late year (Day of Year 150-300). METNO and DMI were only sparsely selected late in the year, reinforcing their role as special-case options.

ECMWF: The Workhorse

This model provided stable performance with the lowest error variance, making it the agent's most frequent and reliable choice across the year.

KNMI: Seasonal Niche

Competitive in mean error, but with slightly higher variance. It found its optimal application during specific periods (mid-to-late year), suggesting a seasonal advantage.

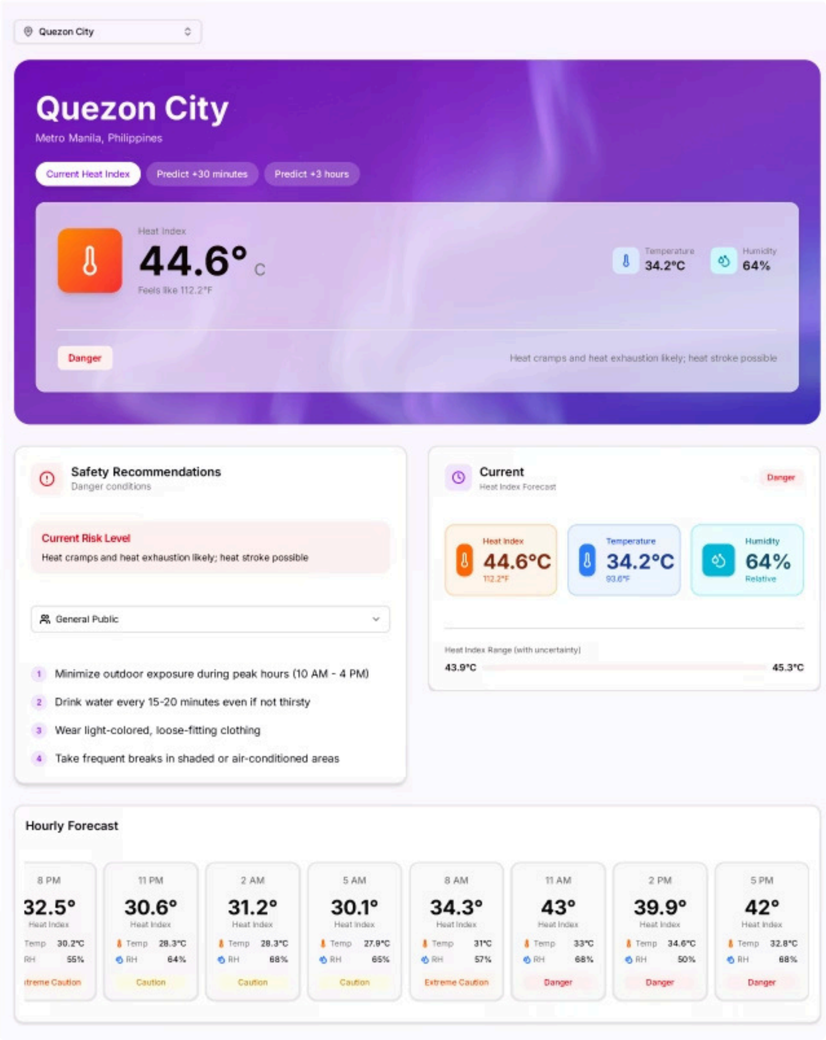
METNO & DMI: Avoided Biases

Characterized by high MSE and heavy error tails, these models were largely avoided by the agent, indicating systematic biases or poorer HI derivation for Manila in 2024.

The PPO selector effectively acts as a risk-aware portfolio manager, consistently prioritizing high-performing models and leveraging context-specific advantages to maximize improvement over the benchmark.

References

- Agrawal, S., & Goyal, N. (2013). Thompson sampling for contextual bandits with linear payoffs. Proceedings of the 30th International Conference on Machine Learning (ICML). PMLR.
- Agarwal, A., Luo, H., Neyshabur, B., & Schapire, R. E. (2017). Corralling a band of bandit algorithms. Proceedings of the 30th Conference on Learning Theory (COLT). PMLR.
- Auer, P., Cesa-Bianchi, N., Freund, Y., & Schapire, R. E. (2002). The nonstochastic multiarmed bandit problem. SIAM Journal on Computing, 32(1), 48–77.
- Ben Bouallègue, Z., Magnusson, L., Bremnes, J. B., et al. (2024). The rise of data-driven weather forecasting: A first statistical assessment of ML-based forecasts in an operational-like context. Bulletin of the American Meteorological Society, 105(6).
- Bi, K., Chen, Y., Deng, X., et al. (2023). Accurate medium-range global weather forecasting with 3D neural networks. Nature.
- de Burgh-Day, C. O., Pook, M. J., Jakob, C., & May, P. T. (2023). Machine learning for numerical weather and climate modelling. Geoscientific Model Development, 16, 6433–6477.
- Dudík, M., Langford, J., & Li, L. (2014). Doubly robust policy evaluation and optimization. Statistical Science, 29(4), 485–511.
- Foster, D. J., Krishnamurthy, A., & Luo, H. (2019). Model selection for contextual bandits. Advances in Neural Information Processing Systems (NeurIPS).
- García, J., & Fernández, F. (2015). A comprehensive survey on safe reinforcement learning. Journal of Machine Learning Research, 16, 1437–1480.
- Iyengar, G. N. (2005). Robust dynamic programming. Mathematics of Operations Research, 30(2), 257–280.
- Jiang, N., & Li, L. (2016). Doubly robust off-policy value evaluation for reinforcement learning. Proceedings of the 33rd International Conference on Machine Learning (ICML). PMLR.
- Kassraie, P., Emmenegger, N., Krause, A., & Pacchiano, A. (2023). Anytime model selection in linear bandits. Advances in Neural Information Processing Systems (NeurIPS).
- Kumar, A., Zhou, A., Tucker, G., & Levine, S. (2020). Conservative Q-learning for offline reinforcement learning. Advances in Neural Information Processing Systems (NeurIPS).
- Lam, R., Das, A., Nguyen, T., et al. (2023). Learning skillful medium-range global weather forecasting. Science. (GraphCast).
- Laroche, R., Trichelair, P., & Tachet des Combes, R. (2019). Safe policy improvement with baseline bootstrapping (SPIBB). Proceedings of the 36th International Conference on Machine Learning (ICML). PMLR.
- Li, L., Chu, W., Langford, J., & Schapire, R. E. (2010). A contextual-bandit approach to personalized news article recommendation. arXiv preprint arXiv:1003.0146.
- Nilim, A., & El Ghaoui, L. (2005). Robust control of Markov decision processes with uncertain transition matrices. Operations Research, 53(5), 780–798.
- Pathak, J., Subramanian, S., Harrington, P., et al. (2022). FourCastNet: A global data-driven high-resolution weather model using adaptive Fourier neural operators. arXiv preprint arXiv:2202.11214.
- Raftery, A. E., Kárný, M., & Ettler, P. (2010). Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill. Technometrics, 52(1), 52–66.
- Timmermann, A. (2006). Forecast combinations. Handbook of Economic Forecasting, Vol. 1. Elsevier.
- Watson-Parris, D. (2021). Machine learning for weather and climate are worlds apart? Philosophical Transactions of the Royal Society A, 379(2194), 20200098.
- Open-Meteo. (n.d.). Weather forecast API documentation. (Accessed 2025).
- Hugging Face. (n.d.). Inference Endpoints API reference. (Accessed 2025).
- Supabase. (n.d.). Supabase Docs: Database & REST API references. (Accessed 2025).
- Department of Education (DepEd), Philippines. (2022). DepEd Order No. 037, s. 2022 – Guidelines on the cancellation or suspension of classes and work in schools in the event of natural disasters, power outages/power interruptions, and other calamities.
- Intergovernmental Panel on Climate Change (IPCC). (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report. Cambridge University Press.
- Philippine Atmospheric, Geophysical and Astronomical Administration (PAGASA). (n.d.). Heat index information. (Accessed 2025).
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347.
- Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press.



Heat Index Forecast Site

This is the Minimum Viable Website Sample of the Proxima Labs as of October 2025