A Cyber-Physical Architecture for Crop Recommendation: Integrating Climate Prediction (XGBoost), IoT Sensing, and Multi-Criteria Decision-Making

Vitor Hugo Barbosa Melo[1],André Vinicius Neves Alves[1],
Saulo Mattheus Ribeiro de Oliveira[1], Waldemiro José Assis Gomes Negreiros[1],
Filipe Corrêa da Silva[1], Nilton Rodolfo Nascimento Melo Rodrigues[1],
Diego Lisboa Cardoso[1], Marcos Cesar da Rocha Seruffo[1]

[1]Institute of Technology – Federal University of Pará (UFPA)
66075-110 – Belém – PA – Brazil
{vitor.melo,andre.neves.alves}@itec.ufpa.br
{saulomattheus,filipecorrea13976}@gmail.com
waldemiro.negreiros@ifpa.edu.br
{niltonrodolfo,diego,seruffo}@ufpa.br

Abstract. *This work proposes a unified cyber-physical framework to evaluate the agronomic suitability of Açaí, Cocoa, and Soybean in the state of Pará, Brazil. The methodology integrates ERA5 climate data, XGBoost forecasting, and a conceptual IoT node for local microclimatic calibration using Edge Computing. Climate projections for 2026–2028 fed a modified SAW (Simple Additive Weighting) decision engine. For Soybean, the system identifies optimal planting windows (January–April), while for Cocoa and Açaí it supports year-round thermal and water stress monitoring. Simulations indicated high suitability during the rainy season and severe penalties in dry months (July–October), highlighting Cocoa's vulnerability to heat stress and Soybean's planting restrictions. Results validate the software layer and support future IoT field deployment.*

1. Introduction

The northeastern and southeastern mesoregions of Pará have established a strategic triad in the national agribusiness sector through the cultivation of soybeans, cocoa, and açaí. According to the 2023 Pará GDP Report [FAPESPA 2024], agriculture and livestock expanded their share to 12.9% of the state's value added. However, increasing climate variability threatens the sector's production efficiency, highlighting the limitations of traditional planning methodologies. Currently, isolated monitoring tools fall short by failing to integrate global climate forecasts with the specific water and nutrient conditions of each field.

To address this technological gap, this study proposes a hybrid and modular cyber-physical architecture for agricultural decision support, structured around three fundamental pillars: (i) predictive macroclimatic modeling, using historical time series from the ERA5 reanalysis and the XGBoost algorithm to project medium-term scenarios (2026–2028) [Oses et al. 2020, Chakri et al. 2025]; (ii) microclimatic calibration via an IoT tower for real-time monitoring of atmospheric and edaphic variables (including NPK) [Garg et al. 2021]; and (iii) a multi-criteria decision engine based on the modified SAW (Simple Additive Weighting) method [Fishburn 1967]. The system aims to translate these projections into direct agronomic suitability indices for açaí, cocoa, and soybeans.

The scope of this WIP is limited to the computational validation of the architecture: the design of the IoT sensor node (ESP32/LoRa) [Nawaz et al. 2025], the Edge Computing architecture using Raspberry Pi [Gong et al. 2025], and the algorithmic PoC integrating XGBoost with the SAW engine.

2. Methodology

The proposed solution is based on a modular cyber-physical architecture designed to integrate global climate modeling with local edaphoclimatic sensing. **Figure 1** illustrates the data flow of this methodological pipeline, which is structured across four interlinked zones: (i) **Zone A (Cyber Layer)** for macroclimatic predictive modeling using XGBoost; (ii) **Zone B (Physical Layer)** for local data acquisition via IoT sensors; (iii) **Zone C (Edge Layer and Blending)** for data fusion and multi-criteria decision-making (SAW); and (iv) **Zone D (Results and Outputs)** for generating the final agronomic suitability reports.

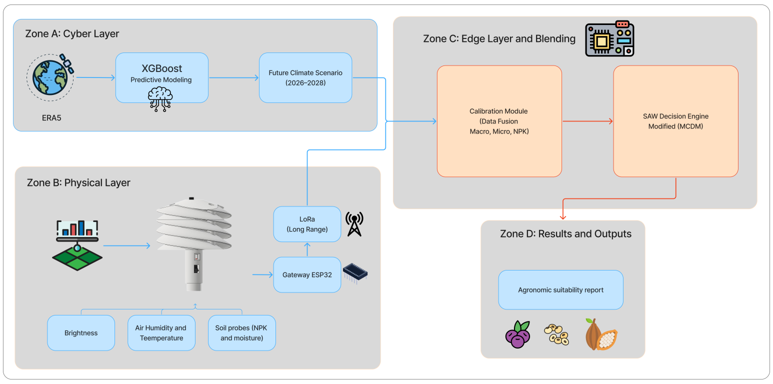

Figure 1. Proposed cyber-physical architecture: integration workflow between macroclimatic predictive modeling (ERA5/XGBoost), local sensing via an IoT tower, and multi-criteria decision-making (SAW) processed at the edge

2.1. Data Collection (Macro and Micro)

At the macroclimatic scale, daily time series from the ERA5 reanalysis (ECMWF) were used, covering the period from 2016 to 2025 for the mesoregions of Pará. The variables considered include maximum air temperature (*t2m_max*), cumulative precipitation (*tp_soma_mm*), and soil moisture (*swvl1_media*), selected because they define the thresholds for heat stress and water availability, critical factors for the phenology of açaí, cocoa, and soybeans [Abreu and Barbosa 2021, Ministério da Agricultura e Pecuária 2024, Serra and Sodré 2021]. For the microclimatic scale, we propose an IoT tower equipped with atmospheric and edaphic sensors, based on ESP32. The data collected *in loco* will be transmitted to a gateway (Raspberry Pi), where data fusion with global information will occur, characterizing an Edge Computing architecture.

2.2. Predictive Modeling

The analytical core uses the XGBoost algorithm to forecast climate variables. The model was trained on ERA5 historical data using feature engineering with time lags (t-1, t-7, and t-30 days) and 30-day sliding-window statistics as predictors, with hyperparameters tuned via 5-fold cross-validation (learning rate = 0.05, max depth = 6, n_estimators = 500). During validation, the model achieved a Mean Absolute Error (MAE) of 0.29°C and a Root Mean Square Error (RMSE) of 0.47°C for mean temperature, demonstrating high fidelity to local tropical seasonality. Inference is autoregressive: generated forecasts are iteratively fed back into the model to project the 2026–2028 horizon, with the known trade-off of progressive bias accumulation over longer forecastings horizons.

2.3. Multi-Criteria Inference

The final layer translates the forecasts into monthly recommendations via a modified SAW method. The base score assigns equal weights (33.3% each) to precipitation, soil moisture, and heat stress (days above 30°C), a simplifying assumption in the absence of field-calibrated data, with dynamic weight adaptation planned for the IoT deployment phase. Stochastic variability of ±3% is applied and the score is capped at 9.6. For soybeans, a 75% penalty is applied outside the January–April planting window, consistent with ZARC guidelines [Abreu and Barbosa 2021]. Cocoa receives an additional penalty when $T_{\max} > 30°\text{C}$ coincides with low humidity, reflecting documented heat-drought sensitivity [Serra and Sodré 2021, Nangoi et al. 2006]. Açaí shows greater seasonal resilience per official ZARC [Ministério da Agricultura e Pecuária 2024]. The system generates realistic predictive suitability rankings for all three crops.

3. Preliminary Results and Discussion

Figure 2 shows the daily forecast horizon (2026–2028). The visualization highlights the effectiveness of XGBoost in projecting local seasonality: peaks in precipitation and humidity in the first half of the year, contrasting with a severe decline in rainfall and extreme heat (temperatures above 30°C) between July and October.

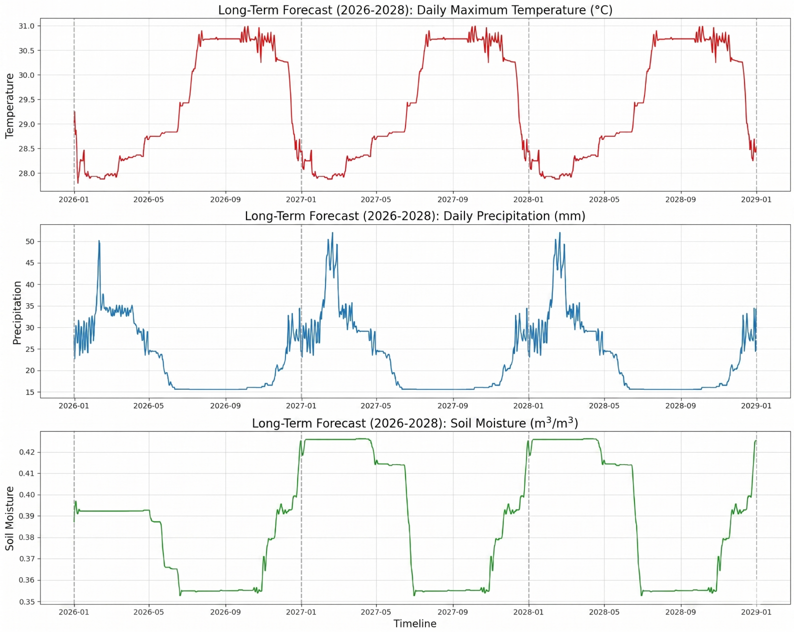

Figure 2. Daily macroclimatic projections (2026–2028) generated by the XGBoost model for the variables of maximum temperature, precipitation, and soil moisture

The SAW method was adopted for its computational efficiency and linearly interpretable fitness index [Fishburn 1967]. Applied to 2026 projections, the results align with the distinct agronomic roles of each crop. Soybean scored above 8.0 during the January–April window and below 2.0 outside it, directly supporting planting-window recommendations. Cocoa reached near-maximum scores during the rainy season but dropped critically to 1.2 in August — indicating thermal stress risk and signaling growers to adopt irrigation or shading strategies. Açaí maintained a minimum score of 3.8 throughout the dry season, confirming its resilience and the system's utility for continuous stress monitoring.

4. Conclusion and Future Work

The PoC validated the effectiveness of the XGBoost-SAW integration in generating agronomic predictive scores. However, software-only implementation has limitations: the autoregressive approach accumulates biases over longer horizons, and the static equal weights of the SAW lack edaphic calibration, both limitations explicitly motivate the IoT deployment phase [Nawaz et al. 2025, Gong et al. 2025]. To overcome this barrier, future work will focus on the physical deployment of the IoT tower. On-site data collection (especially NPK) will allow the gateway to calibrate forecasts at the Edge, bridging computational modeling with field reality.

References

Abreu, A. L. and Barbosa, I. C. (2021). Desenvolvimento da soja para determinação do período de plantio em tomé-açu/pa. Master's thesis, UFRA, Tomé-Açu, PA.

Chakri, A. et al. (2025). Advanced forecasting using era5 and machine learning in tropical regions. *Agricultural and Forest Meteorology*, 320:108900.

FAPESPA (2024). Relatório do PIB do pará – 2023. Technical report, Fundação Amazônia de Amparo a Estudos e Pesquisas, Belém, PA.

Fishburn, P. C. (1967). Additive utilities with incomplete product sets. *Operations Research*, 15(3):537–542.

Garg, R. et al. (2021). Iot-based precision agriculture: A comprehensive review. *IEEE Internet of Things Journal*, 8(15):11850–11875.

Gong, L. et al. (2025). Edge computing architectures for real-time agricultural monitoring. *IEEE Transactions on AgriFood Electronics*, 3(2):112–125.

Ministério da Agricultura e Pecuária (2024). Portaria spa/mapa nº 5/2024: Zarc para a cultura do açaí irrigado. Diário Oficial da União, Seção 1, p. 7. Brasília, DF.

Nangoi, I. I., Luiz, A. J. B., and Sentelhas, P. C. (2006). Determinação dos parâmetros de risco climático para a cultura do cacau. ESALQ/USP.

Nawaz, M. et al. (2025). Iot and ai for smart agriculture in resource-constrained environments. *Discover Agriculture*, 2(1):15.

Oses, N. et al. (2020). Machine learning applications for agricultural planning using era5 reanalysis. *Computers and Electronics in Agriculture*, 175:105500.

Serra, R. and Sodré, G. A. (2021). Manual do cacauicultor: perguntas e respostas. Technical Report Boletim Técnico 221, Embrapa.