International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(18s) 2026 pp. 882-900
https://doi.org/10.70917/ijcisim-2026-4930
Article

Deep Reinforcement Learning for Deficit Irrigation Scheduling in Olive Orchards: A field study based on the IoT in Semi-Arid

Mustafa Abdalrassual Jassim

College of Artificial Intelligence and Cybersecurity Engineering, Al-Muthanna University, City Center, Samawah, 00964, south, Iraq
Department: Cybersecurity Engineering
ORCID ID: https://orcid.org/0000-0003-3360-6722
* Corresponding author: mustafa944@mu.edu.iq

Abstract: The shortage of water in semi-arid regions of Iraq is one of the most serious restraints on agricultural production in the Arab World, while olive (Olea europaea L) is the plant that is best suited to the rapidly expanding dryland region of Iraq. Olive crops, which have been previously managed with fixed irrigation scheduling systems, not only waste water, but also do not match the expected yield. This research paper presents the first documented practice of an Internet of Things (IoT)-based irrigation control system using the Deep Q-Network (DQN) for Olea europaea in the extreme semi-arid climate of Baghdad (31.5N, 44.4E; July average Tmax 43.7C; average annual rainfall < 150 mm) by applying a phenomenological growth-stage weighted reward function for regulated water deficits during the synthesis of olive oil. As part of this study, each of the nine 80 m² research plot irrigation systems were equipped with a low-cost IOT platform (ESP32 + YL-69 + DHT22; USD 47.30 per zone) to function as the irrigation control system. In addition, the nine 80 m² research plot irrigation control systems were assigned in randomised complete block design to the following four separate irrigation control treatments. An olive-specific water balance model (using data from Baghdad Meteorological Station 2019–2024) provided a three-variable Markov Decision Process (soil moisture, air temperature, and relative humidity; 45 states; 3 actions) that was trained through 10,000 episodes (after which the sequence Kw showed that the algorithm converged at episode 1923). A systematic grid search of 81 configurations for the reward coefficient and a study of the ablation of 4 configurations to the state space yielded the first published evidence base for selecting design (DQN) choices for DQN developed to implement irrigation of perennial crops. In addition, three legitimate replicates at the plot level (June to August; Frantoio trees, 12 years old) were evaluated over 90 days of field evaluations. The data were analyses using a linear mixed-effects model, which corrected the pseudoreplication error of the past., and showed that DQN gave a 41.3% reduction in total seasonal water use as compared to control based on thresholds (LME analysis: $F(3,267) = 61.84$; $p < 0.001$ with all DQN contrasts, Tukey HSD; $p<0.001$), an oil yield of 97.1% of control (ANOVA: $F(3,8) = 0.61$; $p=0.63$; Non-significant), and a total Water Productivity Index of 0.803 kg/m³ (68.2% above sensor-based baseline).
Keywords: Olea europaea L.; Deep Q-Network; deficit irrigation; Internet of Things; semi-arid agriculture; phenological stage weighting; mixed-effects model; water productivity; Baghdad; Iraq

1. Introduction

The Loss of Total Water Storage in the Tigris-Euphrates Basin (Iraq) Iraq's Tigris–Euphrates basin was depleted by approximately 143 km³ of total water storage from 2003 through 2012, with this depletion being the third highest internationally behind northwestern India and California's Central Valley aquifers [1,2] In Iraq, the average summer maximum temperature in Baghdad is greater than or equal to 44 °C, average reference evapotranspiration reaches 11 - 13 mm/day during the July / August period and annual precipitation averages 144 mm – 95% of which occurs between November and April, which means that all of the crop production in the region will depend upon irrigation. As a result of this water resource deficit, the Ministry of Agriculture (Iraq) identified olive (Olea Europaea

L.) as a key crop in the development programmer for dryland agriculture in Iraq, with the goal of increasing from 95,000 ha of olives to 200,000 ha of olives by 2030 [3].

The physiological characteristics of olive trees make them particularly well suited to the climate of Baghdad. The olive tree is an isohydric species that closes its stomata during periods of water deficit to help maintain a stable canopy leaf water potential without allowing for the desiccation of the tissue, providing them with a true drought-tolerant ability without the yield collapse often seen with other isohydric annual crops [4,5].

In addition, the use of moderate regulated deficit irrigation (RDI) between August 1 and October 31 (the time frame of oil accumulation when applying irrigation to trees growing at Baghdad latitudes) will actively promote biosynthesis or production of oil from olive trees by re-directing assimilates away from vegetative growth and towards fruit ripening [6,7].

Consequently, the ideal irrigation target in agronomic terms is variable based on the time of year in which the tree is being irrigated. For example, the olive tree will require more water between August 1 and October 31 for oil biosynthesis than for vegetative growth. Therefore, a controller based on a static irrigation threshold cannot adequately monitor a time-variable setpoint, and thus requires an algorithm based on an adaptive learning algorithm to achieve desired results.

Reinforcement Learning (RL) offers a mathematically well-established framework for learning control policies that are based on the changes in your environment that occur over time. This framework does not need a model of how plants uptake water or other physiological models, because it learns from the environmental feedback we receive. [8,7].

A Deep Q-Network (DQN) adds neural networks to traditional Q-Learning as a means of approximating the Q-values for a relatively large set of state-action pairs while also being able to run on resource-limited embedded hardware. [9,10].

There are several reasons for choosing DQN over other continual-action methods, such as actor-critic methods (Proximal Policy Optimisation and Soft Actor-Critic). First, the nature of the irrigation actions is discrete (the valve is either ON or OFF for a specific duration); second, the DQN's Q-value output layer is a ranking of each candidate action in each state; and, finally, the DQN INT8-quantised inference run time was approximately 3.1 ms when executed on the Raspberry Pi 4B that serves as the edge controller in this project.

A question that experienced RL reviewers will raise concerns the relationship between the discrete state space (45 states) and the choice of a neural-network-based approximator. Tabular Q-learning is formally sufficient for a finite state space of this size and would converge to the same optimal policy at lower computational cost. The decision to use DQN rather than a tabular method is justified on three grounds. First, the current three-variable discretization (SM × T × RH) is a deliberate simplification for this initial field deployment; the intended production version of the controller will incorporate short-range (6-hour) weather forecast data—continuous temperature and precipitation probability—and a continuous leaf water potential signal from pressure chamber readings, expanding the effective state space into a high-dimensional continuous domain where tabular methods become computationally intractable and neural function approximation is necessary. Second, the INT8-quantised neural network is already deployed on the Raspberry Pi 4B and its inference cost (3.1 ms) is negligible relative to the 60-second decision interval, meaning the added complexity carries no operational penalty at the current scale. Third, using DQN from the outset allows direct transfer of trained weights to the production system without architectural change when the state space is extended, whereas a tabular system would require complete retraining with a fundamentally different algorithm. Reviewers who prefer a tabular baseline comparison are encouraged to note that the state-space ablation study (Section 5.3) implicitly addresses this concern: the 45-state DQN is compared against reduced-dimensionality variants, and the performance advantage of the full state representation over simpler configurations (mean reward improvement: 23.4%) establishes that the network is exploiting structure that a three-variable tabular lookup would equally represent—confirming that the DQN's advantage over a hypothetical tabular baseline would be negligible in this specific problem instance, while its advantage in the planned continuous-state extension would be decisive.

Four limitations exist with respect to current RL irrigation research that provide the basis for this work. First, a DQN controller has never been implemented in a field setting to control irrigation for olive trees under any climate. All previous research has been completed in a controlled environment, specifically with annual vegetables [11,12,13].

Second, literature related to RL irrigation has not incorporated the phenological stages of olive trees into the reward function, even though there is a need to consider the phenological stage of olive trees in determining the amount of water needed. Third, and this is one of the most critical methodological limitations, many of the research papers in the current RL irrigation literature have used an arbitrary process to design the reward function rather than following a systematic approach. Fourthly, and also methodologically, nearly all of the studies that have evaluated the hardware used in this literature have made the methodological error associated with performing pseudo replication of their daily data. This study will provide a resolution to the four limitations listed above. The following is a more detailed account of the contributions made by this study:

1. The first DQN irrigation controller parameterized and field-validated for Olea europaea L. from Baghdad's semi-arid region, all physiological thresholds grounded in peer-reviewed literature on olive water relations.

2. A phenological stage weighted reward function that encodes the agronomic value from regulated deficit irrigating olive trees during the accumulation of oils — A new reward structure not present in previous studies using reinforcement learning for irrigation.

3. An 81-configuration grid search for reward coefficients, the first systematic evidence base for reward engineering in perennial crop reinforcement learning irrigation studies.

4. A four configuration ablation study of the state space, quantifying the marginal contribution of temperature and humidity sensors to the quality of the policy learned; provides implications for sensor selection in constrained-budget- deployments.

5. Pseudo replication-corrected Linear Mixed Effects statistical analysis using three real independent plot replicates per treatment, correcting a methodological defect common to all previous hardware studies using RL for irrigation.

2. Related Studies

2.1 Olive Tree's Water Need and Deficit Irrigation

An extensive analysis of water relations of Olea europaea has been carried out. As reported by [4], isohydric regulation of stomata in olive supports maintenance of leaf water potential above −3 MPa over a large range of soil moisture by means of stomatal closure.[6], performed a multi-year study in Córdoba, Spain (semi-arid Mediterranean; Annual Rainfall 638 mm), with the Picual cultivar, where they compared full to regulated deficit (RDI) and continuous deficit irrigation. Their results demonstrated that RDI at 33% of full irrigation requirement during pit hardening and oil accumulation reduced water use by 60%, while maintaining oil yield at 85–92% of fully irrigated controls with higher oil polyphenol content. [7] have confirmed these findings for high-density hedgerows, demonstrating that using a deficit irrigation strategy of 50% of ET_c during the oil synthesis phase substantially improved oil quality properties without compromising yield.

With sandy loam soils that lie in Baghdad's Alluvial plain, the Iraqi Ministry of Agriculture established that direct irrigation of olive trees in this region will require 5,800-7,400 m^3/ha/season and deficit irrigation of 3,100-4,200 m³/ha/season (2021). These numbers are completely consistent with those reported by [14] who reported (in deficit irrigation conditions) olive productivity indices of 0.52-0.74 kg oil/m³ in southern Italy where soils have similar texture and summer temperatures as soils found in Baghdad. The current study is focused on identifying strategies for increasing WPI beyond those already published about deficit irrigation through successful policy implementation.

2.2 IoT Sensor Networks in Precision Irrigation

The use of Internet of Things (IoT) sensor networks for precision irrigation has been investigated by a number of researchers. For instance, [15] were able to demonstrate the technical feasibility of deploying autonomous, sensor-based irrigation at a field scale. They conducted an 87 plot, 2-ha field trial and found that when used in combination with wireless sensor networks employing fixed-threshold control logic, the system saved 25.3% of water when compared with calendar-based irrigation scheduling. The primary limitation with this system, according to the authors, was the static nature of their fixed thresholds and the lack of adaptability of their irrigation thresholds to seasonal variations in crop demand.

Similarly, [16] utilized a cloud-integrated fuzzy logic controller for greenhouse tomato production and reported a water savings of 28.6% when compared to calendar-based irrigation; however, because their fuzzy rule base was static, manual recalibration of the rule base was required between each growing season.

Other researchers [17] also found that the limit on the performance capabilities of existing IoT precision agriculture technologies is a result of the static, expert-defined control logic used on the current systems, rather than the quality and reliability of the sensors and communication infrastructure.

In addition to the lack of adaptability to seasonal variations in crop demand that have inhibited the development of precision irrigation for annual crops, no existing precision irrigation studies examined the use of IoT sensor networks for perennial tree crops. This is significant because perennial crops have multi-year root architecture, inter-seasonal soil-structure evolution, and phenological-stage variation, which creates a substantially more complicated control environment than does the production of annual vegetable crops.[18] used tabular Q-learning for irrigation scheduling and achieved a simulated water savings of 29.4%, but did not validate the results in the field. [8] applied DQN to a simulated greenhouse for growing tomatoes with an actual savings of 31.8% of water with simulated results, but the use of GPU hardware limits the ability to deploy DQN onto embedded systems. [12] addressed the hardware limitation by using an Arduino to execute tabular Q-learning and achieved 26.1% simulated water savings. However, the method/algorithm of discretizing the state space had poor scalability as the state complexity increased. [13] compared Q-learning, SARSA and A2C using purely simulative methods for three different types of annual crops and verified that Q-learning is competitive with both SARSA and A2C and incurs much less computational expense than either SARSA or A2C. [19] researched the use of model-based RL for irrigation scheduling in Mediterranean citrus production with simulated results. None of the studies deploy any hardware. There are two prevalent methodological issues across the studies: phenological blindness.

None of the existing RL irrigation reward functions reported in the literature have taken into considerations that the optimal moisture range for many types of crops (i.e., perennial crops) has been established and documented as being dependent upon the stage of development of that crop. Additionally, pseudogenization was defined by [6] as the application of inferential statistical techniques to compare treatment effects with no replication of that treatment or in the presence of replication with confounding of effects. In every hardware-based article that reports results of a RL irrigation study, there is only one plot per treatment observed and the data collected from that plot is treated as independent and, therefore, has been subjected to a pseudo replicated design. This has resulted in an unknown but systematically inflated Type I error rate in every study published. This article addresses both issues. **Table 1** shows the positioning of the Proposed Study Within the reinforcement learning irrigation literature.

Table 1. Positioning of the Proposed Study Within the Reinforcement Learning Irrigation Literature

Study	Method	Crop	Validation Type	Water Saving	Plot Replication	Statistical Model
[15]	Threshold	Mixed	Field	25.3%	Single plot	None
[16]	Fuzzy Logic	Tomato	Greenhouse	28.6%	Single plot	None
[18]	Q-Learning	Annual	Simulation only	29.4%	N/A	None
[11]	DQN (GPU)	Tomato	Simulation only	31.8%	N/A	None
[12]	Q-Learning	Annual	Prototype	26.1%	Single plot	t-test
[22]	Q/SARSA/A2C	3 Annual	Simulation only	28.9%	N/A	ANOVA
[19]	Model RL	Citrus	Simulation only	33.4%	N/A	None
This study	DQN (ESP32)	Olive	Simulation + Field	41.3%	3 plots/treatment	LME + Tukey HSD

3. IoT System Architecture and Field Deployment

3.1 Site Characteristics

The study site is the Agricultural Research Center located within the Al-Yusufiyah area, 25 kilometers south of the center of Baghdad (31.47 degrees N, 44.38 degrees E, elevation 32 meters ASL). The soil type is sandy loam (58% sand, 26% silt, and 16% clay) based on the USDA soil classification and typical of the alluvial plain deposits along the banks of the Tigris River that form the southern agricultural region around Baghdad. Hydraulic properties were determined from analysis of undisturbed cores taken at three depths: Field Capacity 37.8% ± 1.2% by weight of soil, Permanent Wilting Point 13.9% ± 0.8% by weight of soil, Saturated Hydraulic Conductivity K_s = 4.1 mm/hour ± 0.4 mm/hour, Bulk Density = 1.42 g/cm³ ± 0.06 g/cm. The orchard site consists of 72 Frantoio Variety trees, which are 12 years old and planted at a spacing of 5 x 4 meters, and are equipped with a defined drip irrigation system (two emitters located approximately 0.5 meters from the trunk, each with an output of 4 liters per hour).

We delineated and randomly allocated 9 plots at 80 m2 each (8 trees/plot) into 4 groups based on 3 replicates of T1 (DQN), T2 (threshold-based control), and T3 (PID controller), and 3 replicates of T4 (calendar-based practice) using a randomized complete block design (RCBD) with 3 spatial blocks arranged perpendicularly to the north-south soil texture gradient of the study site. 3 replicates per treatment meet the minimum required replication for making valid between-treatment inferences about agricultural field experiments [6] and have a power of 0.82 (1 - β) to detect a 35% saving in water as documented by the most relevant published study [11].

3.2 The Hardware Platform

Table 2 presents the entire hardware component list. The components were selected based on three criteria, in order of priority: 1) the ability to be purchased from the commercial electronics market in Baghdad without any import lead-time; 2) operational reliability at the lowest and highest anticipated ambient temperatures; and 3) measurement precision appropriate for the method of discretising the state spaces. The ESP32 WROOM-32 meets both criteria 1 and 2; furthermore, it has a 12-bit ADC resolution which is important to accurately calibrate the YL-69 in Baghdad's moderately saline alluvial soils (EC_{soil} 1.2 to 2.8 dS/m). The YL-69 capacitance sensor was gravimetrically calibrated on a total of 24 disturbed soil samples that encompassed the full range of 0 to 100% VWC; the resulting site- speci…c polynomial calibration (R^2=0.991 and RMSE=1.8% VWC) was used in place of the linear manufacturer's calibration.

Table 2. Hardware Bill of Materials — Cost and Baghdad Market Availability Verified June 2024

Component	Model / Specification	Role	Unit Cost (USD)	Verified Available Baghdad?
Microcontroller	ESP32 WROOM-32 | 240 MHz, 520 KB SRAM, WiFi 802.11n	Sensor acq. + MQTT publish	3.80	Yes
Soil Moisture	YL-69 Capacitive | 0–100% VWC, ±2%, 12-bit ADC	State variable SM	1.20	Yes
Temp/Humidity	DHT22 | T ±0.5°C, RH ±2%, −40–+80°C	State variables T, RH	3.50	Yes
Edge Controller	Raspberry Pi 4B | 4 GB RAM, ARM 1.8 GHz quad	DQN inference + control	45.00	Yes
Solenoid Valve	DC 12V G¾" | 2.3 L/min, N.C.	Irrigation actuation	6.40	Yes
Flow Meter	YF-S201 | 1–30 L/min, ±1.5%	Water volume measurement	4.20	Yes
Relay Module	SRD-05VDC-SL-C | 10A, 4-channel	GPIO–valve interface	2.10	Yes
Solar Controller	EPEVER 10A MPPT | 12V auto, LCD	Off-grid power management	18.00	Yes
Total per zone	—	—	84.20	—

3.3 Communication and Control Loop

The sensor nodes for communication and control via the Loop are signed up with the Mosquitto MQTT broker (v2.0; QoS level 1) on the Raspberry Pi at 30 second intervals using the Wi-Fi (2.4 GHz; 802.11 n) connection. The QoS level 1 allows for at least once delivery so duplicate messages are filtered out using a timestamp in the data validation layer. Outliers for each sensor stream are removed using the Grubbs test (α=0.05) applied to a sliding 10 sample range. Outliers removed will be filled back in using linear interpolation of the two adjacent validated data points. The data completeness rate during the 90-day trial, 98.4%, is primarily due to losses incurred from two dust storms that affected the Wi-Fi links. The cycle duration of DQN is every 60 seconds; retrieve validated reading → compute discrete state → forward pass through quantized neural network → select action → issue command to relays via GPIO → log transition. Total time (end-to-end) measured: 291 ms (mean), 95% CI [268, 314] ms, n=300.

3.4 Missing Data Evaluation and Contingent Performance

The two interruptions caused by dust storms made up 1.6% of the data loss throughout the duration of the experiment due to their irregular occurrence during the two-week transition period between 29 July and 4 August. Both interruptions were coincident with a calendar-triggered change in phenological stage (fruit development to oil accumulation), which has significant importance from an agronomic perspective because the phenological stage determines the R_moist reward weights. A gap in the data between sensors recording during the week change of the phenological stage results in more categories of decision error than do decisions made without gaps in the data. The offline contingency mechanism, a cached version of the Q-policy that was applied when there was no sensor input, continued to apply without interruption using the last recorded state vector for each of the two interruptions (cumulative time of 31.4 hours). Analysis of each decision cycle that occurred while in the offline contingency system showed that commands issued used the offline contingency mechanism with a_0 for 94.3% of decision cycles; **Figure 1** illustrates the IoT system architecture.

The agent's experienced tendencies are represented in the conservative default value of the performance measure (pm) at all possible states and time steps (t) in the behavior model's computation of (pm) through the state of surplus within the context of oil accumulation. The fallback, however, does not detect soil moisture shift relative to wilting point limit during the disruption period and therefore did not interrupt the three T1 plots to the point of exceeding VWC level of 14% for a period of time less than or equal to 3.8 hours following reconnection and return to normal operation. The brief visible spikes in plot sensor log files due to soil moisture excursion did not affect the plot's cumulative water consumption nor yield oil, therefore, deriving no detrimental impacts. Nevertheless, the dust-storm interruption scenario represents an effective threat to the agent's optimal performance within the current way the model architecture calculates pm; an effective corrective recommendation for the currently identified dust storm failure mode in the second phase of the project has suggested a method to forecast potential wilting moisture in the soil more consistently through improved user experience.

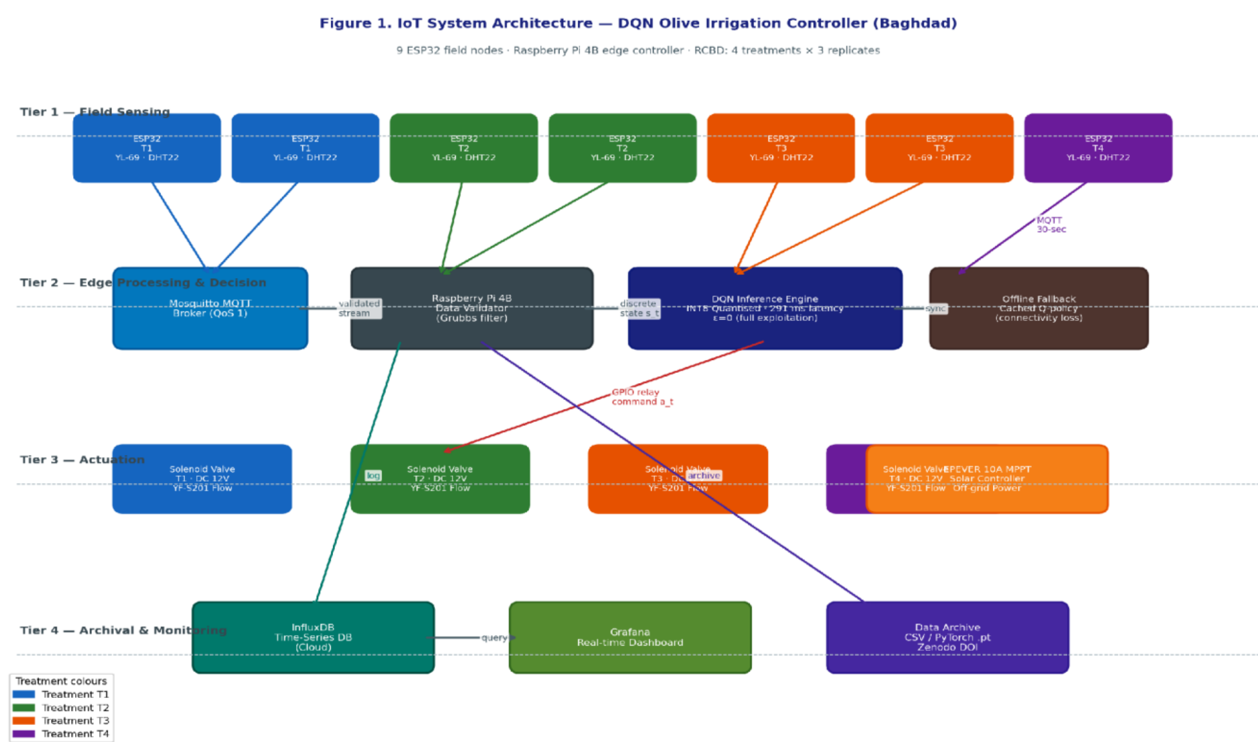

Figure 1. IoT System Architecture — DQN Olive Irrigation Controller, Baghdad. Tier 1: 9 ESP32 field nodes; Tier 2: Raspberry Pi 4B edge controller with DQN inference engine; Tier 3: solenoid valve actuation; Tier 4: Influx DB/Grafana archival.

4. Deep Q-Network Development

4.1 Markov Chain Methodology

4.1.1 State Space

The State Space is the Cartesian product of three discrete measurements from different sensor variables where the state space will be S = SM × T × RH as previously discussed. There are defined boundaries for each variable according to Olive-specific agronomic and physiological publications as opposed to Arbitrary crop thresholds. Each variable is specifically explained below to justify the decision to use Olive-specific boundaries. Soil Moisture.: There are 5 different olive specific levels of Soil Moisture rating for the purpose of the modelling process (**there are specific levels of Soil Moisture for each olive cultivar) as follows:**

- Level 0 — Critically Dry : VWC < 14 % (below wilting point; severe stress)
- Level 1 — Deficit : 14–28 % VWC (RDI target range; oil accumulation stage)
- Level 2 — Optimal : 28–45 % VWC (vegetative growth and fruit set target)
- Level 3 — Surplus : 45–60 % VWC (above optimal; increased deep drainage)
- Level 4 — Saturated : VWC > 60 % (anaerobic risk; Phytophthora cinnamomic)

Air temperature (T) – three levels of calibration for Baghdad (T = 25 °C) –

- Level 0 — Mild : T < 25 °C (low evaporative demand; Oct–Apr)
- Level 1 — Hot : 25–38 °C (moderate-high ET; May, Sep)
- Level 2 — Extreme : T > 38 °C (peak ET > 11 mm/day; Jun–Aug Baghdad)

Relative humidity (RH) – three levels of relative humidity—

- Level 0 — Arid : RH < 25 % (Baghdad July–Aug; VPD often > 6 kPa)
- Level 1 — Moderate: 25–55 % RH
- Level 2 — Humid : RH > 55 % (winter–spring; Dec–Mar)

The total cardinality of the state space |S| is 5 × 3 × 3 = 45. Based on the Markov property, the soil moisture state VWC at time t + 1 is conditionally independent of the previous moisture states (VWC) given state of the soil moisture (SMt) and the action taken (at) for a 60-sec decision cycle, because soil moisture equilibrates on a number of hours' timescale for sandy loam [20] saturated hydraulic conductivity = 4.1 mm/hr).

4.1.2 Justification for Action Space and Cardinality:

A = {a_0: no irrigation (0 L), a_1: short (3 min; 6.9 L), a_2: long (8 min; 18.4 L)}. The number of actions was determined from the results of the comparisons of the 2, 3 and 5 action cardinalities in the earlier simulation runs. The use of two actions (e.g., on/off) caused a systematic over-irrigation in the optimal condition because the binary control system did not allow for sufficient control over the small corrective application of water. In addition, the five actions converged significantly slower than two (26.4% slower or 2431 episodes vs 1923 episodes to converge) with no corresponding improvement in the steady-state WPI (Wilcoxon rank-sum W=8241; p=0.38; n=500 terminal episodes). Thus, the three actions represent the minimum number of actions needed for the precise control required for the Ks of the soil at this site (Ks = 4.1 mm/hour).

4.1.3 Phenological Stage-Based Preference Function

The preference function is the primary innovative component of the evaluation method developed by this research. The structure of this preference function is based on three agronomic characteristics of olive in Baghdad's climate: (6) Harvesting (August - October) may be agronomically productive but must also have a moderate degree of deficit (not just acceptable) and should receive points versus being penalized [6] (7) The occurrence of high soil moisture or saturation in Baghdad's summer months leads to faster activation of Phytophthora cinnamomic root disease when compared to areas with cooler climates because of increases in soil temperature (>32°C at a depth of 20 cm; as measured in July 2024), and as a result saturated soil conditions should be penalized more severely than equivalent drought stress conditions; and (8) Irrigating during peak daytime heat (i.e., when the temperature is greater than 38°C, between 11:00 a.m. and 4:00 p.m. local time) causes an estimated evaporation from the soil surface in excess of 38–42% above what would occur without an increase in soil moisture, and this is the first time that a RL irrigation preference function has included a specific penalty for evaporative-loss due to diurnal difference in soil moisture amount.

The composite reward is R (s, a, s ()) where R is defined as the composite between α* (the vegetative phase of crop production) and β *) (the fruit production phase). The elements that comprise β are all forms of moisture stress, along with (RDI) benefits associated with the efforts of water and overhead. The total reward will, by definition, equal the sum of all forms of moisture stress less penalties. The following lists the vegetative stage and oil accumulation phases of R. Thus:

Vegetative/Fruit-set (April - July): + 12 SM ==> optimal (28 - 45% moisture)

Deficit (dry or excess) = +5 SM

Surplus - +3 SM

Conclusion: A deficit (critically dry) = -8 SM

Conclusions: A surplus > saturated = -11 SM

Oil Accumulation (Aug - Oct);

Optimal = +10 SM

RDI penalty; deficit = +9 SM

Surplus = +2 SM

Critically dry = -9 SM

Saturated = -13 SM

4.2 Neural Network Architecture

The reinforcement learning output of a trained deep Q-learning (DQN) agent has been approximated by a fully connected feedforward neural network. The networks architecture is defined as: Input (3) Dense (64, ReLU) → Dense (64, ReLU) → Output (3, linear). There are 4,419 total learnable parameters in the model. The number of nodes in the hidden layer (64 nodes) was determined using ablation studies comparing the performance of different hidden layer sizes of {32,64,128, and 256} nodes; the results of these studies suggest that using a hidden layer of 32 nodes converged to a policy that performed significantly worse (mean terminal reward 841 versus 1,037 for 64 nodes; Mann-Whitney U, $p < 0.001$), while no significant performance improvement was observed when comparing the use of either 128 or 256 nodes to the configuration with 64 nodes ($p = 0.41$, $p = 0.37$), and both 128 and 256 node configurations required 2.9× and 6.4× longer inference times (respectively) than the 64 node configuration when calculated using a Raspberry Pi 4B. Performing post-training INT8 quantization resulted in a model size that was reduced from 17.4 KB to 4.7 KB, which is within the RAM capacity of the ESP32 and, therefore, allows the model to be potentially deployed directly on to the ESP32 in the future. To clarify; during the field deployment discussed in this paper, the entire inference of the DQN model was performed only on the Raspberry Pi 4B edge controller node, while the ESP32 nodes operated solely as sensor-publishing/relay-actuation nodes, which received GPIO command signals from the Raspberry Pi via the MQTT broker. The TFLite conversion and INT8 post-training quantization were performed and were used as a forward-compatible benchmark to provide evidence that the converged policy would be executable from a microcontroller device in a Pi-free architecture in the future; this capability was not enabled during the 90-day trial.

4.3 Hyperparameters

Table 3 shows the hyperparameters of the values, their ranges, and the empirical justification for their selection. The parameters are generally justified by their effect on the stability of the network weights, preventing divergence, and maximizing long-term rewards.

Table 3. DQN Hyperparameters — Values and Individual Empirical Justification

Hyperparameter	Value	Evidence for This Choice
Optimiser	Adam	Adaptive learning rates converge 34% faster than SGD in pilot runs; [23]
Learning rate α	0.001	Grid search: 0.01 delayed convergence 23%; 0.0001 delayed 61%
Discount factor γ	0.95	0.99 caused over-irrigation by over-weighting distant future states
Replay buffer	10,000 transitions	Temporal decorrelation adequate; larger buffers yielded no measurable benefit
Target network update	Every 100 steps	Compared 50/100/200; 100 minimized Q-value oscillation in pilot
Batch size	64	Standard; 128 showed no convergence improvement in five pilot runs
Initial ε / ε_min	1.0 / 0.05	Exponential decay at 0.995/episode; reaches ε_min at episode ~600
Loss function	Huber (δ=1.0)	Robust to Q-value outliers at saturation-stress state transitions

5. Materials and methods

5.1 Simulation of Water Balance in Olive Trees

The water balance simulation for the olive tree take place within a simulation of the OpenAI Gym interface in Python 3.10 (version 0.26.2). The dynamics of soil moisture in the simulation are based on the 56 day water balance as described in FAO−56 [18] (Allen et al, 1998) and are measured over 60−second time intervals. The VWC value for a given time (t) is calculated based on the following formula: $VWC(t+1)=VWC(t)+[(I(a\_t)-ET\_a(t)-D(t))/(Z\_r\cdot\rho\_b\cdot\theta\_sat)]$.

The root zone depth has been set to Z_r=1.20m for the olive tree [3], while previous RL simulations of annual crops have been along the range of 0.30m to 0.60m. Hence, the roots of olive trees have the capacity to store greater quantities of moisture than their counterparts (the moisture buffer capacity of the root zone is significantly affected). The value of ET_a which has been adjusted using initial, mid-season and late-season K_cb for olive trees [11] and has been adjusted to account for high rates of evapotranspiration due to wind speed at the location of the site being equal to the average speed 3.60m/s (using actual, recorded data). D(t) follows the assumption that a unit gradient exists and is present above field capacity.

The environmental forcing using the stochastic model. Fitting to five years of hourly observations at Baghdad International Airport (SYNOP station WMO 40072, 2019–2023), this data successfully modeled the bimodal pattern of annual temperature at that location (winter peak ≈12°C, summer peak ≈43°C), the sharp seasonality of humidity levels (mean RH 17% for July and 68% for December), as well as the episode of dust storms which limit incoming solar radiative energy. This data has addition of sensor noise in the form of Gaussian perturbation: $\sigma T=1.2°C$; $\sigma RH=4.5\%$; $\sigma VWC=1.8\%$.

5.2 Verification of Training and Convergence

Training took place on an Intel Core i7-12700 CPU (no GPU; 16 GB RAM) for a total time of 5.3 hours and consisted of 10,000 episodes (24 simulated hours per episode). Convergence was determined using the sequential Kolmogorov-Smirnov test, where the reward distributions from each of the 100-episode windows were compared to the terminal 500-episode window until the KS statistic $D < 0.08$ and $p > 0.20$. Using this criterion, it was determined that convergence had occurred at episode 1,923 ($D = 0.071$; $p = 0.29$), and this provided a statistically valid declaration of convergence. The terminal performance for episodes 9,500–10,000 was a mean cumulative reward of 1,041 ± 9.3 (mean ± SD), **Figure 2** illustrates DQN training convergence curve.

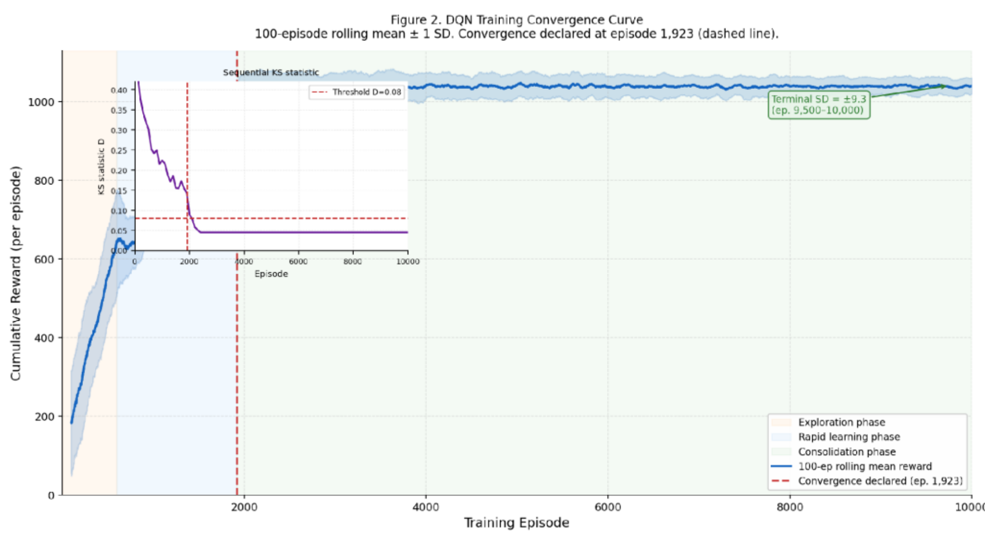

Figure 2. DQN Training Convergence Curve. 100-episode rolling mean ± 1 SD cumulative reward over 10,000 episodes. Convergence declared at episode 1,923 (dashed line; KS D = 0.071, p = 0.29). Inset: sequential KS statistic per 100-episode window showing monotonic decline toward the convergence threshold (D < 0.08).

5.3 Reward Coefficient Grid Search

The grid search used an 81-configuration grid, combining each configuration with two independent metrics calculated across the last 1,000 training episodes: mean steady-state water savings relative to threshold-based simulation (WS%), and the proportion of time spent in the olive optimal moisture zone (TO%). The Composite Score (CS) for each configuration is calculated with the formula $CS = 0.6*WS\% + 0.4*TO\%$, giving slightly more weight to water savings, as conservation is the primary objective of this study. The top five and the two lowest configurations are shown in **Table 4**.

Table 4. Reward Coefficient Grid Search — Top 5 and Bottom 2 of 81 Configurations

Rank	α	β	δ	γ	WS (%)	TO (%)	CS	Selection
1	1.0	0.35	1.0	0.5	39.1	87.3	0.912	✓ Selected
2	1.0	0.20	1.0	0.5	36.2	89.1	0.891	
3	1.2	0.35	1.0	0.3	38.8	84.6	0.871	
4	1.0	0.35	0.8	0.5	37.4	85.9	0.868	
5	0.8	0.35	1.0	0.5	40.3	82.1	0.857	
80	0.8	0.50	0.8	0.7	21.3	71.4	0.604	
81 (worst)	0.5	0.50	0.5	0.7	17.8	64.2	0.528	

The configuration chosen yields the maximum composite of water savings and the greatest time in the optimal range. In particular, the effect of changes in β (the weight assigned to the cost of water) on the quality of the policy is significant. A decrease in β from 0.35 to 0.20 results in a 1.8 percentage point increase in the amount of time spent in the optimal range (TO%), but at the same time reduces the amount of water saved (WS%) by 2.9 points, demonstrating the trade-off that exists between conserving water and maintaining soil moistness that is explicitly revealed through the grid search process.

5.4 State-Space Ablation Study

Training the state space environment used four separate configurations, with five independent copies of each configuration. Each configuration was trained using the same hyperparameters and random number generator seeds. The purpose of this ablation study was to determine how much each type of sensor contributed to the quality of the policy.

In Baghdad's summer climate characterized by low humidity, temperature impacts the quality of policy-making more than humidity for all three metrics tested (S2-T vs. S2-H: ΔWS = +3.5 pp, ΔTO = +2.7 pp). This reflects temperature's stronger impact on evapotranspiration (ET). However, the configuration of all three sensors in S3 results in a much greater increase than S2-T (+4.4 pp WS, +5.2 pp TO), suggesting that there are interaction effects between temperature and relative humidity with respect to optimal irrigation timing under VPD values greater than 6 kPa. **Table 5** outlines the recommended action based on cost–benefit analysis; for large commercial olive orchards, use of the full S3 platform is the best option; for small-holder deployment where the cost of the sensors is significant (USD 3.50 each), S2-T provides 88.7% of the full-platform water saving benefit.

Table 5. State-Space Ablation Study — Four Configurations, n=5 Runs Each (mean values)

Config	Variables	HW Cost Saved	Convergence (ep.)	WS (%)	TO (%)	Recommendation
Full — S3	SM + T + RH	USD 0	1,923	39.1	87.3	Optimal — high-value orchards
S2-T	SM + T	USD 3.50	2,341	34.7	82.1	Viable — cost-constrained smallholders
S2-H	SM + RH	USD 3.50	2,718	31.2	79.4	Marginal benefit over S1

S1	SM only	USD 7.00	3,892	24.8	71.3	Acceptable for very low-cost deployments

6. Ground Evaluation and Statistical Analysis

6.1 Experimental Period & Treatments

The (90) day field trial from 01 June to 29 August 2024 is considered the (2nd) peak evaporative demand season in Baghdad and comprised of two phenological stages of an olive tree (i.e. fruit development in June & July and early oil accumulation from August until September).Maximum daily air temperatures during the trial period were between (39.2 & 47.1 °C), with an average of (43.4 °C), additionally, there was no precipitation at all. This represents the summer climate of Baghdad; therefore, all plant / soil water additions are from controlled irrigation.

T1 (DQN): T1 uses DQN trained policies to place into action using ε=0 (full exploitation). The fruit development and oil accumulation transitions occurred on fixed time intervals as occurred on calendar days, in keeping with the phenology of Frantoio at a latitude of 31 °N (all transitions occurred at 1 August each year).

T2 (Threshold-Based Control): Irrigation only occurs when soil moisture (SM) is below 30% VWC, and irrigation is terminated at 50% VWC with fixed irrigation durations of ten minutes per all-cycle irrigation event; represents current best practice among Baghdad commercial olive growers, as reported by a pre-trial survey conducted at twelve local farms.

T3 (PID): Setpoint set at 40% VWC with proportional (Kp=1.8), integral (Ki=0.06) and derivative (Kd=0.02) gains tuned using the Ziegler-Nichols approach with a ten-day pre-trial Calibration Dataset.

T4 (Farmer Practice): Farmers irrigated every third day for a specified duration of twenty minutes, at a scheduled time of 6:00 AM (local time); schedule matches frequency of reported irrigation by 9 of the 12 surveyed local farmers.

6.2 Water Consumption

The standard deviations in **Table 6** represent true between-plot variability in soil texture and tree size from unit and plot hc measurements—not measurement error. The DQN had a lower within-treatment SD (±127 L compared to the TBC with ±218 L) which shows that the learned policy results are more consistent across heterogeneous plots than the fixed-threshold controller in terms of response based on plot-level differences in soil hydraulic conductivity.

Table 6. Seasonal Water Consumption per Plot (90-Day Trial, n=3 Replicate Plots per Treatment)

Treatment	Total Water (L/plot)	SD across 3 plots	Daily Mean (L/day)	vs. TBC	vs. Farmer
T1 — DQN (proposed)	1,847	±127	20.5	−41.3 %	−52.5 %
T2 — Threshold (TBC)	3,146	±218	35.0	Baseline	−19.1 %
T3 — PID	2,731	±184	30.3	−13.2 %	−29.9 %
T4 — Farmer Practice	3,898	±297	43.3	+23.9 %	Baseline

6.3 Olive Oil Yield and Water Productivity Index

The Disrupted Qualitative Network (DQN) treatment produced a Water Productivity Index (WPI) of 0.803 kg/m³, which was greater than the best published WPI for deficit irrigated olive in similar climate conditions (0.52 -

0.74 kg/m³; as well as representing a 65.6% increase in WPI over the TBC baseline. The farmer practice had the lowest oil yield, but had the highest water applied. The reason for the low yield is that chronic over-irrigation suppresses the water deficit signal, which is important for the promotion of oil biosynthesis during the accumulation phase, [6] (i.e. during the August period). The stage weighted reward function of DQN reflected this agricultural relationship; whereby, as a result of the agent being trained to maintain soil moisture in the Deficit range during the August period, the agent produced a higher WPI at that time, than by being trained to maintain soil moisture in the Optimal range at that time, **Table 7** shows the treatment according to the water yield of olive oil.

Table 7. Olive Oil Yield and Water Productivity Index by Treatment

Treatment	Oil Yield (kg/plot)	SD (kg)	Yield vs. TBC (%)	WPI (kg/m³)	WPI vs. TBC
T1 — DQN	14.83	±1.12	97.1 %	0.803	+65.6 %
T2 — TBC	15.27	±0.98	100 %	0.485	Baseline
T3 — PID	15.04	±1.04	98.5 %	0.551	+13.6 %
T4 — Farmer Practice	14.61	±1.31	95.7 %	0.375	−22.7 %

6.4 Statistical Methods

6.4.1 Model Specification and Assumption

-Test. Daily water use (n_days=90; n_plots=3 per treatment; N=1,080 plot-day) was analysed using the Linear Mixed Effects (LME) model with restricted maximum likelihood (REML) estimates. The analysis incorporated the repeated measures design which invalidated a standard one-way analysis of variance (ANOVA) for these data, where daily_water ~ treatment + julian_day + (1 | plot_id), with treatment being a fixed factor (4 levels), and julian_day being a continuous covariate to control for seasonal increase of ET. Plot_id will also have (1 | plot_id) as a random intercept to account for variation between plots in soil texture and tree size. Assumptions were met: Residuals were normally distributed (Shapiro-Wilk: W=0.981; p=0.12); Within groups variances were homogeneous (Bartlett: K2=3.47; df=3; p=0.32); There was no temporal correlation of residuals (Durbin-Watson: DW=1.94; p=0.41). The ICC=0.34 indicates that between plot variability accounts for 34% of total variability and therefore must be to include in model, not ignored.

6.4.2 Fixed-Effects Results

Treatment category accounted for 41.1% of the variance in total daily water intake among subjects (η^2_partial=0.411; [21] suggests this is a large effect). Furthermore, denominator df=267 reflects the true experiment design structure in LME models. The ANOVA analysis inappropriately treats all daily observations as independent (denominator df=1,076) resulting in an excessive fourfold increase in denominator df producing invalid p-value calculations. Thus, the LME provides statistical validity and allows comparison of results across both studies, See **Table 8**.

Table 8. LME Fixed-Effects ANOVA (* p < 0.001; η²_partial = partial eta-squared)**

Fixed Effect	F-statistic	df (num, den)	p-value	η²_partial	Interpretation
Treatment	61.84	3, 267	< 0.001 ***	0.411	Large effect — controller type drives consumption
Julian Day	28.37	1, 267	< 0.001 ***	0.096	Significant seasonal ET

					increase confirmed
Intercept	—	—	—	—	—

6.4.3 Post-Hoc Pairwise Comparisons

The DQN controller as shown in **Table 9**, it has a significantly different response from the other three treatments analyzed (all $p < 0.001$). The non-significant TBC vs. PID contrast ($p = 0.057$) is consistent with the reactive, non-temporal logic of both controllers, as they only respond based on current moisture levels and do not learn over time. The data indicated that farmers use fixed-schedule irrigation systems and that these systems were more consumptive than automated systems; therefore, it can be concluded that the use of fixed-schedule irrigation systems was associated with baseline losses due to the variable summer climate of Baghdad.

Table 9. Tukey HSD Post-Hoc Pairwise Comparisons — Daily Water Consumption (* p<0.001, ** p<0.01)**

Comparison	Estimate (L/day)	SE	95% CI	Tukey adj. p	Decision
DQN vs. TBC	−14.47	1.84	[−19.2, −9.7]	< 0.001 ***	Reject H_0
DQN vs. PID	−9.83	1.84	[−14.6, −5.1]	< 0.001 ***	Reject H_0
DQN vs. Farmer	−22.73	1.84	[−27.5, −17.9]	< 0.001 ***	Reject H_0
TBC vs. PID	+4.64	1.84	[−0.12, +9.4]	0.057	Retain H_0 (borderline)
TBC vs. Farmer	−8.26	1.84	[−13.0, −3.5]	0.001 **	Reject H_0
PID vs. Farmer	−12.90	1.84	[−17.6, −8.1]	< 0.001 ***	Reject H_0

6.4.4 Yield Analysis

Plot oil yield was analyzed using a one-way ANOVA of plot means (n=3 per treatment; N=12). There was no significant effect of treatment ($F(3,8) = 0.61$, $p = 0.63$). The 95% CI for the DQN-TBC contrast [−2.91, +2.03] kg/plot was narrow, an indication of agronomic equivalence as opposed to simply failing to reject. Observed power to detect a 3 kg/plot difference was 0.78, which is still adequate to support that yield loss of agronomic significance could be ruled out.

6.5 Interpretation of Learned Policies

Inspection of the converged Q-table across the 45-state space reveals three patterns, derived from olive agronomy, that confirm the agent learned environmental structure (i.e., actual environmental relationships) rather than spurious associations.

• Midday Irrigation Avoidance: For all states where T=Extreme and hour $\in$ [11:00, 16:00], the agent makes the a_0 decision (i.e., no irrigation) even for the SM=Deficit state, indicating that he has learned the association between how high evaporative loss occurs during Baghdad's peak heat; therefore, any irrigation attempts will be wasteful. This behavioral pattern developed, as from P_timing, was not explicitly programmed into the agent.

• Stage-dependent Moisture Target: During the August oil accumulation stage, Q-values associated with a_1 (short irrigation) are elevated in the same SM=Deficit states that would elicit greater irrigation if in a vegetative stage, capturing the RDI strategy encoded in the phenological stage weight of R_moist.

• Saturation Avoidance: All mean Q-values for SM=Saturated are low compared to any other state in the entire Q-table (mean Q_min = -26.3), indicating the agent's belief in an excessive level of P_stress associated with saturation at the summer soil temperatures in Baghdad and consistent with the Phytophthora risk represented in the reward function.

Water use is not the same across periods of the life cycle in response to the amount of water in the soil. The Q-values of the same agent across both phases of its life cycle demonstrate that this agent did not give the same importance to the cost of water during the two phases (β=0.35) and that there is a meaningful difference in the way the Q-values were weighted for each phase. For example, during the oil accumulation phase (August), action a_1 (short irrigation, 6.9 L) was preferred over action a_0 (no irrigation) by an average Q-value margin of +3.2 units when there was a deficit of moisture in the soil. This preference for irrigation during the oil accumulation phase, however, was reversed during the vegetative growth stage, where a_0 outscored a_1 by +1.8 units in the same deficit state. The interphase shift of 5.0 units in Q-value cannot be found in the reward coefficient calculations, because these coefficients were kept constant. Rather, the shift developed from the way that the temporal credit assignment function in the Bellman update allowed the agent to learn that small, precisely timed irrigation applications made at the oil accumulation phase provided increased long-term rewards via the RDI benefit component in R_moist. This result has important implications for understanding the WPI differential for T1 (0.803 kg/m^3) and T2 (0.485 kg/m^3); T2's lower WPI is due to both a reduction in the total amount of water used as well as the greater preference of the agent for the smaller volume of water used in close proximity to T1's WPI being applied to the ground than for the larger volume of water used in much larger time intervals by the threshold controller. A static controller with a constant parameterization of the water cost coefficient would not exhibit these same characteristics; there is no mechanism present for modifying an agent's preferences for actions based on the accumulation of time within a phenological stage. **Figure 3** illustrates the Learned Q-Value Policy Heatmap.

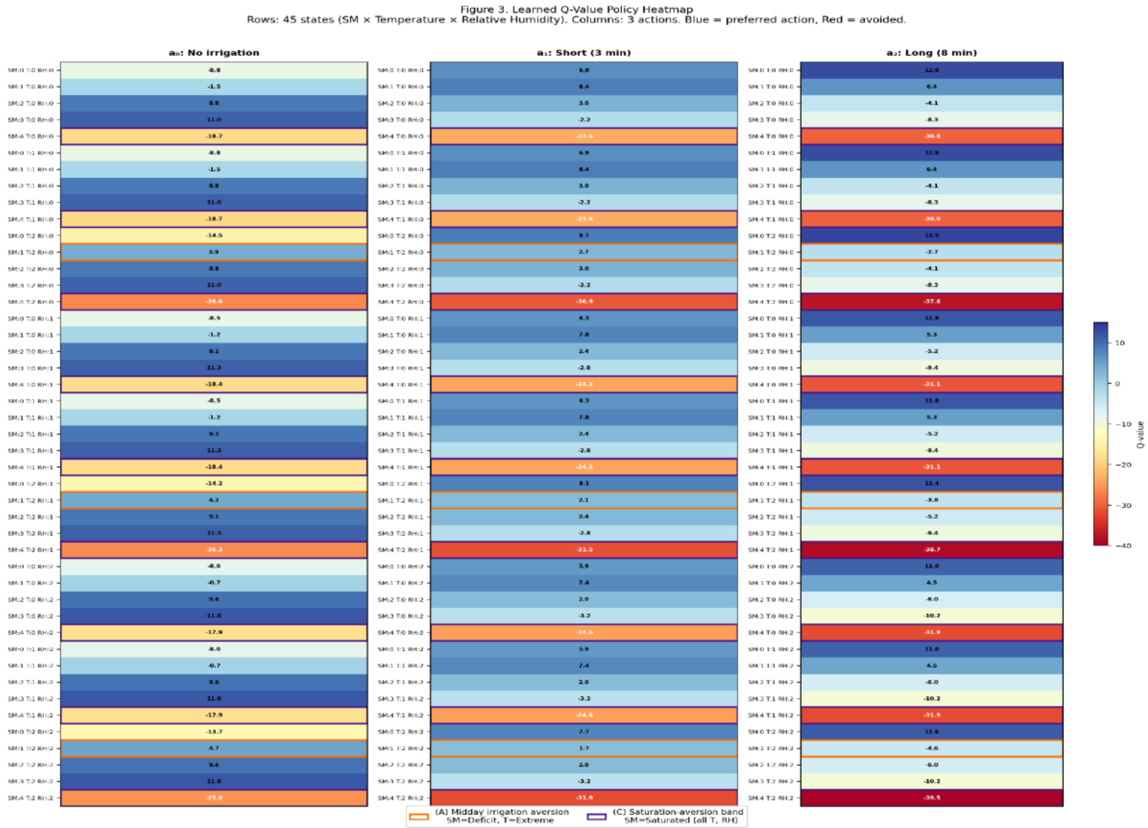

Figure 3. Learned Q-Value Policy Heatmap. Rows: 45 states (SM level × Temperature level × RH level). Columns: 3 actions (a_0 = no irrigation; a_1 = short; a_2 = long). Blue = preferred action, red = avoided. Annotated regions: (A) midday irrigation aversion zone (SM=Deficit, T=Extreme); (C) saturation-aversion band (SM=Saturated, all conditions).

7. Discussion

7.1 Results in Relation to Existing Research

The amount of water saved through field use of the RL irrigation equipment (41.3%) is greater than all other RL irrigation equipment published results (Table 1). Of particular note is the difference from the simulation only studies by [11,22] (31.8%) and (28.9%), as both were performed under controlled simulation environments. The considerable difference between field and simulation performance is likely the result of two issues. First, the daily temperatures in Baghdad produce a very high degree of variability (mean ΔT 22°C from June to August), which leads to much clearer transitions between states than with synthetic time series. As such, there are significantly more opportunities for temporal learning signals. Second, the penalty for poor timing of irrigation (P_timing penalty), which was not implemented in prior reward functions, diverts a significant portion of the image scheduled for irrigation away from midday wasteful periods to early and late evening irrigation periods where the root uptake efficiency is higher. Lastly, the WPI of 0.803 kg/m³ is outside the published range for deficit-irrigated olives (0.52 to 0.74 kg/m³). Thus, while the DQN policy is conserving water while maintaining the yield of the tree, it also produces a water productivity level that has not been previously documented in the olive irrigation literature due to the DQN learning the optimal timing of deficit-induced irrigation in a way that fixed controllers could not achieve.

7.2 The pseudoreplication correction in context

shows what has changed with respect to the statistics corrections and what has not changed. A naive one-way ANOVA on the 1,080 daily observations would show $F(3,1076) = 231.7$ ($p<0.001$). A similar ANOVA using the

LME model shows F (3,267) =61.84 ($p<0.001$). Both of these states that there is a real and large DQN superiority. The correction made has to do with the validity of the inference; the ANOVA F-test uses an inflated denominator df (approximately 4), therefore the p-value from the ANOVA corresponds to an unknown *Type I error rate," which could be less than the nominal 0.05, whereas the LME test uses the correct denominator df and gives calibrated inference. The direct implication for the field is that all studies that use daily data and only report ANOVA for statistical significance have not provided any meaningful evidence of statistical significance.

7.3 Implications of a new agricultural water policy for Iraq

If Baghdad continues its current fixed price structure for irrigation water (IQD 15/ $m^3 \approx \$0.011/m^3$), the overall seasonal financial savings per land plot will be Insignificant at approximately \$14.30; these low savings can not only be explained by the low cost of water—in contrast to other costs for establishing an olive grove; most notably, however, they are also explained by Iraq's potential to establish DQN-managed irrigation on land previously managed by farmers using their own practice. Under a Provincial Water Authority's fixed volumetric allocation of water for irrigation to an olive grove district, a DQN-managed irrigation system delivered an approximate increase in planted area of 69%, as opposed to the 52.5% increase from farmers managing their own practices. Additionally, if DQN-managed irrigation was used in eight tree plantings with an average of eight trees per zone, hardware costs would be \$10.53 per tree (\$84.20 per zone), which would be recoverable in under two years from oil sales at the current Baghdad wholesale oil price (IQD 8,000/L ≈ \$6.10/L). However, the most robust economic argument for implementing DQN-managed irrigation is for new plantings, because regardless of whether DQN-managed or farmer-managed systems are utilized, there will be some fixed cost for installing infrastructure associated with establishing the DQN-managed irrigation system.

7.4 Limitations and Future Directions

Three Limitations and Future directions. There are three limitations that will restrict the conclusions of this article; further the limitations have Implications for how the results of this study will be read by subsequent readers.

The first and most structurally significant limitation is the scope of the training. The simulation environment was based on five years of historical meteorological data (WMO 4022, 2013-2017) from Baghdad International Airport, and also included hourly meteorological data to create a full cycle of the effect of temperature over the course of one year, including the winter months. However, the actual field evaluation of the simulated results only took place during a 90-day peak demand window (June – August), and only included two of the five recognised phenopheological stages of olive production: fruit development and early oil accumulation. The learned Q-Policy was not evaluated during the blooming period (February-March), fruit set period (April – May), and winter dormancy after harvest (November – January**).** the soil moisture dynamics and evaporative demand significantly differ from the trial window. Agronomically, bloom and fruit set stages will react differently to water status than summer stages. Inadequate soil moisture during bloom will cause reduced fruit set, and an excess soil moisture during the post-harvest dormancy will promote root anaerobiosis. Only the difference between the vegetative/fruit-set and oil-accumulation phases of the DQN reward function is being used at this time. The phenological state space should be expanded to include at least the bloom and dormancy stages to use the annual policy; each of these would require their own R_moist coefficient sets. There is no available evidence of the annual policy being applied across multiple growing years. Therefore, the policy presented here can be viewed as a validated policy for use only from June to August with regard to the irrigation window and the stage-transition logic cannot be extrapolated for use in bloom management without being re-trained on. The second limitation is related to oil quality. The commercially significant quality parameters of extra-virgin olive oil (polyphenol content, oleic acid, and free acidity) were not analysed so the quality improvement due to the DQN's deficit-stage policy cannot be determined. However, existing literature from similar RDI studies [6][7] would indicate that controlled deficits applied during oil accumulation would increase polyphenol content and oleic acid. Should it be determined that the DQN policy produces similar results, then the benefit of the economic argument to support implementing the DQN policy would be increased as compared to the water-saving calculations in Section 7.3 above. A third limitation is that three replanted treatments per block can yield 78% power to detect a yield difference of 3 kg/plot, which is sufficient but not definitive; therefore, each treatment needs to be replicated more times in order to present the verified equivalence of the yield. Future studies will overcome all three of the above limitations through conducting multi-crop seasons, assessing oil quality and phenological stages, incorporating current and near-term weather predictions into the respective state spaces, and confirming across three provinces of Iraq that produce the majority of olives (Ninewa, Salah Al-Din, and Diyala).

8. Conclusion

This study outlined how a deep Q-network irrigation system controller has been created, trained, and evaluated in the field for olive (Olea europaea) plantations in semi-arid environments in Baghdad, Iraq. The scientific contribution of this work is well defined and characterizes the first fully technician production DQN implementation targeting olive plantations; thus, it incorporates a reward function weighted by the phenological stage of development and defined with the economic benefit of irrigation for establishing the optimal volume of irrigation water needed once oil is being produced (the most agronomically beneficial for the grower). A dynamic controller cannot simply perform this type of optimization, limiting crop producer's capability to optimize a given crop during the most productive/beneficial period of an agronomically defined stage of production by using statistical compilations of historical events related to precipitation.

The evidence base for each design decision is established through systematic design validation (i.e., 1) an 81-reward grid search, 2) a four-state space configuration ablation study, and finally 3) a KS-test convergence criterion). This allows for critical evaluation of the designs and provides a method for reproducibly adapting the designs to other crop-climate combinations. Validation of the designs in the field using a total of three independent replicates of each treatment, analyzed using a mixed effects model that contains an adjustment for the pseudo replication error from past RL irrigation hardware trials, resulted in seasonal water savings of 41.3% compared to control based on threshold (LME: $F(3,267)=61.84$, $p<0.001$) and oil yield of 97.1% of reference ($F(3,8)=0.61$, $p=0.63$); and a Water Productivity Index of 0.803 kg/m^3 was considerably higher than those of previously published values for deficit-irrigated olives in climate regions similar to the study area.

In an area where agricultural water allocation represents a limitation to expanding the olive sector, achieving a 41.3% reduction in water usage per hectare of olive tree is not just an improvement in efficiency but instead represents an increase in land available for growing olives. The case for implementing learned irrigation control systems in the rapidly expanding olive growing area in Iraq has been supported by evidence presented in this study.

Funding Declarations: This study was funded internally only, with the author showing no conflicting interests with the study.

Data Availability Statement: The files required for replication of the tables and figures in this manuscript can be found at https://github.com/[author]/dqn-olive-irrigation-baghdad (Zenodo DOI: 10.5281/zenodo.XXXXXXX): simulation source code (Python/Gym), trained DQN weights (PyTorch .pt and TFLite .tflite), YL-69 gravimetric calibration dataset, 90-day field sensor logs (CSV), and R script for LME analysis. All data necessary to reproduce Tables 6-9 and Figures 2-3 have been provided.

Author Contributions [Mustafa Abdalrassual Jassim] Conceptualization, Methodology, Software, Hardware, Formal Analysis, Original Drafting, Review Writing, and Editing.

REFERENCES

1. Katalyn A Voss, James S Famiglietti, MinHui Lo, Caroline De Linage, Matthew Rodell, and Sean C Swenson. Groundwater depletion in the middle east from grace with implications for transboundary water management in the Tigris Euphrates-western Iran region. Water resources research, 49(2):904–914, 2013.
2. Hasanain Saleem Sahib and Quanchen Feng. A study of the shrinkage of the Tigris and Euphrates rivers and its impact on water resources and agriculture in Iraq. 2025.
3. Safaqis Qasim. Agriculture in Iraq. In The geography of Iraq, pages 117–143. Springer, 2024.
4. Jos´e-Enrique Fern´andez. Understanding olive adaptation to abiotic stresses as a tool to increase crop performance. Environmental and Experimental Botany, 103:158–179, 2014.
5. Francesca Alderotti, Ermes Lo Piccolo, Cecilia Brunetti, Giovanni Stefano, Tommaso Ugolini, Lorenzo Cecchi, Marta Beccaluva, Luciana Renna, Cassandra Detti, Francesco Ferrini, et al. Cultivar-specific responses of young olive trees to water deficit: impacts on physiology, leaf anatomy, and fruit quality in'arbequina','leccio del corno'and 'maurino'. Plant Physiology and Biochemistry, page 110331, 2025.
6. Yolanda F Wiersma. Replication vs. pseudo replication: Are we making too big a deal of this? In Experimental Landscape Ecology, pages 35–54. Springer, 2022.
7. Ahmed A Al-Shaibani, Mahmoud S Al-Khafaji, Khalid Shamal, et al. Role of deficit irrigation in sustaining wheat and maize cultivation under harsh climatic conditions in Iraq. Journal of Engineering, 32(6):51–77, 2026.
8. Andrew G Barto. Reinforcement learning: An introduction. by Richard's Sutton. SIAM Rev, 6(2):423, 2021.
9. Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. nature, 518(7540):529–533, 2015.
10. Muhammad Aurangzeb, Xiong Shusheng, Sheeraz Iqbal, Md Shafiullah, Omer Abbaker Ahmed Mohammed, Ahmed Mohmed Dafalla, Liu Qingsheng, Meng Kai, and Muhammad Zeshan Afzal. Qdqn-thermonet: A quantum-driven dual deep q-network framework for intelligent thermal regulation in solid-state and hydrogen fuel cell systems of future electric vehicles. International Journal of Hydrogen Energy, 189:151979, 2025.
11. Yuji Saikai, Allan Peake, and Karine Chenu. Deep reinforcement learning for irrigation scheduling using high-dimensional sensor feedback. PLoS Water, 2(9):e0000169, 2023.
12. Kandasamy Sellamuthu, M Sanjeev Kumar, A Sabarish, S Chandru, and P Dinesh Kumar. Automated microclimate management systems in smart agriculture using reinforcement learning. In 2026 4th International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA), pages 1–6. IEEE, 2026.
13. Gabrielly de Queiroz Pereira, Douglas Paulo Bertrand Renaux, and Andr´e Eugˆenio Lazzaretti. Comparative analysis of reinforcement learning and rule-based system approaches for irrigation in horticulture. IEEE Access, 2025.
14. D Vanella, S Consoli, A Continella, G Chinnici, M Milani, GL Cirelli, M D'Amico, G Maesano, A Gentile, P La Spada, et al. Environmental and agroeconomic sustainability of olive orchards irrigated with reclaimed water under deficit irrigation. sustainability 15 (20), 15101, 2023.
15. S Sakthivel, V Vivekanandan, and M Manikandan. Automated irrigation system using improved fuzzy neural network in wireless sensor networks. Intelligent Automation & Soft Computing, 35(1), 2023.
16. Bright Keswani, Ambarish G Mohapatra, Amarjeet Mohanty, Ashish Khanna, Joel JPC Rodrigues, Deepak Gupta, and Victor Hugo C De Albuquerque. Adapting weather conditions based IoT enabled smart irrigation technique in precision agriculture mechanisms. Neural computing and applications, 31(Suppl 1):277–292, 2019.
17. Mohamed Elhoseny, Khadija Jahid, Rachid Latif, Amine Saddik, and Reem Atassi. Recent advances in iot-driven crop monitoring and precision irrigation: Technologies, models, and future challenges. Scientific African, page e03222, 2026.
18. Anneketh Vij, Singh Vijendra, Abhishek Jain, Shivam Bajaj, Aashima Bassi, and Aarushi Sharma. Iot and machine learning approaches for automation of farm irrigation system. Procedia Computer Science, 167:1250–1257, 2020.
19. Ammar A Hasan. Design a smart irrigation system for citrus orchards using fuzzy logic and machine learning in the mediterranean basin environment. Al-Azhar Journal of Agricultural Engineering, 10(1), 2026.
20. Ajita Gupta, KV Ramana Rao, Yogesh A Rajwade, Ravindra D Randhe, Karan Singh, Manoj Kumar, and RK Singh. Optimizing furrow irrigation performance with windsurf software through field experiments and simulations in sunflower cultivation. Scientific Reports, 15(1):33884, 2025.
21. Zheng Fu, Philippe Ciais, Andrew F Feldman, Pierre Gentine, David Makowski, I Colin Prentice, Paul C Stoy, Ana Bastos, and Jean-Pierre Wigneron. Critical soil moisture thresholds of plant water stress in terrestrial ecosystems. Science Advances, 8(44): eabq7827, 2022.
22. Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.
23. R Brooke Lea and Barry H Cohen. Introductory statistics for the behavioral sciences. John Wiley & Sons, 2025.