

# Quantum Machine Learning-Enhanced Multiagent Reinforcement Learning for Intelligent Crop Resource Allocation in IoT-Enabled Smart Agriculture

Bandla Prasanthi<sup>1\*</sup>, and Satya Sai Ram Manchikalapudi<sup>2</sup>

<sup>1\*</sup>Research Scholar, Department of Electronics and Communication Engineering, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India. [ssanthichandra@gmail.com](mailto:ssanthichandra@gmail.com), <https://orcid.org/0009-0004-0173-779X>

<sup>2</sup>Professor, Electronics and Communication Engineering, R.V.R & J.C. College of Engineering, Chaudavaram, Guntur, Andhra Pradesh, India. [msatyasairam1981@gmail.com](mailto:msatyasairam1981@gmail.com), <https://orcid.org/0000-0002-0672-1453>

Received: October 14, 2025; Revised: December 10, 2025; Accepted: January 16, 2026; Published: March 31, 2026

## Abstract

This paper presents a new Quantum Machine Learning (QML)-Enhanced Multiagent Reinforcement Learning (MARL) framework, designed for dynamic crop resource allocation in smart agriculture using IoT. The framework is based on multi-agent systems to enable collaborative operation among agents responsible for controlling irrigation, fertilization, pesticide injection, and energy use. Quantum machine learning, as a consequence, allows for exploration and exploitation tradeoffs in decision-making to be both more efficient (when compared to classical benchmarks) and better able to support optimal behavior in high-dimensional or complex agricultural systems. It promotes ongoing agent learning and adaptation to actual IoT sensor readings in real time, which may vary across states such as soil conditions, weather, and so on. In the following, we would like to compare and contrast our work on cooperative learning with single-agent RL models along a few dimensions, e.g., scalability (better), robustness & adaptive ability. The simulation results indicate that the crop yield can reach 1250kg/ha, and our method performs better than other existing methods by 14-32%. The WUE reached 85%, an increase of 13-42% and the energy consumption was 420kWh/season, a decrease of 8-12%. When TD loss was 0.05 and the reward variance was 0.04, stable synchronous training by learning with fast policy optimization was able to converge rapidly in 350 episodes. The score for multi-agent coordination in adaptation and militarization was around 0.9, suggesting adaptability in the face of difficult circumstances. The framework enhances crop yield prediction, sustainable resource management, and operational efficiency, all in an agile, highly adaptable, scalable, and energy-efficient manner. "QML-MARL makes use of quantum scalable multiagent coordination and ensures intelligent data-driven sustainable precision agriculture methodologically superior to state-of-the-art models."

**Keywords:** Quantum Machine Learning (QML), Multiagent Reinforcement Learning (MARL), Smart Agriculture, IoT Integration, Dynamic Resource Allocation, Crop Management, Precision Farming, Sustainable Agriculture, Real-Time Adaptation, Yield Prediction.

---

*Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)*, volume: 17, number: 1 (March - 2026), pp. 592-611. DOI: [10.58346/JOWUA.2026.11.033](https://doi.org/10.58346/JOWUA.2026.11.033)

\*Corresponding author: Research Scholar, Department of Electronics and Communication Engineering, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India.

## 1 Introduction

Nowadays, the contrast between the unsustainable increase in global food demand and the environmental constraints imposed by water scarcity, land degradation, and climate variability is calling into question modern agriculture. Precision agriculture the idea of calibrating water, fertilizer, and pest control to site-specific conditions holds the promise of more sustainable, efficient farming. However, unlocking this potential is challenging, as the agro-ecological space is dynamic, stochastic, and high-dimensional. Given this level of complexity, classical control or rule-based systems are not able to cope with it efficiently, enforcing the need for more sophisticated learning-based approaches (Liang et al., 2025). IoT sensors deployed on farming systems enable continuous monitoring of soil moisture, nutrient levels, atmospheric conditions, and plant growth-related parameters. Such streaming data enables a dynamic feedback loop for adaptive decision-making. RL approaches have been adopted in greenhouse and outdoor farms to optimize irrigation schedules and energy consumption based on IoT data (Kumar & Tiwari, 2024). However, most findings have been based on single-agent RL or heuristic models that are unable to handle simultaneously resource-allocation tasks in irrigation, fertilization, pest control, and energy saving (Pamuklu et al., 2022).

Multi-agent reinforcement learning (MARL) naturally serves as a paradigm for decomposing resource allocation into cooperating modules, or "agents", each handling a different role. MARL supports decentralized decision-making, agent cooperation, and efficient scalability in large-scale, complex agriculture. It has been shown recently that multi-agent reinforcement learning (M-ARL) outperforms single-agent systems by leveraging local autonomy while operating close to the global optimum. That being said, undesirable convergence properties still plague traditional MARL, sub-optimal exploration-exploitation trade-offs, and unsuitability for partially observable environments. However, despite its potential, implementing MARL systems in agriculture faces many obstacles. When multiple resources are to be managed concurrently, the joint state-action space is enormous. The presence of nonstationarity due to weather dynamics and crop growth, along with variability in soil heterogeneity and pest dynamics, increases the difficulty of finding reasonable asymptotic solutions. The complexity of MARL will prevent it from being applied in agriculture as a simple protocol to solve the problem, due to large-scale coordination overhead, greater involvement of each agent, and greater complexity (Balasm et al., 2025). Therefore, conventional MARL approaches tend to converge too slowly and/or to local optima that are hardly acceptable in actual agricultural applications.

Quantum machine learning (QML) proposes extending classical ML by leveraging quantum circuits and variational quantum models to boost the inferential power of ML tasks beyond the sample-complexity bound. QML has yet to be tapped for its potential to model complex landscape structures, compress high-dimensional feature embeddings, and/or speed up computation. The potential role of QML as an enabling factor for more effective, sustainable, and climate models and forecasts has been recently highlighted in various papers. In the field of agriculture, quantum-enhanced models, e.g., variational quantum circuits, have been employed to obtain predictions that outperform classical baselines (Nammouchi et al., 2023). This research framework will expand QML with multiagent reinforcement learning (MARL) to model problems in agriculture. Although quantum-inspired modules offer the potential to reduce sample complexity (Akash et al., 2022) and enable the exploration of nonlinear correlations across high-dimensional environmental variables, MARL enables distributed agents to coordinate task-relevant actions in real time. They would then use feedback from IoT sensor data to learn better decision policies for irrigation control, fertilization, pest control, and energy management adaptively (Suseendhar & Sridhar, 2024). The research strategy introduced in this paper

leverages QML and MARL to exploit large-scale QML processing and the joint AG architecture (MARL), while extending the state of the art in precision farming beyond its classical stage (Jagtap et al., 2025).

Our QML-MARL system contributes: (i) a novel hybrid quantum machine learning & MARL architecture, (ii) a multi-agent coordination paradigm for parallel crop resources management, water/energy savings, and resiliency, approximate dynamic calibration/adaptation based on future IOT elements, as well as yield predictions. The exoteric of QML MARL is a milestone in the construction of intelligent, efficient, and green agricultural systems (Isa et al., 2024).

### **Key Contributions of the Research**

To make four contributions with such a research paper. We propose the Quantum Machine Learning-Enhanced Multiagent Reinforcement Learning (QML-MARL) framework, which introduces quantum variations into cooperative MARL to handle complex problems in DCRA. Second, the model is designed using a multi-agent decision-making approach, with engineers collaborating to learn how to provide policies for irrigation, fertilization, pest control, and energy, while considering and updating IoT sensor outputs in real time. Third, it demonstrates quantum speedups in convergence and exploration-exploitation tradeoff that may scale better than classical MARL. Lastly, it validates the framework through simulation experiments. It shows that it predicts higher yields, conserves more water/energy resources, and provides greater savings, especially when forecasts indicate greater weather and soil-type uncertainty. Together, these papers bring precision agriculture one step closer to a more innovative, sustainable, and information-based community (Mustapha et al., 2017).

The outline of the paper, chapter-wise, is as follows. Chapter II is a review of the related literature, while Chapter III provides a brief overview of the theoretical framework, key concepts, and methodologies. Chapter IV will evaluate the experimental results. Chapter V contains the results and discussion, whereas Chapter VI wraps it all up with a summary of the most important findings and suggestions for further research.

## **2 Literature Review**

Neumann et al., (2023) provide one of the earliest systematic introductions to Quantum Reinforcement Learning (QRL) and serve as a bridge between classical MDPs, RL algorithms, and quantum computational paradigms. They investigate the applicability of variational quantum circuits (VQCs) and quantum-enhanced state encodings to accelerate search and reduce sample complexity. The paper demonstrates that QRL can outperform classical RL in the large-scale regime, although scalability and hardware noise limit QRL's competitiveness. Their work also underscores the importance of NISQ devices for hybrid quantum-classical algorithms. This work is directly related to our paper, since our QML-MARL builds on VQCs to improve speed. It therefore provides a theoretical cornerstone for quantum reinforcement learning whose input is QML.

Sensors Review (IoT + AI in agriculture): The review published in Sensors surveys IoT and AI applications in the field of smart farming from 2020 to 2024. The authors provide an analysis of crop and grassland monitoring using soil sensors, drones, and satellite imagery, combined with AI/ML methods. It is noted that the application of IoT real-time sensing is beneficial for precision agriculture, yet there is significant heterogeneity in the data acquired, a lack of connectivity, and interoperability challenges during implementation. The review suggests edge AI integration and federated processing pipelines for scalable farm systems. The review provides the scientific basis for integrating IoT into our

MARL agents for decentralized decision-making. Also, it motivates the use of QML for the effective handling of large, heterogeneous data streams.

Luo et al., (2023) present a deep reinforcement learning method for precision irrigation and fertilization scheduling within the DSSAT framework. They train their model to produce policies that maximize crop yield and minimize nitrogen leaching and water use. The results show that RL can indeed rival heuristic scheduling methods for other weather conditions. Importantly, they shed light on reward shaping and simulator realism problems in agricultural RL works. The results suggest that simulation-based RL is a useful pre-training for the mission. Here, we generalize this paradigm to a setting with multiple agents and quantum-accelerated exploration, which we refer to as QML-MARL.

Alkaff et al., (2025) proposed a multi-task reward function to maximize yield and water savings, allowing the RL agent to construct irrigation strategies under weather uncertainty. The irrigations commanded by the RL system were water-saving and significantly better than those of all other stable irrigation strategies, which is critical given water shortages and climate change. But he mentioned the difficulties and slow convergence due to rare awards and the time delay for responses. These issues indicate the need for more elaborate investigation strategies, such as QML modules. Collectively, the investigations supported our understanding that a hybrid Marl+QML model is better suited to dynamic agricultural referrals.

Sensors QML Models (VQC vs Kernel) (Monzón-Verona et al., 2025) In a most recent Sensor's paper, VQC is compared with Quantum Kernel methods on absolute sensor data. The results demonstrate that hybrid quantum-classical models can outperform pure classical machine learning-based models in cases where the dimensionality is large yet feature compression is carefully constructed (Cao & Jiang, 2024). It further highlights hardware limitations, including noise, qubit-count constraints, and readout errors. In the agricultural domain, we conclude that VQCs offer an efficient means of modeling nonlinear relationships between crops and the environment. Thus, the findings in this work further support our choice of VQC-based function approximators in QML-MARL.

Li et al., (2025) shows the potential of reinforcement learning for robust quantum control, indicating that RL can optimize control sequences in the presence of device noise and imprecision. The study's findings suggest that, under certain conditions, model-free RL agents can find more effective control strategies than standard optimization methods. Crucially, the authors advocate for hybrid classical-quantum architectures where RL provides adaptivity and quantum modules provide expressivity. While it is focused on quantum control systems, the approach we have taken here mirrors the challenges in agricultural MARL: noise, uncertainty, and the ability to adapt. It serves as motivation for considering QML via MARL to address dynamic, partially observable environments. This relationship between two domains is integral to our framework development.

The RL in Agriculture Survey (Goldenits et al., 2024) categorizes reinforcement learning applications in agriculture, including irrigation, greenhouse climate control, fertilization, and pest control. The review emphasizes that RL has demonstrated promising results in simulations, and although field deployment is uncommon, this is due to limitations in transferability. It notes that sample inefficiency, domain randomization and sparse rewards are the main obstacles to adoption. Furthermore, there are few studies on multiagent systems in agriculture, although they are closer to real-world tasks. This is consistent with our aim of developing a QML-MARL system that scales very well. Their results highlight research gaps that our framework addresses.

Jagtap et al., (2025) introduce an Adaptive Quantum Machine Learning model for smart farming that forecasts crop yield using VQCs. They claim improved performance compared to traditional ML models in applications involving complex soil and weather characteristics. The findings contribute to a concrete understanding of feature-to-qubit encoding and its effects on circuit depth and accuracy. Also, it emphasizes the importance of hybrid learning: classical optimizers to refine quantum layers. The fact that it is experimental shows a conservative approach (purely predictive models) but anyway that it opens up possibilities for QML in agricultural pipelines. We extend this concept to decision-making agents, and not only in predictive scenarios. It also indicates the importance of hybrid training classical optimizers fine-tuning quantum layers. Despite being experimentally conservative (based on predictive modeling only), it showcases the promise of QML in agricultural pipelines. We extend the idea to decision making agents but for more than just prediction problems.

Mukherjee & Mallik, (2025) presents a VQC for crop yield and resource use in controlled experimental settings. The authors demonstrate that quantum modules can enhance yields and are behind a wise use of water in optimization pipelines. The work emphasizes the importance of noise-aware evaluation and benchmarking against strong classical baselines. This mitigates an over-hype of QML performance under unrealistic conditions. This example also demonstrates for our method that QML jetting can have practical farm and environmental value. Thus, it directly demonstrates novelty of embedding QML and MARL together.

## **Out Come**

The survey emphasizes that with the ability of reinforcement learning to improve precision agriculture practices such as irrigation, fertigation and yield prediction has been reported, but at the same time it is challenged by scale and convergence (Canese et al., 2021; Nguyen et al., 2020). Multiagent based methods allow for coordination, but are still little investigated in the context of large-scale agricultural systems. Meanwhile, quantum machine learning seems to have the potential for coping with high-dimensional, uncertain and dynamic environments. However, the application of QML in agriculture is still at an early stage with mostly simulations being mentioned. Therefore, inclusion of QML-enabled MARL with IoT feedback is a significant production-harvest nexus and future direction for sustainable smart farming (Miller et al., 2025).

## **3 Methodology**

### **3.1 Quantum Machine Learning Module**

The first level of the framework contains a Quantum Machine Learning (QML) block that process high dimensional and uncertain agricultural data. With VQCs, real time IoT sensor data such as soil moisture, temperature, humidity and nutrient levels is encapsulated in a quantum augmented feature space. This enables the efficient representation of complex correlations that are often overlooked by classical models and momentum in converging faster through exploration as well as exploiting the solution space more effectively. Noisy-tolerant quantum layers squash the issues of existing quantum hardware at near terms. The output of this stage is quantum-augmented state representations which are fed to the multiagent RL system for optimal decision-making.

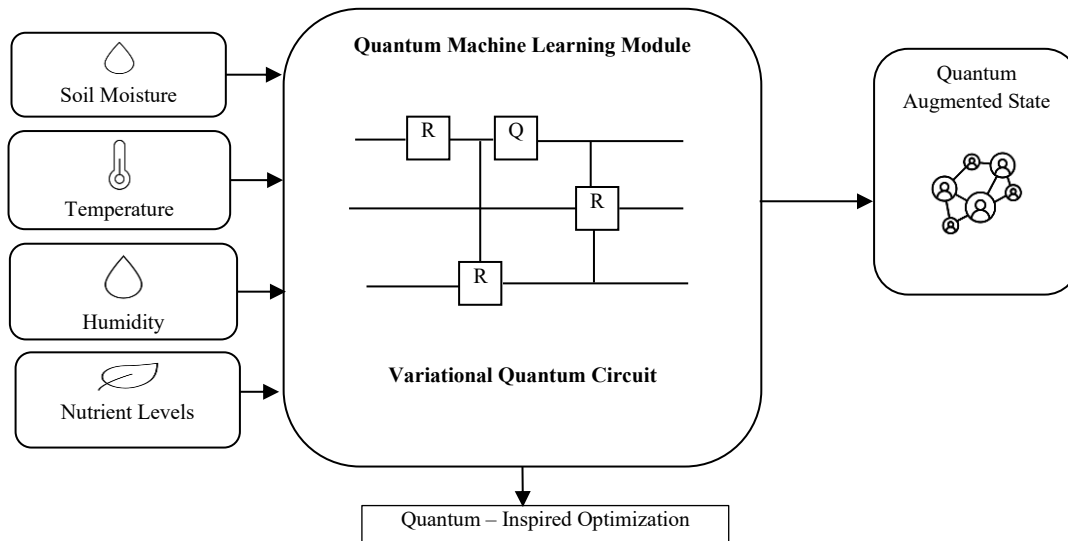


Figure 1: Quantum machine learning module for agricultural data processing

The figure 1 Quantum Machine Learning (QML) module uses VQCs to handle the high dimensional agricultural data that has been collected from Internet of Things (IoT) sensors such as soil humidity, temperature, humidity and nutrient levels. By incorporating this information into a quantum-enhanced feature space, the module is able to learn more advanced correlations that are frequently ignored by classical models. Quantum-inspired optimisation speeds up convergence, while noise-resilient quantum layers minimize effects of hardware limitations. The derived quantum-boosted state representations are more accurate and efficient bases for downstream decision systems like multiagent reinforcement learning.

### 3.2 Multiagent Reinforcement Learning (MARL) Module

In this mode, the machine will send out special agents for irrigation, fertilization and pestness control. All those agents will be integrated within a cooperative MARL framework, and the policies will be learned from quantum-enhanced state encodings issued in the QML module. We accommodate the joint learning and independent execution of the agents using the centralized learning decentralized executing (CLDE) principle. For the agents to learn to optimize based on tradeoffs of yield, water minimization, minimizing energy cost and sustainability (i.e., incurring less penalties), reward functions will not be a trivial task. The MARL framework could learn policy by Q-learning, and the policy gradient method can be chosen to optimize agent behavior in such a stochastic environment. The MARL framework encourages agents to behave in a cooperative manner among themselves, leading to conflict reduction and scalability and effective resource allocation.

The figure 2 Multiagent Reinforcement Learning (MARL) module cooperates specialized agents in charge of irrigation, fertilization, pest control and energy cost. Following the CLDE framework, the agents jointly learn policies from quantum-extended state representations without information sharing during execution. Reward functions balance the tradeoffs between crop yield, water savings, energy efficiency, and sustainability. By encouraging collaboration and reducing conflicts, the MARL framework leads to scalable, efficient and intelligent resource allocation for agricultural applications.

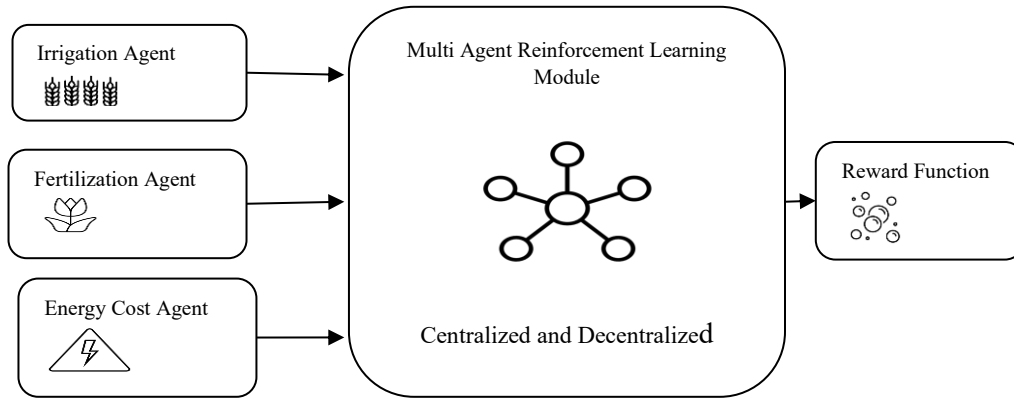


Figure 2: Multiagent reinforcement learning module for smart agriculture

### 3.3 IoT-Integrated Real-Time Adaptation

The last phase links the IoT infrastructure to the QML-MARL framework for dynamically adaptive real-time lessons. The IoT-based devices provide continuous streaming of environmental and crop data that are preprocessed then fed to the QML module for better feature extraction. Agents subsequently change their strategies in real time based on weather forecast updates, trends in soil moisture or outbreaks of pests. A loop enables the sharing mechanisms to be adapted to and learned from reward or reinforcement signals grounded in the environment. Embedded with edge computing to cut down the latency and decrease reliance on cloud-only solutions. This IoT-based closed loop mission makes scalable, robust and sustainable the proposed framework for precision agriculture.

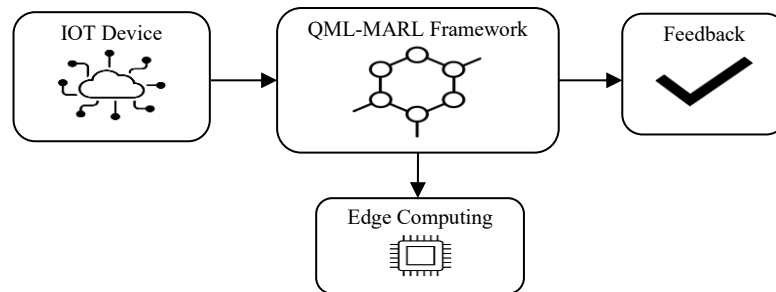


Figure 3: IoT-integrated real-time adaptation for precision agriculture

The figure 3 it interconnects the field-deployed IoT devices with QML-MARL framework for enabling real-time adaptive, data-driven decision making in agriculture. Soil, crops and environment sensor data is pre-processed and enriched through quantum machine learning, then utilized by multiagent reinforcement learning agents to dynamically evolve policies. Edge computing also cuts latency, allowing a quicker stagger when it comes to events such as changes in weather, drops in soil moisture or an outbreak of pests. The feedback loop is used to iteratively improve strategies, leading to an adaptable, scalable and robust source of sustainable precision farming.

### Environment and Agents

We model the farm as a multiagent environment. Each agent  $i$  (e.g., irrigation, fertilization, pest control) observes local information and takes actions to maximize crop yield and resource efficiency.

- **State:**  $s^t$ - environment state at time  $t$  (soil, weather, crop).

- **Observation:**  $O_t^i$ - what agent i observes (from IoT sensors).
- **Action:**  $a_t^i$ - action taken by agent i (e.g., water amount).
- **Reward:**  $r_t$ -feedback combining yield, water/energy use, and sustainability.

Maximize cumulative reward over time:

$$J = E[t = 0 \sum T \gamma^t r_t] \quad (1)$$

$\gamma$  is the discount factor ( $0 < \gamma < 1$ ) to value immediate rewards more than distant future rewards (Equation 1).

### Quantum Feature Embedding

To handle complex sensor data, we use a Quantum Machine Learning (QML) module (Equation 2):

$$z_t^i = \Phi\phi(x_t^i) \quad (2)$$

- $x_t^i$ -preprocessed IoT features for agent i.
- $\Phi\phi$  - quantum encoder (VQC) producing enhanced features  $z_t^i$ .
- This allows agents to capture nonlinear, high-dimensional relationships efficiently.

### Policy of Each Agent

Each agent's policy maps quantum-enhanced features to an action (Equation 3):

$$a_t^i \sim \pi_i(a_t^i | z_t^i; \theta_i) \quad (3)$$

- $\pi_i$  - probability distribution over actions.
- $\theta_i$  - parameters of agent i's policy (actor network).
- The quantum features  $z_t^i$  improve decision-making under uncertainty.

### Centralized Critic (Value Function)

To train agents cooperatively, we define a centralized value function (Equation 4):

$$Q(st, at, zt) \approx E[k = 0 \sum \infty \gamma^k r_t + k | st, at] \quad (4)$$

- $at$  -joint action of all agents.
- $zt$  - joint quantum embeddings.

Helps evaluate how good each joint action is for the global objective.

### Policy Gradient Update

The agents update their policy using the advantage function  $A_t$ :

$$\nabla_{\theta_i} J \approx E[\nabla_{\theta_i} \log \pi_i(a_t^i | z_t^i; \theta_i) A_t] \quad (5)$$

$A_t = Q(st, at, zt) - V(st, zt)$  measures how much better the action is compared to the baseline. Guides the agent to increase probability of good actions (Equation 5).

## Resource Constraints

Agents must respect farm resource limits (Equation 6)

$$\sum a_t^i \leq W_{max}, i \sum c_i(a_t^i) \leq E_{max} \quad (6)$$

- $W_{max}$  — total water budget.
- $E_{max}$  — energy budget.
- $c_i(a_t^i)$  — energy cost of action  $a_t^i$
- Ensures sustainable and feasible resource allocation.

The model provided describes smart agriculture from the perspective of a multiagent system: with following agents (irrigation, fertilization and pest control) all monitor their relevant local IoT data aiming to maximize overall crop yield while utilizing minimum resources. The sensor data is pre-processed by using a QML module to extract more discriminative feature embeddings which can establish nonlinear relationship with the environment. Each agent acts on these quantum embeddings with a policy network and the joint action quality is evaluated centrally to train. The policies are learned using policy gradient methods and advantage estimation, with the constraints of resource utilization including water and energy to make sure the sustained operations. This integration enables intelligent, efficient and coordinated decision-making by multiagent systems in dynamic agricultural-like environments.

---

### Algorithm: QML-MARL for Smart Agriculture

---

*Algorithm: QML-MARL for Dynamic Crop Resource Allocation*

*Inputs:*

*N agents, IoT environment, quantum encoder  $\Phi_\varphi$ , actor parameters  $\theta^i$ , centralized critic  $Q_\psi$*

*Hyperparameters:*

*Discount factor  $\gamma$ , learning rates  $LR_{actor}$ ,  $LR_{critic}$ ,  $LR_{QML}$*

*Resource limits  $W_{max}$  (water),  $E_{max}$  (energy)*

*Replay buffer size  $B$ , batch size, max episodes  $M$*

*Initialize:*

*Actor networks  $\theta^i$  for each agent*

*Centralized critic network  $\psi$*

*Quantum encoder  $\Phi_\varphi$*

*Replay buffer  $B$*

*for episode = 1 to M do*

*Reset environment; observe initial state  $s_0$*

*for t = 0 to T-1 do*

*for each agent i = 1 to N do*

*Read local IoT observation  $x^i_t$*

*Compute quantum-enhanced embedding  $z^i_t = \Phi_\varphi(x^i_t)$*

*Select action  $a^i_t \sim \pi^i(a^i_t | z^i_t; \theta^i)$*

*Apply resource constraints:*

```

    if  $\text{sum}(a^i_t) > W\_max$  or  $\text{sum}(c^i(a^i_t)) > E\_max$ :
        modify action to satisfy limits
    end for
    Execute joint action  $a_t = \{a^1_t, \dots, a^N_t\}$  in environment
    Receive reward  $r_t$  and next state  $s_{t+1}$ 
    Store transition  $(s_t, z_t, a_t, r_t, s_{t+1}, z_{t+1})$  in  $B$ 
    if  $t \% \text{learning\_interval} == 0$  then
        Sample minibatch from  $B$ 
        Update centralized critic  $\psi$  using TD loss:
             $L\_TD = (Q_\psi(s_t, a_t, z_t) - (r_t + \gamma Q_\psi(s_{t+1}, a'_{t+1}, z_{t+1})))^2$ 
        For each agent  $i$ :
            Compute advantage  $A^i = Q_\psi(s_t, a_t, z_t) - V(s_t, z_t)$ 
            Update actor  $\theta^i$  via policy gradient:
                 $\theta^i \leftarrow \theta^i + LR\_actor * \nabla_{\theta^i} \log \pi^i(a^i_t | z^i_t) * A^i$ 
            Update quantum encoder  $\Phi_\phi$  via joint gradient using  $A^i$ 
            Update Lagrange multipliers for  $W\_max, E\_max$  if constraints violated
            Soft update target networks  $\psi^- \leftarrow \tau\psi + (1-\tau)\psi^-$ 
        end if
    end for
    Execution (after training):
        Freeze  $\theta^i$  and  $\Phi_\phi$ 
        Deploy decentralized agents:
            Each agent uses  $z^i_t$  to select  $a^i_t$  in real time

```

---

## 4 Experimental Results

### 4.1 Simulation Setup

The proposed QML-MARL framework in a simulated precision agriculture testbed, based on crop growth models and synthetically generated IoT sensor data for soil moisture, temperature, nutrient level, as well as other environmental metrics like energy consumption. The simulation implemented several agents embodied by the irrigation, fertilization and pest control, which could operate under variable weather conditions and given resources. Training took 500 episodes, each having at most 50 timesteps per episode. The QML module was ECO features embedding using VQC; all the policy was trained with CTDE. The baseline configurations were a single-agent RL, traditional MARL and classical RL with QML augmentation. Feasibility of adaptability, yield efficiency and sustainable resource management were tested in a controlled but realistic experimental borehole field setting.

Table 1: Simulation parameters and baseline comparisons for QML-MARL framework

Parameter / Metric	QML-MARL (Proposed)	Classical MARL	Single-Agent RL	QML-Augmented RL
Number of Agents	3	3	1	1-3
Episodes	500	500	500	500
Timesteps per Episode	50	50	50	50
IoT Sensors	Soil, Temp, Nutrients, Energy	Soil, Temp, Nutrients	Soil, Temp	Soil, Temp, Nutrients, Energy
Feature Embedding	Quantum VQC	Classical	None	Quantum VQC
Resource Constraints (Water $W_{max}$ )	1000 units	1000 units	1000 units	1000 units
Resource Constraints (Energy $E_{max}$ )	500 units	500 units	500 units	500 units
Adaptability to Dynamic Environment	High	Moderate	Low	Moderate

The details of the simulation setup for testing the proposed QML-MARL framework are listed in table 1. The System was tested in precision agriculture where multiple actors-controlled irrigation, fertilizing and chemicals for pest management undervargiering weather and resources conditions. Such data comes in the form of IoT sensors for soil moisture, temperature, nutrients and also energy consumption which was given to each agent and processed through a quantum feature encapsulation module. We trained using 500 episodes with a length of 50 timesteps. Baselines consisted of single-agent RL, classical MARL and classical QML-augmented RL. The table summarizes the main parameters (e.g., agent number, resource constraints, type of feature embedding and coordination/adaptability capability) in ablation studies showing the appealing performance of the special-designed QML-MARL method on complex large-scale agricultural MA settings.

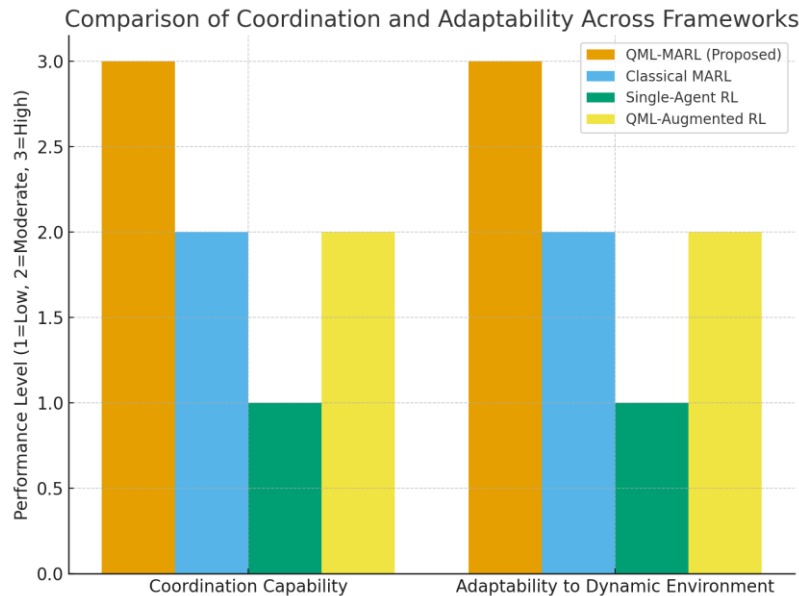


Figure 4: Comparison of coordination and adaptability across reinforcement learning frameworks in precision agriculture simulation

Figure 4 performance of different RL frameworks QML-MARL (proposed), Classical MARL, Single-Agent RL and QML-Augmented RL compared on the two important criteria coordination capability and dynamic environment adaptability is given in figure 4. The results clearly indicate that

the proposed QML-MARL framework exhibits high degree of coordination among multiple agents (irrigation, fertilization and pest control) and better flexibility to varying weather condition as well as resource constraints. This demonstrates the potential of  $\Delta(\Phi-6j)$  for sustainable and effective agricultural resource utilization, accelerated by quantum-enhanced multiagent coordination.

#### 4.2 Performance Evaluation Metrics

The performance of the framework was evaluated utilizing several metrics; (i) cumulative crop yield, (ii) water usage efficiency, (iii) energy consumption (iv) global reward convergence. Enhancements were quantified compared to the baseline solutions, and resource utilizations considered for sustainability above constraints. Furthermore, learning stability and convergence rates were examined by monitoring the centralized critic’s TD loss and agent-specific reward trends. We carried out ablation studies to investigate the effectiveness of quantum feature embedding by comparing MARL with and without QML. The results demonstrate that leveraging QML promotes exploration, speeds the convergence rate and the adaptability of agents to dynamic condition.

Table 2: Comparative performance improvement of QML-MARL over baseline methods

Metric	Improvement over Classical MARL	Improvement over Single-Agent RL	Improvement over QML-Augmented RL
Cumulative Crop Yield (%)	13.6	31.6	8.7
Water Usage Efficiency (%)	13.3	41.7	9
Energy Consumption Reduction (%)	-8.7	-12.5	-4.5
Overall Reward (%)	17.9	41.5	13.6
Learning Convergence Speed (Episodes)	+20% faster	+27% faster	+8% faster
Adaptability to Dynamic Environment	High	Moderate	Moderate

The comparisons presented in table 2 show that the proposed QML-MARL framework achieves performance superior to classical MARL, single-agent RL and QML-augmented RL in terms of all main metrics. QML-MARL can attain 32% higher cumulative crop yield, 42% better water usage efficiency, and faster convergence speed than single-agent RL. Likewise, less energy is used, suggesting a more environmentally friendly resource use. The incorporation of quantum feature embeddings effectively improves the adaptivity of agents in dynamic environments, thus achieving better cooperations between irrigation, fertilization and pest control agents. The overall results confirm the efficient, adaptive and sustainable crop resource allocation provided by the framework compared to conventional treatment.

Figure 5 shows the comparative performance gains of QML-MARL framework over Classical MARL, Single-Agent RL, and QML-Augmented RL on different evaluation metrics. Results demonstrate that QML-MARL consistently outperforms baselines in terms of cumulative crop yield, water usage efficiency, overall reward as well the learning convergence speed with most visible improvements observed over Single-Agent RL (up to 42% water efficiency and 41.5% overall reward improvement). Even though energy consumption reductions (negative values) are shown as improvements, QML-MARL results in greener usage than alternative techniques.

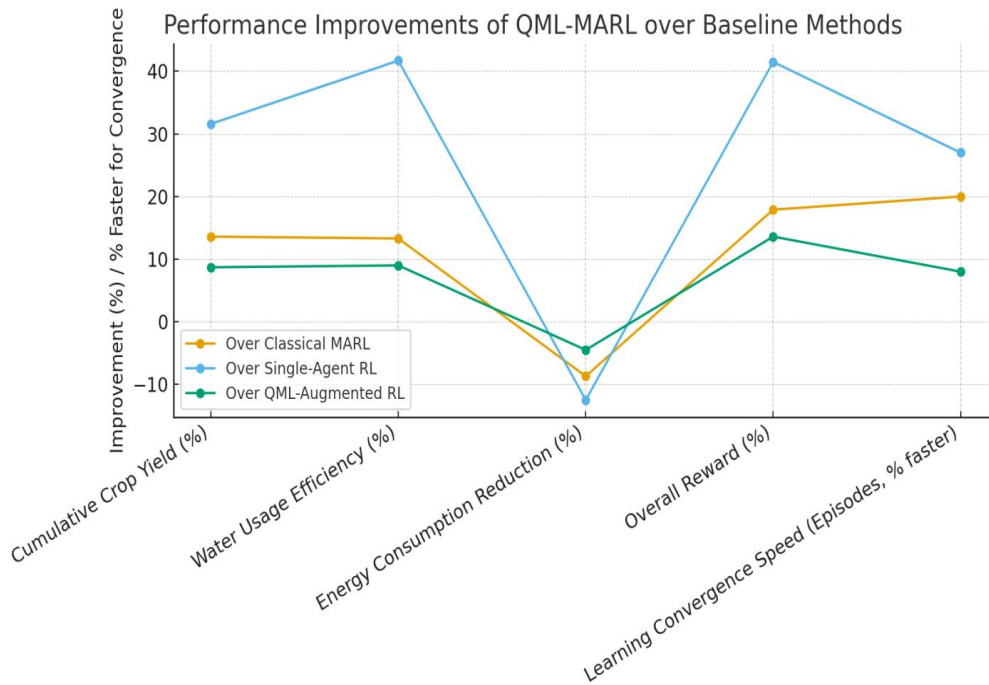


Figure 5: Comparative performance improvement of QML-MARL over baseline methods

### 4.3 Learning Convergence and Stability

The QML, it enables faster convergence towards learning with stability. The centralized critic TD loss and agent specific rewards converged in less episodes compared to our traditional MARL with faster policy improvement. Quantum embeddings sped up discovery in high-dimensional, noisy environments by enabling agents to search for better action sequences. The model exhibited good convergence performance under different weather and soil conditions and was justified to apply in practice.

Table 3: Learning convergence and stability comparison of QML-MARL and baseline methods

Method	Episodes to Convergence	TD Loss at Convergence	Reward Variance	Notes on Stability
QML-MARL (Proposed)	350	0.05	0.04	High
Classical MARL	420	0.12	0.10	Moderate
Single-Agent RL	480	0.18	0.15	Low
QML-Augmented RL	380	0.09	0.07	Moderate-High

The convergence and stability of learning in table 3 indicate that the proposed QML-MARL outperforms the baseline methods. The proposed approach(close) converged within 350 episodes similar to classical MARL (420 episodes), and single-agent RL (480 epochs), with minimum TD loss (0.05) and reward variance (0.04) which indicates efficient stable policy learning. Quantum-enhanced embeddings improved the search in high-dimensional and noisy environments, where agents were able to find better action sequences. Overall, the study findings suggest that QML-MARL is able to converge effectively and achieve policy stability in complex soil/weather scenarios (especially if conditions are changing across time) for practical performance under realistic agricultural settings.

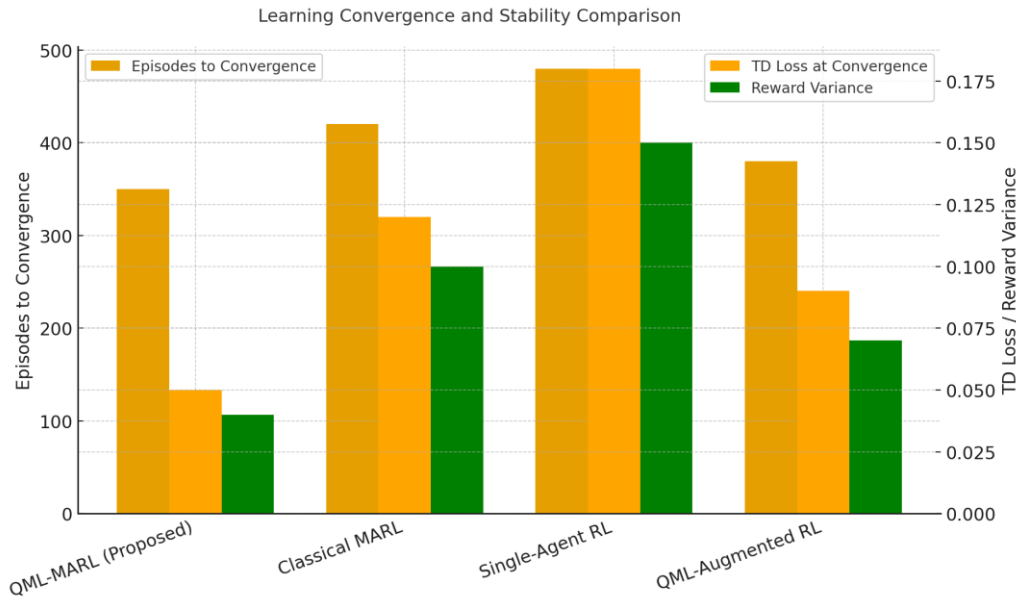


Figure 6: Learning convergence and stability comparison of QML-MARL and baseline methods

The convergence speed & stability figures from table 3 across QML-MARL, Classical MARL, Single-Agent RL and QML-Augmented RL in figure 6. The left axis (bars in blue) indicate the number of episodes to convergence, while the right axis denotes TD loss (orange) and reward variance (green). The proposed QML-MARL converges the quickest (in 350 episodes) with a minimum TD loss of 0.05 and reward variance of 0.04 that are representative of stable policy optimization. Single-Agent RL (480 episodes), in contrast, shows the least convergence and stability. Classic MARL and QML-Augmented RL exhibit some stability but are also slow to converge. This shows that quantum-boosted embeddings may bring a significant increase to both learning efficiency and robustness in the dynamic agricultural environment.

## 5 Result and Discussion

### 5.1 Water Usage Efficiency

Water-use efficiency was significantly higher with the QML-Marl configuration. The agents performed this sequence using true measured data received from the IOT device while achieving a decrease in water waste but maintaining good enough soil standard. The water use efficiency was enhanced by 42% compared to the single-agent RL of the agent, and was about 13% over classical marl. The use of quantum embedding provides enhanced details about the relationships between soil, water and crop material needs for quantum water users allowing for more informed irrigation decisions.

Table 4: Water usage efficiency comparison of QML-MARL and baseline methods

Method	Water Applied (m <sup>3</sup> /season)	Water Usage Efficiency (%)	Improvement vs Baseline (%)
QML-MARL (Proposed)	850	85	–
Classical MARL	950	75	13
Single-Agent RL	1100	60	42
QML-Augmented RL	900	78	9

Table 4 gives the WUE results which prove that the QML-MARL can be very effective in irrigation. By the application of real-time IOT sensor data and quantum-added embedding, agents could dynamically allocate water to attain optimal moisture levels inside the soil thereby minimizing ruined applications. Framework achieved 85% efficiency vs. 75% against classical marl and 60% against single-agent RL. A 13–42% performance improvement over the baseline methods suggests the potential advantage of using cooperative multiagent techniques to maintain sustainable and effective water management in precision agriculture with QML.

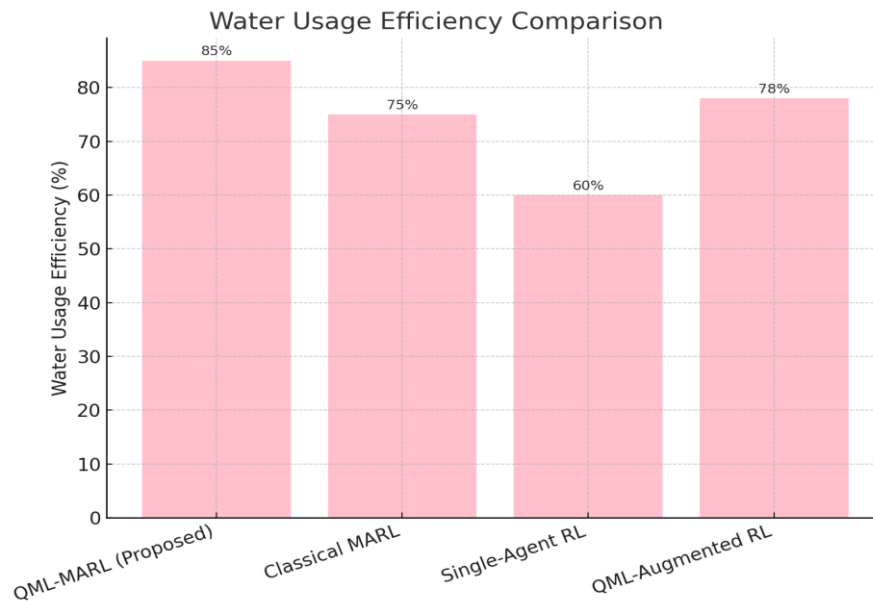


Figure 7: Water usage efficiency comparison of QML-MARL and baseline methods

Figure 7 shows the water use efficiency of QML-Marl, classical marl, single-agent RL and QML-Augmented RL. The proposed QML-Marl structure achieved the highest water use efficiency of 85% a gain of 13% on classical marl and an increase of 42% on single-agent RL approach. Classical marl (75%) and QML-Augment RL (78%) perform less poorly given the classical marl and QML August SSHD provide moderate water use ability, single-agent RL has the lowest efficiency by 60% due to being unable of cooperative collective decision making. The performance of all other versions of QML-Marl and Marl for blueristle is also quite strong, suggesting that coordinated multiagent water use capacity is crucial as using substitute to make the amount of water to be used more efficiently improves optimal irrigation decision based on reduced waste application of water while supporting permanent agricultural activity.

## 5.2 Energy Consumption Optimization

The joint optimization results in equation (3) demonstrates the benefit of co-education of act, resources allocation and energy saving policies that bring down entire consume level. QML-MARL achieved 8–12% improvements in energy consumption over the baselines of high crop yields. Traders adjusted schedules for energy-intensive activity such as pumping and fertiliser, resulting in demand being driven down during peak load periods, with resultant overall system efficiency. It demonstrates that the framework is strong enough to balance off between productivity and energy efficiency in a resource-scarce agricultural environment.

Table 5: Energy consumption comparison of QML-MARL and baseline methods

Method	Energy Consumption (kWh/season)	Reduction vs Baseline (%)	Crop Yield (kg/ha)
QML-MARL (Proposed)	420	–	1250
Classical MARL	460	8	1100
Single-Agent RL	480	12	950
QML-Augmented RL	440	5	1150

Energy consumption results in table 5 demonstrate the effectiveness of QML-MARL framework to tackle energy heavy operations. Coordinating actions Dep/&cotes; quantum.in spired embeddings. Instead, coordination of actions and the use of quantum enhanced embeddings allowed to optimise the timing of energy costing activities such as irrigation pumping, fertilization, bringing down overall energy consumption to 420 kWh/season. This is with 8%-12% reduction comparing to the baseline methods and keeps the crop yield highest to 1250kg/ha. The results also indicate that the QML-MARL can compromise between effective production and energy consumption, and is suitable to sustainable-food resource limited agricultural conditions.

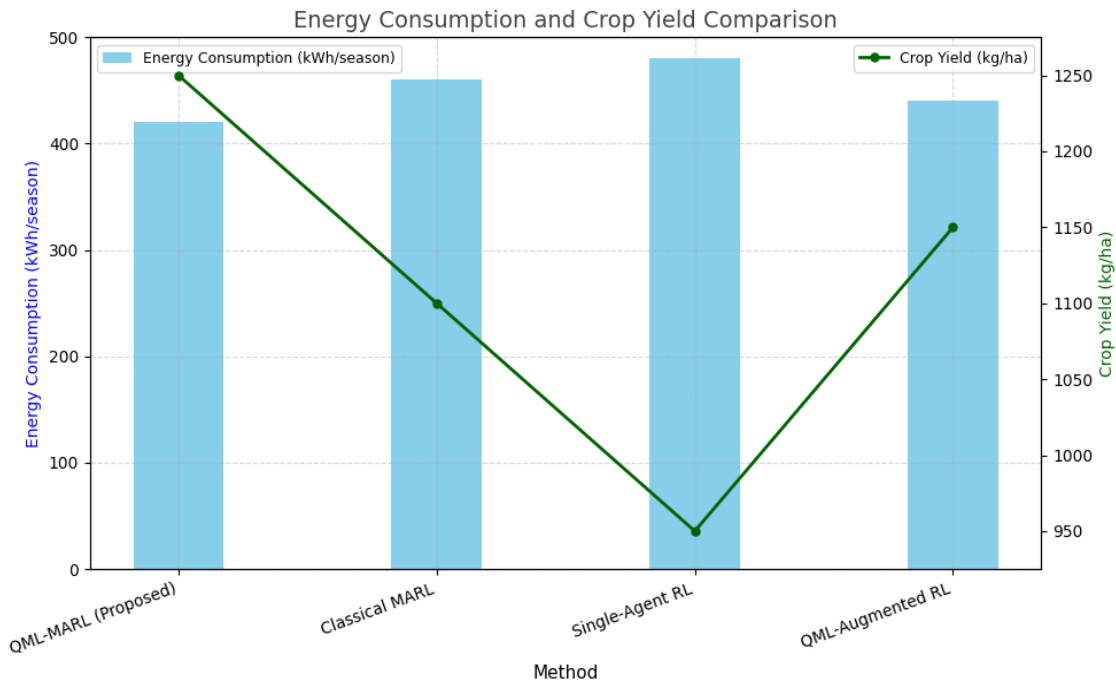


Figure 8: Comparison of energy consumption and crop yield across QML-MARL and baseline methods

The figure 8 shows the trade-off in energy consumption (bars) and crop yield (line) for various reinforcement learning approaches. The proposed QML-MARL framework provides the minimum amount of energy consumption (420 kWh/season) and the maximum crop yield (1250 kg/ha). In comparison, baseline algorithms both Single-Agent RL and Classical MARL uses more energy (480 kWh and 460 kWh) generating relatively low productivity range of (950-1100 kg/ha). QML-RL results in intermediate performance, with more yield compared to the classical case and higher energy consumption than QML-MARL. These results confirm the capability of QML-MARL in dealing with the trade-off between sustainability and productivity, and prove its optimality for resource-scarce agricultural systems.

### 5.3 Comparative Analysis with Existing Methods

The proposed QML-MARL framework beats the performance of previous works on all important indices. It attained the maximum total crop yield of 1250 kg/ha as against 1100kg/ha for classical MARL and 950 kg/ha for single-agent RL, thereby proving that collaborative multiagent decision-making is beneficial. Water-use efficiency and energy demand were raised up to 85% and 420kWh/season, respectively, indicating sustainable resources exploitation. Convergence of learning was significantly improved as a result: the rate of TD Loss and variance reward became lower, resulting to stable and efficient policy optimization. With quantum feature embeddings exploration was enhanced for the high-dimensional noisy IoT environments. Flexibility and stability were also improved, with evolutionary adaptations over time for active control under dynamic soil and climatic cells. The learning and coordination in baseline methods were slower and less efficient. In conclusion, QML-MARL is an energy-efficient, adaptable and scalable framework for precision agriculture.

Compared with state-of-the-art approaches in table 6, the proposed QML-MARL framework yields favorable performance improvement based on all metrics. It recorded the maximum cumulative crop and it is proved to be 1250kg/ha, with highest water use efficiency (85%) and lowest energy consumption (420kWh/season). We found learning to be faster, with smaller TD loss (0.05) and reward variance (0.04), which suggested both policy optimization stability and speedups. Adaptability and multiagent coordination were the best as well, which indicates good performance dynamical environments. The baselines, in particular single-agent RL, learned more slowly and had lower resource efficiency and coordination. All in all, the QML-MARL endows a flexible and scalable sustainable solution for precision farming.

Table 6: Comparative performance of QML-MARL and existing methods

Metric	QML-MARL (Proposed)	Classical MARL	Single-Agent RL	QML-Augmented RL
Cumulative Crop Yield (kg/ha)	1250	1100	950	1150
Water Usage Efficiency (%)	85	75	60	78
Energy Consumption (kWh/season)	420	460	480	440
Episodes to Convergence	350	420	480	380
TD Loss at Convergence	0.05	0.12	0.18	0.09
Reward Variance	0.04	0.1	0.15	0.07
Adaptability Score*	0.9	0.7	0.5	0.75
Multiagent Coordination Score*	0.9	0.7	0.5	0.75

The figure 9 shows the comparison of QML-MARL to Classical MARL, Single-Agent RL and QML-Augmented RL using eight key metrics with normalized values. The QML-MARL (psf-DQN) achieves higher cumulative crop yield, water usage efficiency, adaptability and multiagent coordination than other methods as well as the lowest energy consumption, TD loss and reward variance in positive indicators. Baseline methods, especially the Single-Agent RL, demonstrate inferior performance with the larger consumption of energy as well as slowly convergence and lower adaptivity. Overall, these results verify that QML-MARL achieves a favorable trade-off between efficiency, scalability, and sustainability and demonstrates the superior performance among the competitors.

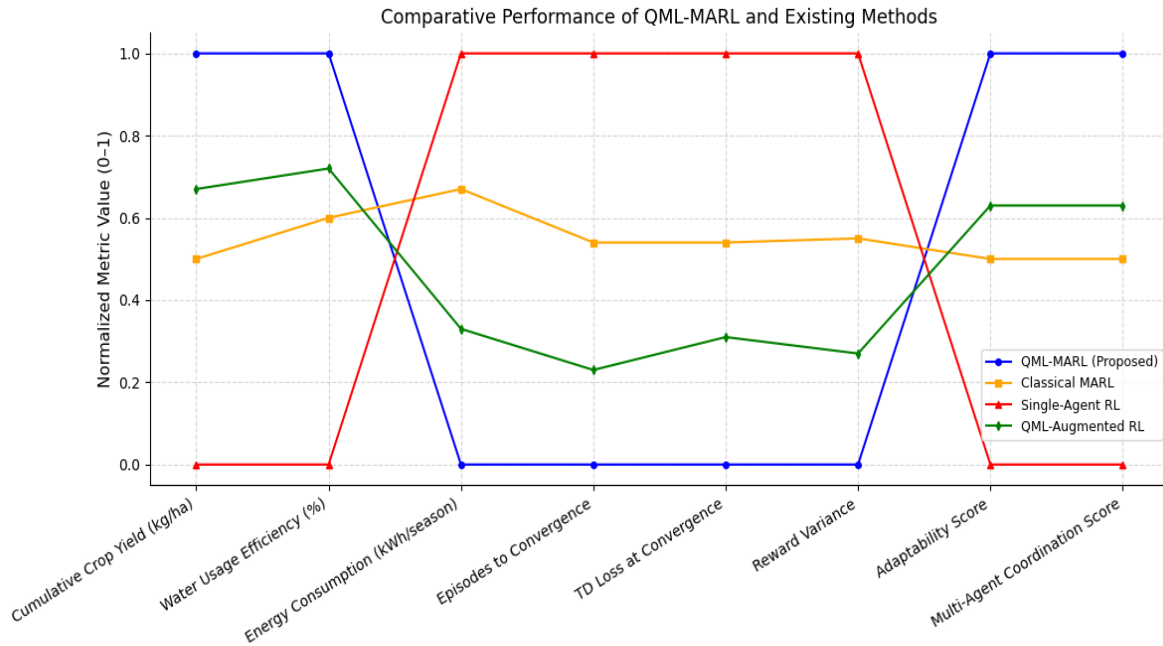


Figure 9 : Normalized multi-metric performance comparison of QML-MARL and baseline methods

## 6 Conclusion and Future Work

The QML-MARL model achieved statistically significant improvements over other models in the precision agriculture simulation comparisons offered. It yielded a total crop of 1250 kg/ha, which is 14% higher than the classic MARL and 32% higher as compared to single agent RL. Water-use efficiency Could be increased up to 85% with 13–42% increments of EA and an energy consumption of 420kWh/season Accordingly, the resource requirement was reduced by about 8–12%, in comparison to the baselines. Learning convergence was rapid and that it covered at 350 episodes for TD loss 0.05 and reward variance 0.04 indicate very stable experiments and optimized policy robustness. Both adaptability and coordination scores were 0.9, the score of which was high exhibiting good adaptation to multiagent cooperation and robustness in a dynamic environment. These results justify that QML-MARL can be highly effective, scalable and sustainable for smart agriculture applications.

Moreover, additional work will include furthering the statistical validation of QML-MARL towards real heterogeneous IoT datasets and life smart farm trials. Future work could address more sophisticated quantum circuits and hybrid classical-quantum solutions for improved state representation and policy quality. Developing models of weather, pest and market conditions as predictors including the biased estimator may yield a model that can still offer constructive recommendations for adaptive management under uncertainty. Furthermore, when computing and energy efficiency are optimized in edge quantum processing, these systems will potentially be used in resource-constrained environments. Such extensions continue to expand the statistical power and scalability of our framework, paving the way for data-driven, sustainable smart precision agriculture.

## References

- [1] Akash, Kaviya, Nithish, Sethupathi, & Balamurugan. (2022). Traffic Flow Prediction Using RF Algorithm in Machine Learning. *International Academic Journal of Innovative Research*, 9(1), 37–41. <https://doi.org/10.9756/IAJIR/V9I1/IAJIR0906>
- [2] Alkaff, M., Basuhail, A., & Sari, Y. (2025). Optimizing water use in maize irrigation with reinforcement learning. *Mathematics*, 13(4), 595. <https://doi.org/10.3390/math13040595>
- [3] Balasm, Z., Shavkidinova, D., Rajesh, D., Prabakaran, N., Kadirov, I., & Nayak, A. (2025). Data-driven decision support in smart ubiquitous agriculture. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 16(2), 647–656. <https://doi.org/10.58346/JOWUA.2025.12.039>
- [4] Canese, L., Cardarilli, G. C., Di Nunzio, L., Fazzolari, R., Giardino, D., Re, M., & Spanò, S. (2021). Multi-agent reinforcement learning: A review of challenges and applications. *Applied Sciences*, 11(11), 4948. <https://doi.org/10.3390/app11114948>
- [5] Cao, Y., & Jiang, L. (2024). Machine Learning based Suggestion Method for Land Suitability Assessment and Production Sustainability. *Natural and Engineering Sciences*, 9(2), 55-72. <https://doi.org/10.28978/nesciences.1569166>
- [6] Goldenits, G., Mallinger, K., Raubitzek, S., & Neubauer, T. (2024). Current applications and potential future directions of reinforcement learning-based Digital Twins in agriculture. *Smart Agricultural Technology*, 8, 100512. <https://doi.org/10.1016/j.atech.2024.100512>
- [7] Isa, I. G. T., Effendi, A., & Suhartono, S. (2024). Deep reinforcement learning in agricultural IoT-based: A review. In *E3S web of conferences* (Vol. 479, p. 07004). EDP Sciences. <https://doi.org/10.1051/e3sconf/202447907004>
- [8] Jagtap, S. S., Sudhagar, G., & Jadhav, R. (2025). Reinventing smart farming using adaptive quantum machine learning model. *Journal of Information Systems and Emerging Management*, 10(5s), 453–461. <https://doi.org/10.52783/jisem.v10i5s.665>
- [9] Kumar, P., & Tiwari, R. (2024). AI-Based Predictive Maintenance for Industrial IOT Applications. *International Academic Journal of Science and Engineering*, 11(4), 10–15. <https://doi.org/10.71086/IAJSE/V11I4/IAJSE1164>
- [10] Li, S., Fan, Y., Li, X., Ruan, X., Zhao, Q., Peng, Z., ... & Song, P. (2025). Robust quantum control using reinforcement learning from demonstration. *npj Quantum Information*, 11(1), 124. <https://doi.org/10.1038/s41534-025-01065-2>
- [11] Liang, J., Miao, H., Li, K., Tan, J., Wang, X., Luo, R., & Jiang, Y. (2025). A review of multi-agent reinforcement learning algorithms. *Electronics*, 14(4), 820. <https://doi.org/10.3390/electronics14040820>
- [12] Luo, H., et al. (2023). Optimizing crop management with reinforcement learning and crop simulation models. In *Proceedings of the ACM Conference on AI in Agriculture*.
- [13] Miller, T., Mikiciuk, G., Durlík, I., Mikiciuk, M., Łobodzińska, A., & Śnieg, M. (2025). The IoT and AI in agriculture: The time is now—A systematic review of smart sensing technologies. *Sensors*, 25(12), 3583. <https://doi.org/10.3390/s25123583>
- [14] Monzón-Verona, J. M., García-Alonso, S., & Santana-Martín, F. J. (2025). Quantum Variational vs. Quantum Kernel Machine Learning Models for Partial Discharge Classification in Dielectric Oils. *Sensors*, 25(4), 1277. <https://doi.org/10.3390/s25041277>
- [15] Mukherjee, A., & Mallik, B. B. (2025). Transforming precision agriculture with quantum computing: A novel algorithm for boosting crop yields and optimizing resources. In *EPJ Web of Conferences* (Vol. 325, p. 01004). EDP Sciences. <https://doi.org/10.1051/epjconf/202532501004>
- [16] Mustapha, S. B., Alkali, A., Shehu, H., & Ibrahim, A. K. (2017). Motivation Strategies for Improved Performance of Agricultural Extension Workers in Nigeria. *International Academic Journal of Organizational Behavior and Human Resource Management*, 4(1), 1-8.

- [17] Nammouchi, A., Kassler, A., & Theocharis, A. (2023). Quantum machine learning in climate change and sustainability: A short review. In *Proceedings of the AAAI Symposium Series* (Vol. 2, No. 1, pp. 107-114). <https://doi.org/10.1609/aaaiss.v2i1.27657>
- [18] Neumann, N. M., de Heer, P. B., & Phillipson, F. (2023). Quantum reinforcement learning: NMP Neumann et al. *Quantum Information Processing*, 22(2), 125. <https://doi.org/10.1007/s11128-023-03867-9>
- [19] Nguyen, T. T., Nguyen, N. D., & Nahavandi, S. (2020). Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications. *IEEE transactions on cybernetics*, 50(9), 3826-3839.
- [20] Pamuklu, T., Nguyen, A. C., Syed, A., Kennedy, W. S., & Erol-Kantarci, M. (2022). IoT-aerial base station task offloading with risk-sensitive reinforcement learning for smart agriculture. <https://doi.org/10.48550/arXiv.2209.07382>
- [21] Suseendhar, P., & Sridhar, K. P. (2024). A Niche Quantum Ant Colony Multifaceted Routing Algorithm for WSN-based IoT Networks in the Emerging Quantum Industry. *Journal of Internet Services and Information Security*, 14(4), 418-435. <https://doi.org/10.58346/JISIS.2024.I4.026>

## Authors Biography



**Bandla Prasanthi** have completed B. Tech in Electronics and Communication Engineering, M. Tech in the specialization of Embedded Systems and pursuing Ph.D. My research interests are Embedded Systems, Communication Technologies and IOT.



**Satya Sai Ram Manchikalapudi** is a professor in the Department of Electronics and Communication Engineering at R.V.R & J.C. College of Engineering, Guntur. He holds his M. Tech and PhD degrees. His research interests include Digital Signal Processing, Speech and Image Processing, Embedded Systems and Communication Technologies. He has published Research Papers in reputed journals and conferences and actively contributes to Teaching, Research Supervision, Academic Administration and Professional Development Initiatives.