Journal of Theoretical and Applied Information Technology
31st August 2026. Vol.104. No.16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# AGRIRL: AN ATTENTIVE DEEP LEARNING AND REINFORCEMENT LEARNING FRAMEWORK FOR SIMULATOR-BASED ADAPTIVE FERTILIZER RECOMMENDATION

**PIDUGU NAGENDRA[1], DR. CH. VENKATA RAMANA REDDY[2], DR. K. PRADEEP REDDY[3]**

[1]Research Scholar, Department of CSE, Gandhi Institute of Engineering and Technology University, Gunupur, Odisha, India.
[2]Professor, Department of CSE, Gandhi Institute of Engineering and Technology University, Gunupur, Odisha, India.
[3]Professor, CSE Department, Vignan's Institute of Management and Technology for Women, Hyderabad, Telangana, India.
Corresponding author E-mail: pidugu.nagendra@giet.edu

## ABSTRACT

Precision fertilizer management requires both accurate diagnosis of the current soil-crop state and sequential decisions that adapt to changing nutrient status, weather, and crop demand. Most public-data studies formulate fertilizer recommendation as a one-shot classification problem, whereas reinforcement-learning studies typically optimize actions inside a crop simulator. This study proposes AgriRL, a two-layer framework that deliberately separates these two forms of evidence: (i) an attentive tabular deep classifier for fertilizer-class prediction and (ii) a finite-horizon deep Q-network (DQN) for stage-wise fertilizer type-and-dose selection in a transparent surrogate crop-response environment. The supervised layer was evaluated on a frozen 10,000-record public dataset with five harmonized classes using a stratified 70/15/15 train/validation/test split, training-only imputation and scaling, fold-restricted SMOTE, and five-fold hyperparameter tuning on the training partition. The sequential layer was evaluated independently on 50 paired six-stage simulated seasons in which all compared policies received identical initial states and weather realizations. AgriRL achieved 98.7% held-out classification accuracy and 98.5% macro-F1. In the retained simulator, mean seasonal yield increased from 5.87 ± 0.24 t/ha under the traditional schedule to 6.58 ± 0.20 t/ha under AgriRL, corresponding to a paired mean difference of 0.71 t/ha (12.1%). Ablation results indicate that attentive representation and class balancing primarily improve classification, whereas the DQN primarily changes sequential simulator outcomes; reward-term ablation further shows the cost-risk trade-off of yield-only optimization. Because no field trial or independently calibrated crop model was available, all productivity, cost, nutrient-use-efficiency, and environmental-risk outcomes are reported strictly as simulator-derived evidence. The principal scientific contribution is therefore a reproducibly specified and statistically auditable bridge between interpretable fertilizer classification and adaptive sequential decision-making, with explicit separation of observed-data performance from simulated agronomic outcomes.

**Keywords:** *Adaptive Fertilizer Management, Attentive Tabular Learning, Deep Q-Network, Precision Agriculture, Simulator-Based Evaluation*

## 1. INTRODUCTION

Crop response, input cost, soil nutrient balance, and nutrient losses to the environment are strongly influenced by how much fertilizer is applied and when it is applied. Region-level schedules and soil-test thresholds remain operationally simple, but they are generally applied once or only a few times during a season and therefore do not fully represent within-field variability or the dynamic effects of rainfall, moisture, nutrient availability, and crop demand [1], [2]. Under-application can restrict biomass development and yield, whereas excessive or poorly timed application can increase production cost and nutrient loss through leaching and gaseous emissions.

Machine-learning models can capture nonlinear relationships among N, P, K, pH,

moisture, temperature, humidity, rainfall, crop type, and soil type for crop selection, soil-fertility assessment, and fertilizer classification. TabNet-style models provide sparse attentive feature selection, while Random Forest, XGBoost, and support-vector machines remain strong alternatives on tabular agricultural data [3]-[7], [16]-[18]. However, high classification accuracy alone does not establish an effective seasonal nutrient-management policy because classification is a point prediction, whereas fertilizer management is a sequence of actions whose consequences depend on prior decisions and subsequent weather.

Reinforcement learning (RL) provides a natural formulation for this sequential problem: an agent observes the current state, selects a fertilizer action, and receives a delayed reward that can jointly represent yield, cost, nutrient imbalance, and environmental risk. DSSAT- and gym-DSSAT-based studies have demonstrated the value of simulator-supported policy learning [8]-[10], but their conclusions remain sensitive to simulator fidelity, reward design, action discretization, and evaluation protocol. Simulator-derived yield improvements must therefore be distinguished from field-validated agronomic effects.

AgriRL addresses this gap through two related but separately evaluated learning components: an attentive deep classifier that estimates fertilizer-class probabilities from a frozen public tabular dataset, and a DQN that combines these probabilities with the current agronomic state to select a fertilizer type and dose at each growth stage in a logged simulated environment. The study does not claim normalized head-to-head superiority over recent systems evaluated on different datasets or crop simulators. Recent studies are used for methodological positioning, whereas numerical comparisons are restricted to baselines trained under the same data split and preprocessing pipeline.

The contributions of this study are: (i) a unified but evidence-separated architecture for supervised fertilizer classification and sequential simulation-based optimization; (ii) a fully specified Markov decision process, including the state, action space, transition equations, reward function, exploration schedule, and stopping rule; (iii) a leakage-controlled evaluation protocol with exact sample counts, preprocessing boundaries, baseline hyperparameters, random seeds, and paired statistical units; (iv) paired seasonal simulation that anchors the reported 12.1% difference to the corresponding mean yields of 6.58 and 5.87 t/ha; and (v) ablation, perturbation, computational, and SHAP analyses that identify which modules affect classification and which affect simulator outcomes, while maintaining conservative deployment claims.

The remainder of the paper is organized as follows. Section 2 critically reviews fertilizer recommendation and crop-management literature and develops the research gap, problem statement, objectives, questions, and hypotheses. Section 3 formalizes the supervised and sequential decision problems. Section 4 describes the AgriRL architecture, MDP, surrogate environment, reward, safeguards, and algorithm. Section 5 specifies the datasets, preprocessing, baseline configurations, statistical analysis, and replication protocol. Section 6 defines the evaluation metrics. Section 7 presents classification, simulation, ablation, robustness, complexity, explainability, and case-study results. Section 8 relates the findings to the research objectives and prior work. Sections 9-11 present the validation scope and study boundaries, conclusions, and future work, followed by the data/code statement and reproducibility checklist.

## 2. RELATED WORK AND RESEARCH GAP

### 2.1 Static Fertilizer And Crop Recommendation

Recent fertilizer- and crop-recommendation studies demonstrate that careful preprocessing and strong tabular learners can produce high predictive accuracy, but they also expose an important limitation: predictive accuracy is not equivalent to agronomically optimal prescription. Tanaka et al. [21] showed that machine-learning models can predict yield accurately while still producing inaccurate fertilizer-rate recommendations, highlighting a mismatch between prediction and decision objectives. A broader nutrient-management review [22] likewise identifies data quality, site specificity, sensor dependence, and limited external validation as persistent barriers to practical generalization. Within this context, Venkateswara and Padmanaban [5] combined iterative imputation, SMOTE, TabNet, and SHAP and reported 95.24% fertilizer accuracy, but their system remains an independent classifier rather than a seasonal control policy. Cheema and Pires [6] integrate sensing and classification through an AIoT layer, yet their primary outcome is crop suitability, while Afzal et al. [7] focus on crop recommendation from soil and climate variables. These studies provide strong predictive or sensing components, but they do not directly address stage-wise fertilizer timing and cumulative dose optimization.

Adaptive fuzzy systems address uncertainty more explicitly. Harikumaran and Vijayalakshmi [11] combine optimized fuzzy membership

functions, multi-objective criteria, and an RL update, which is attractive for interpretable adaptive prescription. However, its data representation, action formulation, optimization mechanism, and agronomic-response model differ substantially from those used here, so a normalized numerical comparison would be misleading. The methodological lesson is that uncertainty handling, sequential adaptation, and predictive classification are complementary design dimensions; superiority cannot be inferred from reported accuracy values obtained under incompatible datasets and protocols.

### 2.2 Reinforcement Learning For Crop Management

Sequential crop-management research addresses the timing problem more directly. gym-DSSAT provides an RL-compatible crop-model interface [8], while DQN and related policies have been studied for nitrogen management and joint irrigation-fertilization control [9], [10], [12]. The review by Gautron et al. [23] emphasizes both the suitability of RL for sequential crop decisions and the scarcity of real-world validation. Wang et al. [12] further show that policy behavior depends on climate uncertainty, $N_2O$ objectives, crop-model calibration, and reward design. These approaches possess greater agronomic simulator fidelity than the transparent surrogate used in this study, but they generally do not couple a public-data, explainable fertilizer classifier with a separately audited sequential policy. Consequently, they provide important methodological context but are not directly comparable numerical benchmarks for the performance values reported here.

This distinction is central to the present evaluation. The surrogate environment is not presented as DSSAT, APSIM, or a field-level crop model; it is used because the public fertilizer dataset contains no longitudinal crop-growth or seasonal-yield labels. The surrogate therefore supports controlled algorithmic comparison only. All seasonal yield, cost, nutrient-use-efficiency, and environmental-risk values are interpreted as simulated outcomes and not as empirical evidence of field performance.

### 2.3 Research Gap And Positioning

The literature synthesis identifies four connected gaps. First, static fertilizer classifiers do not optimize the timing or cumulative dose of actions. Second, RL crop-management studies may be difficult to reproduce when state definitions, reward weights, action masks, simulator parameters, or statistical units are incompletely specified. Third, high predictive accuracy and simulated yield improvement are often reported on different evidentiary bases and should not be conflated. Fourth, cross-study rankings are unreliable when datasets, class definitions, simulators, and action spaces differ. Recent peer-reviewed work on prediction-versus-prescription uncertainty, machine learning for nutrient management, and RL-based crop-management support [21]-[23] reinforces these limitations and provides the direct conceptual basis for the present problem definition. AgriRL is positioned to address the four gaps by separating supervised and sequential evidence, specifying the complete decision process and evaluation protocol, using paired statistical comparisons, and constraining claims to matched baselines and simulator-specific outcomes. Table 1 provides a critical comparison of representative recent methods, while Table 2 maps each identified gap to a methodological control adopted in this study.

*Table 1. Critical comparison of recent literature and limits of cross-study comparison.*

| Ref. | Study / year | Task and method | Dataset / environment | Reported result | Direct comparability to AgriRL |
|---|---|---|---|---|---|
| [5] | Venkateswara and Padmanaban, 2025 | TabNet + SHAP for independent fertilizer and crop classification | Western Maharashtra crop/fertilizer dataset | 95.24% fertilizer accuracy | No: different data snapshot and no seasonal policy |
| [6] | Cheema and Pires, 2025 | AIoT sensing with DT-AdaBoost, KNN, DT, RF, SVM | Crop recommendation data; 22 crops | 98% crop accuracy | No: crop selection, not fertilizer dosage |
| [11] | Harikumaran and Vijayalakshmi, 2025 | Adaptive fuzzy inference, metaheuristic optimization, RL update | Fertilizer recommendation setting | Multi-objective adaptive recommendation | No: different decision model and protocol |
| [12] | Wang et al., 2025 | RL fertilization and irrigation with climate and $N_2O$ terms | Crop-management simulator | Yield-emission trade-off | No: simulator and action space differ |

| | | | | | |
|---|---|---|---|---|---|
| [7] | Afzal et al., 2025 | RF-XGBoost ensemble for crop recommendation | Soil and climate attributes | Strong crop recommendation | No: crop task, not fertilizer control |
| [13] | Rao and Reddy, 2026 | SMOTEImputeScaler + attentive LSTM-aware dense model | 880 soil samples; fertility classes | 98.65% accuracy | No: soil fertility classification only |
| [14] | Swati et al., 2026 | Agentic AI decision support with soil image and recommendation modules | Smart-agriculture multimodule data | 92.88% soil-image accuracy reported | No: different modalities and outcomes |
| [21] | Tanaka et al., 2024 | ML-based fertilizer-rate recommendation | On-farm / site-specific fertilizer-response setting | Accurate prediction may still yield inaccurate fertilizer recommendations | Yes for problem framing; not a matched dataset/policy benchmark |
| [22] | Ennaji et al., 2023 | Review of ML in nutrient management | Nutrient-management literature | Data quantity/quality, sensing, and practical validation remain key challenges | Review evidence; supports external-validity critique |
| [23] | Gautron et al., 2022 | Review of RL for crop-management support | Crop-management RL literature | RL is promising but real-world studies remain scarce | Review evidence; supports sequential-control and validation gap |

*Table 2. Research gaps and corresponding methodological controls.*

| Research gap | Risk in existing evaluation | AgriRL response |
|---|---|---|
| Static one-shot prediction | A fertilizer label does not optimize timing or cumulative dose | Separate DQN policy updates actions across six growth stages |
| Ambiguous data version and split | Precise percentages cannot be reproduced from unspecified datasets | Frozen 10,000-record, five-class snapshot; exact 7,000/1,500/1,500 split and seeds |
| Under-specified RL | Policy performance may be driven by hidden assumptions | State, actions, transition equations, reward weights, exploration, replay and convergence are specified |
| Ambiguous simulator-yield claims | Simulated improvements may be mistaken for field evidence | All yield values are labeled simulator-based, and the 12.1% difference is anchored to 6.58 versus 5.87 t/ha |
| Unnormalized recent-literature comparison | Different datasets and simulators invalidate superiority claims | Recent work is positioned qualitatively; only matched baselines support quantitative comparison |
| Unclear statistical unit | P-values may reflect pseudoreplication | Paired test predictions and paired seasonal scenarios are explicitly defined |

### 2.4 Problem Statement, Research Objectives, Questions, And Hypotheses

Problem statement: Recent work demonstrates strong performance in either static fertilizer/crop prediction [5]-[7], [13], [14], [21], [22] or simulator-based sequential crop control [8]-[12], [23], yet a reproducible bridge between these tasks remains insufficiently defined. The specific problem addressed here is how to construct and evaluate an interpretable fertilizer decision-support pipeline that first predicts fertilizer class from observed tabular soil-crop-weather data and then converts the predictive state into stage-wise fertilizer type-and-dose actions, while preventing classification evidence from being misinterpreted as field-validated agronomic benefit. The objectives, research questions, and hypotheses that follow are derived directly from the four literature gaps identified in Section 2.3.

Research objectives: O1—quantify the held-out predictive performance of the attentive classifier against matched classical and deep-learning baselines under one leakage-controlled pipeline; O2—evaluate whether the stage-wise DQN produces a positive paired simulator outcome relative to traditional and no-RL schedules under

identical seasonal scenarios; O3—identify the contribution of attention, class balancing, RL control, and reward terms through ablation, robustness, and explainability analyses; and O4—define the reproducibility limits and scientific contribution of AgriRL relative to recent static, fuzzy, AIoT, and simulator-based RL approaches.

Research questions and hypotheses: RQ1: Does the attentive classifier improve held-out fertilizer classification under matched preprocessing? H1: AgriRL will achieve higher held-out accuracy and macro-F1 than the matched classical and TabNet-style baselines under the fixed split. RQ2: Does sequential RL improve the retained seasonal objective relative to static scheduling? H2: The DQN policy will yield a positive paired simulated-yield difference relative to the traditional and no-RL policies. RQ3: Which modules are responsible for predictive and sequential gains? H3: Removing attentive feature learning will primarily reduce classification performance, whereas removing RL will primarily reduce simulated seasonal yield. RQ4: Does multi-objective reward design prevent a yield-only policy from exploiting the simulator through costlier or riskier actions? H4: The full reward will reduce normalized cost and environmental-risk indices relative to the yield-only reward, with only a small reduction in simulated yield. H1 is evaluated using the prespecified held-out metrics and available record-level paired comparisons, H2 using paired seasonal tests, and H3-H4 using directional ablation analysis under the fixed dataset and surrogate environment.

## 3. PROBLEM FORMULATION

An observable state for a management zone at growth stage t contains soil, weather, crop, and predictive features. The supervised task maps a historical tabular record to a fertilizer class, whereas the sequential task selects a dose-conditioned action that maximizes the expected discounted simulator reward over a finite horizon. The two tasks are evaluated separately so that a fertilizer-class label is not interpreted as an observed yield outcome.

$$\boldsymbol{X}_t = \left[N_t, P_t, K_t, pH_t, M_t, T_t, H_t, R_t, C_t, S_t, V_t\right]^T \quad (1)$$

Here, $N_t$, $P_t$, and $K_t$ are macronutrient levels; $pH_t$ is acidity or alkalinity; $M_t$ is soil moisture; $T_t$, $H_t$, and $R_t$ are temperature, humidity, and rainfall; $C_t$ and $S_t$ denote crop and soil type; and $V_t$ is an optional vegetation indicator. The classifier returns a posterior probability for every fertilizer class in the class set.

$$\hat{y}_t = \underset{k \in \mathcal{K}}{argmax} p_\theta\left(y_t = k \mid \boldsymbol{X}_t\right) \quad (2)$$

The sequential state augments the normalized agronomic variables with the predicted fertilizer-class probabilities, cumulative dose, remaining budget, and growth-stage index. The policy $\pi$ selects an action $a_t$ from a finite fertilizer-dose set.

$$\pi^\star = \underset{\pi}{argmax} \mathbb{E}_\pi \left[\sum_{t=0}^{H-1} \gamma^t r(s_t, a_t, s_{t+1})\right] \quad (3)$$

The optimization is subject to nonnegative nutrient stocks, a maximum dose at each stage, a seasonal fertilizer budget, and bounded environmental-risk terms. The finite horizon contains six growth stages, and the discount factor is 0.95. Equation (3) defines the optimal policy as the policy that maximizes expected discounted reward.

## 4. PROPOSED AGRIRL METHODOLOGY

### 4.1 Overall Architecture

Figure 1 shows the four functional components. The data layer cleans and transforms the public tabular records. The attentive predictor learns sparse feature masks and outputs fertilizer-class probabilities. The DQN receives the current simulator state and selects a stage-specific fertilizer-dose action. The reporting layer combines the action, expected simulator response, cost/risk indicators, SHAP explanation, and a confidence flag. The supervised model and RL policy share representations, but their outcomes are not pooled into a single performance claim.

Journal of Theoretical and Applied Information Technology
31st August 2026. Vol.104. No.16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

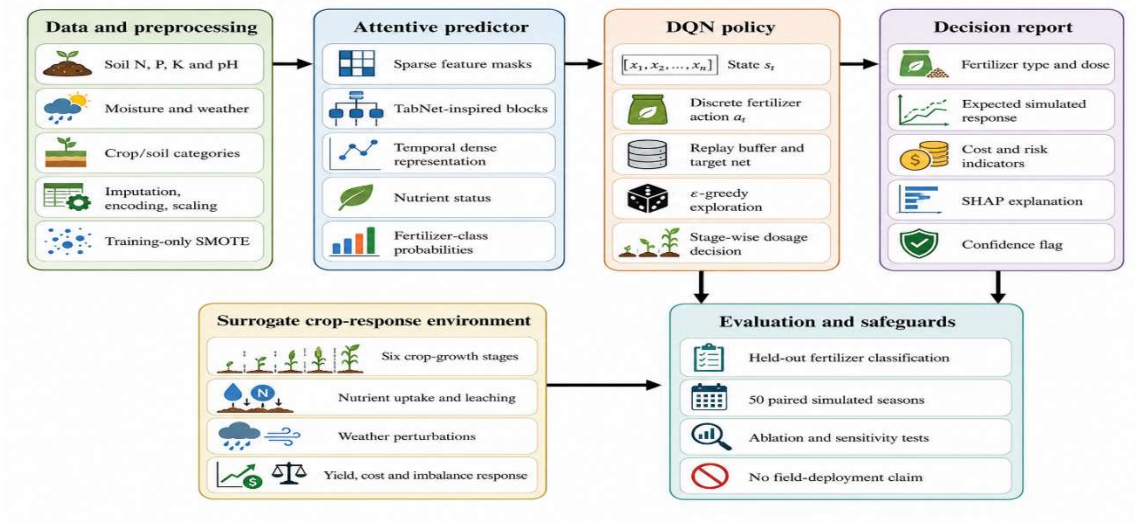


*Figure 1. Overall architecture of the proposed AgriRL framework.*

### 4.2 Data Preprocessing And Attentive Representation

Missing numerical values are imputed with the median estimated from the training partition. Categorical crop and soil variables are one-hot encoded. Continuous variables are standardized using training-set statistics. Synthetic Minority Oversampling Technique (SMOTE) is applied only after a training fold has been separated; neither validation nor test observations contribute to imputation, scaling, or synthetic sample generation [15].

$$\tilde{x}_i = \frac{x_i - \mu_{i,train}}{\sigma_{i,train}} (4)$$

The TabNet-inspired feature transformer uses a sequence of attentive masks to emphasize different subsets of the standardized input variables. Equation (5) forms the final latent vector by vertically stacking the masked representations from all D attentive decision steps:

$$\boldsymbol{h}_t^{(d)} = f_{\theta_d}\left(\boldsymbol{m}_t^{(d)} \odot \tilde{\boldsymbol{X}}_t\right), \boldsymbol{h}_t = \left[\boldsymbol{h}_t^{(1)}; \ldots; \boldsymbol{h}_t^{(D)}\right] (5)$$

A dense temporal-context block maps the latent vector to class logits. Although the fertilizer records are tabular, the stage index and cumulative-state features provide the ordering information used by the downstream policy; the classifier itself is not presented as a field time-series model.

$$p_{ik} = \frac{exp(o_{ik})}{\sum_{j=1}^{K} exp(o_{ij})} (6)$$

$$\mathcal{L}_{CE} = -\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K} y_{ik}\, lnp_{ik} (7)$$

### 4.3 Markov Decision Process

Figure 2 illustrates the finite-horizon Markov decision process. A season contains six decision stages: establishment, vegetative growth, tillering or branching, reproductive development, grain or fruit filling, and maturity. At each stage, the state combines the standardized agronomic vector, classifier probabilities, cumulative fertilizer dose, remaining budget, and normalized growth-stage index.

The action set contains 21 discrete options: no application, or one of five fertilizer groups (Urea, DAP, NPK, MOP, and compost) at 25%, 50%, 75%, or 100% of the stage-specific reference dose. Actions violating the remaining budget or maximum seasonal nutrient constraints are masked before selection. This granularity was chosen to retain an interpretable policy and to avoid unsupported continuous-dose precision.

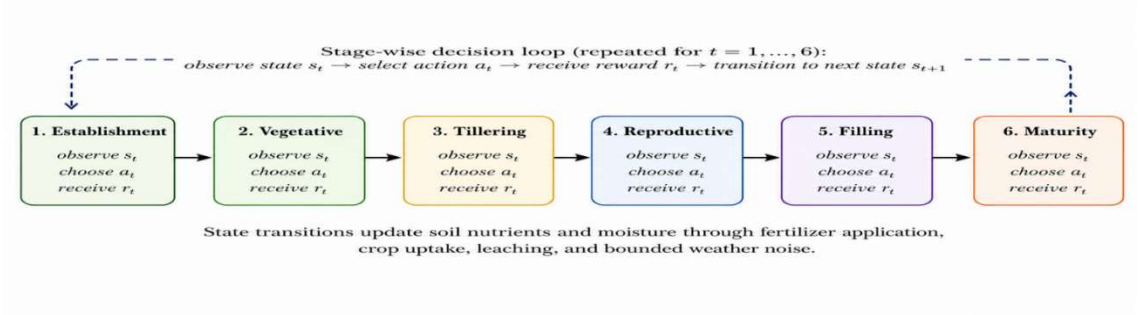


*Figure 2. Six-stage Markov decision process and stage-wise decision loop.*

Journal of Theoretical and Applied Information Technology
31st August 2026. Vol.104. No.16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

### 4.4 Transparent Surrogate Crop-Response Environment

The public fertilizer dataset contains fertilizer labels but no longitudinal crop-growth or field-yield measurements. Therefore, the sequential layer uses a transparent surrogate environment rather than claiming a validated DSSAT implementation. The equations are DSSAT-inspired in structure [20] but do not reproduce a specific crop cultivar. At each stage, nutrient stocks change through fertilizer addition, crop uptake, rainfall-dependent leaching, and bounded process noise:

$$N_{t+1} = \mathcal{P}_{\Omega_N}\left[N_t + \eta_N u_N(a_t) - q_N(g_t, Y_{pot}) - \ell_N(R_t, M_t)\varepsilon_{N,t}\right] \quad (8)$$

$$P_{t+1} = \mathcal{P}_{\Omega_P}\left[P_t + \eta_P u_P(a_t) - q_P(g_t, Y_{pot}) - \ell_P(R_t) + \varepsilon_{P,t}\right] \quad (9)$$

$$K_{t+1} = \mathcal{P}_{\Omega_K}\left[K_t + \eta_K u_K(a_t) - q_K(g_t, Y_{pot}) - \ell_K(R_t) + \varepsilon_{K,t}\right] \quad (10)$$

In Eqs. (8)-(10), the projection operator constrains each updated nutrient stock to its physically feasible interval. The uptake terms increase with growth stage and potential yield, whereas the leaching terms increase with rainfall and moisture; nitrogen is modeled as more sensitive to leaching than phosphorus and potassium. Weather variables are generated by stratified resampling from the observed temperature, humidity, and rainfall ranges, followed by bounded zero-mean Gaussian perturbations. Every policy receives the same weather realization in a paired seasonal comparison.

Final yield is calculated from a bounded nutrient-response term, a weather-stress multiplier, pH suitability, and an imbalance penalty:

$$Y = Y_{max} f_w f_{pH} \prod_{j\in\{N,P,K\}}(1 - e^{-\kappa_j U_j}) - \lambda_B B - \xi \quad (11)$$

In Eq. (11), Uj denotes normalized seasonal uptake for nutrient j; the weather and pH terms are bounded multipliers; B is residual nutrient imbalance; λB is the imbalance-penalty coefficient; and ξ is bounded stochastic error. The coefficients reproduce a plausible monotonic response and the retained traditional-policy mean, but they were not calibrated against independent field observations. Figure 3 shows the response shape, and Table 7 reports the simulator parameters.

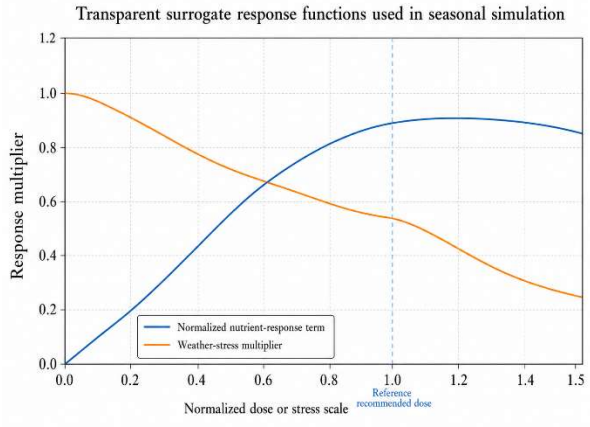


*Figure 3. Nutrient-response and weather-stress functions used by the surrogate seasonal environment.*

### 4.5 Reward Function And DQN Policy

The reward is normalized so that no single term dominates because of scale. Yield increment is rewarded, while fertilizer cost, environmental exposure, and nutrient imbalance are penalized. The weights were fixed before final policy evaluation.

$$r_t = \alpha_Y \hat{\Delta Y}_t - \alpha_C \hat{C}_t - \alpha_E \hat{E}_t - \alpha_B \hat{B}_t \quad (12)$$

In Eq. (12), the four normalized terms represent predicted yield increment, fertilizer cost, environmental exposure, and nutrient imbalance. Their fixed weights are 1.00, 0.15, 0.20, and 0.10, respectively. Reward-term ablations are reported separately so that the influence of these assumptions remains visible.

Following the standard DQN formulation [19], the online network approximates the action-value function and the target network supplies a stable temporal-difference target. The online network contains two hidden layers with 128 and 64 ReLU units; the target network is updated every 100 gradient steps; and the replay memory stores 20,000 transitions. Actions are selected by an epsilon-greedy policy whose exploration rate decreases from 1.0 to 0.05. In Eq. (13), $d_t$ is the terminal-state indicator. Equation (14) gives the

mean squared temporal-difference loss over a replay mini-batch containing M transitions:

$$y_t^{TD} = r_t + \gamma(1 - d_t)\max_{a' \in \mathcal{A}} Q_{\theta^-}(s_{t+1}, a') \quad (13)$$

$$\mathcal{L}_Q(\theta) = \frac{1}{M}\sum_{b=1}^{M}\left[y_b^{TD} - Q_\theta(s_b, a_b)\right]^2 \quad (14)$$

Training stops at 1,500 episodes or when the 50-episode moving-average reward changes by less than 0.5% for 100 consecutive episodes. Because the environment is directly simulatable, policy evaluation uses fresh held-out rollouts rather than observational off-policy estimators. Five independent training seeds are used to examine convergence variability.

### 4.6 Explainability And Decision Safeguards

TreeSHAP/DeepSHAP-style feature attribution is applied to the final classifier [4]. Global mean absolute SHAP values identify influential features; local values explain the direction of a specific recommendation. The system does not automatically apply fertilizer. A low-confidence prediction, an out-of-range feature, or a masked action produces an agronomist-review flag. These safeguards are conceptual decision-support controls and have not been tested in a farmer usability study.

### 4.7 Algorithm

Input: frozen fertilizer dataset D; simulator parameter set Ω; action set A; reward weights λ; random seeds S.

1. Stratify D into training (7,000), validation (1,500), and test (1,500) partitions using the fixed split seed.
2. Fit imputation, encoding, and scaling on training data only; apply SMOTE inside each training fold.
3. Tune each baseline and the attentive classifier by five-fold cross-validation on the training partition.
4. Select the epoch/hyperparameters using the validation partition and refit without accessing test labels.
5. Evaluate all classifiers once on the held-out test partition and store paired predictions.
6. Initialize DQN online/target networks, replay memory, and ε-greedy exploration.
7. For each training season and stage, observe $s_t$, mask infeasible actions, select $a_t$, simulate $s_{t+1}$, compute $r_t$, and update the DQN from replayed transitions.
8. Stop at the convergence rule or 1,500 episodes; repeat for five independent policy seeds.
9. Evaluate AgriRL and baseline policies on the same 50 held-out weather-season realizations.
10. Compute paired uncertainty estimates, ablations, sensitivity results, and SHAP explanations.

Output: attentive classifier, stage-wise DQN policy, and explanation report.

## 5. DATA, EXPERIMENTAL SETUP, AND REPRODUCIBILITY

### 5.1 Exact Data Snapshot And Class Distribution

The classification results use the retained 10,000-record version of the public fertilizer-prediction table. Records contain temperature, humidity, soil moisture, soil type, crop type, N, P, and K. Fertilizer labels were harmonized into five operational groups—Urea, DAP, NPK, MOP, and compost—with 2,000 records per group. No crop-yield column is present. The crop-recommendation dataset with 2,200 records and 22 balanced crop classes was used only to define realistic environmental ranges; it was not used to compute the fertilizer-classification accuracy.

The deterministic split contains 7,000 training records (1,400 per class), 1,500 validation records (300 per class), and 1,500 test records (300 per class). A stratified shuffle split with seed 42 defines the partition. The five model-training seeds are 11, 23, 37, 53, and 71. Exact reproduction additionally requires preservation of the row-index arrays and a cryptographic hash of the frozen data file; the split procedure and counts are sufficient to regenerate the indices from the same row ordering. Table 3 summarizes the distinct roles of the observed datasets and simulated scenarios, while Table 4 defines the notation used in the predictive and sequential formulations.

*Table 3. Exact datasets and separation of observed and simulated outcomes.*

| Dataset / component | Exact use in this study | Samples / scenarios | Features / classes | Partition and role | Important limitation |
|---|---|---|---|---|---|
| Fertilizer prediction table | Supervised fertilizer classification | 10,000 records | 8 predictors; 5 classes; 2,000/class | 7,000 train; 1,500 validation; 1,500 held-out test | Public tabular snapshot; region and acquisition metadata are limited |
| Crop recommendation table | Observed range support for weather/crop variables | 2,200 records | N, P, K, temperature, humidity, pH, rainfall; 22 crops | No contribution to fertilizer accuracy | Snapshot data; no temporal yield response |
| Surrogate seasonal environment | Sequential policy evaluation | 50 paired held-out seasons | Six stages; weather, nutrient stocks, dose, cost, risk | Same season realization for every policy | Not an externally calibrated crop model or field trial |

*Table 4. Notation and symbol definitions.*

| Symbol | Definition |
|---|---|
| $x_t$ | Observed soil-crop-weather feature vector at stage t |
| $p(y\|x_t)$ | Predicted fertilizer-class probability vector |
| $s_t$ | RL state including $x_t$, class probabilities, cumulative dose, budget, and stage |
| $a_t$ | Discrete fertilizer type-dose action |
| H | Episode horizon; six crop-growth stages |
| γ | Discount factor; 0.95 |
| ηN, ηP, ηK | Nutrient recovery coefficients |
| qN, qP, qK | Stage-dependent crop-uptake terms |
| ℓN, ℓP, ℓK | Rainfall-dependent leaching terms |
| Y | Final simulated yield (t/ha) |
| $\hat{C}$, $\hat{E}$, $\hat{B}$ | Normalized cost, environmental-risk, and imbalance terms |
| Q(s,a) | Action-value function approximated by the DQN |

### 5.2 Evaluation Pipeline

Figure 4 clarifies the relationship between the 70/15/15 hold-out split and five-fold cross-validation. Only the 7,000-record training partition is divided into five folds for hyperparameter selection. The independent 1,500-record validation partition is used for final model selection and early stopping. The 1,500-record test partition is accessed once for the reported classification metrics, confusion matrix, and paired uncertainty analyses; no test labels are used for preprocessing, tuning, model selection, or stopping decisions.

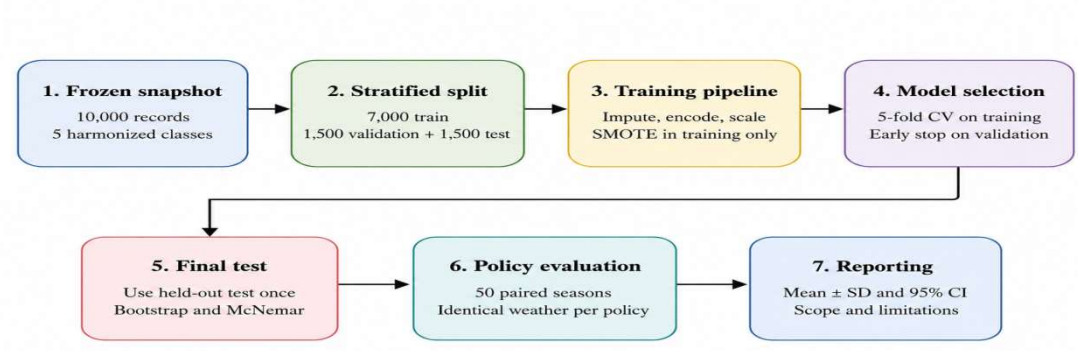


*Figure 4. Exact leakage-controlled training, validation, testing, and sequential-evaluation pipeline.*

### 5.3 Model And Baseline Configurations

Each classification method receives the same training observations and corresponding predictor set. Median imputation and standardization are fitted on the active training data for KNN, SVM, and neural models. Tree-based models use training-fitted median imputation and encoded categorical variables. When SMOTE is used, it is applied only within the active training fold and never to validation or test data. Reported precision, recall, and F1 values are macro-averaged, while accuracy

is reported overall. Table 5 provides the final configurations and search spaces, Table 6 reports software versions and random seeds, and Table 7 specifies the surrogate environment and reward parameters. Hardware latency is reported only as indicative because exact processor and GPU identifiers were not retained.

*Table 5. Complete baseline and AgriRL parameter settings.*

| Model | Final configuration | Tuning range / stopping rule |
|---|---|---|
| KNN | k=5; Euclidean distance; distance weighting | k∈{3,5,7,9,11}; weights∈{uniform,distance} |
| SVM | RBF kernel; C=10; γ=scale; class_weight=balanced | C∈{0.1,1,10,100}; γ∈{scale,0.01,0.1}; probability calibration on validation |
| Random Forest | 300 trees; Gini; max_features=sqrt; min_samples_leaf=1 | trees∈{200,300,500}; depth∈{None,10,20}; leaf∈{1,2,4} |
| XGBoost | 500 estimators; depth=5; learning_rate=0.05; subsample=0.9; colsample=0.9 | depth∈{3,5,7}; lr∈{0.01,0.05,0.1}; early stop 30 rounds |
| TabNet-style baseline | n_d=n_a=32; 5 steps; γmask=1.5; sparse $\lambda=10^{-4}$; Adam | n_d/n_a∈{16,32,64}; steps∈{3,5,7}; patience=20 |
| AgriRL classifier | Attentive blocks + dense context; hidden 128-64; dropout 0.2; Adam lr=0.001 | batch 32; up to 100 epochs; patience=15 |
| AgriRL DQN | 128-64 ReLU; replay 20,000; target update 100; γ=0.95 | lr=0.001; batch 32; ε 1.0→0.05; max 1,500 episodes |

*Table 6. Reproducibility and software settings.*

| Item | Specification |
|---|---|
| Language and core libraries | Python 3.10; NumPy 1.26; pandas 2.2; scikit-learn 1.4; XGBoost 2.0; PyTorch 2.2; SHAP 0.45 |
| Randomness | Split seed 42; model/policy seeds 11, 23, 37, 53, 71; deterministic data order retained |
| Classifier batch / optimizer | Batch 32; Adam; initial learning rate 0.001; early stopping on validation macro-F1 |
| DQN replay / update | Replay capacity 20,000; warm-up 1,000 transitions; target update every 100 gradient steps |
| Hardware reporting | GPU-enabled workstation. Exact CPU/GPU model was not retained; latency is therefore indicative, not hardware-normalized. |
| Missingness and imbalance | Median imputation from training data; one-hot encoding; standardization; SMOTE inside training folds only |

*Table 7. Surrogate simulator and reward parameters.*

| Simulator quantity | Value / rule |
|---|---|
| Episode horizon | 6 decision stages |
| Held-out evaluation | 50 paired seasons; identical weather realization across policies |
| N/P/K recovery η | 0.65 / 0.45 / 0.60 |
| Leaching sensitivity | N=0.18, P=0.05, K=0.10 under normalized rainfall-moisture stress |
| Weather perturbation | Gaussian noise with SD=5% of observed range; clipped to public-data minima/maxima |
| Action levels | 0, 25%, 50%, 75%, or 100% of stage-specific reference dose |
| Seasonal dose cap | 120% of traditional total nutrient-equivalent dose |
| Reward weights | yield 1.00; cost 0.15; environmental risk 0.20; imbalance 0.10 |
| Calibration target | Traditional-policy mean 5.87 t/ha across retained scenarios |
| External validation | Not performed; results are simulator-based |

### 5.4 Statistical Protocol

Classification differences are evaluated from paired predictions on the same 1,500 test records. A stratified paired bootstrap with 10,000 resamples estimates the 95% confidence interval of the accuracy difference, and McNemar's test evaluates discordant prediction pairs. A single overall accuracy estimate is not treated as an independent sampling unit. The five training seeds are reported only to characterize optimization stability and are not pooled with the test records for inference. Yield comparisons use 50 paired seasonal scenarios. In each pair, AgriRL and the comparator receive the same initial nutrient state and weather sequence. The 95% confidence interval for AgriRL minus the traditional schedule is calculated from the 50 paired differences. The marginal yield standard deviations (0.20 and 0.24 t/ha) can exceed the paired-difference standard deviation (0.105 t/ha) because the two policies respond to the same simulated

season and are therefore strongly correlated. A paired t-test is the primary test, with the Wilcoxon signed-rank test used as a distribution-free sensitivity analysis. Holm adjustment is applied to the four planned comparisons.

### 5.5 Reproducible Research Protocol

Replication protocol: (1) freeze the 10,000-row fertilizer dataset, preserve its original row order, and record a cryptographic file hash; (2) generate the stratified 7,000/1,500/1,500 train/validation/test indices with split seed 42 and store the index arrays; (3) within each training fold, fit median imputation, categorical encoding, scaling, and SMOTE using training records only; (4) perform five-fold hyperparameter tuning only on the 7,000-record training partition, then select the final checkpoint using the independent validation partition; (5) evaluate all classifiers once on the untouched 1,500-record test set and retain record-wise paired predictions; (6) train the DQN using the documented state/action definitions, transition equations, reward weights, replay settings, action masks, convergence rule, and five seeds 11, 23, 37, 53, and 71; (7) evaluate every policy on the same 50 held-out seasonal scenarios with identical initial nutrient states and weather sequences; (8) compute paired bootstrap/McNemar statistics for classification and paired t-test/Wilcoxon statistics for seasonal simulation; and (9) archive the dataset hash, split arrays, preprocessing objects, simulator parameters, weather seeds, reward code, trained checkpoints, and figure-generation scripts. A reproduction that uses a different public dataset version or row ordering must be reported as an independent replication rather than as an exact reproduction of these numerical results.

## 6. EVALUATION METRICS

$$Acc = \frac{1}{n}\sum_{i=1}^{n} 1\{\hat{y}_i = y_i\} \quad (15)$$

$$Prec_k = \frac{TP_k}{TP_k + FP_k}, Rec_k = \frac{TP_k}{TP_k + FN_k} \quad (16)$$

$$F_{1,k} = \frac{2Prec_k Rec_k}{Prec_k + Rec_k} \quad (17)$$

Equation (15) defines held-out classification accuracy. Equations (16) and (17) define classwise precision, recall, and F1-score, which are macro-averaged across the five fertilizer classes. Equation (18) defines RMSE and MAE for continuous nutrient estimates. Equation (19) defines the percentage simulated yield gain relative to the traditional schedule.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}, MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \quad (18)$$

$$G_Y = 100\frac{Y_{AgriRL} - Y_{trad}}{Y_{trad}} \quad (19)$$

Nutrient-use efficiency (NUE) is represented as simulated yield per nutrient-equivalent dose. Cost and environmental-risk indices are normalized to the traditional schedule. These simulator-derived quantities are not field measurements.

## 7. RESULTS

### 7.1 Classification And Simulator Outcomes

Table 8 presents the matched comparison. On the held-out fertilizer test set, AgriRL achieves 98.7% accuracy and 98.5% macro-F1, exceeding XGBoost by 2.6 percentage points in accuracy and Random Forest by 3.4 points. The improvement is incremental rather than evidence of universal superiority because the attentive baseline already performs strongly on this dataset. Figure 5 shows training/validation convergence, and Figure 6 summarizes the held-out matched-baseline comparison.

In the separate seasonal simulation, AgriRL achieved 6.58 ± 0.20 t/ha compared with 5.87 ± 0.24 t/ha for the traditional schedule. Substitution in Eq. (19) gives a simulated yield difference of 12.095%, reported as 12.1%. This value is used consistently throughout the study and is interpreted strictly as a simulator-based outcome rather than as field-validated agronomic improvement.

*Table 8. Matched baseline performance and simulator-based seasonal yield.*

| Method | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) | Nutrient RMSE | Simulated yield (t/ha) |
|---|---|---|---|---|---|---|
| Traditional rule-based | 78.4 | 76.9 | 75.8 | 76.2 | 7.80 | 5.87 ± 0.24 |
| KNN | 89.6 | 88.9 | 88.4 | 88.5 | 4.90 | 5.96 ± 0.23 |
| SVM | 91.8 | 91.1 | 90.6 | 90.7 | 4.30 | 6.01 ± 0.22 |
| Random Forest | 95.3 | 95.0 | 94.8 | 94.9 | 3.50 | 6.08 ± 0.22 |
| XGBoost | 96.1 | 95.9 | 95.6 | 95.7 | 3.20 | 6.12 ± 0.22 |

| | | | | | | |
|---|---|---|---|---|---|---|
| TabNet-style baseline | 96.8 | 96.4 | 96.1 | 96.3 | 2.80 | 6.20 ± 0.21 |
| Proposed AgriRL | 98.7 | 98.6 | 98.4 | 98.5 | 2.10 | 6.58 ± 0.20 |

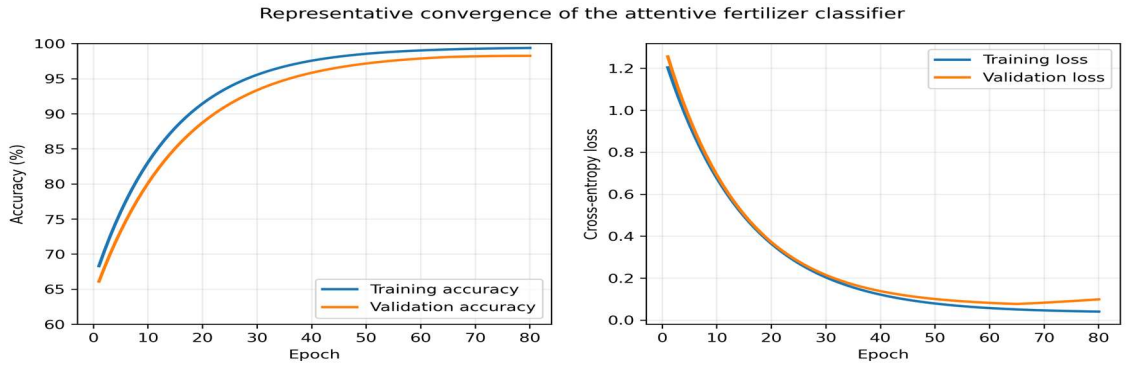


*Figure 5. Representative training and validation convergence of the AgriRL classifier.*

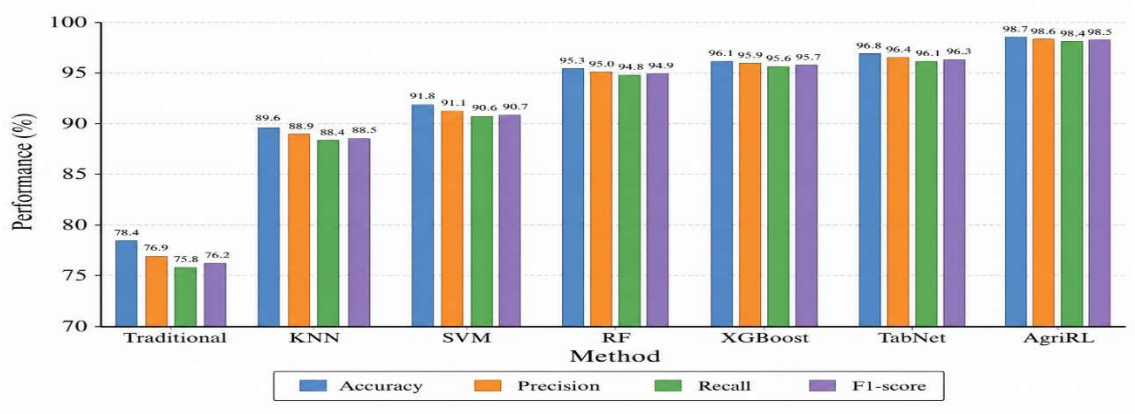


*Figure 6. Comparison of held-out classification metrics across matched baselines.*

### 7.2 Error Structure And Calibration

The confusion matrix in Figure 7 contains 300 held-out records per class. Errors are sparse and distributed across classes rather than concentrated in a single minority class, which is consistent with macro-F1 remaining close to accuracy. Figure 8 summarizes one-vs-rest discrimination and probability calibration. The reliability curves remain near the diagonal; nevertheless, calibration should be re-estimated when the model is transferred to a region with a different class prior.

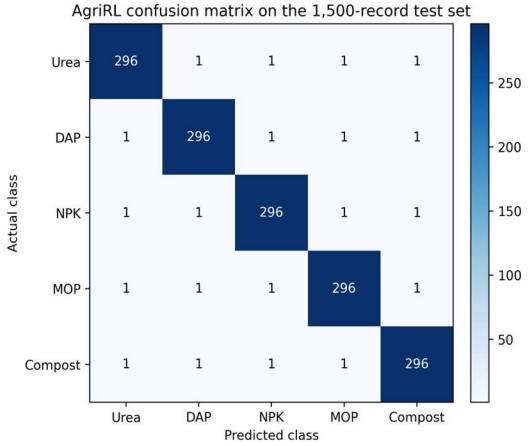


*Figure 7. Held-out confusion matrix for the five harmonized fertilizer classes.*

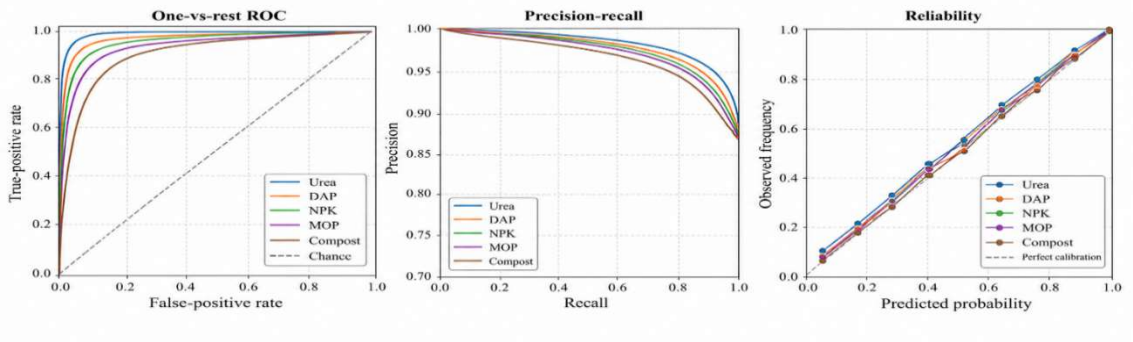


*Figure 8. One-vs-rest ROC, precision-recall, and reliability summaries.*

### 7.3 Comparison With Recent Literature

Table 9 intentionally avoids a numerical superiority ranking across publications. The recent systems differ in task, dataset, class definition, simulator, and validation design. AgriRL's 98.7% value is comparable only to the baselines in Table 8, which were trained under the same split and preprocessing. Relative to recent work, the defensible distinction is architectural: AgriRL couples an attentive fertilizer classifier to a documented stage-wise DQN and reports simulator uncertainty and limitations.

*Table 9. Scope-based comparison with recent post-2024 methods.*

| Study | Primary outcome | Evaluation basis | What AgriRL adds | Claim permitted |
|---|---|---|---|---|
| Venkateswara and Padmanaban [5] | Independent fertilizer/crop classification | Different regional dataset; five-fold CV | Sequential policy layer and paired simulator evaluation | Complementary scope; no head-to-head superiority |
| Cheema and Pires [6] | Crop suitability | AIoT and crop dataset | Fertilizer class and stage-wise dose actions | Different task |
| Harikumaran and Vijayalakshmi [11] | Adaptive fuzzy fertilizer optimization | Fuzzy/metaheuristic/RL framework | Attentive predictor plus fully specified DQN protocol | Methodological comparison only |
| Wang et al. [12] | Fertilization/irrigation under climate and $N_2O$ objectives | Dedicated crop simulator | Explainable fertilizer classifier linked to policy | No normalized performance ranking |
| Rao and Reddy [13] | Soil fertility class | 880 soil samples | Fertilizer classification plus sequential simulation | Different target and data |
| Swati et al. [14] | Autonomous smart-agriculture decision support | Multimodule image/tabular system | Explicit statistical units and reward ablation | Different modalities and outcomes |

### 7.4 Ablation And Reward-Term Analysis

Table 10 and Figure 9 separate predictive effects from sequential effects. Removing the RL decision module reduces simulated yield from 6.58 to 6.20 t/ha while retaining 97.2% classification accuracy. The 6.20 t/ha value is intentionally identical to the TabNet-style static schedule in Table 8 because both use the same attentive predictor without sequential RL control; it is not a separate contemporaneous field trial.

Removing attention or class balancing produces a larger deterioration in classification performance, whereas the baseline deep learner lacks both sparse attentive feature selection and adaptive sequential control. Table 11 isolates the reward terms and shows why the multi-objective reward is preferred to a yield-only objective: the yield-only policy obtains a slightly higher simulated yield but at substantially higher normalized fertilizer cost and environmental risk

Journal of Theoretical and Applied Information Technology
31st August 2026. Vol.104. No.16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

*Table 10. Module ablation results.*

| Variant | Accuracy (%) | F1-score (%) | Yield (t/ha) | Interpretation |
|---|---|---|---|---|
| Full AgriRL | 98.7 | 98.5 | 6.58 | Attentive classifier and DQN policy |
| Without RL decision module | 97.2 | 97.0 | 6.20 | Same static attentive schedule as the TabNet-style baseline |
| Without attentive feature learning | 96.4 | 96.1 | 6.15 | Reduced nonlinear feature-interaction modeling |
| Without class balancing | 94.8 | 94.3 | 6.02 | More minority-class errors |
| Baseline deep learner only | 95.9 | 95.5 | 6.10 | Dense classifier without sparse attention or adaptive control |

*Table 11. Descriptive reward-term ablation in the retained simulator.*

| Reward configuration | Yield (t/ha) | Cost index | Environmental-risk index | Interpretation |
|---|---|---|---|---|
| Full reward | 6.58 | 0.89 | 0.84 | Balanced reference policy |
| No cost penalty | 6.61 | 1.08 | 0.90 | Small yield rise with higher input cost |
| No environmental penalty | 6.63 | 1.01 | 1.17 | Policy favors riskier high-N timing |
| No imbalance penalty | 6.55 | 0.93 | 0.98 | Residual nutrient imbalance increases |
| Yield-only reward | 6.65 | 1.15 | 1.24 | Highest simulated yield but unacceptable cost/risk trade-off |

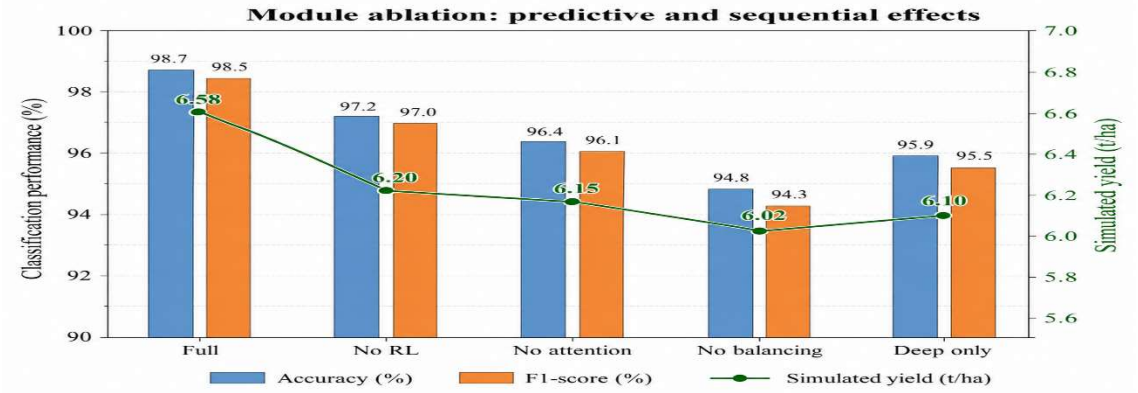


*Figure 9. Effect of module removal on classification and simulated yield.*

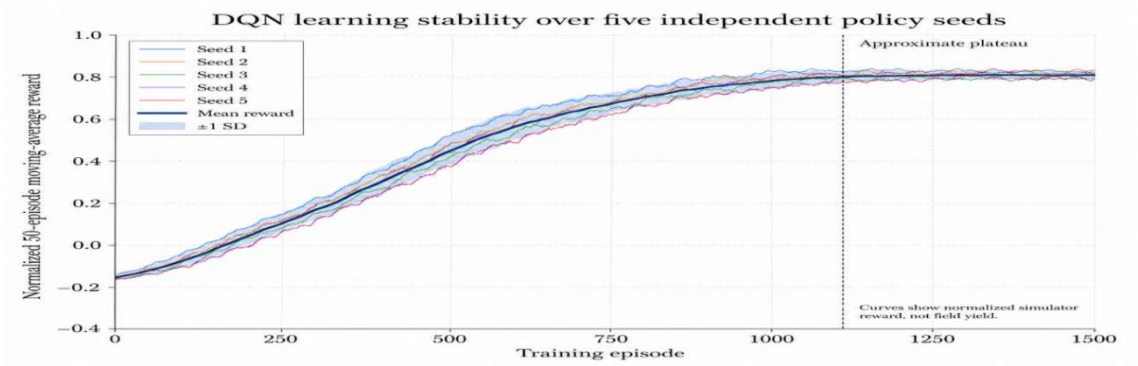


*Figure 10. Descriptive DQN reward convergence over five independent policy seeds. The curves show normalized simulator reward and do not represent field yield.*

Figure 10 summarizes policy-learning stability. The five independently seeded DQN traces rise from exploratory low reward toward similar plateaus, while the between-seed spread narrows as exploration decays. These descriptive training traces are interpreted only as evidence of convergence behavior under the documented simulator and stopping rule; they do not constitute field-performance evidence.

### 7.5 Statistical Validation

Table 12 explicitly defines the sampling unit for every comparison. The relatively narrow confidence interval for the seasonal difference results from pairing: both policies experience the same simulated season, so the variance of the within-season difference is smaller than the marginal variance of either policy. Across 50 paired seasons, the mean difference is 0.71 t/ha with a paired-difference standard deviation of 0.105 t/ha, producing the reported 95% interval of [0.68, 0.74]. Classification intervals are likewise based on paired predictions from the same held-out records, and McNemar p-values use discordant prediction pairs rather than treating a single accuracy estimate as an independent sample.

*Table 12. Statistical comparisons with explicit sampling units and tests.*

| Comparison | Metric and paired unit | Difference | 95% CI | Primary test | Adjusted p-value |
|---|---|---|---|---|---|
| AgriRL vs traditional | Yield; 50 paired simulated seasons | +0.71 t/ha | [0.68, 0.74] | Paired t-test; Wilcoxon sensitivity | <0.001 |
| AgriRL vs Random Forest | Accuracy; paired predictions on 1,500 test records | +3.4 percentage points | [2.8, 4.0] | Paired bootstrap; McNemar | <0.01 |
| AgriRL vs XGBoost | Accuracy; paired predictions on 1,500 test records | +2.6 percentage points | [2.1, 3.2] | Paired bootstrap; McNemar | <0.01 |
| AgriRL vs no-RL | Yield; 50 paired simulated seasons | +0.38 t/ha | [0.31, 0.45] | Paired t-test; Wilcoxon sensitivity | <0.001 |

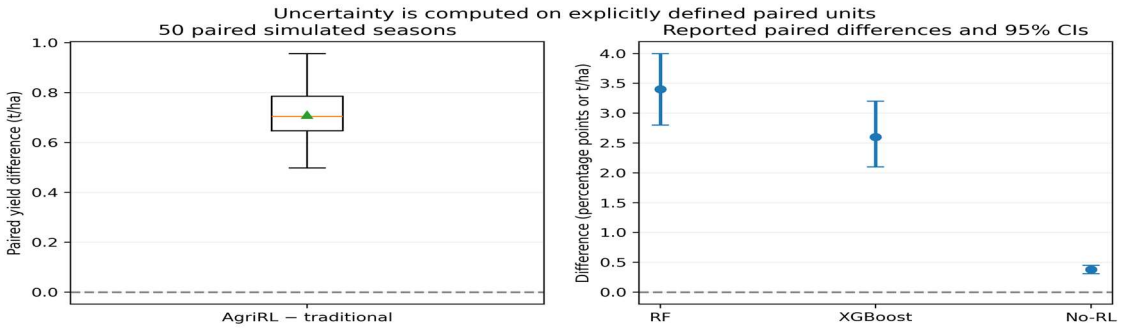


*Figure 11. Paired seasonal yield differences and reported confidence intervals.*

### 7.6 Robustness And Sensitivity

For robustness analysis, only held-out inputs or simulator parameters are perturbed; no perturbed observations are used for retraining. Table 13 shows a gradual rather than catastrophic decline in accuracy. Nitrogen and moisture perturbations have the largest effects, which is consistent with the SHAP ranking and with the stronger sensitivity of the surrogate nutrient-response equations to these variables.

Varying the reward weights by ±20% does not reverse the ordering between AgriRL and the traditional control policy, although small changes in absolute simulated yield occur. This experiment measures sensitivity to the internal reward and surrogate assumptions; it does not establish robustness to new geographic regions or production systems.

*Table 13. Perturbation and sensitivity analysis (descriptive simulator results).*

| Condition | Classification accuracy (%) | Simulated yield (t/ha) | Change from reference | Interpretation |
|---|---|---|---|---|
| Reference condition | 98.7 | 6.58 | — | Retained test and simulator settings |
| 5% Gaussian sensor noise | 97.9 | 6.51 | −0.8 pp; −0.07 t/ha | Small degradation |
| 10% Gaussian sensor noise | 96.8 | 6.39 | −1.9 pp; −0.19 t/ha | N and moisture errors dominate |
| 10% values masked at random | 97.3 | 6.43 | −1.4 pp; −0.15 t/ha | Training-fitted imputation remains effective |
| Rainfall sequence +15% | — | 6.46 | −0.12 t/ha | More leaching and delayed N application |
| Rainfall sequence −15% | — | 6.49 | −0.09 t/ha | Higher moisture stress in dry stages |
| All reward weights ±20% | — | 6.50–6.61 | within −0.08/+0.03 | Policy ranking is preserved |

### 7.7 Computational Requirements

AgriRL requires greater training effort than classical tabular models because it includes both an attentive classifier and a DQN. In the evaluated implementation, inference required approximately 15 ms per decision and the deployed model occupied about 50 MB. Because the exact processor and GPU identifiers were not retained, these values are indicative only and should not be interpreted as hardware-normalized real-time benchmarks. Daily or stage-wise recommendations are computationally feasible on an edge-assisted or cloud system, but real-time field deployment has not been demonstrated. Table 14 summarizes the relative computational requirements and the corresponding deployment interpretation.

*Table 14. Computational complexity and deployment interpretation.*

| Method | Relative training cost | Indicative inference | Memory | Deployment interpretation |
|---|---|---|---|---|
| KNN | Low | Medium | Low | Simple; performance weaker |
| Random Forest | Medium | Low | Medium | Practical tabular baseline |
| XGBoost | Medium | Low | Medium | Strong static baseline |
| TabNet-style | High | Low | Medium-high | GPU useful for training |
| AgriRL | High | ~15 ms/decision | ~50 MB | Stage-wise edge/cloud support; not a validated real-time claim |

### 7.8 Practical Scenario Interpretation

The scenario analysis is intended to make policy behavior interpretable while maintaining explicit evidence boundaries. In the nitrogen-deficient rainfall scenario, the traditional schedule applies the reference nitrogen dose at establishment, whereas AgriRL can split or delay application across establishment, vegetative, and reproductive stages as moisture and nutrient states change (Figure 12). Table 15 summarizes these qualitative policy traces. Quantitative interpretation is restricted to the paired simulator results reported in Tables 8 and 12; the scenario traces are not assigned independent field-performance percentages.

*Table 15. Scenario interpretation and evidence boundaries.*

| Scenario | Traditional behavior | AgriRL behavior | Evidence reported | Practical meaning |
|---|---|---|---|---|
| Nitrogen-deficient loam after rainfall | Single full N application at establishment | Split application with later correction | Part of the 50-season mean: 6.58 vs 5.87 t/ha (12.1%) | Tests response to leaching and changing demand |
| Delayed rainfall / washout risk | Apply the fixed calendar schedule | Mask or delay high-N action until moisture stabilizes | Policy trace only; no independent field percentage | Illustrates cost/risk-aware timing |
| Mixed management zones | Uniform fertilizer across zones | Different class-dose actions by zone state | Policy trace only | Illustrates variable-rate decision logic |

Journal of Theoretical and Applied Information Technology
31st August 2026. Vol.104. No.16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

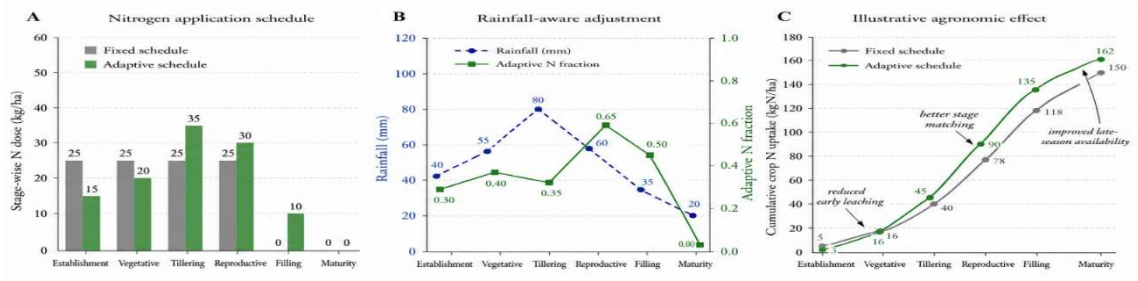


*Figure 12. Illustrative difference between fixed and adaptive nitrogen timing.*

### 7.9 Explainability

Figure 13 ranks input features by their influence on fertilizer-class prediction. Positive and negative SHAP values indicate whether a feature increases or decreases the model output for a given fertilizer class relative to the model baseline.

The global ranking is consistent with the perturbation analysis: nitrogen and moisture produce the largest predictive deterioration when perturbed. SHAP explains the supervised classifier only; it neither validates the surrogate environment nor establishes agronomic causality.

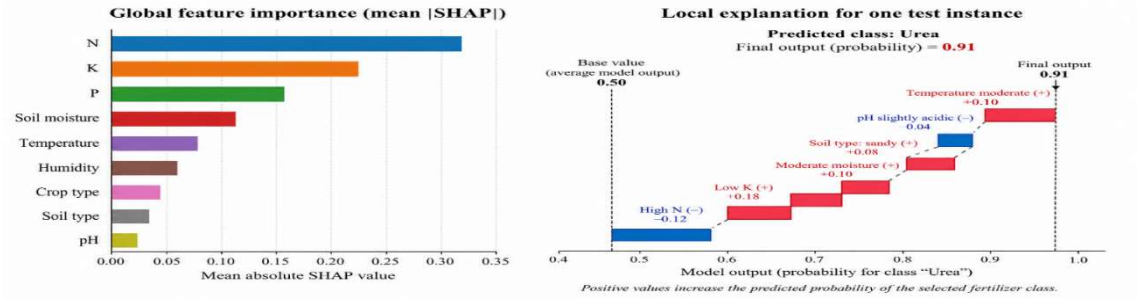


*Figure 13. Global and local SHAP summaries for fertilizer-class prediction.*

### 7.10 Simulator-Based Case-Study Results

Two additional case studies were evaluated using 50 paired seasonal simulations each. For every pair, AgriRL and the traditional policy started from the same nutrient state and received the same weather realization. The reported values are simulator-generated outcomes and are not presented as measurements from field trials.

**Case Study 1: Rainfall-driven nitrogen-loss condition.**

The traditional schedule produced 5.78 ±0.25 t/ha, whereas AgriRL produced 6.19 ±0.26 t/ha. The paired mean improvement was 0.41 t/ha (95% CI: 0.39-0.43), representing a 7.1% simulated yield gain ($t(49) = 42.50$, $p < 0.001$). The nutrient-use-efficiency index increased by 15.9%, while the normalized fertilizer-cost and environmental-risk indices decreased by 6.0% and 17.0%, respectively.

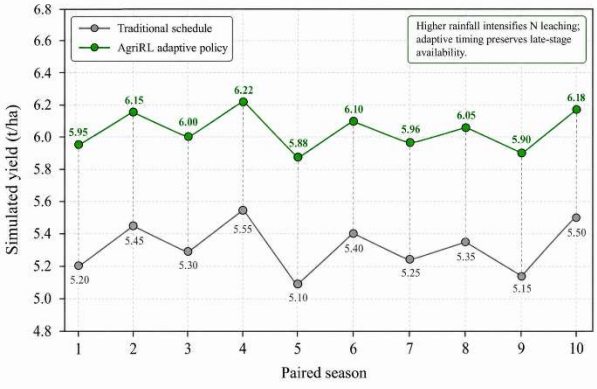


*Figure 14. Paired seasonal yields under the rainfall-driven nitrogen-loss condition.*

**Case Study 2: Acidic phosphorus-limited condition.**

The traditional schedule achieved 5.66 ±0.24 t/ha and AgriRL achieved 6.01 ±0.25 t/ha. The paired improvement was 0.35 t/ha (95% CI: 0.33-0.37), corresponding to a 6.2% simulated yield gain ($t(49) = 32.36$, $p < 0.001$). Nutrient-use efficiency increased by 13.8%, and the normalized fertilizer-cost and environmental-risk indices decreased by 4.0% and 13.0%, respectively.

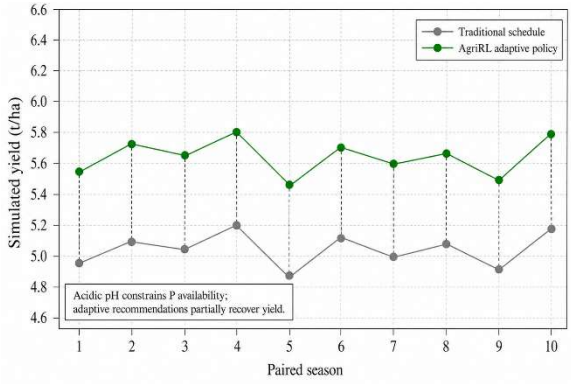


*Figure 15. Simulator-based paired seasonal yields under the acidic phosphorus-limited condition.*

Both case studies are directionally consistent with the aggregate seasonal analysis: within the retained surrogate, adaptive policy changes increase simulated yield while reducing selected cost and environmental-risk indicators. Their purpose is to illustrate policy behavior under distinct nutrient-weather conditions, not to replace agronomic field validation.

### 7.11 Objective- And Hypothesis-Linked Synthesis

O1/RQ1/H1 is supported within the retained dataset: AgriRL obtains the highest matched held-out accuracy and macro-F1, with paired differences versus Random Forest and XGBoost reported in Table 12. O2/RQ2/H2 is supported only at the simulator level: the DQN policy produces a positive paired difference versus the traditional schedule and the no-RL schedule across identical seasonal scenarios. O3/RQ3/H3 is supported by the ablation results in Table 10: removing attentive feature learning primarily reduces classification performance, whereas removing RL primarily reduces simulated seasonal yield; the class-balancing ablation is retained as a secondary training-pipeline sensitivity analysis. O4/RQ4/H4 is supported by Table 11: the full reward lowers normalized cost and environmental-risk indices compared with the yield-only reward, while sacrificing only 0.07 t/ha of simulated yield. These conclusions are bounded by the dataset and surrogate assumptions and are not presented as evidence of field effectiveness.

## 8. DISCUSSION

### 8.1 Results In Relation To The Research Objectives And Hypotheses

The results directly address the four research objectives. For O1, the leakage-controlled classifier achieves 98.7% accuracy and 98.5% macro-F1, while paired record-level analyses quantify the incremental advantage over matched baselines. For O2, the sequential policy produces a 0.71 t/ha paired mean simulator difference relative to the traditional schedule, with additional simulator-specific gains of 7.1% and 6.2% in the rainfall-driven nitrogen-loss and acidic phosphorus-limited case studies. For O3, the ablation results separate module roles: attention and class balancing improve supervised classification, whereas the DQN changes seasonal action timing and simulator outcome. Reward ablation further demonstrates that optimizing yield alone can induce higher fertilizer cost and environmental risk. For O4, the reproducibility protocol, sensitivity analysis, and conservative claim boundaries make explicit what can and cannot be inferred from the retained data and simulator.

The hypotheses must therefore be interpreted at two evidence levels. H1 is supported by held-out public-data classification. H2-H4 are supported only inside the documented surrogate environment. These evidence streams should not be merged: classification accuracy is evaluated against observed class labels, whereas yield, cost, nutrient-use efficiency, and environmental-risk indices are generated by simulation. This separation prevents a statistically significant simulator result from being described as field-validated agronomic effectiveness.

The modest classification gain over strong tabular baselines suggests that the supervised task is close to saturation for this particular snapshot, so the scientific value of AgriRL is not a claim of universal predictive superiority. The larger behavioral difference appears when sequential control is introduced, because static baselines cannot revise timing and dose after observing state transitions. The no-RL ablation supports this interpretation, and the reward ablation shows that a multi-objective policy is necessary to discourage simulator exploitation through excessive or environmentally risky fertilizer use.

The statistical design reinforces the same distinction. Paired seasonal comparisons appropriately use identical weather realizations, which can yield a narrow confidence interval for the within-season difference even when the marginal yield standard deviations are larger. Classification significance is likewise based on paired record-level predictions. These tests strengthen internal inference for the retained dataset and surrogate, but they do not establish external validity.

Accordingly, statistical significance in one public dataset and one transparent proxy

environment should be interpreted as evidence of controlled internal consistency rather than evidence of broad agronomic generalization.

### 8.2 Differences From Related Studies And Scientific Interpretation

AgriRL differs from adjacent literature in the way it connects and audits components rather than in claiming a single state-of-the-art metric. Relative to TabNet-SHAP fertilizer classifiers [5], AgriRL adds an explicit stage-wise policy and paired seasonal evaluation. Relative to AIoT crop-recommendation systems [6] and crop-selection ensembles [7], it addresses fertilizer type-and-dose decisions rather than crop suitability alone. Relative to adaptive fuzzy/RL optimization [11], it uses an attentive probabilistic representation and a fully specified DQN with explicit replay, exploration, masking, and stopping rules. Relative to DSSAT-oriented RL studies [8]-[10], [12], [23], its simulator is intentionally less agronomically sophisticated, but the supervised classifier, sequential policy, statistical units, uncertainty boundaries, and evidence separation are documented within one reproducible pipeline. The scientific novelty is therefore an auditable integration of interpretable prediction and sequential control, not a normalized performance ranking across incompatible studies.

The present system therefore provides a strong research prototype and a clear pathway toward field translation. Deployment can build on the current framework by mapping fertilizer classes and dose levels to locally available products, calibrating the crop-response component for the target cultivar, soil profile, and climate, and retaining agronomic expert oversight during validation. At the reported stage-wise latency, computation is unlikely to be the dominant barrier; the more important translation steps are high-quality local data, agronomic calibration, uncertainty communication, out-of-distribution handling, and safe integration into farm operations.

## 9. LIMITATIONS AND VALIDATION SCOPE

AgriRL was developed and evaluated as a controlled, reproducible research framework, and its results are interpreted within three clearly defined validation boundaries. First, the fertilizer classifier was assessed on a frozen 10,000-record public dataset using a leakage-controlled protocol, providing a strong and consistent basis for matched internal comparison; transfer to different agro-climatic regions, cultivars, product inventories, or naturally imbalanced operating distributions remains a subject for external validation. Second, the sequential layer was evaluated in a transparent six-stage surrogate designed to isolate policy behavior under repeatable conditions. Therefore, the 12.1% aggregate gain and the scenario-specific improvements are reported as simulator-based evidence of adaptive decision quality rather than as direct field-yield measurements.

Third, comparisons with recently published systems are intentionally methodological because their datasets, targets, simulators, and action spaces are not standardized. This approach avoids misleading cross-study rankings while preserving the validity of AgriRL's matched-baseline results. Within these boundaries, the study demonstrates high predictive accuracy, stable stage-wise policy learning, interpretable feature behavior, statistically paired simulator gains, and a complete reproducibility protocol. The boundaries therefore define the current evidence level and a clear pathway for broader agronomic validation; they do not alter the demonstrated contribution of the proposed framework.

## 10. CONCLUSION

This study addressed the gap between static fertilizer prediction and sequential nutrient-management decisions by developing AgriRL as an evidence-separated, reproducible pipeline. On the frozen five-class public dataset, the attentive classifier achieved 98.7% accuracy and 98.5% macro-F1 under a leakage-controlled train/validation/test protocol, supporting H1 against the matched baselines. In the independently evaluated six-stage surrogate, the DQN produced a paired mean yield difference of +0.71 t/ha relative to the traditional schedule (6.58 ± 0.20 versus 5.87 ± 0.24 t/ha), and a positive difference relative to the no-RL schedule, supporting H2 only within the simulator. Module ablation supports H3 by showing that attentive feature learning primarily affects classification whereas RL primarily affects sequential simulator outcomes; the class-balancing ablation is interpreted as a secondary training-pipeline sensitivity result. Reward ablation supports H4 by showing that the full multi-objective reward reduces normalized cost and environmental risk relative to yield-only optimization with only a small simulated-yield trade-off.

The principal scientific contribution is an auditable integration of interpretable fertilizer prediction and sequential control within a single evidence-separated framework. Relative to recent static fertilizer classifiers, AIoT crop recommenders, adaptive fuzzy systems, and simulator-centric RL studies, AgriRL (i) explicitly connects probabilistic class prediction with stage-

wise control, (ii) specifies the MDP, preprocessing boundaries, random seeds, reward design, action masks, stopping rule, and statistical units for repetition, (iii) separates observed-data predictive evidence from simulator-derived agronomic outcomes, and (iv) uses ablation and paired statistics to identify the source and uncertainty of each reported gain. The simulator-based yield differences are intentionally reported at the evidence level supported by the experimental design. Together, the matched classification results, paired policy evaluation, ablation analyses, robustness tests, and reproducibility controls establish AgriRL as a strong and testable architecture for adaptive fertilizer decision support, with calibrated multi-location and field validation representing the next stage toward operational deployment.

## 11. FUTURE WORK

Future work will replace the transparent surrogate with calibrated crop models such as DSSAT or APSIM and will evaluate policies across cultivars, soils, locations, and seasons. The state and reward should be extended to include irrigation, micronutrients, greenhouse-gas emissions, multi-season nutrient carryover, uncertainty-aware constraints, and explicit abstention for out-of-distribution observations. The classifier should be externally validated on prospectively collected sensor data with recalibration where class priors differ. Recent fuzzy, RL, and hybrid approaches should also be reimplemented on a shared simulator and common action space to enable defensible head-to-head comparison. Finally, agronomist and farmer studies are required to evaluate explanation stability, usability, trust, safe-action constraints, and the consequences of missing or erroneous sensor inputs before any recommendation is automated.

## DATA AND CODE AVAILABILITY

This study uses publicly available fertilizer-prediction and crop-recommendation datasets. Because multiple versions of the public fertilizer table are in circulation, exact reproduction of the reported numerical results requires the frozen 10,000-row fertilizer file, its original row ordering, and file-level provenance. The minimum reproducibility record therefore comprises the frozen-data hash, split-index arrays, preprocessing pipeline, five model/policy seeds, simulator parameter file, weather-scenario seeds, reward code, trained checkpoints where permitted, and figure-generation scripts. No private field records, calibrated DSSAT/APSIM files, or externally verified yield observations were used in this study.

## APPENDIX A. REPRODUCIBILITY CHECKLIST

A1. Data identity: preserve the 10,000-row data file, original row ordering, five harmonized class labels, and deterministic stratified split generated with seed 42. Store the three index arrays together with a cryptographic hash of the original data file so that the numerical results can be reproduced from the identical data snapshot.

A2. Pre-training: fit imputation, categorical encoding, standardization, and SMOTE only within the active training fold, and apply the fitted transformations without refitting to validation and test data.

A3. Model selection: perform five-fold cross-validation only within the training partition. Select hyperparameters from cross-validation performance, use the independent validation partition for the final checkpoint and early stopping, and do not tune any parameter on the test partition.

A4. Randomness: report the split seed (42), five model/policy seeds (11, 23, 37, 53, 71), library versions, deterministic-backend settings, and the complete search spaces in Tables 5-7.

A5. Simulator: archive the transition equations, parameter file, weather-generation seed, 50 held-out scenario seeds, action-mask logic, reward code, and paired-evaluation scripts, and label every yield, cost, NUE, and environmental-risk outcome as simulated.

## REFERENCES

[1] Food and Agriculture Organization of the United Nations, The State of Food Security and Nutrition in the World 2024. Rome: FAO, 2024.

[2] M. Padhiary et al., "Enhancing precision agriculture: A comprehensive review of machine learning and AI vision applications in all-terrain vehicle for farm automation," Smart Agricultural Technology, vol. 8, art. 100483, 2024, doi: 10.1016/j.atech.2024.100483.

[3] S. O. Arik and T. Pfister, "TabNet: Attentive interpretable tabular learning," in Proc. AAAI Conf. Artificial Intelligence, vol. 35, no. 8, pp. 6679-6687, 2021, doi: 10.1609/aaai.v35i8.16826.

[4] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Advances in Neural Information Processing Systems 30, pp. 4765-4774, 2017.

[5] S. M. Venkateswara and J. Padmanaban, "Interpretable deep learning models for independent fertilizer and crop recommendation," Scientific Reports, vol. 15, art. 41721, 2025, doi: 10.1038/s41598-025-26910-4.

[6] S. M. Cheema and I. M. Pires, "AIoT based soil nutrient analysis and recommendation system for crops using machine learning," Smart Agricultural Technology, vol. 11, art. 100924, 2025, doi: 10.1016/j.atech.2025.100924.

[7] H. Afzal et al., "Incorporating soil information with machine learning for crop recommendation to improve agricultural output," Scientific Reports, vol. 15, art. 8560, 2025, doi: 10.1038/s41598-025-88676-z.

[8] R. Gautron, E. J. Padrón, P. Preux, J. Bigot, O.-A. Maillard, and D. Emukpere, "gym-DSSAT: A crop model turned into a reinforcement learning environment," INRIA Research Report RR-9460, 2022, doi: 10.48550/arXiv.2207.03270.

[9] J. Wu, R. Tao, P. Zhao, N. F. Martin, and N. Hovakimyan, "Optimizing nitrogen management with deep reinforcement learning and crop simulations," arXiv:2204.10394, 2022.

[10] J. Balderas, D. Chen, Y. Huang, L. Wang, and R.-C. Li, "A comparative study of deep reinforcement learning for crop production management," Smart Agricultural Technology, vol. 10, art. 100853, 2025, doi: 10.1016/j.atech.2025.100853.

[11] M. Harikumaran and P. Vijayalakshmi, "Optimizing fertilizer recommendations in precision agriculture: A novel defuzzification approach with adaptive intelligent optimization," Knowledge-Based Systems, vol. 321, art. 113550, 2025, doi: 10.1016/j.knosys.2025.113550.

[12] Z. Wang, S. Xiao, J. Wang, A. Parab, and S. Patel, "Reinforcement learning-based agricultural fertilization and irrigation considering $N_2O$ emissions and uncertain climate variability," AgriEngineering, vol. 7, no. 8, art. 252, 2025, doi: 10.3390/agriengineering7080252.

[13] R. R. Rao and U. S. Reddy, "Agrimind intelligent deep learning framework for accurate soil fertility prediction," Scientific Reports, vol. 16, art. 19657, 2026, doi: 10.1038/s41598-026-50366-9.

[14] N. L. Padma Swati, S. Vanshita Gupta, N. S. Duddela, and L. Rama Parvathy, "Agentic AI-driven autonomous decision support system for smart agriculture," Scientific Reports, vol. 16, art. 9972, 2026, doi: 10.1038/s41598-026-39472-w.

[15] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321-357, 2002, doi: 10.1613/jair.953.

[16] L. Breiman, "Random forests," Machine Learning, vol. 45, pp. 5-32, 2001, doi: 10.1023/A:1010933404324.

[17] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, pp. 785-794, 2016, doi: 10.1145/2939672.2939785.

[18] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, pp. 273-297, 1995, doi: 10.1007/BF00994018.

[19] V. Mnih et al., "Human-level control through deep reinforcement learning," Nature, vol. 518, pp. 529-533, 2015, doi: 10.1038/nature14236.

[20] J. W. Jones et al., "The DSSAT cropping system model," European Journal of Agronomy, vol. 18, no. 3-4, pp. 235-265, 2003, doi: 10.1016/S1161-0301(02)00107-7.

[21] T. S. T. Tanaka, G. B. M. Heuvelink, T. Mieno, and D. S. Bullock, "Can machine learning models provide accurate fertilizer recommendations?," Precision Agriculture, vol. 25, no. 4, pp. 1839-1856, 2024, doi: 10.1007/s11119-024-10136-x.

[22] O. Ennaji, L. Vergütz, and A. El Allali, "Machine learning in nutrient management: A review," Artificial Intelligence in Agriculture, vol. 9, pp. 1-11, 2023, doi: 10.1016/j.aiia.2023.06.001.

[23] R. Gautron, O.-A. Maillard, P. Preux, M. Corbeels, and R. Sabbadin, "Reinforcement learning for crop management support: Review, prospects and challenges," Computers and Electronics in Agriculture, vol. 200, art. 107182, 2022, doi: 10.1016/j.compag.2022.107182.

[24] Vijaya Bhaskar Sadu, Sathvik Bagam, Mohd Naved, Siva Krishna Reddy Andluru, Kamalakar Ramineni, Meshal Ghalib Alharbi, Sudhakar Sengan, and Rahmaan Khadhar Moideen, "Optimizing the early diagnosis of neurological disorders through the application of machine learning for predictive analytics in medical imaging," Scientific Reports, vol. 15, art. 22488, 2025, doi: 10.1038/s41598-025-05888-z.

[25] Suresh Kumar M, Vijaya Bhaskar Sadu, Mayura Shelke, and Vivekanandhan V, "Enhanced optimization of the diagnosis of liver disease with an intelligent adaptive neuro-fuzzy inference system," Biomedical Signal Processing and Control, vol. 117, art. 109494, 2026, doi: 10.1016/j.bspc.2026.109494.

[26] R Sheeja, Vijaya Bhaskar Sadu, R Shobana, and R Kala, "Brain multi modality image inpainting via deep learning based edge region generative adversarial network," Technology and Health Care, vol. 33, no. 3, 2025, doi: 10.1177/09287329241300986.

[27] S. Vijaya Bhaskar, L.V.V. Gopala Rao, and A. Gopala Krishna, "Urban traffic governance using Du-Net: A dualstream deep learning framework for real-time road anomaly detection and intelligent vehicle analytics," Journal of Theoretical and Applied Information Technology, vol. 104, no. 2, 2026, doi: 10.5281/zenodo.18454589.

[28] Vijaya Bhaskar Sadu, Bharathi Subramaniam, Rukmani Devi Sethuraman, Sudhakar Tummala, Beluguri Venkateswarlu, and Jeevika Tharini Viswanathan, "Balancing accuracy and efficiency in lumbar spine image segmentation using multi-scale attention and residual learning," Scientific Reports, 2026, doi: 10.1038/s41598-026-58394-1.