Volume 2, 2026, Pages 225-238 *Journal of Smart and Sustainable Farming*
| Regular Article | Open Access

Adaptive Multi Agent Intelligence for Coordinated Resource Management in Sustainable Farming Systems

Muhammad Farooq
Department of Agriculture, American University of Beirut, Riad El Solh, Beirut 1107-2020, Lebanon.
fa58955@hotmail.com

Article Info
Journal of Smart and Sustainable Farming
https://www.ansispublications.com/jssf/jssf.html

© The Author(s), 2026.
https://doi.org/10.64026/JSSF/2026019

Received 23 April 2026
Revised from 26 August 2026
Accepted 15 September 2026
Available online 03 October 2026
Published by Ansis Publications.

Corresponding author(s):
Muhammad Farooq, Department of Agriculture, American University of Beirut, Riad El Solh, Lebanon.
Email: fa58955@hotmail.com

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Abstract – Farms must share scarce resources, groundwater, pumping capacity, solar energy and irrigation labour, among fields that differ in crop, soil and economic value, under weather and market conditions that change from season to season. Centralised optimisation requires complete information and a single decision maker, whereas uncoordinated, field-by-field decisions can waste capacity at peak demand. This paper develops an adaptive multi-agent system for farm resource management in which parcel agents learn irrigation policies by Q-learning, resource agents (water, energy, crew) publish adaptive prices that reflect congestion and the pace of seasonal water use, and requests are allocated within pump, crew and seasonal caps. The system was evaluated on a simulated 16-parcel farm with three crops, a seasonal groundwater allocation of 70 % of unconstrained demand and a limited pump and crew capacity, in six scenarios (normal, dry, wet, pump failure, price shock and a combined dry–failure–shock case) with three independent training seeds. The proposed system achieved 95 % of the net value of an omniscient central planner in a normal year and 84 % in a dry year, with a yield Gini of 0.025 against 0.085 for the planner and a lowest parcel yield of 0.82 against 0.66, while independent Q-learners reached only 80 % with large variation between training runs. Ablation showed that the adaptive price signal was the essential component (net value fell to 75 % without it) whereas the bid-based priority ranking added nothing. Simple heuristics were competitive when capacity was not scarce, and the central planner remained clearly better in dry years. After a severe combined shock, continued learning with drift-triggered exploration did not outperform the frozen policy, so most of the robustness came from the daily price adaptation.

Keywords – Multi-Agent Systems, Reinforcement Learning, Resource Coordination, Irrigation, Water–Energy Nexus, Adaptive Pricing, Sustainable Farming.

I. INTRODUCTION

Sustainable farming depends on how a limited set of shared resources is allocated: groundwater and the pumps that lift it, solar and grid electricity, labour for operating valves and machinery, and the nutrients and inputs bought for each field. Digital technologies for smart farming, from sensors and drones to decision support, now generate the data to manage these resources at a fine scale [1], and machine learning is increasingly used to turn the data into decisions [2]. However, most decision tools are designed for a single field or a single resource. On a real farm or in a producer group, many fields draw on the same well, the same crew and the same energy supply, so that the decision for one field changes what is available for the others. Shared resources of this kind are classical commons problems, in which independent, individually rational use can deplete the resource or waste capacity unless suitable rules or institutions are in place [3].

Multi-agent systems offer a natural framework for such problems. In a multi-agent system, autonomous agents perceive their environment, decide and act, and coordinate with others through communication, negotiation or shared incentives [4]. Representing each field as an agent, and each shared resource as a resource agent that grants access under rules, mirrors the distributed structure of farms and allows decision making to remain local, an advantage when communication is limited,

data are private, or components fail. At the same time, farm conditions are non-stationary: weather differs between seasons, pumps fail, and crop prices shift, so that a fixed coordination rule tuned for one situation may perform poorly in another. The agents therefore need to adapt, both quickly (within days) and over longer time scales (between seasons) [5].

This paper develops and evaluates an adaptive multi-agent system for coordinated farm resource management. Its contributions are: (1) a decentralised architecture in which parcel agents learn irrigation requests by reinforcement learning and resource agents manage scarce capacity through adaptive prices and priority allocation; (2) an environment that couples soil-water dynamics, crop yield response, solar and grid energy, pump and crew capacities and a seasonal groundwater allocation; (3) a comparison with uncoordinated, calendar, market-based, centralised and independent-learning alternatives in six scenarios, including resource failures and price shocks; (4) an ablation of the coordination components; and (5) an analysis of adaptation after a severe shock, which tests whether continued learning adds value beyond the adaptive price signal. The results are reported with their weaknesses, including the cases in which simple heuristics or the central planner outperform the learned system.

II. RELATED WORKS

Agent-Based Modelling and Multi-Agent Systems in Agriculture

Agent-based modelling has long been used in agriculture to represent the heterogeneous decisions of farms, households and institutions, mainly for the simulation of policy effects such as subsidy schemes and land-use change; a review of agent-based models for agricultural policy evaluation shows that most applications are explanatory simulations at regional scale, with relatively few models that embed learning or support operational decisions on a farm [6]. Operational multi-agent systems on farms have been proposed for tasks such as coordinating machinery, monitoring and irrigation with sensor networks, but these are mostly rule-based and are rarely evaluated against centralised or learning-based baselines under changing conditions.

Market-Based and Negotiation-Based Coordination

Coordination of distributed agents has been studied through negotiation protocols and through markets. The contract net protocol lets a manager agent announce tasks and award them to bidding contractors [7], and market-oriented programming allocates resources in distributed systems by letting agents trade at equilibrium prices computed by auctioneers, which can solve resource allocation problems without a central planner [8]. Price-based coordination also has a firm basis in optimisation: dual decomposition and the alternating direction method of multipliers decompose a large problem into local subproblems that are coordinated by prices (dual variables) updated from the violation of shared constraints [9]. These ideas underlie the resource agents of the present work; whose prices rise when demand exceeds capacity or the seasonal water budget is consumed too fast. Their classical forms assume that agents know their own value functions; in farming, such values depend on stochastic weather, crop stage and soil water, and must be estimated or learned.

Multi-Agent Reinforcement Learning

Reinforcement learning allows agents to learn policies from interaction. The simplest multi-agent approach, in which every agent runs its own learner and treats the others as part of the environment, was found to work in some settings but to suffer when the agents' interactions make the environment non-stationary, and communication or shared information helps [10]. Tabular Q-learning, which updates action values from the temporal-difference error [11], remains a practical baseline because it is simple and robust with small state spaces. Surveys of multi-agent reinforcement learning describe the challenges of non-stationarity, credit assignment and scalability and the main families of methods, such as team learning, independent learners and centralised training with decentralised execution [12], and actor–critic methods with centralised critics have been proposed for mixed cooperative and competitive settings [13]. These methods are mostly evaluated on games and robotics benchmarks, and their use in resource management on farms is less explored.

Learning-Based Irrigation and Resource Management

In agriculture, reinforcement learning has been applied to irrigation and greenhouse control: for instance, deep reinforcement learning has been used in IoT-based smart agriculture systems to decide irrigation actions from sensor data [14], and deep reinforcement learning for irrigation scheduling has been shown to use less water than conventional schedules while maintaining yield in simulation and field-based studies [15]. These studies treat a single field or a single controller and an unconstrained water supply. They do not address the competition among fields for limited pump capacity, crew time and a seasonal allocation, or what happens when the capacity changes unexpectedly.

Positioning of This Work

Table 1 summarises the main lines of work and how the present study relates to them. The distinctive features here are the combination of learning parcel agents with adaptive resource prices, the explicit modelling of several shared capacities, the evaluation under resource failures and price shocks, and an ablation that separates the contribution of prices, priority allocation and learning.

Table 1. Comparison of Related Approaches with the Present Study

Line of work	Coordination	Learning	Shared resources	Main limitation for farms
Agent-based policy models [6]	Simulated interaction	Rarely	Land, markets	Explanatory, regional scale
Contract net / market-oriented [7,8]	Bidding, equilibrium prices	None	General	Needs known value functions
Dual decomposition, ADMM [9]	Price (dual) updates	None	General	Static problems, central solver
Independent learners [10,11]	None	Q-learning	General	Non-stationarity, instability
MARL surveys and methods [12,13]	Centralised critic, communication	Deep RL	Games, robotics	Data hungry; few farm studies
RL for irrigation [14,15]	Single controller	Deep RL	Unconstrained water	No competition for capacity
This study	Adaptive prices + priority allocation	Tabular Q-learning, drift reset	Water, pump, crew, energy	Simulated farm only

III. MATERIALS AND METHODS

System Architecture

The system (**Fig. 1**) has two kinds of agents. Parcel agents, one per field, observe the soil-water state, crop stage and a weather forecast, and request an irrigation event of none, 12 mm or 24 mm each day. Resource agents represent the water source (a borewell and tank with a daily pump capacity and a seasonal allocation), the energy supply (solar generation and grid purchases), the irrigation crew (a limited number of valve operations per day) and the market (crop values). The resource agents receive the requests, allocate the capacities and broadcast prices and capacity status to the parcel agents (**Fig. 3**). The farm layout is shown in **Fig. 2**.

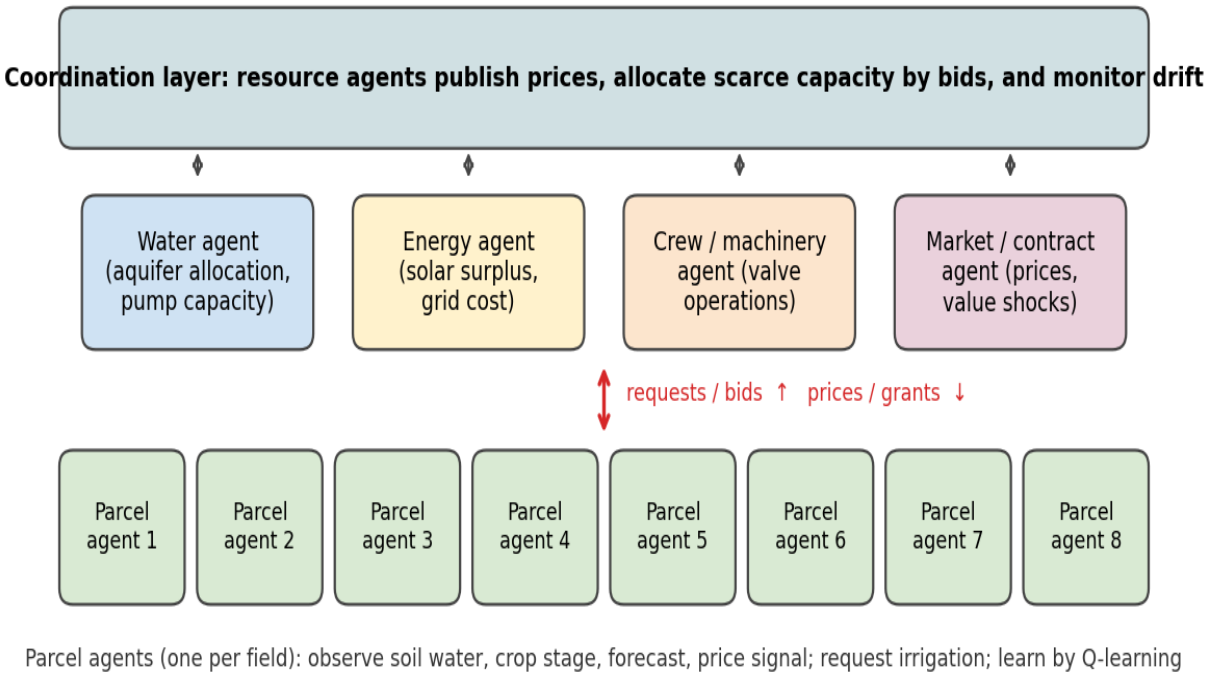


Fig 1. Architecture of the Adaptive Multi-Agent System: Resource Agents (Water, Energy, Crew, Market) Coordinate Parcel Agents Through Prices and Allocation.

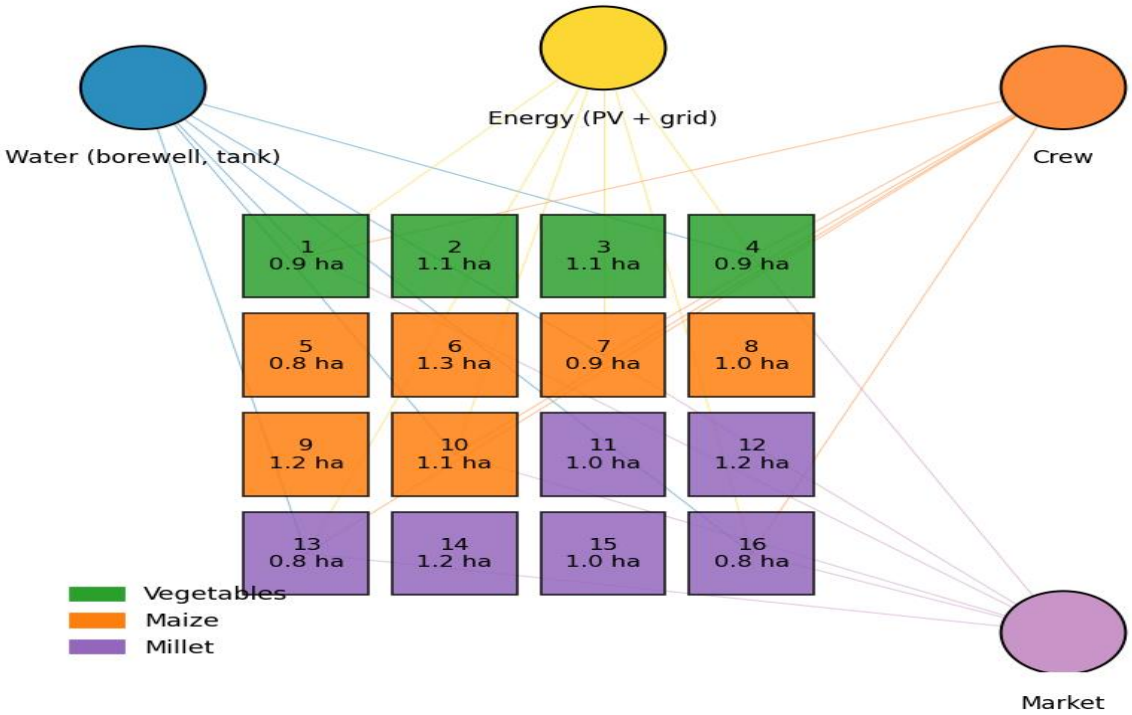


Fig 2. Simulated Farm of 16 Parcels with Three Crops, Shared Resource Agents (Circles) and Communication Links (Areas in Hectares).

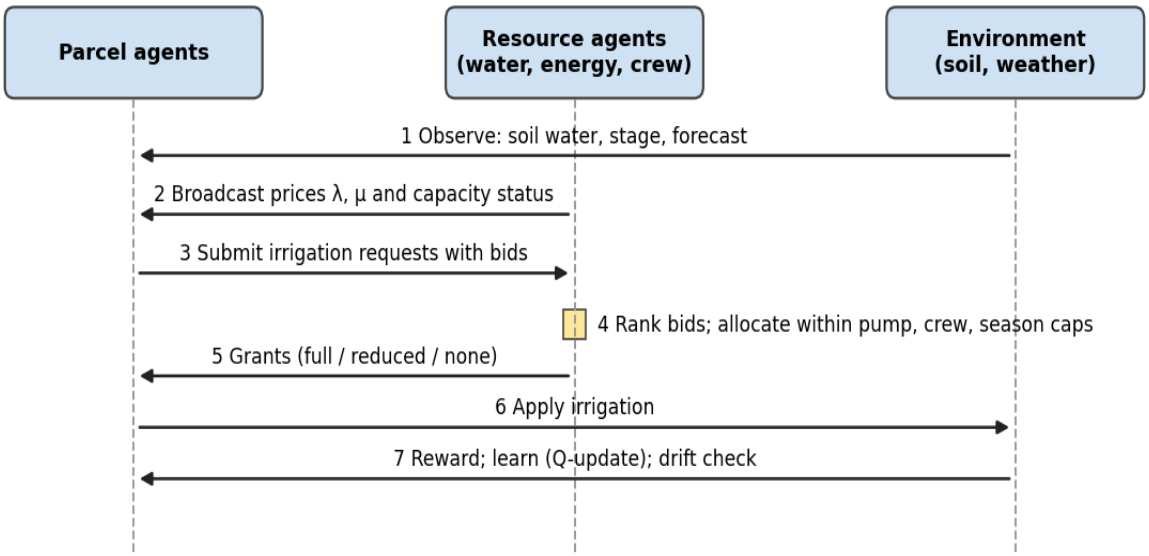


Fig 3. Daily Interaction Protocol Between Parcel Agents, Resource Agents and the Environment.

Environment Model

The farm has 16 parcels of 0.8–1.3 ha with vegetables (4 parcels), maize (6) and millet (6). A 120-day season was simulated daily. Reference evapotranspiration, rainfall and solar energy were generated stochastically for each season, with scenario-specific rainfall and evaporative demand. The root-zone water S of each parcel follows the FAO-56 water balance, in which actual evapotranspiration is reduced by a stress coefficient K_s when soil water falls below the readily available fraction,

$$S_{t+1} = \min\left(TAW, \max(0, S_t + P_e + \eta_{irr} I_t - K_s K_c ET_0)\right), \quad K_s = \min\left(1, \frac{S_t}{(1-p)TAW}\right) \tag{1}$$

where P_e is effective rainfall (80 % of rain), I the applied net irrigation, η_{irr} = 0.9 the application efficiency, p the depletion fraction (0.5 for vegetables and maize, 0.6 for millet) and TAW the total available water (60, 80 and 90 mm for

the three crops, with ±15 % variation between parcels). The relative yield of each parcel at the end of the season follows the linear water-production function of FAO Irrigation and Drainage,

$$\mathrm{RY} = 1 - \mathrm{K_y}\left(1 - \frac{\sum_t \mathrm{ET_a}}{\sum_t \mathrm{ET_c}}\right) \tag{2}$$

with yield-response factors K_y of 1.05, 1.25 and 0.85 for vegetables, maize and millet. **Table 2** lists the crop parameters and values. The resources are limited as follows: a seasonal groundwater allocation of 44.2 thousand m³ (70 % of the 63.1 thousand m³ that the uncoordinated rule would use without limits in a normal year), a pump capacity of 1321 m³ per day (30 % of the volume that would be needed if all parcels irrigated 24 mm on the same day), and a crew that can irrigate at most 6 parcels per day. Pumping needs electricity, e_p = 0.34 kWh per m³ for a head of 65 m, supplied first by solar generation of about 360 kWh per day on clear days (half on rainy days) and then by the grid at ₹7 per kWh, so that grid energy is

$$\mathrm{E_{grid}} = \mathrm{e_p}\max\left(0, \mathrm{V} - \frac{\mathrm{E_{PV}}}{\mathrm{e_p}}\right) \tag{3}$$

where V is the volume pumped in the day and E_{PV} the solar energy. Water also carries a charge of ₹1.5 per m³. The farm's seasonal net value, the main performance measure, is the crop revenue minus water and energy costs,

$$\mathrm{J} = \sum_i \mathrm{v_i A_i RY_i} - \mathrm{c_w}\sum_i \mathrm{V_i} - \tau \mathrm{E_{grid}} \tag{4}$$

with v_i the value of a fully yielding hectare (₹4,00,000, ₹1,20,000 and ₹70,000 for vegetables, maize and millet) and A_i the area.

Table 2. Crop and Farm Parameters of the Simulated Environment (Indicative Values)

Parameter	Vegetables	Maize	Millet	Notes
Parcels	4	6	6	16 parcels, 0.8–1.3 ha
TAW (mm)	60	80	90	±15 % between parcels
Kc (initial/mid/end)	0.6/1.15/0.8	0.5/1.2/0.6	0.4/1.0/0.5	25/35/35/25-day stages
Ky	1.05	1.25	0.85	FAO-33 [16]
Value (₹/ha, full yield)	4,00,000	1,20,000	70,000	Price-shock case: ×0.5, ×1.4, ×1.0
Irrigation levels	0 / 12 / 24 mm	0 / 12 / 24 mm	0 / 12 / 24 mm	Gross volume = depth / 0.9

Parcel Agents and Learning

Each parcel agent chooses an irrigation level at the start of each day from its observation: the depletion of its root zone relative to the readily available water (six bins), the crop stage (three bins) and a noisy forecast of rain tomorrow (two bins, 80 % accurate). In the proposed system, two coordination signals extend this local observation: the ratio of the remaining seasonal water to the amount planned for the remaining days (three bins), and the current resource price (four bins). Agents of the same crop share a Q-table to speed up learning. Action values are updated by Q-learning [11],

$$\mathrm{Q(s,a)} \leftarrow \mathrm{Q(s,a)} + \alpha[\mathrm{r} + \gamma \max_{\mathrm{a'}} \mathrm{Q(s',a')} - \mathrm{Q(s,a)}] \tag{5}$$

with α = 0.15 and γ = 0.9, and ε-greedy exploration that decays from 0.35 to 0.03 over 320 training episodes. Training used normal, dry and wet seasons in proportions 50, 25 and 25 %, and the pump failure, price shock and combined scenarios were never seen in training. The daily reward of parcel i is the negative of the yield value lost to water stress that day and the water and energy cost of the irrigation it received,

$$\mathrm{r_{i,t}} = -\frac{\mathrm{v_i K_{y,i} A_i}}{\sum_t \mathrm{ET_{c,i}}}\left(\mathrm{ET_{c,t}} - \mathrm{ET_{a,t}}\right) - \mathrm{V_{i,t}}(\mathrm{c_w} + \mathrm{c_e}) \tag{6}$$

where c_e is the energy cost per m³ (grid share times tariff). In the proposed system, the learning signal is shaped to promote cooperation and to internalise scarcity, by adding half of the mean reward of all agents and subtracting the cost of the resource price,

$$\tilde{\mathrm{r}}_\mathrm{i} = \mathrm{r_i} + \beta \bar{\mathrm{r}} - (\lambda + \mu)\mathrm{V_i}, \quad \beta = 0.5 \tag{7}$$

Resource Agents: Allocation and Adaptive Prices

Each day the resource agents collect requests and grant them in priority order until the pump capacity, crew limit or remaining seasonal allocation is reached (a request that does not fit as a 24 mm event is downgraded to 12 mm before being rejected). In the proposed system, priority follows the bid of each agent, its learned marginal value per m³,

$$b_i = \frac{Q(s,a_i)-Q(s,0)}{V_i(a_i)}, \quad \max \sum_i b_i x_i \;\; \text{s.t.} \;\; \sum_i V_i x_i \le Q_t, \;\; \sum_i x_i \le K \tag{8}$$

The water agent keeps two prices. The seasonal price λ rises when cumulative withdrawal U_t runs ahead of a planned path proportional to the cumulative evaporative demand, and the congestion price μ rises when the requested volume D_t exceeds the pump capacity Q_t, in the manner of dual (subgradient) updates in distributed optimisation [9],

$$\lambda_{t+1} = \mathrm{clip}\left(0.8\lambda_t + 30\,\frac{U_t - W_{cap} s_t}{W_{cap}}, 0, 6\right), \quad \mu_{t+1} = \mathrm{clip}\left(0.7\mu_t + 2\,\frac{D_t - Q_t}{Q_t}, 0, 4\right) \tag{9}$$

where s_t is the fraction of the season's evaporative demand expected up to day t. Prices are broadcast to the parcel agents, which observe them in discretised form the next day; they are also part of the reward (Eq. 8) so that the agents learn to reduce requests when water is expensive and to prefer the days when capacity is free. The prices thus adapt on a daily time scale to the situation of the whole farm, without the agents knowing each other's states.

Compared Methods and Ablations

Eight methods were compared (**Table 3**). The uncoordinated rule has each parcel request 24 mm when its depletion reaches 75 % of the readily available water and serves requests in random order until capacities run out. The fixed calendar irrigates 24 mm every five days. The market method requests 12 or 24 mm at depletion of 75 or 100 %, and serves requests in order of the myopic bid $v_i K_{y,i}/\Sigma ET_{c,i}$, the value of each millimetre of water, without pacing. The central planner is an omniscient benchmark with perfect knowledge of the state of all parcels: it projects the stress over the next five days, using the rain forecast, ranks parcels by marginal value per m³ and meters the daily volume so that the seasonal allocation lasts to the end of the season. Independent Q-learning uses only local observations and random service order. The proposed system uses the coordination signals, shaped rewards and bid-based priority described above, and two ablations remove the price signal (and the price term in the reward) or the bid-based priority (random service order). The three learning methods and the two ablations were trained with three independent seeds (random initialisation and training seasons), and the results combine the episodes of all seeds.

Table 3. Compared Methods

Method	Learns	Information used	Allocation of scarce capacity
Uncoordinated rule	No	Own soil water	Random order
Fixed calendar	No	Calendar	Random order
Market (myopic bids)	No	Own soil water, crop value	Myopic bid order
Central planner	No (omniscient)	All states, forecast, plan	Marginal-value ranking with metered budget
Independent Q-learning	Yes	Own observation	Random order
Adaptive MAS (proposed)	Yes	Own observation + prices + budget status	Learned-bid order
Proposed w/o price signal	Yes	As proposed, without prices	Learned-bid order
Proposed w/o bid priority	Yes	As proposed	Random order

Scenarios, Metrics and Adaptation

Each method was tested on 16 seasons per scenario (the same weather for all methods): *Normal*; *Dry* (40 % of normal rainfall, 10 % higher evaporative demand); *Wet*; *Pump failure* (pump capacity halved from day 60); *Price shock* (vegetable value ×0.5, maize ×1.4, which changes the relative importance of the crops compared with training); and *Dry + failure + shock*. Metrics were net value *J*, groundwater withdrawn, grid electricity, mean relative yield, the lowest parcel yield, the day on which the seasonal allocation was used up, and the inequality of relative yields among parcels, measured by the Gini coefficient,

$$G = \frac{\sum_i \sum_j |RY_i - RY_j|}{2n^2 \bar{RY}_j}, \quad J_{rel} = \frac{J}{J_{central}} \times 100\,\% \tag{10}$$

Adaptation was tested by running 50 consecutive seasons of the hardest scenario (dry, pump failure and price shock together) with three variants: the trained proposed system with learning switched off (*frozen*); the proposed system with continued learning and a drift detector (*adaptive*); and independent Q-learning with continued learning. The detector is a one-sided CUSUM on the seasonal net value that raises an alarm when the cumulative shortfall relative to the normal-season mean μ_0 (in units of its standard deviation σ_0) exceeds a threshold, which resets exploration to 0.30 before it decays again,

$$g_k = \max\left(0, g_{k-1} + \frac{\mu_0 - J_k}{\sigma_0} - 0.5\right), \quad \text{alarm if } g_k > 3 \tag{11}$$

IV. RESULTS AND DISCUSSION

Learning

Fig. 4 shows the training progress as the mean over the three seeds. The learners improved from low initial returns to a plateau within about 150 episodes. The proposed system learned to a higher level than independent Q-learning and then the version without the price signal, and these two reached lower returns with much larger variation between training runs: the standard deviation of the normal-season net value over seeds was ₹2.15 lakh for independent Q-learning and ₹2.07 lakh without prices, against ₹0.35 lakh for the proposed system (**Table 4**). The coordination signals therefore both raised and stabilised learning, which is consistent with the finding that independent learners suffer from non-stationarity [10]: without a signal that reflects the state of the shared resource, each agent sees an environment whose behaviour changes as others learn.

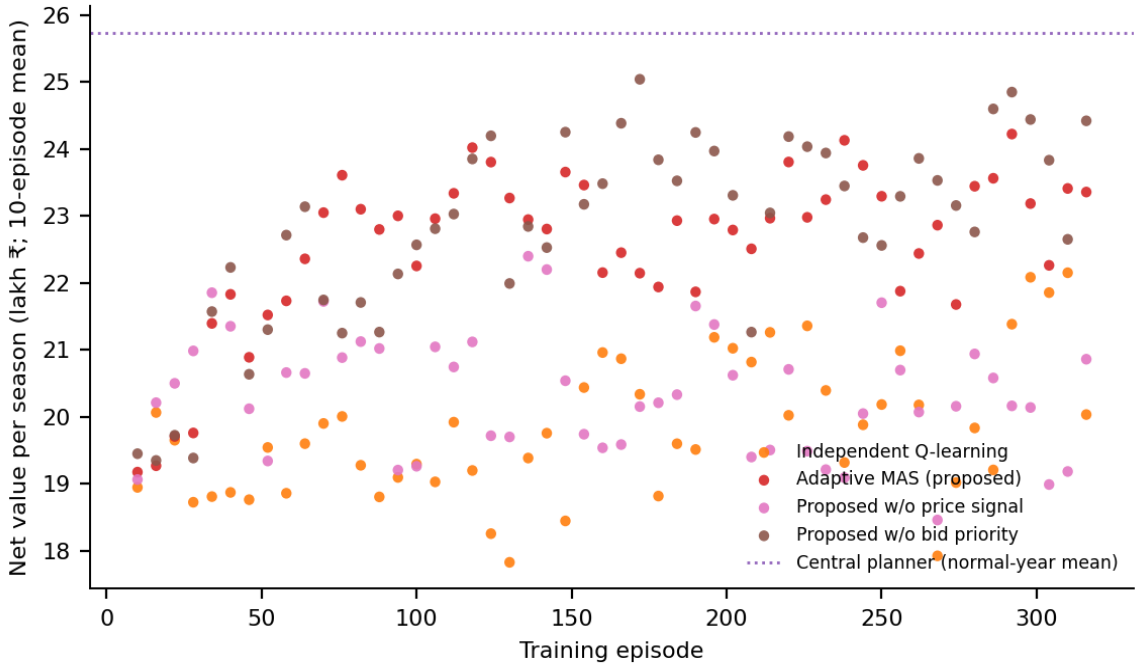


Fig 4. Training Progress (Mean of Three Seeds, 10-Episode Moving Average, Every Sixth Point Shown) of the Learning Methods; The Dotted Line Marks the Mean Net Value of the Central Planner in a Normal Year.

Normal-Year Performance

Table 4 and **Fig. 5** give the normal-year results. The central planner achieved the highest net value (₹25.74 lakh), by concentrating water on the high-value crops and metering it through the season (the allocation lasted until day 118); but this came at the cost of equity, with a lowest parcel yield of 0.66 and a Gini coefficient of 0.085. The market method (99 % of the planner), the uncoordinated rule (97 %) and the proposed system (95 %) came close in net value and gave more equal yields (Gini 0.013, 0.019 and 0.025) and higher lowest-parcel yields (0.88, 0.87 and 0.82). That the simple uncoordinated rule did so well shows that in a normal season, with moderate scarcity, coordination has limited value; the pump capacity and crew limit only occasionally prevent parcels from being irrigated, and the seasonal allocation runs out late (day 90). The proposed system did not beat the market heuristic or the uncoordinated rule in net value; it matched them while learning from experience, without a hand-written rule for each crop. The fixed calendar was clearly worse (69 %) because it ignores soil-water state, and independent Q-learning reached only 80 %. In the electricity use, the planner purchased the least grid power (175 kWh per season), since metering the water evenly keeps the daily pumping within the solar capacity, whereas the uncoordinated rule bought 2223 kWh and the proposed system 1517 kWh: the learned agents

internalise the energy cost but do not coordinate the timing of pumping with solar availability, a missing feature that an energy agent with forecasts could add.

Table 4. Normal-Year Results

Method	J	J (% central)	Grid (MWh)	Mean RY	Lowest RY	Gini	Cap used up (day)
Uncoordinated rule	24.85	97	2.22	0.91	0.87	0.019	90
Fixed calendar	17.82	69	1.09	0.68	0.57	0.082	49
Market (myopic bids)	25.49	99	0.96	0.92	0.88	0.013	95
Central planner	25.74	100	0.18	0.85	0.66	0.085	118
Independent Q-learning	20.66	80	1.74	0.77	0.68	0.045	64
Adaptive MAS (proposed)	24.42	95	1.52	0.88	0.82	0.025	89
Proposed w/o price signal	19.24	75	1.74	0.72	0.63	0.060	56
Proposed w/o bid priority	24.91	97	1.33	0.90	0.82	0.024	92

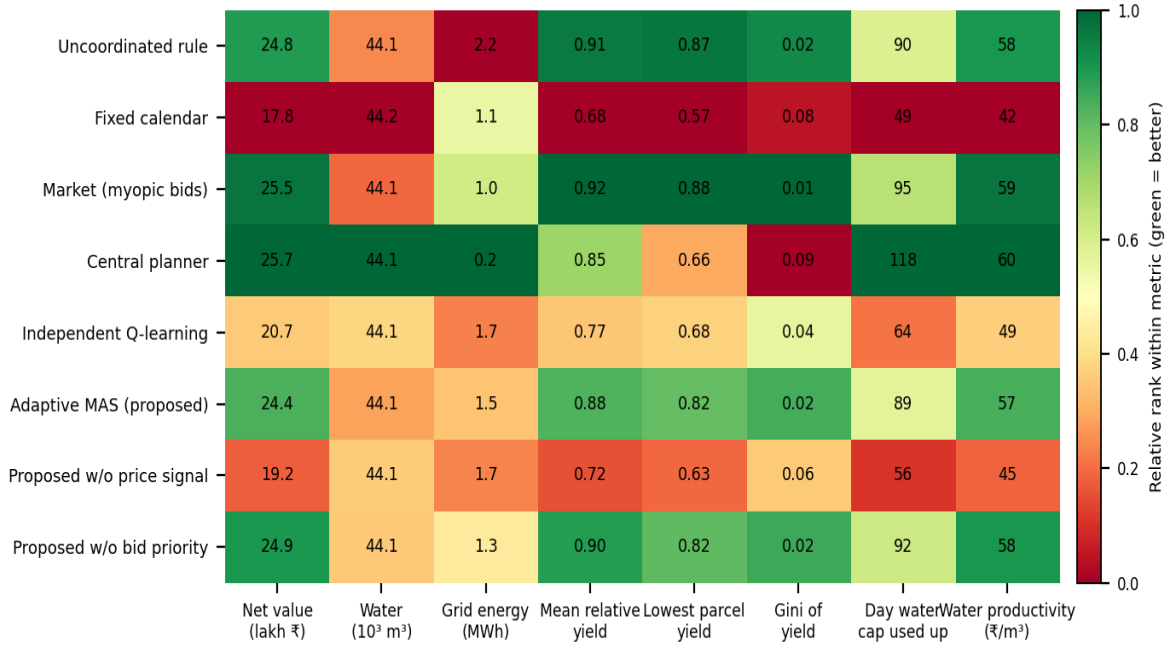


Fig 5. Normal-Year Performance of the Methods on Eight Metrics (Cell Values; Colour: Rank Within Each Metric, Green = Better).

Scenarios and Robustness

Fig. 6 shows net value relative to the central planner in the six scenarios. In the wet season all methods except the calendar and the weakly learning variants reached 99–101 % of the planner, since water was not scarce. In the dry season, differences widened: the planner remained best, the market method reached 85 %, the proposed system 84 %, the version without bid priority 86 %, the uncoordinated rule 80 % and independent learning 69 %. The planner's advantage in scarcity comes from information and foresight, not from coordination as such: it knows the state of every parcel, uses the forecast to avoid irrigating before rain and meters the allocation to the end of the season. The proposed system recovered 15 percentage points over independent learning and 4 points over the uncoordinated rule, winning against the uncoordinated rule in 79 % of dry seasons but in 29 % of normal and 0 % of wet seasons. The benefit of the learned coordination therefore appears mainly under scarcity. After a pump failure the proposed system reached 97 % of the planner, and after a price shock, which it had not seen in training, 97 %; both are close to the market method (101 % and 101 %), and in the combined case it reached 95 % against 97 % for the market method, 90 % for the uncoordinated rule and 76 % for independent learning. The price-shock result is notable because the learned Q-values were obtained with the original crop values; the robustness comes from the observation that the agents act on their own soil state and on prices that are set by the actual congestion, not on a fixed ranking of crops. Independent Q-learning lost ground in every scenario.



Fig 6. Net Value of Each Method Relative to the Central Planner (%) in Six Scenarios (Colour Scale 40–105 %).

Ablation: What Makes the Coordination Work

The ablations are informative. Removing the price signal (and its term in the reward) reduced the normal-year net value from 95 % to 75 % of the planner and in the combined scenario from 95 % to 71 %, and also increased the variation between seeds to ₹2.07 lakh; the variant without prices performed similarly to independent learning (80 %), which shows that the price signal and the budget status that go with it carry the benefit of coordination. In contrast, replacing the bid-based priority by random order did not hurt (97 % in a normal year against 95 %, and 86 % against 84 % in a dry year): the priority ranking by learned marginal value gave no measurable advantage. A plausible reason is that when prices already reduce the requests of agents with low marginal value, the number of competing requests is small enough that the order matters little. This negative result suggests that the simpler version, which uses prices and random order, may be preferable in practice; the ranking would be more important when demand exceeds capacity by a larger factor.

Trade-Offs Between Value, Grid Energy and Equity

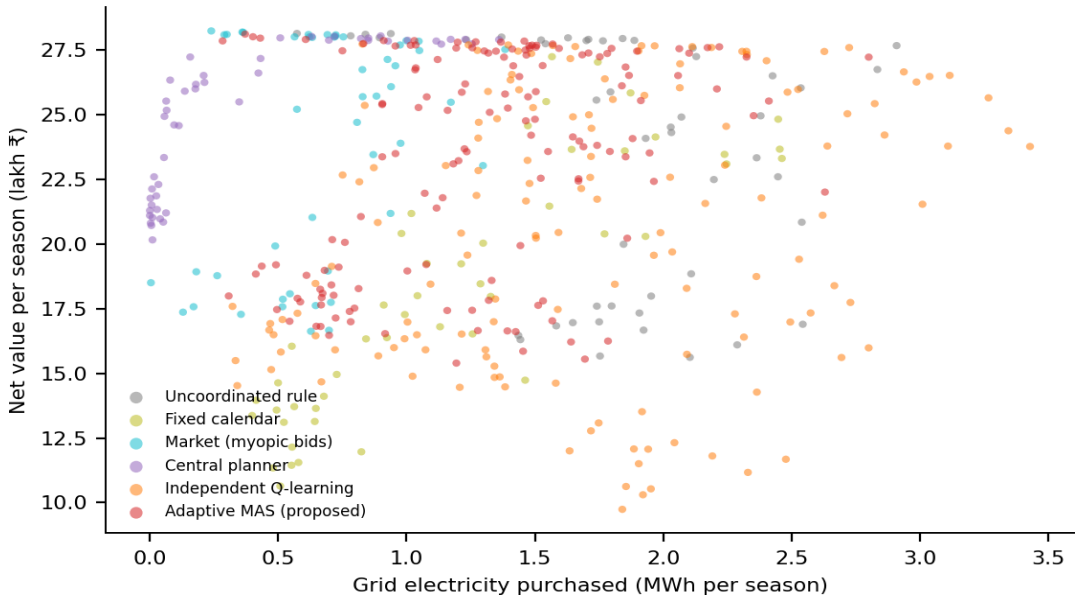


Fig 7. Net Value Against Purchased Grid Electricity for All Test Seasons in Normal, Dry and Wet Scenarios (Each Point is One Season).

Allocation

Figs. 8 and **9** show the irrigation depth of each parcel and day in the same normal season for the proposed system and for the uncoordinated rule (rows sorted by crop). In this season the seasonal allocation was used up on day 85 under the proposed system and on day 78 under the uncoordinated rule, after which no parcel could be irrigated. The maps allow the coordination to be inspected: irrigation events of different parcels are spread over the days under the proposed system so that the pump and crew limits are not hit repeatedly, whereas the uncoordinated rule lets parcels with the same crop and soil reach their threshold on the same days and compete for capacity. The net value of this single season was ₹23.77 lakh for the proposed system and ₹23.64 lakh for the uncoordinated rule, which illustrates how little separates them in a normal year.

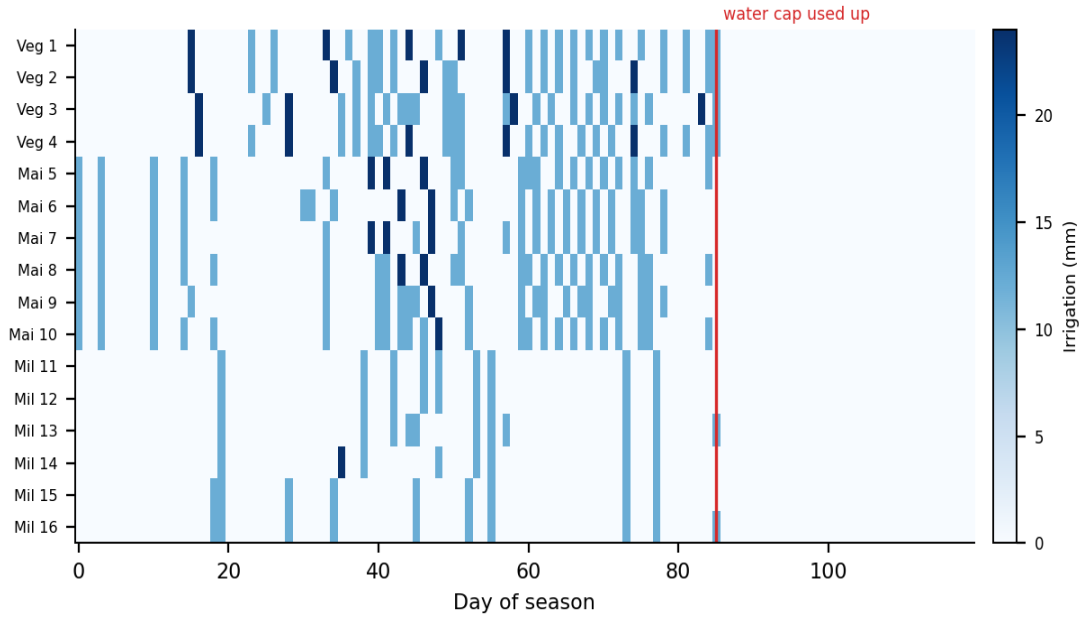


Fig 8. Irrigation Depth (Mm) By Parcel and Day in a Normal Season Under the Proposed System; The Vertical Line Marks the Day the Seasonal Water Allocation Was Used Up.

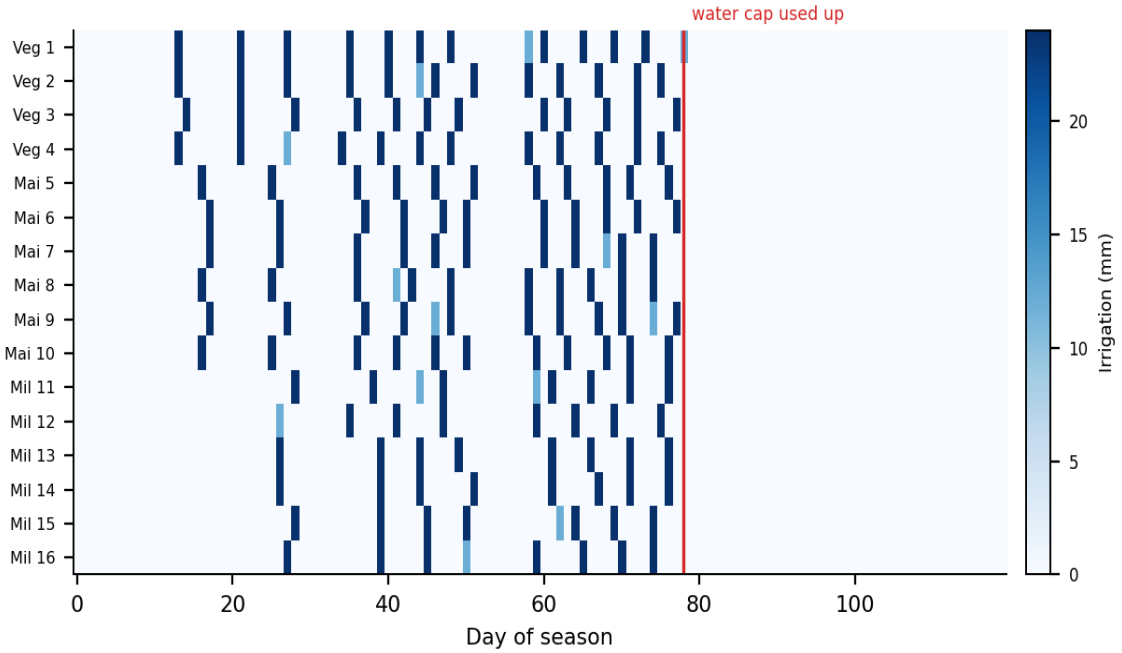


Fig 9. Irrigation Depth (Mm) By Parcel and Day in The Same Season Under the Uncoordinated Rule.

Gains and Losses in a Dry Year

Fig. 10 maps the change in the relative yield of each parcel between the proposed system and the uncoordinated rule, averaged over the dry seasons, and **Fig. 13** compares the parcel yields directly. On average, the proposed system changed the yield by +4.6 percentage points for vegetables, +4.1 for maize and -8.3 for millet. The learned coordination thus shifts scarce water towards the higher-value crops and away from the low-value millet, which is the efficient allocation under scarcity, but it means that individual smallholders with low-value crops bear the cost. In an institution in which parcels belong to different farmers, such shifts require compensation or a rule for fair sharing; the price-based mechanism makes the trade-off explicit and could be combined with transfers.

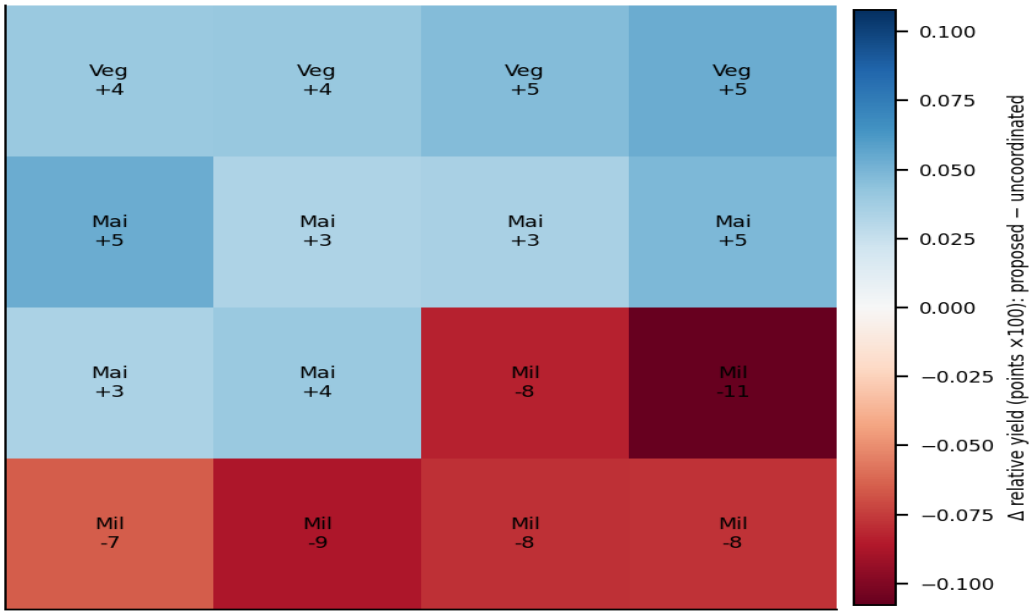


Fig 10. Change in Relative Yield of Each Parcel (Proposed System Minus Uncoordinated Rule, Percentage Points) in Dry Seasons, With Crop Labels.

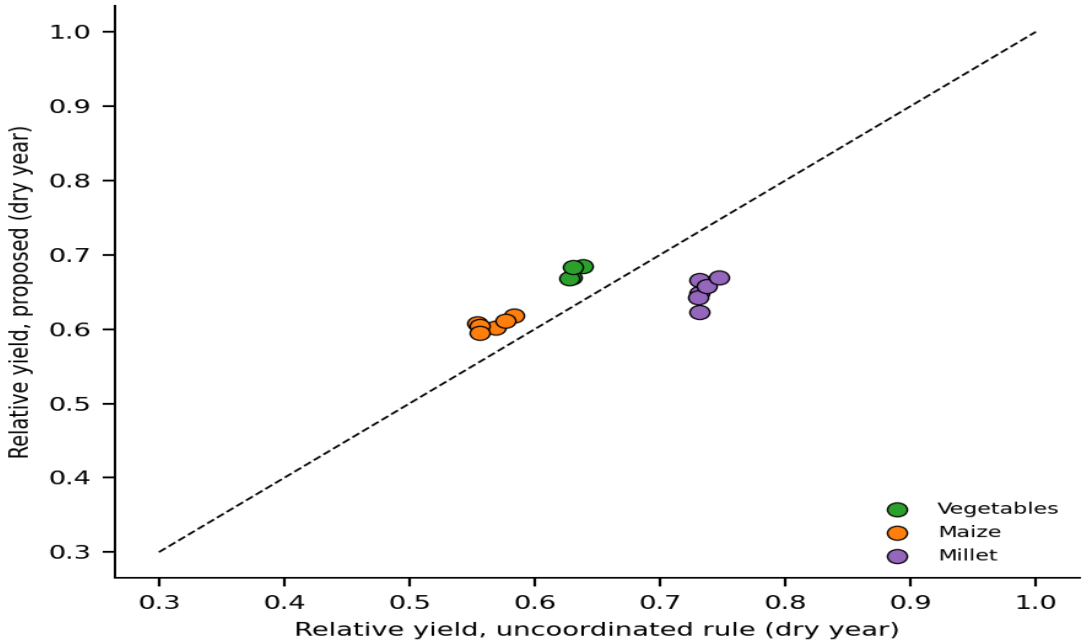


Fig 11. Relative Yield of Each Parcel Under the Proposed System Versus the Uncoordinated Rule in Dry Seasons (Dashed Line: Equal Yield).

Price Signal and Congestion

Fig. 12 plots the resource price against the volume requested in a dry season. The price tends to rise with demand (correlation 0.78) and with the progress of the season, as the colour scale shows: days with requests above the pump capacity (dotted line) carry high prices, and in the late season the price remains elevated because withdrawal runs ahead of the planned path. This is the signal that the parcel agents learn to respond to: they postpone or reduce irrigation when it is high. The price is a compact summary of the state of the shared resources that has the same form for all agents, which is why it scales to more parcels without increasing the information each agent must process.

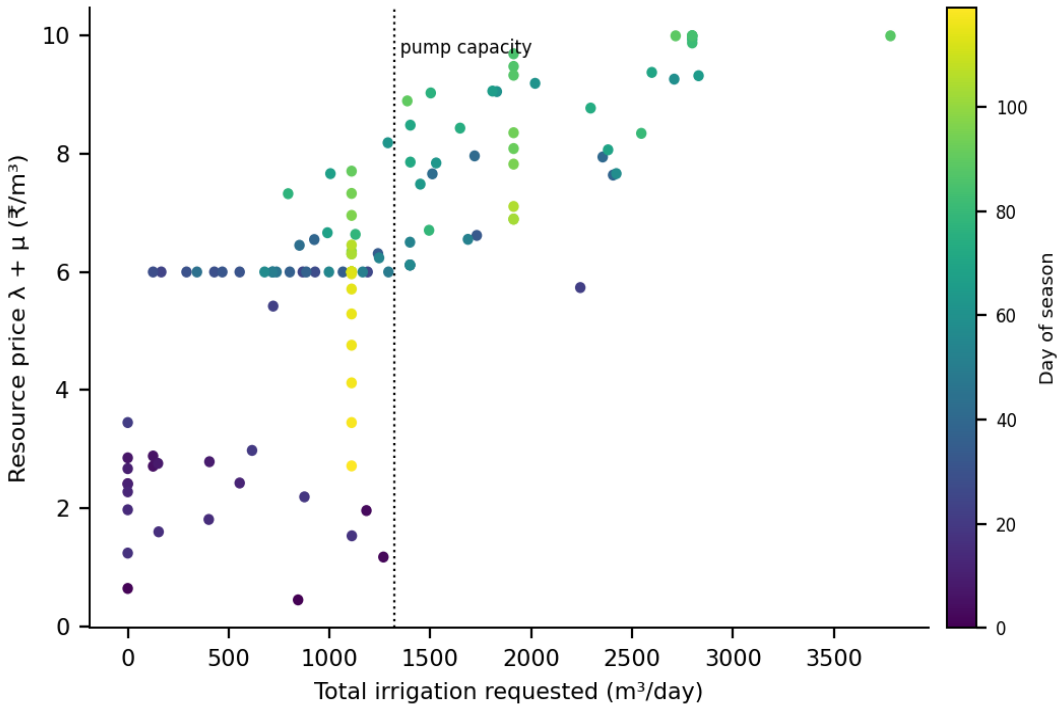


Fig 12. Resource Price Against the Total Volume Requested Per Day in a Dry Season (Colour: Day of Season; Dotted Line: Pump Capacity).

Adaptation After a Severe Shock

Fig. 13 and **Table 5** show what happened in 50 consecutive seasons of the combined shock. The central planner and the market method reached ₹15.04 and ₹14.56 lakh on average in this scenario. The frozen proposed policy, without any retraining, achieved ₹14.75 lakh in the first ten seasons and ₹14.08 lakh in the last ten, i.e. 98 % of the planner immediately after the shock. The adaptive variant, with continued learning, raised an alarm in the first season (in all three seeds) and reset exploration, which cost value initially (₹12.78 lakh in the first ten seasons) and recovered to ₹13.74 lakh in the last ten, an improvement of 7 % over its own start, but still below the frozen policy (₹14.08 lakh). Independent Q-learning with continued learning started lower (₹11.19 lakh) and deteriorated (₹9.97 lakh), because learning in a non-stationary multi-agent environment without a coordinating signal is unstable. We conclude that, in this study, the main source of adaptivity is the daily adjustment of prices, which is available immediately and needs no learning, and not the slow retraining of the agents; continued learning had a cost (exploration) and in 50 seasons did not repay it. This is consistent with the general finding that exploration in a changed environment can be costly, and it suggests safer designs: retraining on logged data, conservative exploration, or limiting adaptation to the parcels that are most affected by the shock.

Table 5. Adaptation to the Combined Shock

Variant	First 10 seasons	Last 10 seasons	Drift alarm (season)
Frozen proposed policy	14.75 ± 0.36	14.08 ± 0.43	—
Adaptive proposed (continued learning)	12.78 ± 0.06	13.74 ± 0.10	0, 0, 0
Independent Q-learning (continued)	11.19 ± 0.10	9.97 ± 0.61	—
Central planner (reference)	15.04	—	—
Market method (reference)	14.56	—	—

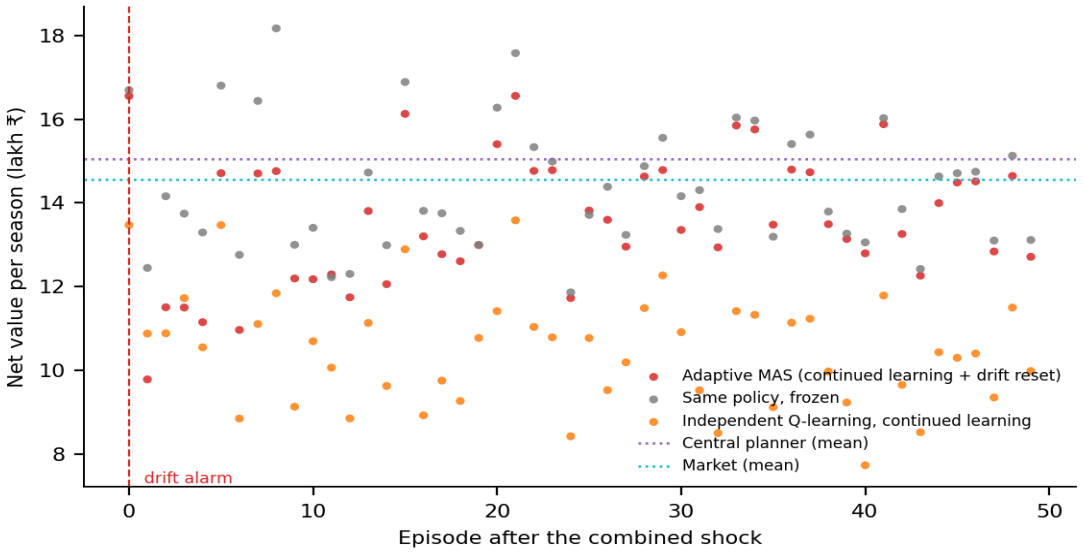


Fig 13. Net Value in Each of 50 Seasons After the Combined Shock for The Adaptive Proposed System, The Frozen Policy and Independent Q-Learning (Means of Three Seeds); Dotted Lines: Planner and Market References.

Implications for Farm Management

Several points emerge for the design of coordinated farm resource management. First, coordination matters most when resources are scarce or fail: the uncoordinated rule was competitive in normal and wet seasons but fell behind in dry seasons and after failures, so the value of a coordination layer should be judged by its performance in stress conditions, not in average ones. Second, a small amount of shared information, the resource price and the budget status, is enough to coordinate learning agents; no agent needs to know the state of the others. Third, the benefits of a central planner come from information and foresight and can be approached only partly by decentralised mechanisms; a farm that can collect sensor data on all parcels and has a single decision maker should consider a planner or a model predictive controller, and the multi-agent system is most attractive where decision rights are distributed among farmers, where communication or data sharing is restricted, or where components fail. Fourth, efficiency and equity pull in different directions, and the choice of reward function is a governance decision. Fifth, adaptation should rely first on fast, safe mechanisms, such as price updates, before slow learning, which is costly and may degrade performance after a shock.

Limitations and Future Work

The study has several limitations. The farm is simulated, with indicative parameters and a daily time step, and a single configuration of 16 parcels and three crops was studied. The agents use tabular Q-learning with discretised observations; deep or actor–critic methods could handle richer observations but need more data and tuning. The energy agent does not yet schedule pumping around solar availability, which explains the higher grid purchases of the learned system compared with the planner, and nutrients, machinery and labour beyond the irrigation crew were not modelled. The weather forecast was a simple noisy indicator. Prices are set by a rule and not by a market with strategic agents, so incentive compatibility and strategic behaviour were not examined. Only three training seeds were used, and the number of test seasons was modest. Future work will include a multi-resource energy agent, deep multi-agent reinforcement learning with communication, safe-adaptation methods, strategic behaviour and fairness-aware rewards, and the application of the framework to a real farm or water-user association with sensor data.

V. CONCLUSION

An adaptive multi-agent system for coordinated resource management was developed in which learning parcel agents and resource agents that publish adaptive prices share water, pump capacity, crew time and energy. On a simulated farm it reached 95 % of the net value of an omniscient central planner in a normal year and 84 % in a dry year, with a more equal distribution of yields (Gini 0.025 against 0.085), and clearly outperformed independent Q-learners (80 % and 69 %), which were also much less stable across training runs. The adaptive price signal was the essential component, the bid-based priority was not, simple heuristics were competitive when capacity was not scarce, and the central planner was clearly better under scarcity because of its information and foresight. After a severe combined shock, the frozen policy performed as well as or better than the retrained one, so that the fast price adaptation provided most of the robustness. The results suggest that lightweight, price-based coordination of learning agents is a practical option for farms and water-user groups with distributed decision rights, but they depend on a simulated environment and should be tested with field data, richer agents and strategic behaviour.

CRediT Author Statement
The author reviewed the results and approved the final version of the manuscript.

Data Availability
No data was used to support this study.

Conflicts of Interests
The authors declare no conflict of interest.

Funding
No funding agency is associated with this research.

Competing Interests
There are no competing interests.

References

[1]. B. Paudel, S. Riaz, S. W. Teng, R. R. Kolluri, and H. Sandhu, "The digital future of farming: A bibliometric analysis of big data in smart farming research," Cleaner and Circular Bioeconomy, vol. 10, p. 100132, Mar. 2025, doi: 10.1016/j.clcb.2024.100132.
[2]. T. Ali et al., "Smart agriculture: utilizing machine learning and deep learning for drought stress identification in crops," Scientific Reports, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-74127-8.
[3]. E.-S. M. El-Kenawy, A. A. Alhussan, N. Khodadadi, S. Mirjalili, and M. M. Eid, "Predicting Potato Crop Yield with Machine Learning and Deep Learning for Sustainable Agriculture," Potato Research, Jul. 2024, doi: 10.1007/s11540-024-09753-w.
[4]. J. Sang, D. Ma, X. Xie, and X. Hu, "Group-aggregation of Hierarchical Containment Control for Homogeneous Multi-agent Systems in Precision Agriculture," International Journal of Control, Automation and Systems, vol. 22, no. 4, pp. 1400–1408, Mar. 2024, doi: 10.1007/s12555-022-1071-y.
[5]. W. Tao, Q. Sang, C. Chen, and Q. Mao, "Intelligent Agents for Smart Agriculture: Architectures, Applications, and Future Challenges," Agriculture, vol. 16, no. 17, p. 1808, Aug. 2026, doi: 10.3390/agriculture16171808.
[6]. D. Bazzana, J. Foltz, and Y. Zhang, "Impact of climate smart agriculture on food security: An agent-based analysis," Food Policy, vol. 111, p. 102304, Aug. 2022, doi: 10.1016/j.foodpol.2022.102304.
[7]. A. Tzachor et al., "Large language models and agricultural extension services," Nature Food, vol. 4, no. 11, pp. 941–948, Nov. 2023, doi: 10.1038/s43016-023-00867-x.
[8]. S. Bushaj, X. Yin, A. Beqiri, D. Andrews, and İ. E. Büyüktahtakın, "A simulation-deep reinforcement learning (SiRL) approach for epidemic control optimization," Annals of Operations Research, vol. 328, no. 1, pp. 245–277, Sep. 2022, doi: 10.1007/s10479-022-04926-7.
[9]. G. G. Devarajan, S. M. Nagarajan, R. T.V., V. T., U. Ghosh, and W. Alnumay, "DDNSAS: Deep reinforcement learning based deep Q-learning network for smart agriculture system," Sustainable Computing: Informatics and Systems, vol. 39, p. 100890, Sep. 2023, doi: 10.1016/j.suscom.2023.100890.
[10]. M. Aldossary, H. A. Alharbi, and C. Anwar Ul Hassan, "Internet of Things (IoT)-Enabled Machine Learning Models for Efficient Monitoring of Smart Agriculture," IEEE Access, vol. 12, pp. 75718–75734, 2024, doi: 10.1109/access.2024.3404651.
[11]. S. K. S. Durai and M. D. Shamili, "Smart farming using Machine Learning and Deep Learning techniques," Decision Analytics Journal, vol. 3, p. 100041, Jun. 2022, doi: 10.1016/j.dajour.2022.100041.
[12]. N. Marwah, V. K. Singh, G. S. Kashyap, and S. Wazir, "An analysis of the robustness of UAV agriculture field coverage using multi-agent reinforcement learning," International Journal of Information Technology, vol. 15, no. 4, pp. 2317–2327, Apr. 2023, doi: 10.1007/s41870-023-01264-0.
[13]. A. Choudhari, D. Bhoyar, and W. P. Badole, "Design of an IoT based Monitoring System Utilizing Deep Reinforcement Learning and Ensemble Learning for Precision Agriculture," 2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS), pp. 845–852, Jul. 2024, doi: 10.1109/icscss60660.2024.10625266.
[14]. A. S. Assiri, F. k. Alsheref, M. A. Ghamdi, and M. Abbassy, "MADRL-SI: a multi-agent deep reinforcement learning-based swarm intelligence approach for coordinating multiple UAVs in smart agriculture systems," International Journal of Information Technology, Nov. 2025, doi: 10.1007/s41870-025-02807-3.
[15]. A. Din, M. Y. Ismail, B. Shah, M. Babar, F. Ali, and S. U. Baig, "A deep reinforcement learning-based multi-agent area coverage control for smart agriculture," Computers and Electrical Engineering, vol. 101, p. 108089, Jul. 2022, doi: 10.1016/j.compeleceng.2022.108089.

Publisher's note: The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. The content is solely the responsibility of the authors and does not necessarily reflect the views of the publisher.

ISSN: 3104-4654