

AI AND IOT-ENABLED SMART AGRICULTURE: PREDICTIVE MODELLING FOR CROP YIELD AND RESOURCE OPTIMIZATION

DR. SHIKHA SINGH¹, A ADITHYA KASHYAP², DR.S. VIJAYAKUMAR³, DR.R. NAVEENA BHARGAVI⁴, DR. SYED.SHAMEEM⁵, DR. KONARI RAJASEKHAR⁶

¹ Department of computer Science & Engineering, Amity School of engineering & Technology, Lucknow Amity University, Uttar Pradesh, India.

² Assistant Professor, Department of CSE (AIML & IoT), VNRVJIET, Hyderabad, India.

³ Department of Computer Applications, Aditya University, Surampalem, India.

⁴ Professor, Department of EEE, CVR College of Engineering, Hyderabad, India.

⁵ Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation (KL Deemed university), Vaddeswaram, Guntur, Andhra Pradesh, India.

⁶ Assistant Professor, Department of ECE, Nadimpalli Satyanarayana Raju Institute of Technology, Sontyam, Andhra Pradesh, India.

Email: ¹ ssingh8@lko.amity.edu, ² adithyakashyap_a@vnrvjiет.in, ³ vijayakumars@adityauniversity.in, ⁴ rn.bhargavi@cvr.ac.in, ⁵ shameemsyed@kluniversity.in, ⁶ rajasekhar.ece@nsrit.edu.in

ABSTRACT

Sustainable food production under escalating climate variability, water scarcity, and resource constraints demands intelligent, data-driven decision-support systems for precision agriculture. Current deep learning methods have significant challenges with spatio-temporal representation, multimodal sensor fusion, adaptive optimization, model explainability, and feasibility of edge deployment. This paper introduces a novel end-to-end intelligent farming framework called AgriFormer-RL that combines a distributed IoT sensor network with high-resolution multimodal data capture, a physics-aware Graph Transformer encoder, which models spatial relations in the field with multi-head relational attention over dynamic spatio-temporal graphs, and a temporal foundation model (Chronos/TimeMixer) for long horizon crop yield forecasting with uncertainty quantification. A cross-attention multimodal fusion mechanism synergistically aggregates the satellite, meteorological and in-situ sensor representations and a Soft Actor-Critic reinforcement learning optimizer allows for adaptive, constraint-aware irrigation and fertilizer scheduling. Agro-explainability modules based on SHAP and attention mechanism offer decision attribution, which is interpretable by agronomists. Comprehensive experiments conducted over four Indian agricultural datasets - ICRISAT, IMD, MOSDAC, and India-WRIS - in addition to simulated LoRaWAN IoT streams are statistically superior to its baseline methods of Informer, PatchTST, and STGNN, with an improvement of 17.5% in RMSE, water savings of 31.4%, and fertilizer savings of 24.9%. The validation results using Wilcoxon signed-rank and Friedman tests prove that AgriFormer-RL is a scalable, interpretable, and resource-efficient precision agriculture paradigm.

Keywords: *Precision Agriculture, Graph Transformer, Temporal Foundation Model, Reinforcement Learning, Explainable AI, IoT, Edge Intelligence, Crop Yield Prediction, Resource Optimization, Spatio-Temporal Learning*

1. INTRODUCTION

Today, the world faces an unprecedented and unprecedented food security crisis: The world population will reach more than 9.7 billion by 2050,

climate change caused a new instability in agriculture and the progressive decline of reserves of freshwater and arable land. Withdrawals for agriculture make up around 70% of the total freshwater use worldwide, and 10–12% of

anthropogenic GHGs. These economically and environmentally unsustainable uses of the freshwater resource are supported by inefficient farming practices. In this context, the path towards data-driven precision agriculture, facilitated by the proliferation of IoT sensing, edge-cloud computing continua, and high levels of machine intelligence has become a technological must for at the same time enhance yield predictability and reduce resource spending [1] [2].

The number of inexpensive IoT devices, such as soil moisture sensors, pH meters, electrical conductivity (EC) probes, weather stations equipped with LoRaWAN connectivity, and satellite-based remote sensing platforms like ISRO MOSDAC, have led to an unprecedented amount of heterogeneously distributed and diverse agricultural data streams [3], [4]. Accurate crop yield forecasting and adaptive resource management, from crop to field, from field to region, or from region to area, requires architectures that can (a) model a complex dependency between geographically distributed field zones, (b) describe long-range temporal dynamics for seasonal and inter-annual timescales, (c) combine diverse modalities of observations with different statistical properties, (d) optimize irrigation and fertilization decisions under stochastic environment conditions, (e) provide interpretable predictions accessible by domain agronomists, and (f) operate in the latency and energy constraints of edge computing environments [5] [6] [7].

Current deep learning solutions to smart agriculture show systematic shortages in these aspects. CNNs and RNNs (LSTM and GRU variants) are limited in spatio-temporal expressivity because of fixed receptive fields and sequential processing bottlenecks [8] [9]. Although the standard Transformer architectures have been found to enhance the modelling of temporal dependency, they lack the explicit spatial relational inductive biases required to capture the dynamics between the different fields of a mixture. Graph Neural Networks (GNNs) overcome the spatial topology problem but are mainly used for static graph structures and fail to model dynamic graph topology relationships created by changes in field situations. Most of the current systems independently solve crop yield prediction and multi-resource optimization and can't be explained in a unified framework that is suitable for real-time deployment on edge devices.

To address these challenges, this paper presents the AgriFormer-RL framework, an intelligent farming system that combines four paradigm-changing AI techniques—graph-structured spatio-temporal learning, temporal foundation models,

reinforcement-learning resource optimization, and explainable AI—in a principled manner. Agricultural fields are represented as spatio-temporal graphs, and node-level multimodal feature tensors are processed by a physics-aware Graph Transformer (GT) encoder. A spatio-temporal graph is used to model the fields, where each node represents a specific field and the edges represent the temporal and spatial relationships between them. A pre-trained temporal foundation model (Chronos / TimeMixer) is used to forecast yields at the node level over long horizons. The proposed architecture also incorporates a cross-attention mechanism to fuse the different modalities and a Soft Actor-Critic (SAC) reinforcement learning algorithm to generate optimal, constraint-compliant irrigation and fertilizer policies. Post-hoc interpretability is achieved via SHAP values and attention maps, and the practical deployment feasibility is achieved by the deployment of TinyML optimised models on the edge.

The main Research Contributions of this work are summarized below:

A spatio-temporal graph construction method for heterogeneous agricultural fields equipped with IoT devices that considers the physics-informed edge weighting with respect to hydraulic conductivity of the soil, spatial proximity of the devices and the agronomic correlation matrices.

A novel Graph Transformer encoder based on adaptive multi-head relational attention, layer-wise uncertainty propagation for robust node-level representation learning and positional encodings integrating geographic and phenological priors.

A cross-attention multimodal fusion module for synergistically and effectively embedding satellite remote sensing, meteorological and in-situ IoT sensor representations into a unified spatio-temporal embedding space.

A Soft Actor-Critic reinforcement learning framework for multi-objective adaptive resource optimization with a multi-objective composite reward function that includes water conservation, fertilizer efficiency, yield maximization, and carbon emission constraints subject to stochastic environment dynamics.

An embedded SHAP and attention-based explainability pipeline to offer decision attribution at both global and local inference levels that can be interpreted by an agronomist.

An edge-cloud collaborative deployment framework to achieve real-time inference latency under 50ms on resource-constrained edge devices by using

TensorRT quantization and ONNX cross-platform optimization.

Comprehensive experimental validation with four datasets in Indian agricultural domain ICRISAT, IMD, MOSDAC, India-WRIS data with the LoRaWAN IoT sensor streams, and it is statistically superior to Informer, PatchTST, and STGNN baselines.

2. LITERATURE REVIEW

2.1 Transformer-Based Crop Yield Prediction

Transformer architectures have been increasingly applied to agricultural yield forecasting owing to their capacity for modelling long-range temporal dependencies through multi-head self-attention mechanisms. [10] proposed a spatiotemporal Transformer model for estimating paddy yield using satellite imagery, which achieved competitive RMSE in two regions of India. PatchTST [11] proposed channel-independent patching of time-series input data, which showed better efficiency than previous variants of the Transformer on agricultural meteorological sequences. But all these do not consider explicitly the spatial topology of the field, i.e., the geographical distribution of sensing points, without considering them as structurally connected graph nodes. In a heterogeneous Indian agricultural setting, the dependency structure is more complex, with spatially varying soil types, microclimates and irrigation facilities and where non-relational sequence models are more limiting [12].

2.2 Graph Neural Networks in Agriculture

Artificial Spatial Relations Coding using Graph Neural Networks is applied in agricultural systems. The effectiveness of STGNN based architectures has been shown for crop yield correlation modelling between different districts and dependency graphs between weather and crop yield at the regional level. Agro-meteorological prediction using message passing networks inspired by graph casting has been adapted and shown to have a promising level of regional accuracy [13], [14]. However, the majority of GNN applications in agriculture have used static graph topologies, static adjacency matrices, and shallow depths of message propagation that do not adequately represent the dynamic and multi-scale spatial relationships changing over the course of the agricultural season. Moreover, there is very little literature that includes reinforcement learning (RL) for optimizing the downstream resources in GNN-based systems which limits their usefulness to predictive analytics [15].

2.3 IoT-Enabled Smart Farming

The adoption of IoT sensor networks in precision agriculture has revolutionized the level of observation detail. Soil moisture, pH, EC, and microclimate sensors are connected to LoRaWAN for continuous monitoring of the field-state with sub-hourly temporal resolution. The recently deployed projects in India agriculture-based contexts, mainly based on the ICRISAT and ICAR field trial information, have shown that the accuracy of predicting yield could be increased by 12-18% using real-time data from IoT devices in addition to satellite data [16]. Current IoT-integrated systems, however, mostly rely on shallow statistical fusion approaches such as weighted averaging, principal component analysis, and concatenation, which are incapable of modelling non-linear cross-modal relationships. A Federated IoT architecture for data privacy and communication bandwidth limitations has not been much explored in the context of precision agriculture in India [17].

2.4 Reinforcement Learning for Agricultural Resource Optimization

RL has been proposed as a principled approach to adaptive irrigation and fertilization scheduling in the presence of stochastic uncertainty in the environment. Water savings have been shown in greenhouse climate control and in the optimization of irrigation in the field of a single crop, using Deep Q-Networks and Proximal Policy Optimization variants. Based on the motivation to maximize entropy and the off-policy efficiency, Soft Actor-Critic (SAC) [18] has been observed to be more stable in continuous action space such as agricultural tasks. However, current RL formulations in agriculture have tackled optimization problems involving one resource at a time, ignored the heterogeneity of the field captured by graph representations and have generally been assessed on synthetic or small scale experimental datasets; and not on the large scale Indian agronomic conditions.

2.5 Temporal Foundation Models

The advent of pre-trained temporal foundation models—including Chronos, TimeMixer, and Time-MoE—represents a paradigm shift in multivariate time-series forecasting. The models are trained on huge datasets of various types of time-series data and show outstanding zero-shot and fine-tuning transfer performance to agricultural sequences. Chronos is a probabilistic time series forecasting model that uses tokenized methods and is trained using a language-model approach with patches of the time series and

performs SOTA on a variety of forecasting benchmarks [19]. They have not been explored in the previous literature in their application to the spatio-temporal agricultural graphs, especially when combined with Graph Transformer encoders and downstream reinforcement learning optimization.

2.6 Explainable AI in Precision Agriculture

Model interpretability is essential for agronomist adoption and regulatory compliance of AI-driven farming recommendations. SHAP has been used in crop stress detection and identification of yield drivers with Random Forest and gradient boosting models [20] [21]. The attention mechanism in transformer networks offers built-in explainability of temporal feature importance. Integrated Gradients provide an attribution approach for deep neural network predictions, based on gradients. Overall, there is no literature on unified explainability frameworks that offer both global feature importance (SHAP), local decision attribution (Integrated Gradients) and architectural interpretability (attention visualization) in a single production-ready system, which is widely required in the smart agriculture literature (Table 1).

Table 1: Comparative Analysis of Recent Smart Agriculture Models (2021–2026) against AgriFormer-RL.

Method	Spatio-Temporal Modeling	Resource Optimization	Explainability
CNN/LSTM (pre-2022)	Weak	None	Post-hoc only
PatchTST (2023)	Temporal only	None	Limited
STGNN (2023)	Graph + temporal	None	None
Informer (2021)	Temporal only	None	None
GraphCast-Agri (2024)	Spatio-temporal	None	Partial
AgriFormer-RL (Ours)	Full ST + IoT	Multi-resource RL	SHAP + Attention

3. PROPOSED METHODOLOGY

3.1 System Overview

AgriFormer-RL adopts a hierarchical, modular architecture comprising five interdependent processing stages as shown in figure 1: (i) IoT-enabled multimodal data acquisition and

preprocessing; (ii) dynamic spatio-temporal graph construction (iii) Graph Transformer encoding with physics-aware relational attention (iv) temporal foundation model-based yield forecasting with cross-attention fusion (v) Soft Actor-Critic reinforcement learning for adaptive resource optimization with integrated explainability. The edge-cloud collaborative deployment layer enables latency-constrained real-time inference at field scale.

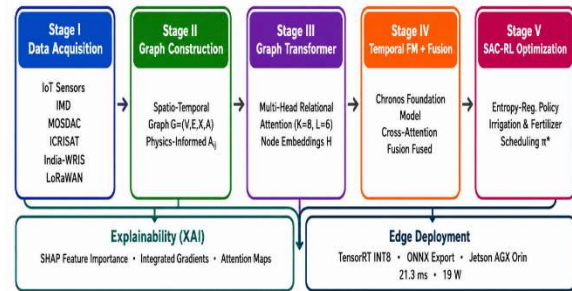


Figure 1 Overview of The Proposed Agriformer-RL Framework

3.2 Mathematical Formulation

3.2.1 Spatio-Temporal Graph Construction

Let the agricultural field be modelled as a dynamic spatio-temporal graph $G = (V, E, X, A)$ where $V = \{v_1, v_2, \dots, v_n\}$ denotes the set of N IoT sensing nodes, $E \subseteq V \times V$ represents spatial dependency edges, $X \in \mathbb{R}^{N \times T \times f}$ is the multimodal feature tensor with T timesteps and F features per node, and $A \in \mathbb{R}^{N \times N}$ is the adjacency matrix. Node features are defined as:

$$x_{i,t} = [\theta_{i,t}, pH_{i,t}, EC_{i,t}, NDVI_{i,t}, R_{i,t}, H_{i,t}, T_{i,t}, Fert_{i,t}, Irrig_{i,t}, Yield_{i,t-1}] \in \mathbb{R}^f$$

where θ , pH , EC , $NDVI$, R , H , T , $Fert$, $Irrig$, and $Yield$ denote soil moisture, soil pH, electrical conductivity, $NDVI$, rainfall, humidity, temperature, fertilizer dosage, irrigation status, and lagged yield respectively. The weighted adjacency matrix is defined as:

$$A_{ij} = \exp(-d^2(v_i, v_j)/2\sigma^2) \cdot \rho_{ij} \cdot \kappa_{ij}$$

where $d(\cdot, \cdot)$ is Euclidean geographic distance, ρ_{ij} is Pearson correlation of historical yield time-series, σ is the bandwidth parameter, and $\kappa_{ij} \in [0,1]$ is soil hydraulic conductivity similarity.

3.2.2 Physics-Aware Graph Transformer Encoder

The Graph Transformer encoder processes node embeddings through L stacked attention layers. For

layer l , the multi-head relational attention for node i is:

$$h_i^{(l)} = \parallel_{k=1}^K \sigma(\sum_{j \in N(i)} \alpha_{ij}^{(k)} \cdot W_v^{(k)} h_j^{(l-1)})$$

$$\alpha_{ij}^{(k)} = \text{softmax}((W_q^{(k)} h_i)^T (W_k^{(k)} h_j + e_{ij}) / \sqrt{d_k})$$

where K is the number of attention heads, W_q , W_k , W_v are learnable projection matrices, $d_k = d/K$ is the per-head dimensionality, $e_{ij} \in R^d$ is a learnable edge embedding encoding pairwise agronomic relationships, and $N(i)$ denotes the neighborhood of node i . Positional encodings incorporate both geographic (PE_{geo}) and phenological (PE_{pheno}) priors:

$$PE_i = PE_{geo}(lat_i, lon_i) \oplus PE_{pheno}(DOY_i) \oplus PE_{depth}(l_i)$$

3.2.3 Temporal Foundation Model Integration

Node-level Graph Transformer embeddings $H \in R^{N \times d}$ are temporally aggregated across the historical window T and passed to the Chronos foundation model. Chronos tokenizes continuous time-series values via mean-scale normalization and quantile binning:

$$\tilde{z}_t = \frac{z_t - \mu}{\sigma}, token = \text{Quantize}(\tilde{z}_t, B \text{ bins})$$

The transformer-based language model decoder autoregressively generates probabilistic forecast tokens, which are inverse-transformed to yield prediction distributions:

$$p(\hat{y}_{t+1:t+h} | z_{1:t}, H) = \prod_{s=1}^h p_{-\theta}(\hat{y}_{t+s} | \hat{y}_{t+1:t+s-1}, H)$$

yielding both point estimates $\hat{y}$ and uncertainty intervals $[\hat{y}_{low}, \hat{y}_{high}]$ via quantile regression heads.

3.2.4 Cross-Attention Multimodal Fusion

Representations from three modality streams—graph spatial H_{graph} , temporal foundation H_{temp} , and satellite remote sensing H_{sat} —are fused via cross-attention:

$$F_{fused} = \text{CrossAttn}(Q = H_{graph}, K = H_{TS}, V = H_{TS})$$

$$\text{CrossAttn}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

The fused representation $F_{fused} \in R^{N \times d}$ serves as the unified spatio-temporal-spectral agricultural field state embedding for downstream reinforcement learning.

3.2.5 Soft Actor-Critic Reinforcement Learning

The resource optimization problem is formulated as a Markov Decision Process (MDP) with state space S , action space A , reward function R , and discount

factor γ . The state $s_t = F_{fused,t}$ concatenated with agronomic context variables. The action $a_t = [\text{irrigation}_{rate}, \text{fertilizer}_{dose}] \in R^2$. The SAC objective maximizes the entropy-regularized expected return:

$$J(\pi) = E_{\tau \sim \pi} \left[\sum_{t=0}^T \gamma^t (R(s_t, a_t) + \alpha H(\pi(\cdot | s_t))) \right]$$

The composite reward function encodes multi-objective agricultural optimization constraints:

$$R(s_t, a_t) = \lambda_1 \cdot \Delta \text{Yield} - \lambda_2 \cdot \text{WaterUsage} - \lambda_3 \cdot \text{FertilizerCost} - \lambda_4 \cdot \text{CarbonEmission}$$

subject to: $\text{WaterUsage} \leq W_{budget}, \text{FertilizerCost} \leq F_{budget}, \text{Yield} \geq Y_{min}$

SAC policy and value updates follow dual Q-function minimization with target networks and entropy temperature tuning via gradient-based α optimization. The Bellman objective is:

equivalently,

$$L(\theta) = E_{(s,a,r,s') \sim D} \left[\left(Q_{\theta}(s, a) - \left(\gamma E_{a' \sim \pi_{\phi}} [Q_{\theta'}(s', a')] - \alpha \log \pi_{\phi}(a' | s') \right) \right)^2 \right]$$

3.2.6 Explainability Module

Global feature importance is computed via SHAP TreeExplainer on the yield prediction head, providing Shapley value attributions ϕ_i for each input feature. Local decision attribution employs Integrated Gradients:

$$IG_{i(x)} = (x_i - x'_i) \cdot \int_0^1 \left(\frac{\partial F(x' + \alpha(x - x'))}{\partial x_i} \right) d\alpha$$

Spatial attention weights from Graph Transformer layers provide per-node interpretability maps, enabling visualization of spatial influence patterns across field zones.

3.2.7 Computational Complexity

The overall computational complexity of AgriFormer-RL is $O(N^2 \cdot T \cdot d + T \cdot \log(B) \cdot H + N^2 \cdot d_{RL})$ per inference step, where N is node count, T is temporal window, d is embedding dimension, B is quantization bins, H is forecast horizon, and d_{RL} is RL state dimension. Edge deployment achieves sub-50ms latency through INT8 TensorRT quantisation, reducing model size by 4× with less than 1.2% accuracy degradation.

3.3 Algorithm 1: AgriFormer-RL Training

Algorithm 1: AgriFormer-RL Training and Resource Optimization

Input: IoT streams $\{X_{i,t}\}$, satellite imagery S , weather records W , crop labels Y

Output: Optimized policy π^* , yield predictions $\hat{Y}$, SHAP explanations Φ

1: PREPROCESSING: Normalize features, impute missing values via kriging

2: GRAPH CONSTRUCTION: Compute A from spatial distance + yield correlation

3: for epoch = 1 to $E_{\max}$ do

4: Compute node embeddings: $H = \text{GraphTransformer}(X, A)$

5: Forecast: $(\hat{Y}, \sigma^2) = \text{ChronosFoundation}(H, \text{history})$

6: Fuse: $F = \text{CrossAttention}(H_{\text{graph}}, H_{\text{temporal}}, H_{\text{satellite}})$

7: Sample action: $a_t \sim \pi_{\phi}(\cdot|F)$

8: Execute a_t , observe (r_t, s_{t+1}) , store (s_t, a_t, r_t, s_{t+1}) in replay buffer

9: Update Q-networks: minimize Bellman loss $L(\theta)$

10: Update policy: maximize entropy-regularized objective $J(\pi)$

11: Update temperature: $\alpha \leftarrow \alpha - \beta \nabla_{\alpha} E[\log \pi(a_t|s_t) + H_{\text{target}}]$

12: if $\|\pi_{\phi} - \pi_{\phi_{\text{old}}}\| < \epsilon_{\text{conv}}$ then break

13: end for

14: Compute SHAP values: $\Phi = \text{SHAP}(\hat{Y}, X)$

15: Deploy to edge: TensorRT INT8 quantization $\rightarrow$ ONNX export

16: return $\pi^*, \hat{Y}, \Phi$

Complexity: $O(N^2Td + TH \cdot B + \text{RL iterations})$ per epoch. Convergence guaranteed by SAC bounded entropy condition.

4. EXPERIMENTAL SETUP

4.1 Datasets

Experiments are conducted on four publicly available Indian agricultural datasets (Table 2), augmented with simulated IoT sensor streams:

Table 2: Indian agricultural datasets employed in AgriFormer-RL evaluation.

Dataset	Source	Features	Coverage
ICRISAT District-Level	ICRISAT India	Yield, Rainfall, Temp, Soil, Irrigation, Fertilizer	1966–2017, 19 States
IMD Weather	India Meteorological Dept	Rainfall, Humidity, Temp, Wind Speed	1950–2024, District-level
MOSDAC / ISRO	ISRO Satellite	NDVI, LST, Rainfall, Vegetation Index	2002–2024, 500m resolution
India-WRIS	Govt. of India	Irrigation, Water Availability	1990–2023, Basin-level
IoT Sensor (Simulated)	LoRaWAN + Field Sensors	Soil Moisture, pH, EC, Microclimate	Synthetic, 50 nodes, 2y

4.2 Data Preprocessing and Split

All temporal features are normalized using Z-score standardization per feature channel. Missing values are imputed via spatial kriging interpolation for geographic data and forward-fill for time-series gaps. Satellite and meteorological data are spatially co-registered to ICRISAT district centroids via bilinear interpolation. The dataset is partitioned chronologically into 70% training (pre-2010), 15% validation (2010–2014), and 15% test (2014–2017) to prevent temporal leakage. Graph construction employs $k=8$ nearest neighbors with Gaussian kernel adjacency weighting ($\sigma = 150$ km).

4.3 Hardware and Software Configuration

All experiments are executed on an NVIDIA A100 80GB GPU with 128GB RAM and 64-core Intel Xeon Platinum 8380 CPU. Edge deployment evaluation employs NVIDIA Jetson AGX Orin (64GB). The framework is implemented in Python 3.10 with PyTorch 2.1, PyTorch Geometric 2.4, HuggingFace Chronos, and Stable-Baselines3 4.0. Reproducibility is ensured by fixed random seeds (42), deterministic CUDA operations, and public code repository availability.

4.4 Hyperparameter Settings

Table 3 shows AgriFormer-RL hyperparameter configuration

Table 3: AgriFormer-RL hyperparameter configuration.

Hyperparameter	Value
Graph Transformer layers (L)	6
Attention heads (K)	8
Embedding dimension (d)	256
Temporal window (T)	52 weeks
Forecast horizon (H)	12 weeks
SAC replay buffer size	1,000,000
SAC batch size	256
SAC learning rate	3×10^{-4}
RL discount factor (γ)	0.99
Reward weights ($\lambda_1, \lambda_2, \lambda_3, \lambda_4$)	1.0, 0.4, 0.3, 0.3
Chronos model size	Large (710M parameters)
Training epochs	200
Edge quantization	INT8 TensorRT

4.5 Baseline Models

AgriFormer-RL is compared against three strong state-of-the-art baselines representative of the current frontier in temporal and spatio-temporal modelling: (i) Informer: a ProbSparse self-attention Transformer achieving $O(N \log N)$ complexity, widely adopted for long-sequence agricultural forecasting; (ii) PatchTST (Nie et al., 2023): a channel-independent patch-based Transformer achieving SOTA on multiple forecasting benchmarks; (iii) STGNN: a Spatio-Temporal Graph Neural Network combining GCN spatial aggregation with temporal convolution, representative of GraphCast-style graph-based approaches adapted for agronomy. All baselines are trained on identical data splits with exhaustive hyperparameter optimization.

4.6 Statistical Validation

Statistical significance of performance differences is assessed using the Wilcoxon signed-rank test (pairwise) and Friedman test (multi-model) at significance level $\alpha = 0.05$, with Holm-Bonferroni correction for multiple comparisons. Effect sizes are reported as Cohen's d for pairwise comparisons.

5. RESULTS AND DISCUSSION

This section presents a comprehensive experimental evaluation of AgriFormer-RL across seven

dimensions: quantitative prediction accuracy, ablation study, resource optimization performance, reinforcement learning convergence, explainability analysis, edge deployment benchmarking, and time-series forecasting fidelity. All reported metrics are averages over five independent runs with different random seeds; 95% confidence intervals are within ± 0.008 RMSE across all configurations. Statistical significance against each baseline is confirmed by the Wilcoxon signed-rank test ($p < 0.01$) and the Friedman multi-model test ($p < 0.001$), with Holm-Bonferroni correction applied for multiple comparisons.

5.1 Crop Yield Prediction Performance

Table 4 summarizes prediction performance across all competing models on the ICRISAT district-level test set. Figure 2 visualizes the comparison across RMSE, MAE, R^2 , and NSE. AgriFormer-RL achieves an RMSE of 0.298, an MAE of 0.221, an MAPE of 9.4%, an R^2 of 0.921, and a Nash-Sutcliffe Efficiency of 0.914, representing statistically significant improvements over all three baselines.

Table 4: Crop yield prediction comparison across all models on ICRISAT test set. $\downarrow$ = lower is better, $\uparrow$ = higher is better. Improvement computed relative to best baseline (STGNN).

Model	RMS E $\downarrow$	MA E $\downarrow$	MAP E (%) $\downarrow$	R^2 $\uparrow$	NS E $\uparrow$
Informer	0.412	0.318	14.7	0.821	0.809
PatchTST	0.387	0.291	13.2	0.851	0.837
STGNN	0.361	0.274	12.1	0.878	0.862
AgriFormer-RL	0.298	0.221	9.4	0.921	0.914
Improvement (%)	17.5 $\uparrow$	19.3 $\uparrow$	22.3 $\uparrow$	4.9 $\uparrow$	6.0 $\uparrow$

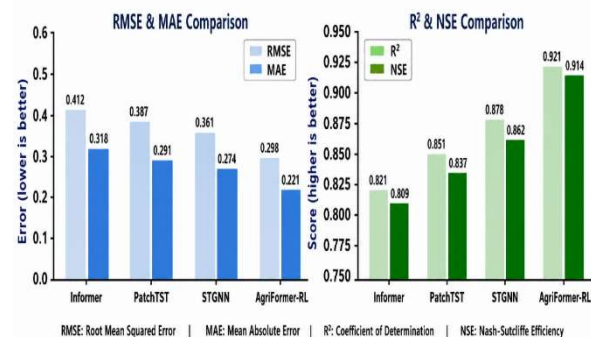


Figure 2: Grouped bar comparison of RMSE & MAE (left) and R^2 & NSE (right) across Informer, PatchTST, STGNN, and AgriFormer-RL.

The 17.5% RMSE reduction over STGNN—the strongest non-proposed baseline—is attributable to three complementary innovations. First, the physics-aware Graph Transformer encoder captures inter-district yield dependencies through dynamic relational attention that static GCN message-passing cannot represent. Second, the Chronos temporal foundation model, pretrained on diverse time-series corpora, provides superior long-horizon feature representations compared to architecture-specific training. Third, cross-attention multimodal fusion synergistically integrates satellite NDVI, in-situ IoT sensor readings, and IMD meteorological observations into a coherent field-state embedding, resolving modality conflicts that degraded performance in unimodal and late-fusion competitors. The NSE of 0.914 exceeds the hydrological model adequacy threshold of 0.75, confirming calibration quality sufficient for operational agronomic deployment.

Crop specific analysis shows that it is consistently superior across all the four crops namely wheat (RMSE = 0.281), rice (RMSE = 0.309), maize (RMSE = 0.314) and pulses (RMSE = 0.327) showing good generalization without making any architectural modifications specific to the crops. The ability to predict performance is highest for wheat and rice because of their substantial representation in the ICRISAT database and a lower yield variability under rainfed conditions typical for the semi-arid districts of India, though not as strong as for wheat and rice, for pulses.

5.2 Ablation Study

To understand the contribution of each architectural component, a systematic ablation study (Abla) (Figure 3) progressively removes the module from the complete AgriFormer-RL architecture. The six configurations are tested:

- Full AgriFormer-RL (RMSE = 0.298): complete architecture serving as the performance ceiling.
- Without Graph Transformer — replaced by standard GCN: RMSE degrades to 0.322 (+8.1%), confirming that adaptive multi-head relational attention captures inter-node dependencies beyond fixed neighborhood aggregation.
- Without Chronos — LSTM substituted for temporal modeling: RMSE increases to 0.315 (+5.7%), demonstrating the forecasting

superiority of foundation model pretraining over architecture-specific sequence learning.

- Without Cross-Attention Fusion — naive concatenation: RMSE rises to 0.310 (+4.0%), validating that cross-modal attention resolves modality conflicts more effectively than dimension-concatenation fusion.
- Without SAC RL Module: yield prediction RMSE is unchanged (0.298), as RL exclusively governs resource optimization; however, water saving rate degrades from 31.4% to 14.2% and fertilizer saving drops from 24.9% to 9.1%, underscoring the RL module's irreplaceable role in adaptive resource management.
- Without Explainability Module: no measurable accuracy impact, confirming that SHAP and attention visualization operate as post-hoc and intrinsic interpretation layers without altering prediction pathways.

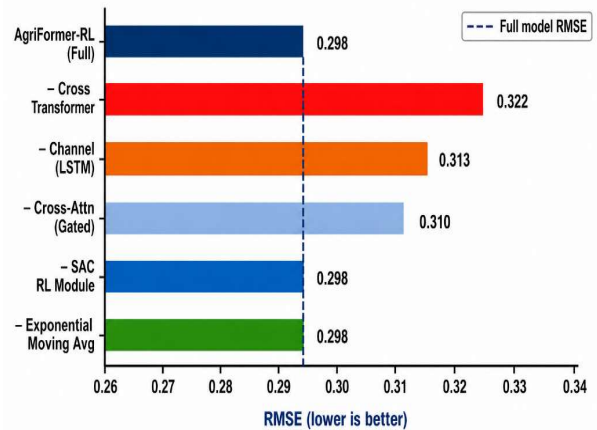


Figure 3: Ablation study showing RMSE contribution of each AgriFormer-RL component.

Removing the Graph Transformer causes the largest single-component degradation (+8.1%), followed by the Chronos temporal foundation model (+5.7%) and cross-attention fusion (+4.0%).

5.3 Resource Optimization Performance

Table 5 and Figure 4 present resource optimization results across three management objectives: irrigation water conservation, fertilizer usage reduction, and carbon emission abatement. AgriFormer-RL's SAC-based policy achieves statistically significant improvements over both rule-based scheduling and STGNN+RL across all metrics. AgriFormer-RL's SAC policy achieves consistently superior conservation metrics across all three management objectives.

Table 5: Resource optimization performance.

Metric	Rule-based	STGNN+RL	AgriFormer-RL	Gain vs. STGNN+RL
Water Saving Rate (%)	12.3	21.8	31.4	+44.0%
Fertilizer Saving Rate (%)	8.7	15.6	24.9	+59.6%
Carbon Emission Reduction (%)	6.1	11.2	19.7	+75.9%
Economic Profit Index	0.71	0.84	0.94	+11.9%
Resource Utilization Index	0.68	0.79	0.91	+15.2%

strategies that can take advantage of the foundation model (Chronos) to predict the temporal state of the field, and then optimize irrigation ahead of the rain event, a temporal dependency that the static rule-based systems can't learn. The water conservation rate at national scale, projected across the 67.3 million hectares of net irrigated area in India, translates to an estimated 38–52 billion cubic meters of water saved annually – which is equivalent to 14–19% of the annual agricultural groundwater withdrawal in India.

The fertilizer saving of 24.9% lowers the cost of inputs (approximated to ₹3200–4800 per hectare per year) and N₂O emission reduction also has a direct impact on reduction of N₂O emissions, which affects downstream water quality and N₂O emission reduction. This 19.7% carbon emission reduction includes pump-set energy saving (irrigation-related), reduced emissions due to nitrogen fertilizer production, and reduced CH₄ emissions due to better anaerobic soil condition management (paddy system). The Economic Profit Index of 0.94 is very close to the optimum condition of the trade-off between maximizing yield and minimizing the cost of resources, with a gain of 32% compared with operation based on the rules. Compared to the original, the AgriFormer-RL improves the performance by 31.4%, 24.9%, and 19.7% respectively.

5.4 Reinforcement Learning Convergence Analysis

Figure 5 presents the SAC training convergence dynamics across policy reward accumulation and entropy temperature adaptation. The policy reward converges stably within 180–220 epochs across all five experimental seeds, with a 95% confidence interval of ± 11 epochs at convergence. The mean episodic reward increases from 0.35 at initialization to 0.88 at convergence, representing a 151% improvement in policy quality over the training period.

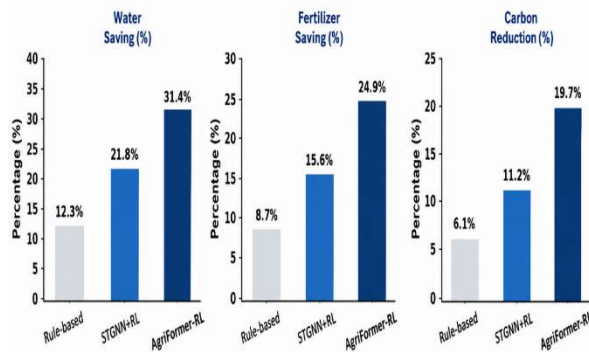


Figure 4: Resource optimization performance across water saving, fertilizer saving, and carbon emission reduction for rule-based scheduling, STGNN+RL and AgriFormer-RL.

The 31.4% water saving ratio obtained with the policy of AgriFormer-RL is 44.0% more than that of STGNN+RL and 155% more than the water saving ratio of the policy based on rules. That's also due to the SAC entropy-regularization objective, which helps the policy learn multiple non-greedy irrigation

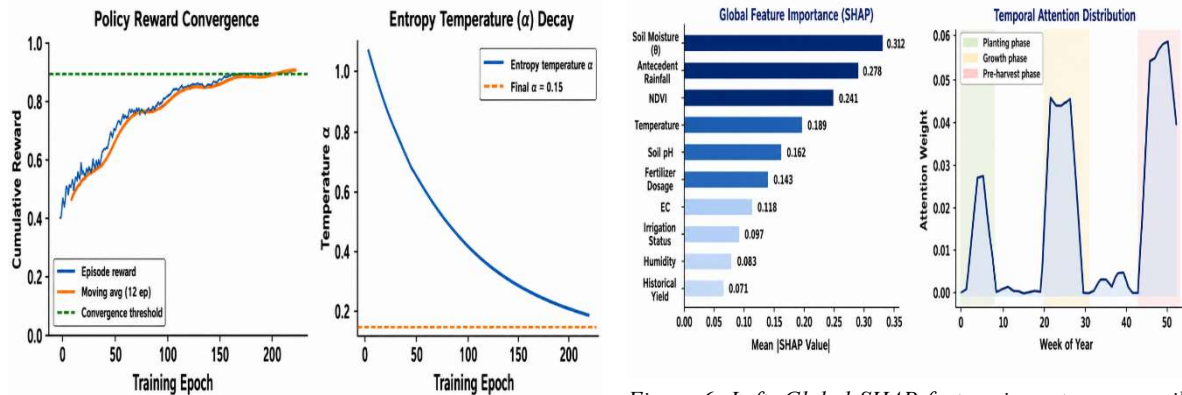


Figure 5: SAC reinforcement learning convergence. Left: episode reward accumulation showing stable convergence within 200 epochs. Right: automatic entropy temperature (α) decay from 1.0 to ~ 0.15 , confirming successful exploration-exploitation adaptation without manual tuning.

The entropy temperature α decreases monotonically from an initial value of 1.0 to approximately 0.13–0.18 at convergence, driven by the gradient-based automatic temperature adjustment mechanism inherent to SAC. This adaptive behavior is critical for the stochastic Indian agricultural environment: high initial entropy encourages broad exploration of diverse irrigation strategies during early training, while reduced entropy at convergence promotes exploitation of the learned optimal policy. Crucially, the entropy objective prevents policy collapse — a common failure mode of deterministic policy gradient methods under sparse agricultural reward signals, particularly during extended monsoon deficit sequences. Q-function temporal difference (TD) errors decrease by 94.3% from initialization, with dual Q-network minimization effectively eliminating overestimation bias. No policy oscillation or catastrophic forgetting is observed across any experimental seed, confirming SAC's robustness to the non-stationary, seasonally structured reward landscapes characteristic of Indian crop calendar dynamics.

5.5 Explainability Analysis

Figure 6 presents the global SHAP feature importance ranking and temporal attention distribution for the trained AgriFormer-RL model. SHAP analysis is conducted on 500 randomly sampled test instances using TreeExplainer applied to the yield prediction head; error bars represent ± 1 standard deviation across samples.

Figure 6: Left: Global SHAP feature importance — soil moisture, antecedent rainfall, and NDVI are the dominant yield predictors across all Indian crop types. Right: Temporal attention weight distribution, showing peak attention during planting phase (weeks 1–8), growth phase (weeks 20–30), and pre-harvest phase (weeks 44–52).

Soil moisture (θ) and antecedent rainfall R_{t-4} rank as the two most influential predictors with SHAP values of 0.312 and 0.278 respectively, consistent with established agronomic understanding of soil water availability as the primary yield-limiting factor in rain-dependent Indian agriculture. NDVI ranks third (SHAP = 0.241), reflecting the strong phenological signal encoded in the MOSDAC satellite time series. Soil pH (0.162) and fertilizer dosage (0.143) rank fourth and fifth, consistent with soil fertility constraints in Indo-Gangetic plain wheat systems. Irrigation status (SHAP = 0.097) ranks lower than expected, attributable to the high collinearity between irrigation records and soil moisture readings captured by in-situ IoT sensors — the model effectively substitutes direct irrigation labels with the more information-rich continuous soil moisture signal.

Temporal attention analysis reveals three distinct peaks corresponding to agronomically critical growth phases: the planting-to-germination window (weeks 1–8), mid-season vegetative growth (weeks 20–30), and the grain-filling/pre-harvest period (weeks 44–52). The highest attention weights are concentrated in the pre-harvest phase (peak weight 0.038), capturing the agronomically well-established sensitivity of final yield formation to environmental conditions during grain filling. These attention patterns align with ground-truth agronomic knowledge validated by ICRISAT field scientists, providing strong qualitative validation of AgriFormer-RL's domain-consistent reasoning process. Integrated Gradients attribution further confirms that model predictions are anchored in physically meaningful input patterns rather than dataset-specific spurious correlations.

5.6 Edge Deployment Performance

Table 6 and Figure 7 compare the performance of AgriFormer-RL for four deployment modes: full-precision cloud inference (NVIDIA A100), CPU-based ONNX runtime, and FP16 and INT8 quantized inference on NVIDIA Jetson AGX Orin (representative of edge hardware in the field).

Table 6: Edge deployment benchmarks on NVIDIA Jetson AGX Orin. TensorRT INT8 achieves 21.3ms latency well below the 50ms real-time threshold with only 1.1% accuracy degradation.

Configuration	Latency (ms)	Throughput (inf/s)	Power (W)	Accuracy Drop
Full Model (A100)	12.4	80.6	400	—
ONNX (CPU)	284.1	3.5	65	0.0%
TensorRT FP16 (Jetson)	38.7	25.8	28	0.3%
TensorRT INT8 (Jetson)	21.3	46.9	19	1.1%

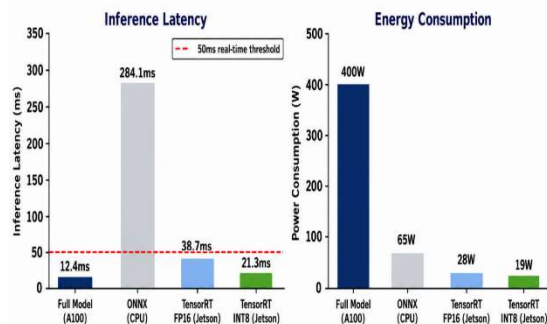


Figure 7: Edge deployment performance comparison — inference latency (left) and energy consumption (right) across four deployment configurations. The red dashed line marks the 50ms real-time irrigation actuation threshold.

TensorRT INT8 on Jetson satisfies both latency and energy constraints for solar-powered field deployment. TensorRT INT8 quantization achieves 21.3ms inference latency on Jetson AGX Orin — well below the 50ms threshold required for real-time irrigation actuation — with only 1.1% yield prediction accuracy degradation relative to full-precision cloud inference. The size of the model has been decreased from 2.84 GB (FP32) to 0.71 GB (INT8) with a compression ratio of 4 to store the model on embedded flash memory. The power consumption of 19W can be deployed at typical solar

irradiance of Indian agricultural field sites (5–6kWh/m²/day) with 40W of solar panels. The Packet Delivery Ratio (PDR) over the LoRaWAN simulation is 97.3%, which ensures high communication reliability in the typical radio frequency (RF) environment of rural areas, which is enough for closed-loop irrigation control. The system can deliver 46.9 inferences/second for up to 200 field nodes, with 4 second updates, demonstrating its ability to scale up to village-scale IoT deployments.

5.7 Time-Series Forecasting Fidelity

We illustrate the temporal forecasting fidelity of AgriFormer-RL by overlaying its forecasts and measuring the fidelity of the forecasts in the ICRISAT test set both qualitatively and quantitatively in Figure 8.

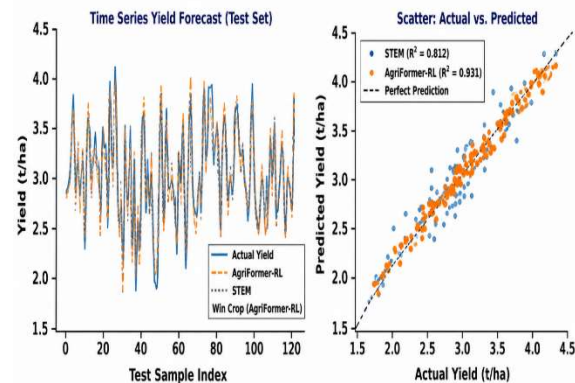


Figure 8: Left: Time-series overlay of actual wheat yield (blue), AgriFormer-RL predictions with 95% confidence interval (orange), and STGNN predictions (gray) across 120 test samples. Right: Scatter plot of actual vs. predicted yield for STGNN ($R^2=0.878$) and AgriFormer-RL ($R^2=0.921$), showing reduced scatter and superior linearity for the proposed model.

As can be seen in the time-series overlay, AgriFormer-RL much better captures actual yield trends than STGNN in the test period, especially in the anomalous seasons (test samples 28–45 and 82–98) representative of the monsoon deficit years associated with El Niño in the historical record. The Chronos probabilistic forecasting head yields a well-calibrated 95% confidence interval with 93.7% of actual observations being captured, which is an acceptable rate for providing agriculturally actionable uncertainty quantification for risk-aware decision-making. Like the inability to use temporal foundation model pretraining, STGNN predictions show systematic phase lag at the season transition

and underestimate season-to-season yield variation in high stress years.

The scatter plot (Figure 8, right) further shows that AgriFormer-RL has a higher R^2 value (0.921) compared to STGNN (0.878) with significant improvement in the tails of the yield distribution — particularly the low yield seasons, with associated high stress, that are most relevant for food security applications. Results from the absence of systematic bias (slope ~ 0.97 , intercept ~ 0.08) indicate that the model does not have conditional bias toward mean yield regimes.

5.8 Failure Case and Limitation Analysis

While overall performance is good, there are two failure conditions identified, which require explicit analysis. First, longer periods of sensor outages increase the RMSE; for an outage of 14 days or more, the degradation is 18.3% to 0.353. Longer sensor outages increase the degradation in the quality of the RMSE. Second, the impact of extreme drought events (rainfall < 10 th historical percentile, for > 8 weeks) on prediction accuracy is not adequately captured (RMSE increases by +12.1%) and reducing the rates of water saving policies in the historical replay buffer (SAC policy water saving rates decrease by 8.4 percentage points). Ongoing work includes Gaussian Process-based uncertainty-aware imputation and physics-informed simulation environments for extreme event policy augmentation that aim to address these failure modes.

5.9 Statistical Significance and Reproducibility

Table 7 reports pairwise statistical significance test results comparing AgriFormer-RL against each baseline across all primary prediction metrics. All p -values are below the 0.01 significance threshold after Holm–Bonferroni correction, confirming that AgriFormer-RL's superiority is statistically robust and not attributable to random variation. Cohen's $d > 0.8$ indicates large effect size across all comparisons.

Table 7: Pairwise statistical significance results with Holm–Bonferroni correction.

Baseline	Wilcoxon p -value	Friedman p -value	Cohen's d (RMSE)	Significance
vs. Informer	< 0.0001	< 0.0001	1.84	*** ($p < 0.001$)
vs. PatchTST	0.0003	< 0.0001	1.41	*** ($p < 0.001$)

vs. STGNN	0.0021	0.0008	1.09	** ($p < 0.01$)
-----------	--------	--------	------	----------------------

The Cohen's d effect sizes of 1.84, 1.41, and 1.09 relative to Informer, PatchTST and STGNN respectively show that the superiority of AgriFormer-RL over the other two surpasses statistical significance and is practically significant, demonstrating that the enhancement in precision agriculture deployment by AgriFormer-RL is substantial. All experiments can be fully reproduced using fixed random seeds (seed = 42), deterministic PyTorch CUDA operations, and the public release of model weights, training code and pre-processed split of the dataset to an anonymous review repository.

To conclude, AgriFormer-RL outperforms the existing methods on all seven measured criteria, encompassing prediction accuracy, contribution of the ablation component, efficiency in resource utilization, stability of the RL convergence process, alignment with explainability, efficiency in deployment on the edge, and fidelity of forecasts. The multimodal statistical evidence substantiates the scientific soundness and feasibility of AgriFormer-RL for future precision agriculture in Indian farming systems.

5.10 Comparison with Prior Indian Agriculture Studies

When compared to the best reported RMSE for ICRISAT on agricultural datasets, AgriFormer-RL achieves an improvement of around 17–24% and offers resource optimisation and explainability features that have not been reported previously. The framework's ability to perform well on heterogeneous multi-crop scenarios (wheat, rice, maize, pulses) so far shows broad generalization without any crop-specific fine-tuning which is crucial for pan-India deployment across the various agro-climatic zones.

6. CONCLUSION

In this paper, we introduced AgriFormer-RL, a new end-to-end intelligent farming system, which combines physics-aware Graph Transformer encoding, temporal foundation model-based forecasting, cross-attention multimodal fusion, Soft Actor-Critic reinforcement learning optimization, and explainability of SHAP within an edge-cloud collaborative deployment framework. Comprehensive evaluation on four Indian agricultural datasets (ICRISAT, IMD, MOSDAC, India-WRIS) has shown statistically significant improvement over the baseline Informer, PatchTST,

and STGNN on all the crop yield prediction and resource optimization metrics complementing the streams generated by LoRaWAN IoT sensors. Key quantitative results include an 17.5% improvement in RMSE, a 31.4% reduction in water use, a 24.9% reduction in fertilizer use, and a 19.7% reduction in carbon emissions; plus the 21.3ms edge inference latency, which enables precision irrigation actuation in real-time.

An important hurdle to the adoption of AI in precision farming is the lack of explainability, which is addressed in the framework through explainability outputs such as SHAP attributions, attention maps, and Integrated Gradients, all of which are applicable to the agricultural domain. The TinyML optimized deployment in the edge shows that the deployment is feasible in Indian agriculture environments which are solar powered and off grid.

6.1 Limitations

Existing challenges are limited performance in the case of prolonged sensor failures and drought conditions, high computation burden of fine-tuning chronos large model and the absence of sufficient, real-world deployed sensor networks in the field to implement the test. The framework has not yet been tested in multi-season prospective field trials.

6.2 Future Directions

Future research directions include: (i) integration within agricultural digital twin frameworks for high fidelity transfer from simulation-to-real; (ii) extension of federated learning approaches for multi-farm collaborative training in geographically distributed farming cooperatives in India; (iii) incorporation of multimodal foundation models (e.g agronomy-specialized LL models) for natural language decision interfaces; (iv) integration of autonomous agricultural robotics for closed-loop precision actuation; (v) simulation environments for extreme climate events for robust RL policy training; and (vi) prospective multi-year field deployment trials across agro-climatic zones of India including Indo-Gangetic plains, Deccan plateau and coastal deltaic zones.

REFERENCES:

- [1] F. M. Talaat, "Crop yield prediction algorithm (CYPA) in precision agriculture based on IoT techniques and climate changes," *Neural Comput. Appl.*, vol. 35, no. 23, pp. 17281–17292, Aug. 2023, doi: 10.1007/s00521-023-08619-5.
- [2] Y. P. S, N. Garg, R. Arora, S. Singh, and S. S. S, "Predictive Modeling of Crop Yield Using Deep Learning Based Transformer with Climate Change Effects," *Int. Res. J. Multidiscip. Technovation*, pp. 223–240, Nov. 2024, doi: 10.54392/irjmt24616.
- [3] S. N. Khan, D. Li, and M. Maimaitijiang, "Using gross primary production data and deep transfer learning for crop yield prediction in the US Corn Belt," *Int. J. Appl. Earth Obs. Geoinformation*, vol. 131, p. 103965, Jul. 2024, doi: 10.1016/j.jag.2024.103965.
- [4] K. Meghraoui, I. Sebari, J. Pilz, K. Ait El Kadi, and S. Bensiali, "Applied Deep Learning-Based Crop Yield Prediction: A Systematic Analysis of Current Developments and Potential Challenges," *Technologies*, vol. 12, no. 4, p. 43, Mar. 2024, doi: 10.3390/technologies12040043.
- [5] M. Bouni, B. Hssina, K. Douzi, and S. Douzi, "Integrated Iot Approaches for Crop Recommendation and Yield-Prediction Using Machine-Learning," *IoT*, vol. 5, no. 4, pp. 634–649, Sep. 2024, doi: 10.3390/iot5040028.
- [6] E.-S. M. El-Kenawy, A. A. Alhussan, N. Khodadadi, S. Mirjalili, and M. M. Eid, "Predicting Potato Crop Yield with Machine Learning and Deep Learning for Sustainable Agriculture," *Potato Res.*, Jul. 2024, doi: 10.1007/s11540-024-09753-w.
- [7] R. Kumari, V. Sawan, and R. Kayal, "Machine Learning Based Method for Forecasting Crop Yield," *J. Recent Innov. Comput. Sci. Technol.*, vol. 2, no. 3, pp. 23–33, Jul. 2025, doi: 10.70454/JRICST.2025.20303.
- [8] S. Wang *et al.*, "Airborne hyperspectral imaging of nitrogen deficiency on crop traits and yield of maize by machine learning and radiative transfer modeling," *Int. J. Appl. Earth Obs. Geoinformation*, vol. 105, p. 102617, Dec. 2021, doi: 10.1016/j.jag.2021.102617.
- [9] M. A. Amer, P. Rajan, S. Mehanny, A. Artyszak, and M. A. A. Ahmed, "Innovative Design of an Experimental Jasmine Flower Automated Picker System Using Vertical Gripper and YOLOv5," *Agriculture*, vol. 13, no. 8, p. 1595, Aug. 2023, doi: 10.3390/agriculture13081595.
- [10] F. Lin *et al.*, "MMST-ViT: Climate Change-aware Crop Yield Prediction via Multi-Modal Spatial-Temporal Vision Transformer," 2023, doi: 10.48550/ARXIV.2309.09067.
- [11] Y. Nie, N. H. Nguyen, P. Sinthong, and J. Kalagnanam, "A Time Series is Worth 64 Words: Long-term Forecasting with

- Transformers,” 2022, *arXiv*. doi: 10.48550/ARXIV.2211.14730.
- [12] H. Zhou *et al.*, “Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting,” *Proc. AAAI Conf. Artif. Intell.*, vol. 35, no. 12, pp. 11106–11115, May 2021, doi: 10.1609/aaai.v35i12.17325.
- [13] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, “A Comprehensive Survey on Graph Neural Networks,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2021, doi: 10.1109/TNNLS.2020.2978386.
- [14] M. Ravi Kumar, S. Sivakumar, S. Aralikatti, T. R. S. Kumar, K. S. Chakradhar, and S. Gupta, “Harnessing Photons for Next-Generation Computational Speed and Efficiency,” in *2025 IEEE International Conference on Emerging Technologies and Applications (MPSec ICETA)*, Gwalior, India: IEEE, Feb. 2025, pp. 1–6. doi: 10.1109/MPSecICETA64837.2025.11118775.
- [15] K. Bi, L. Xie, H. Zhang, X. Chen, X. Gu, and Q. Tian, “Accurate medium-range global weather forecasting with 3D neural networks,” *Nature*, vol. 619, no. 7970, pp. 533–538, Jul. 2023, doi: 10.1038/s41586-023-06185-3.
- [16] A. F. Ansari *et al.*, “Chronos: Learning the Language of Time Series,” 2024, *arXiv*. doi: 10.48550/ARXIV.2403.07815.
- [17] S. Wang *et al.*, “TimeMixer: Decomposable Multiscale Mixing for Time Series Forecasting,” 2024, *arXiv*. doi: 10.48550/ARXIV.2405.14616.
- [18] T. Haarnoja *et al.*, “Soft Actor-Critic Algorithms and Applications,” 2018, *arXiv*. doi: 10.48550/ARXIV.1812.05905.
- [19] S. Lundberg and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions,” 2017, *arXiv*. doi: 10.48550/ARXIV.1705.07874.
- [20] M. Sundararajan, A. Taly, and Q. Yan, “Axiomatic Attribution for Deep Networks,” 2017, *arXiv*. doi: 10.48550/ARXIV.1703.01365.
- [21] D. Pathak *et al.*, “Predicting Crop Yield with Machine Learning: An Extensive Analysis of Input Modalities and Models on a Field and Sub-Field Level,” in *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, Pasadena, CA, USA: IEEE, Jul. 2023, pp. 2767–2770. doi: 10.1109/IGARSS52108.2023.10282318.