sustainability

MDPI

Article

Sustainable Multi-Zone Greenhouse Fertigation Through Reliability-Aware Sensing, Spatiotemporal Graph Learning, and Multi-Agent Reinforcement Learning

Xijue Zhang [1,†], Moxian Lin [1,2,†], Yadan Bai [1,†], Haoxiang Yang [1,2], Jinyu Liu [2], Jiahong Pu [1,2] and Yihong Song [1,*]

[1] China Agricultural University, Beijing 100083, China
[2] Peking University, Beijing 100871, China
* Correspondence: songyih2019@163.com
† These authors contributed equally to this work.

Abstract

With protected agriculture increasingly developing toward large-scale, multi-zone, and intelligent production, reducing water and fertilizer consumption, drainage losses, and energy waste while maintaining crop yield, product quality, and long-term economic profitability has become an important challenge for sustainable greenhouse production. To address sensor abnormalities, spatial heterogeneity among irrigation zones, competition for shared hydraulic resources, and the limited ability of conventional control strategies to jointly optimize resource efficiency and economic performance, a sensor-driven spatiotemporal graph multi-agent reinforcement learning framework, termed SensorGraph-FertiMARL, is proposed for state sensing, response prediction, and sustainable coordinated fertigation optimization in multi-zone greenhouses. Reliable zone-level states are first obtained through sensor-reliability gating, missing-state reconstruction, and multi-relational graph fusion incorporating spatial, hydraulic, and microclimatic relationships. An action-conditioned fertigation–yield–quality–profit digital twin is then constructed to jointly predict root-zone responses, resource consumption, marketable yield, product quality, and net profit associated with candidate control actions. Finally, graph-based multi-agent proximal policy optimization (Graph-MAPPO) with capacity, safety, and fairness constraints is employed to coordinate fertigation allocation among multiple zones under shared pump, main-pipeline, and nutrient-solution resource constraints. Evaluated through rigorous offline digital-twin simulation and replay testing calibrated against complete multi-stage historical dataset records from an eight-zone commercial greenhouse, experimental results show that the proposed method reduced water and fertilizer use to 75.8% and 78.3% (achieving water and fertilizer saving rates of 24.2% and 21.7%), respectively, relative to the uniform timer-based control baseline, while decreasing the drainage rate to 7.2%. The Grade-A product rate increased to 84.6%, net profit reached 138.9% of the baseline level (a +38.9% net-profit improvement), return on investment (ROI) reached 1.94, and the stress-violation rate was limited to 2.9%. For state sensing and digital-twin prediction, the sensor-state reconstruction mean absolute error (MAE) and coefficient of determination (R^2) were 0.036 and 0.918, respectively; the water-content prediction MAE was 0.041, the electrical conductivity (EC) prediction root mean squared error (RMSE) was 0.119, and the drainage prediction mean absolute percentage error (MAPE) was 8.7%. In addition, the R^2 for marketable-yield prediction reached 0.912, the macro-averaged F1 (Macro-F1) score for quality-grade prediction reached 0.851, the R^2 for net-profit prediction reached 0.904, the mean prediction interval width (MPIW) was 0.142, and the prediction interval coverage probability (PICP) reached 0.931. These results demonstrate that SensorGraph-FertiMARL can simultaneously achieve water and fertilizer conservation, drainage reduction, operational-risk control, and

Check for updates

Academic Editor: Teodor Rusu

Received: 10 September 2026
Revised: 22 October 2026
Accepted: 28 October 2026
Published: 31 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

economic improvement while maintaining marketable production and quality, thereby providing an effective intelligent optimization approach for resource-efficient, environmentally sustainable, and economically viable greenhouse production.

Keywords: sustainable agriculture; smart greenhouse; precision fertigation; resource-use efficiency; economic sustainability

1. Introduction

With the continuous expansion, intensification, and intelligent transformation of protected agriculture, multi-span greenhouses are gradually shifting from conventional single-zone production toward fine-grained multi-zone management [1]. Fertigation management directly affects crop yield, product quality, resource-use efficiency, and economic returns, making control accuracy a key determinant of greenhouse production efficiency [2]. However, even when the same crop cultivar is grown, different irrigation zones may exhibit substantial spatial heterogeneity because of roof shading, edge effects, ventilation pathways, solar radiation, crop developmental status, fruit load, substrate water-holding capacity, and emitter performance [3]. Consequently, persistent differences arise in air temperature and humidity, light intensity, evapotranspiration, root-zone water content, and electrical conductivity (EC), making uniform fertigation strategies insufficient for accurately satisfying zone-specific water and nutrient requirements [4]. Advances in the Internet of Things, agricultural sensing, and automated control now enable continuous acquisition of high-frequency measurements, including air temperature and humidity, light intensity, CO_2, root-zone water content, EC, pH, flow rate, pressure, drainage volume, and equipment operating conditions, thereby promoting the transition from experience-driven toward data-driven fertigation management [5]. Accordingly, accurately characterizing crop demand from heterogeneous sensing data, coordinating limited resources under shared-equipment constraints, and simultaneously optimizing marketable yield, product quality, and economic return have become important challenges in precision greenhouse management [6]. These challenges are particularly pronounced in typical arid and semi-arid protected cultivation regions such as the Hetao Plain of Bayannur City, Inner Mongolia Autonomous Region, China, where our experimental facility is located. In this region, protected agriculture faces severe environmental and operational pain points: high ambient solar radiation and intense diurnal temperature swings create extreme microclimatic gradients across large greenhouse bays; natural soil and groundwater exhibit substantial secondary salinization risks, requiring strict root-zone electrical conductivity (EC) regulation; and severe water resource scarcity demands high fertigation efficiency. Furthermore, in commercial multi-span facilities, multiple irrigation sectors inevitably share centralized high-power booster pump stations and main supply pipelines. During peak morning transpiration periods, concurrent irrigation demands from peripheral and central zones readily trigger hydraulic pressure collapse and uneven delivery at terminal drip emitters, making coordinated water–nutrient balancing an urgent agronomic necessity. Conventional greenhouse fertigation management mainly includes timer-based irrigation, moisture-threshold control, EC/pH feedback control, evapotranspiration-based control, fuzzy control, and model predictive control [7]. Timer-based and threshold-based strategies are simple, inexpensive, and interpretable, but their predefined empirical rules cannot adequately adapt to rapidly changing environments or pronounced inter-zone heterogeneity [8]. Under heterogeneous evapotranspiration demand, uniform irrigation may simultaneously cause water deficiency in high-demand zones and over-irrigation, nutrient leaching, and salt accumulation in low-demand zones [9]. Model predictive control improves decision quality through future-state

forecasting and receding-horizon optimization, but its effectiveness depends heavily on accurate environment–crop models, whereas water transport, nutrient uptake, evapotranspiration, crop growth, and yield formation are strongly nonlinear, delayed, and uncertain in real greenhouses [10]. Moreover, conventional approaches often treat irrigation zones independently and therefore inadequately represent hydraulic coupling caused by shared pumps, main pipelines, and nutrient-solution tanks [11]. Simultaneous high irrigation demand across multiple zones may consequently cause required flow, pressure, or fertigation capacity to exceed equipment limits [12]. Existing approaches also focus mainly on maintaining agronomic states or maximizing yield, while water, fertilizer, electricity, labor, drainage-treatment costs, and product-grade price differences are rarely considered together; therefore, higher yield does not necessarily result in higher economic return [13].

Deep learning has recently provided new approaches for fertigation decision making in complex greenhouse environments [14]. Long short-term memory networks (LSTMs), temporal convolutional networks (TCNs), and Transformers can learn nonlinear temporal relationships among environmental variables, root-zone states, and irrigation actions, while convolutional neural networks (CNNs) and Vision Transformers can extract canopy coverage, leaf color, wilting severity, fruit load, maturity, and other phenotypic features from crop images [15]. Graph neural networks can further represent irrigation zones as nodes connected through spatial adjacency, hydraulic sharing, and microclimatic similarity, thereby explicitly modeling inter-zone information propagation and resource coupling [16]. Reinforcement learning can optimize long-term control policies through repeated environmental interaction and is therefore suitable for continuous and dynamic fertigation decisions [17]. In particular, multi-agent reinforcement learning can model individual irrigation zones as separate agents and coordinate shared-resource competition through centralized training and decentralized execution [18]. Despite extensive research, existing methodologies fall into three primary categories, each possessing inherent technical limitations that impede practical, sustainable deployment. First, conventional control methods (e.g., timer schedules, threshold triggers, and model predictive control (MPC)) rely on rigid empirical rules or precise mechanistic crop models. However, static rules cannot adapt to rapid microclimatic fluctuations, while physics-based models struggle with strongly nonlinear root-zone nutrient transport, unmodeled disturbances, and delayed physiological responses; furthermore, they treat zones independently, ignoring hydraulic coupling in shared main pipelines. Second, deep-learning-based prediction methods (e.g., long short-term memory networks (LSTMs), Transformers, and spatiotemporal graph neural networks (GNNs)) successfully capture temporal dependencies and spatial correlations, but they predominantly operate in an open-loop predictive paradigm without counterfactual action evaluation. Crucially, they assume pristine sensor observations, leaving them vulnerable to severe state bias under probe drift, stuck-at faults, and packet loss, while failing to link control decisions directly to marketable fruit grades, operating costs, and marginal net profit. Third, reinforcement learning (RL), including multi-agent reinforcement learning (MARL), optimizes continuous control via trial-and-error, but unconstrained exploration risks catastrophic root-zone water stress and salinization; moreover, standard reward functions prioritize gross harvest yield or moisture setpoint tracking while neglecting shared pump/pipe capacity constraints and inter-zone fairness, frequently causing hydraulic overload or chronic starvation of economically lower-margin sectors [19]. Banerjee et al. [20] proposed a five-stage techno-economic digital transformation framework integrating artificial intelligence, the Internet of Things, blockchain, financial technology, and multi-agent simulation; under full digitalization, eight-year cumulative yield increased by 45%, farmers' income doubled, environmental degradation decreased by 28%, and digital literacy improved by more than 60%. Anas Abouaomar et al. [21] developed a game-theoretic multi-agent digital-twin framework in

which reinforcement learning coordinated irrigation, fertilization, and pest-management agents, with cooperative strategies consistently outperforming non-cooperative strategies in crop yield. Zhengjie Fu et al. [22] proposed the multimodal agricultural input–output optimization network MAION, which combines cross-modal spatiotemporal alignment, causal return modeling, and profit-driven reinforcement learning, achieving a 17.68% increase in net profit, a 15.46% improvement in cost efficiency, an input–output ratio of 1.71, and $R^2 = 0.883$ for net-profit prediction.

To resolve these fundamental limitations, each of the four innovations in this paper is strictly necessitated by a distinct operational challenge: (i) an integrated multi-modal sensor–crop–economic data system is necessary to align heterogeneous sensing, phenotypic, and market signals; (ii) a reliability-aware spatiotemporal graph representation is necessary to filter upstream sensor faults and model physical/hydraulic couplings; (iii) an action-conditioned digital twin is necessary to provide counterfactual evaluation of marginal net returns and risks before executing actions; and (iv) a capacity-, safety-, and fairness-constrained Graph-MAPPO policy is necessary to safely and equitably arbitrate shared hydraulic resources among competing zones. By coupling these components into a unified closed loop, the framework achieves both individual module efficacy and synergistic multi-stage gains, wherein upstream reliability gating insulates the digital twin from corrupt observations, the digital twin furnishes risk-aware marginal economic gradients to the policy, and constrained policy projection enforces hard physical safety guarantees. To address these limitations, a sensor-driven spatiotemporal graph multi-agent reinforcement learning framework, termed SensorGraph-FertiMARL, is proposed for coordinated fertigation control and economic-return optimization. Rather than treating sensor measurements as conventional inputs, the framework integrates sensor reliability, inter-zone spatial relationships, hydraulic-sharing constraints, crop responses, and agricultural economic outcomes throughout state perception, response prediction, and collaborative decision making. Environmental, root-zone, supply-network, and equipment measurements are first subjected to anomaly detection, quality assessment, missing-value reconstruction, and multi-frequency temporal alignment, and they are then integrated with canopy imagery to construct a dynamic multi-relational graph containing physical-adjacency, hydraulic-sharing, and microclimatic-similarity edges. An action-conditioned fertigation–yield–quality–profit digital twin subsequently predicts future root-zone water content, *EC*, drainage, resource consumption, marketable yield, quality grade, and net profit under candidate fertigation actions, while water-balance, fertilizer-consistency, and economic-consistency constraints improve prediction credibility. Finally, each irrigation zone is modeled as an agent within a graph-structured centralized-Critic and zone-specific-Actor framework, in which irrigation amount, irrigation duration, and nutrient-solution *EC* are jointly optimized under shared pump capacity, main-pipeline flow, nutrient-solution inventory, and root-zone safety constraints. Fertigation management is therefore transformed from conventional demand-based supply toward coordinated allocation according to the marginal economic value of additional water and fertilizer inputs. The principal contributions and innovations of this study can be summarized as follows:

1. An integrated multi-zone sensor–crop–fertigation–economic data system is constructed, in which environmental measurements, root-zone states, hydraulic-network information, equipment operating data, crop imagery, yield and quality observations, and economic records are jointly aligned by irrigation zone and time, providing a unified basis for extending fertigation control from agronomic optimization toward economic optimization;
2. A reliability-aware multi-relational spatiotemporal graph state representation method is proposed by integrating sensor-quality gating, missing-state reconstruction,

and graph attention over spatial, hydraulic, and microclimatic relationships, thereby reducing state-estimation bias caused by sensor failures and direct feature concatenation.

3. An action-conditioned fertigation–yield–quality–profit digital twin is developed to jointly predict short-term root-zone responses and long-term marketable yield, product quality, resource costs, and net profit, while physical and economic consistency constraints are incorporated to estimate marginal net returns and associated uncertainty for alternative fertigation actions.
4. A graph-based multi-agent reinforcement learning controller with capacity, safety, and inter-zone fairness constraints is proposed to dynamically allocate limited water and fertilizer resources according to zone-specific demand and marginal economic value while maintaining agronomic safety, thereby reducing water and fertilizer consumption, drainage losses, and energy use and improving greenhouse-level net profit per unit area and overall input–output efficiency.

2. Related Work

2.1. Sensor-Based Greenhouse Fertigation Monitoring and Zone-Level Control

Greenhouse fertigation control relies on continuous sensing of the crop environment, root-zone water and nutrient status, and irrigation-system operation, followed by adjustment of irrigation volume, duration, and nutrient-solution concentration to maintain appropriate root-zone conditions [23]. Early systems mainly adopted timer-based irrigation and quantitative fertilization according to predefined schedules [24]. Although simple to implement, such strategies cannot adequately respond to changing crop requirements across developmental stages and environmental conditions [25]. With advances in sensing technologies, feedback control based on substrate water content, *EC*, and pH thresholds has become increasingly important in protected agriculture [26]. When measurements deviate from target ranges, irrigation or nutrient-solution adjustment is automatically triggered, forming a closed-loop "sensing–decision–execution" process [27]. More advanced approaches include evapotranspiration-based estimation using climatic and crop-state information, fuzzy control for nonlinear relationships, and model predictive control for future-state prediction and receding-horizon fertigation optimization [28].

Air temperature, humidity, light intensity, and CO_2 sensors characterize greenhouse microclimate and evapotranspiration drivers, while substrate water content, *EC*, and pH sensors describe root-zone conditions; flow, pressure, irrigation-input, and drainage sensors further quantify fertigation input, hydraulic operation, and drainage losses [29]. These heterogeneous measurements provide the sensing foundation for precision fertigation management [30]. However, many existing systems still rely on only a few representative sensors, limiting their ability to capture the strong spatial heterogeneity of large multi-span greenhouses [31]. Long-term sensor operation may also introduce drift, stuck-at faults, abnormal fluctuations, communication losses, and inconsistencies among probes installed at different locations [32]. Moreover, environmental, root-zone, equipment, and image data are often collected at different frequencies, making direct feature concatenation insufficient for constructing reliable control states [33]. In multi-zone greenhouses, independently controlled zones commonly share pumps, main pipelines, and nutrient-solution tanks, meaning that locally optimal decisions may not produce globally optimal fertigation allocation, thereby motivating coordinated multi-zone control [34]. In summary, while existing sensor-based monitoring provides the empirical foundation for automated fertigation, existing systems predominantly operate under an idealized assumption of perfect sensor integrity, leaving state estimation vulnerable to sensor degradation, fouling, and missing packets. Furthermore, treating multi-zone systems as independent single zones neglects

the physical reality of shared hydraulic infrastructure, highlighting the urgent need for reliability-aware state estimation coupled with coordinated, network-wide allocation.

2.2. Multi-Source Sensor Fusion, Graph Neural Networks, and Crop Response Prediction

Multi-source sensor fusion exploits complementary information from heterogeneous modalities to provide a more complete representation of greenhouse environmental conditions, root-zone states, and crop-growth processes [35]. Conventional direct feature concatenation is limited in capturing temporal dependencies, spatial relationships, and complex cross-modal interactions [36]. Deep learning has enabled convolutional neural networks (CNNs) and Vision Transformers to extract phenotypic characteristics such as leaf coverage, leaf color, wilting severity, fruit number, and maturity from crop canopy images [37]. Meanwhile, LSTMs, temporal convolutional networks (TCNs), and Transformers can respectively model long-term sequence dependencies, multi-scale temporal dynamics, and long-range associations among environmental conditions, fertigation actions, and root-zone responses, and they have been increasingly applied to greenhouse forecasting, substrate-moisture prediction, and crop-yield estimation [38].

Multi-zone greenhouse systems exhibit not only temporal dependence but also pronounced spatial interactions and equipment connectivity [39]. Graph neural networks provide a natural representation by modeling different entities as nodes and their relationships as edges [40]. Graph convolutional networks (GCNs) capture inter-regional dependencies through neighborhood aggregation, graph attention networks (GATs) adaptively weight neighboring nodes, and spatiotemporal graph neural networks jointly model spatial propagation and temporal dynamics [41]. These methods have been applied to spatial sensor interpolation, microclimate forecasting, agricultural yield estimation, and resource allocation [42]. For multi-zone greenhouses, graph structures can simultaneously represent physical adjacency, hydraulic sharing, and microclimatic similarity among irrigation zones [43], providing a more realistic description of inter-zone coupling than direct concatenation or regular grid-based representations [44]. Nevertheless, most existing graph-based approaches remain focused on environmental-state or yield prediction and rarely estimate how specific fertigation actions affect subsequent resource consumption, product quality, and economic returns [45]. Therefore, an action-conditioned fertigation–yield–quality–profit digital twin is developed to extend graph-based state modeling toward control-oriented prediction of agronomic and economic outcomes under alternative actions. Critically, while modern spatiotemporal GNNs excel at environmental state forecasting and spatial interpolation, they remain fundamentally decoupled from downstream control decision making. Existing models predict passive system evolution under observed conditions but cannot evaluate counterfactual outcomes under hypothetical fertigation inputs. Bridging this gap requires an action-conditioned digital-twin architecture that couples multi-relational graph topologies with multi-task prediction of short-term root-zone dynamics and long-term economic outcomes under strict physical consistency constraints.

2.3. Reinforcement Learning for Fertigation Control and Agricultural Economic Optimization

Reinforcement learning formulates control as a sequential decision-making process in which an agent selects actions from the current state, receives environmental feedback, and learns a policy by maximizing long-term cumulative reward [46]. Compared with fixed-threshold and deterministic-model-based optimization methods, reinforcement learning does not require exact analytical models for all environmental relationships and is therefore suitable for greenhouse fertigation problems characterized by nonlinearity, delayed responses, and long-term cumulative effects [47]. The deep deterministic policy gradient (DDPG) and twin delayed deep deterministic policy gradient (TD3) can optimize

continuous control variables, soft actor–critic (SAC) improves exploration and stability through maximum-entropy learning, and proximal policy optimization (PPO) stabilizes policy updates through clipping mechanisms [48]. Because direct trial-and-error exploration in real greenhouses may cause over-irrigation, water stress, or salinity damage, offline reinforcement learning has received increasing attention, allowing policies to be trained from historical data or digital-twin environments before controlled deployment [49]. This strategy can reduce operational risks during real-world policy learning [50].

When greenhouses are divided into multiple irrigation zones, the rapid growth of the joint state and action spaces makes multi-agent reinforcement learning more suitable than single-agent formulations [51]. Methods such as multi-agent deep deterministic policy gradient (MADDPG) and multi-agent proximal policy optimization (MAPPO) adopt centralized training and decentralized execution, allowing individual zones to act from local observations while exploiting global information during training to learn coordination relationships [52]. This mechanism is well suited to greenhouse systems in which multiple zones share pumps, main pipelines, and nutrient-solution resources [53]. However, reward functions in existing reinforcement-learning-based fertigation studies are commonly based on water-status deviations, resource consumption, or crop yield, while product-grade prices, water and fertilizer costs, energy use, drainage losses, and equipment capacity are rarely integrated into a unified economic objective [54]. Moreover, maximizing total return alone may continuously prioritize high-profit zones and increase long-term stress risks in lower-value zones [55]. Therefore, graph-based multi-agent reinforcement learning is integrated with agricultural economic optimization in this study, while shared-capacity, root-zone safety, and inter-zone fairness constraints are introduced to extend the objective from conventional "maintenance of optimal fertigation states" or "maximization of crop yield" toward "coordinated allocation of limited water and fertilizer resources according to marginal economic value while maintaining agronomic safety and maximizing long-term greenhouse net profit." In recent years, cutting-edge studies have begun exploring the intersection of digital twins and multi-agent reinforcement learning (DT-MARL) for agricultural water–energy management. However, existing DT-MARL frameworks still encounter critical hurdles: they rely on unconstrained cooperative policy gradients that cannot guarantee instantaneous physical safety against pump overload, and their reward formulations focus predominantly on aggregate biomass or volumetric savings without considering tiered product market prices or long-term inter-zone fairness. Our work directly addresses these open frontiers by embedding hard convex safety projections and Jain-fairness regularization into a graph-structured centralized-Critic multi-agent architecture.

3. Materials and Method

3.1. Data Collection

Field data collection was conducted in a modern commercial Venlo-type multi-span glass greenhouse at the Smart Agriculture Fertigation Experimental Base in Wuyuan County, Bayannur City, Inner Mongolia Autonomous Region, China, spanning 184 continuous production days from 10 April 2024 to 10 October 2024. The greenhouse covers a total footprint of 1800 m^2 (45.0 m length × 40.0 m width, with a ridge height of 5.8 m and gutter height of 4.5 m), oriented North–South. The greenhouse is partitioned into eight independently controlled irrigation zones (designated as Zones 1–8), each measuring 22.5 m × 10.0 m (225.0 m^2). Each zone accommodates six standardized coir/rockwool substrate cultivation gutters with a high-density planting arrangement of 3.2 plants/m^2, cultivating indeterminate pink tomato (*Solanum lycopersicum* L., cv. 'Flanders'). All eight zones share a centralized water–fertilizer integration pumping station equipped with a variable-frequency multi-stage booster pump (15.0 m^3/h rated capacity at 0.4 MPa), a main supply pipeline (DN50 PVC),

three concentrated stock-solution tanks (Tank A for calcium/potassium nitrate, Tank B for phosphate/sulfates/micronutrients, and Tank C for nitric/phosphoric acid pH regulation), and independent branch solenoid valves (DN25) paired with pulse flow meters for precise volumetric metering.

Figure 1 summarizes the eight-zone spatial arrangement, shared fertigation network, and monitoring positions.

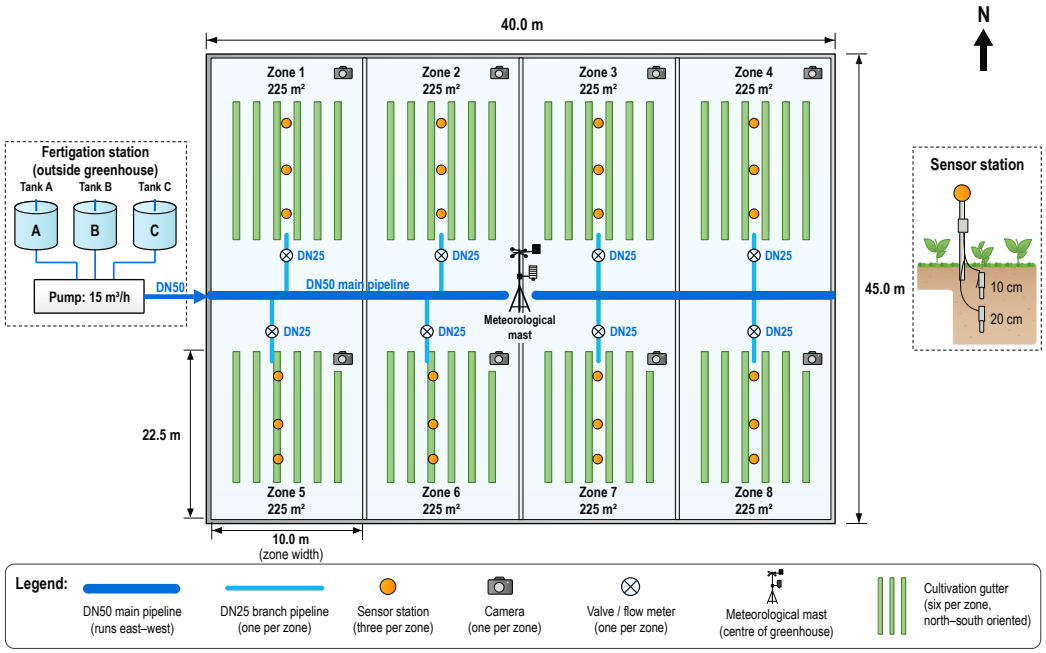


Figure 1. Spatial layout of the eight-zone greenhouse and shared fertigation system. The 40.0 m × 45.0 m greenhouse contains eight 225 m^2 zones, each 10.0 m × 22.5 m with six cultivation gutters. A 15 m^3/h pump and three stock-solution tanks supply a DN50 main pipeline and DN25 zone branches; DN denotes nominal pipe diameter in millimeters. The diagram marks zone valves/flow meters, one fixed camera and three monitoring stations per zone, the central meteorological mast, and north orientation. The inset shows root-zone probe depths of 10 and 20 cm. Equipment symbols and positions are shown schematically.

Data were collected across the complete crop ontogeny, covering the seedling, vegetative growth, flowering and fruit-setting, fruit-expansion, and final harvesting stages, as shown in Table 1. The experimental greenhouse was divided into multiple irrigation zones according to independently controlled solenoid-valve units. A unified identification system was established for all zones, allowing sensors, irrigation branches, control actions, crop images, and harvest records to be temporally and spatially associated with their corresponding irrigation zones. Greenhouse microclimatic data were continuously acquired using an environmental monitoring system and primarily included air temperature, relative humidity (RH), solar radiation, light intensity, and other relevant environmental variables. These measurements were used to characterize evapotranspiration drivers and spatially varying environmental conditions across different irrigation zones. Root-zone dynamics were monitored using Decagon 5TE (METER Group, Pullman, WA, USA) and ZD-6 sensors. Specifically, Decagon 5TE probes measure volumetric water content (VWC, range 0.0–100.0%, resolution 0.08% VWC over 0–50% VWC, typical generic-calibration accuracy ±3.0% VWC and medium-specific calibration accuracy ±2.0% VWC), bulk electrical conductivity (*EC*, range 0.0–23.0 dS/m, resolution 0.01 dS/m over 0–7 dS/m and 0.05 dS/m over 7–23 dS/m, accuracy ±10.0% over 0–7 dS/m with user calibration required over 7–23 dS/m), and substrate temperature (range −40.0–60.0 °C, resolution 1.0 °C, accuracy ±1.0 °C). ZD-6 probes simultaneously measure substrate pH (range 3.0–14.0 pH, accuracy ±0.2 pH)

and pore-water EC. In each irrigation zone, three representative monitoring stations were deployed along the cultivation gutters, with dual-depth probes permanently installed at depths of 10 cm (upper root activity zone) and 20 cm (drainage buffer zone) within substrate slabs. Sensor uniformity was rigorously verified prior to installation through standardized multi-point calibration in controlled environmental chambers and standard buffer solutions (1.413 dS/m KCl and pH 4.01/7.00 buffers), yielding an inter-probe coefficient of variation (CV) of less than 2.4% under identical conditions. The 5TE supports Decagon Digital Interface (DDI) serial or Serial Digital Interface at 1200 baud (SDI-12) communication and a 3.6–15.0 V direct-current (DC) supply, compatible with the described 12 V supply. Its native digital interface connects to a compatible data-logger input; the RS-485/Modbus links serve the corresponding industrial equipment interfaces. Sensor clusters were supplied by isolated 12 V DC solar-battery uninterruptible power systems, with Campbell Scientific CR1000X data loggers (Campbell Scientific, Inc., Logan, UT, USA) used for acquisition. Manufacturer specifications are summarized in the 5TE integrator guide [56]. Logged signals were packetized and wirelessly transmitted to an on-site edge industrial server via LoRaWAN and industrial 4G cellular links at a 10-min acquisition frequency.

Table 1. Comprehensive specification of data acquisition equipment, measurement parameters, installation, and acquisition counts. Continuous-stream counts are planned totals over 184 days; event and harvest counts are logged records.

Data Modality	Sensor/Hardware Model or Equipment Type	Measurement Range	Accuracy/Resolution	Installation/ Interface	Sampling	Planned/Logged Records
Microclimate	Davis Vantage Pro2/SQ-500	T: −40–65 °C, RH: 0–100%, PAR	±0.3 °C, ±2%, ±5%	Central mast, 2.5 m height/Modbus	10 min	26,496
Root-Zone State	Decagon 5TE/ZD-6/CR1000X	VWC: 0–100%, EC: 0–23 dS/m, pH: 3–14	VWC: ±3% generic; EC: ±10% at 0–7 dS/m; pH: ±0.2	Slabs (10, 20 cm); 5TE: DDI serial/SDI-12	10 min	1,271,808
Canopy RGB	Hikvision DS-2CD3T46WDV3-I3	2560 × 1440 pixels, 4MP CMOS	Visual phenotyping, color/load	Truss mounted, 3.2 m, 45° tilt/IP	30 min	70,656
Irrigation Flow	TUF-2000B-type Ultrasonic/Pulse Flow Meter	Flow: 0.1–20.0 m^3/h, Pres: 0–0.6 MPa	±1.0% reading, ±0.5%	DN50 main & DN25 branch lines/ 4–20 mA	Event/10 m	9846
Fertilization	Multi-channel Venturi Injector	Dosing: 0–300 L/h, EC target: 1–4 dS/m	±1.5% dosing concentration	Fertigation mixing unit/Modbus	Event/10 m	6122
Pump & Energy	Acrel DTSD1352 Smart Meter	Power: 0–30 kW, Voltage: 380 V	Class 0.5S active energy accuracy	Control cabinet bus/RS-485	10 min	26,496
Marketable Yield	Mettler Toledo Industrial Scale	Weight: 0.01–100.0 kg	±5 g certified precision	Harvest batch weighing per zone	Batch	2160
Fruit Quality	ATAGO PAL-1 Refractometer/Size	Brix: 0–53%, Fruit diameter: 0–150 mm	±0.2% Brix, ±0.1 mm	Post-harvest laboratory assay	Batch	2160

Manufacturer locations: Acrel DTSD1352—Jiangsu Acrel Electrical Manufacturing Co., Ltd., Jiangyin, China; Mettler-Toledo industrial scale—Mettler-Toledo International Inc., Greifensee, Switzerland; ATAGO PAL-1—ATAGO Co., Ltd., Tokyo, Japan. The flow-meter designation identifies a model family/equipment type; the Venturi injector identifies an equipment type.

Greenhouse environmental conditions were synchronously monitored using Davis Vantage Pro2 (Davis Instruments Corporation, Hayward, CA, USA), HOBO U30 (Onset Computer Corporation, Bourne, MA, USA), SHT31 (Sensirion AG, Staefa, Switzerland), and Apogee SQ-500 (Apogee Instruments, Inc., Logan, UT, USA) devices installed at a central meteorological mast 2.5 m above the floor (measuring air temperature with ±0.3 °C accuracy, relative humidity with ±2.0% accuracy, and photosynthetically active radiation (PAR) with ±5.0% calibration uncertainty). Measurements obtained from different sources

were first calibrated and aligned according to unified timestamps and were subsequently aggregated into the 30-minute control windows used by the proposed model.

Crop visual data were continuously acquired using fixed Hikvision DS-2CD3T46WDV3-I3 red–green–blue (RGB) cameras (Hangzhou Hikvision Digital Technology Co., Ltd., Hangzhou, China). The cameras were installed at fixed positions above or obliquely above the crop canopy in each irrigation zone, and canopy images were captured at predefined intervals, while the imaging distance, viewing angle, and background were maintained as consistently as possible. Additional close-range images were periodically collected using a DJI Osmo Action 4 camera (SZ DJI Technology Co., Ltd., Shenzhen, China) to supplement the fixed-view observations. The resulting image dataset covered different illumination conditions and crop developmental stages and was used to extract phenotypic characteristics such as canopy coverage, leaf color, wilting severity, fruit load, and maturity status. Fertigation-control data were directly obtained from the operational logs of the integrated fertigation control system. Irrigation records included valve-opening time, irrigation duration, and actual irrigation volume, whereas fertilization records contained fertilization timestamps, nutrient-solution formulations, and corresponding fertilizer-management operations. Equipment-operation data were synchronously recorded to characterize the operating states and energy consumption of pumps, valves, and control devices, thereby enabling quantitative associations to be established between fertigation inputs and energy expenditure. Yield and quality measurements were manually collected according to irrigation zone and harvest batch during the harvesting period. Yield-related records included actual harvested weight and marketable yield, whereas quality measurements included fruit grade, individual fruit weight, and representative quality attributes. To establish a rigorous, transparent foundation for the economic optimization objective, local resource prices and cost conversion parameters were standardized based on actual Bayannur agricultural accounting records and utility rate schedules: (i) agricultural water tariff $C_{water} = 2.80$ CNY/m^3; (ii) soluble compound fertilizer blend $C_{fert} = 4.20$ CNY/kg; (iii) grid electricity rates under local time-of-use (TOU) pricing, comprising peak hours (08:00–11:00, 18:00–21:00) at 0.98 CNY/kWh, shoulder hours (07:00–08:00, 11:00–18:00, 21:00–23:00) at 0.62 CNY/kWh, and valley hours (23:00–07:00) at 0.31 CNY/kWh; (iv) routine operational and maintenance labor wage converted at 180.0 CNY/person-day (22.50 CNY/person-hour); and (v) tiered wholesale market tomato prices categorized by commercial quality grading, with Grade-A (premium uniform fruit, high soluble solids > 5.5°Brix) at $P_A = 8.50$ CNY/kg, Grade-B (standard commercial fruit) at $P_B = 5.20$ CNY/kg, and Grade-C (sub-standard/culled fruit) at $P_C = 2.00$ CNY/kg. Water cost, fertilizer cost, electricity cost, labor input, and product selling prices were obtained from production accounting records, equipment-operation logs, and contemporaneous sales records maintained at the experimental base. These economic variables were aligned with sensor and control observations according to date and harvest batch. For the 184-day acquisition schedule, a 10-minute interval gives $184 \times 24 \times 6 = 26{,}496$ planned time points per continuously sampled unit. The eight zones, three monitoring stations per zone, and two probe depths give $48 \times 26{,}496 = 1{,}271{,}808$ planned root-zone records, with each record grouping the measurements at one station–depth and timestamp. The central microclimate station and central energy meter each correspond to 26,496 planned multivariable records. With one fixed camera per zone and a 30-minute interval, the fixed-view image schedule contains $8 \times 184 \times 24 \times 2 = 70{,}656$ planned captures. These continuous-stream counts describe the acquisition schedule before missing-data handling and quality screening; supplementary handheld images are counted separately. Event-based production logs contain 9846 irrigation records, 6122 fertilization records, and 2160 observations each for crop yield and quality. These aligned sensor and production streams provide

a unified data foundation for subsequent sensor-reliability modeling, action-conditioned digital-twin construction, and multi-agent coordinated fertigation control.

3.2. Data Augmentation

The fundamental objective of data augmentation is to expand the state distribution of the training data by constructing perturbed samples that remain consistent with realistic greenhouse operating conditions, while preserving the underlying agronomic semantics and control labels of the original samples. In this manner, robust feature representations can be learned against sensor errors, missing observations, temporal fluctuations, topological variations, and equipment abnormalities. Because the inputs considered in this study simultaneously include environmental sensors, root-zone fertigation sensors, hydraulic-network and equipment information, crop images, and multi-zone graph structures, random perturbation of only a single modality may disrupt the inherent physical correspondence among different data sources. Therefore, data augmentation is conducted according to the principles of physical plausibility, label preservation, and scenario consistency. Sensor measurements, missing-data patterns, temporal sequences, graph structures, control scenarios, and crop images are jointly augmented to improve the generalization capability and robustness of SensorGraph-FertiMARL under complex real-world greenhouse conditions.

For continuous sensor measurements, noise augmentation based on the measurement accuracy of individual devices is first applied to simulate practical errors in temperature, humidity, water content, *EC*, pH, flow rate, and pressure sensors. Let the original sensor observation at time t be denoted by x_t, and let the augmented observation be denoted by $\tilde{x}_t$. Gaussian noise augmentation is formulated as

$$\tilde{x}_t = x_t + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, \sigma_s^2), \tag{1}$$

where σ_s is determined according to the nominal measurement accuracy or historical calibration error of the corresponding sensor rather than by arbitrarily imposing excessive random perturbations. For devices with explicitly specified maximum measurement errors, bounded perturbations are adopted, with $\epsilon_t \in [-\delta_s, \delta_s]$, where δ_s denotes the maximum allowable error range. This strategy enables realistic sensor noise to be reproduced while preventing the generation of anomalous samples that clearly violate greenhouse physical constraints. Considering that wireless communication interruptions, probe failures, and equipment aging may result in missing measurements, stuck-at faults, or gradual drift, mask-based fault augmentation is further introduced. Let the original multi-sensor state vector be denoted by x_t, and let a binary mask m_t be randomly generated. The augmented state is expressed as

$$\tilde{x}_t = m_t \odot x_t + (1 - m_t) \odot x_{mask}, \tag{2}$$

where $\odot$ denotes element-wise multiplication, and x_{mask} represents the placeholder state used for missing observations. During training, not only are individual sensors randomly masked, but continuous missing intervals involving homogeneous sensor groups and complete local sensor failures within individual irrigation zones are also simulated. This design enables the reliability-gating and spatiotemporal graph reconstruction modules to learn how historical information, and neighboring-zone states can be exploited to recover missing observations. For slowly developing sensor-drift faults, the augmented observation can be further defined as

$$\tilde{x}_t = x_t + \alpha(t - t_0) + \epsilon_t, \tag{3}$$

where α denotes the drift rate per unit time, and t_0 represents the onset time of the drift. This formulation is used to emulate systematic deviations that may gradually develop in *EC*, pH, and water-content sensors during long-term operation. Greenhouse water and nutrient

dynamics exhibit pronounced temporal continuity and delayed responses. Therefore, constrained temporal augmentation is applied rather than arbitrary temporal reordering. For a sensor sequence of length T, denoted by $X = x_1, \ldots, x_T$, a continuous interval $[t_s, t_e]$ is randomly selected for temporal masking to simulate short-term communication loss. In addition, window cropping and limited temporal scaling are performed only within the same crop developmental stage. The temporal transformation is formulated as

$$\tilde{X}(t) = X(\gamma t), \qquad \gamma \in [1-\eta, 1+\eta], \tag{4}$$

where γ denotes the temporal scaling factor, and η controls the maximum allowable degree of temporal variation. By constraining η to a limited range, fundamental physiological response patterns, such as the increase in root-zone water content immediately after irrigation followed by its gradual decline due to evapotranspiration, can be preserved. Consequently, augmented sequences that contradict realistic fertigation-response dynamics are avoided. For the graph structure of the multi-zone greenhouse, constrained graph augmentation is further performed. Let the dynamic greenhouse graph be denoted by $\mathcal{G} = (\mathcal{U}, \mathcal{E}, A)$, where A represents the adjacency matrix. Low-weight spatial edges or microclimatic-similarity edges are randomly removed with probability p_e, while small perturbations are introduced into non-critical edge weights:

$$\tilde{A}_{ij} = M_{ij}(A_{ij} + \xi_{ij}), \qquad \xi_{ij} \sim \mathcal{N}(0, \sigma_e^2), \tag{5}$$

where $M_{ij} \in 0, 1$ denotes the edge-retention mask. This procedure is intended to simulate scenarios in which information from local irrigation zones becomes temporarily unavailable or spatial relationships vary over time. However, critical hydraulic connections associated with shared main pipelines, pumps, and nutrient-solution tanks are excluded from ordinary random edge removal, thereby preventing disruption of the actual equipment topology and resource constraints. In addition to data-level perturbations, control-scenario augmentation is conducted using the action-conditioned digital twin. External operating conditions, including available pump capacity C_t, environmental load, sensor operating states, and water and fertilizer prices, are modified while maintaining physiologically plausible crop responses. For example, the degraded equipment capacity can be represented as

$$\tilde{C}_t = \rho C_t, \qquad \rho \in [\rho_{min}, 1], \tag{6}$$

where ρ characterizes the degree of capacity reduction in the pump or main pipeline. This setting is combined with simulated scenarios involving simultaneous irrigation requests from multiple zones, localized high-temperature conditions, elevated solar radiation, insufficient nutrient-solution inventory, and fluctuations in water, fertilizer, or energy prices. Such scenario augmentation allows the reinforcement-learning policy to learn stable and coordinated resource-allocation behavior under resource scarcity and abnormal operating conditions. Finally, lightweight visual augmentation is applied to canopy RGB images, including adjustments in brightness and contrast, slight rotations, local occlusion, and moderate blurring. These transformations are designed to simulate variations in natural greenhouse illumination, minor camera displacement, and partial leaf occlusion. All visual perturbations are constrained to ranges that do not alter key biological labels, including canopy coverage, wilting status, fruit load, and maturity. The hyperparameters governing data augmentation are selected based on strict physical and empirical criteria to ensure biological and operational plausibility: (i) Gaussian noise variance σ_s matches the empirical calibration limits of the respective sensors (± 0.3 °C for temperature, $\pm 2.0\%$ for RH, $\pm 3.0\%$ for VWC, ± 0.1 dS/m for EC), with bounded clip $[\pm\delta_s]$ preventing negative water content or non-physical light intensity; (ii) the mask-based missing ratio is set between 5% and 20% based on historical wireless packet-drop logs observed during extreme weather events;

(iii) the drift rate α is bounded at 0.005 dS/(m · day) to replicate standard electrochemical probe polarization rates before weekly automated cleaning; (iv) the temporal scaling range $\eta \leq 0.15$ strictly preserves physiological water uptake and drainage kinetics without distorting the characteristic 2–4 h root absorption curves; (v) the non-critical edge dropout probability $p_e = 0.10$ reflects intermittent microclimate similarities while preserving permanent hydraulic topologies; and (vi) the capacity degradation factor $\rho \in [0.70, 1.00]$ accurately reflects pump impeller wear and drip-filter clogging before backwashing.

Through the multi-level augmentation strategies described above, the training process is not limited to simply increasing the number of samples. Instead, state distributions corresponding to realistic greenhouse sensor failures, environmental fluctuations, equipment degradation, and resource competition are deliberately constructed. Consequently, the robustness and generalization capability of the spatiotemporal graph state representation, action-conditioned digital twin, and multi-agent fertigation control policy can be improved under sensor anomalies, cross-seasonal conditions, and complex resource constraints.

3.3. Proposed Method

3.3.1. Overall

SensorGraph-FertiMARL takes temporally aligned and feature-constructed multi-source greenhouse data as input and establishes a three-stage closed-loop architecture consisting of reliable state representation, action-response prediction, and collaborative policy optimization. First, environmental, root-zone, hydraulic-network and equipment, crop-visual, and economic features are fed into their corresponding temporal or visual encoders to obtain initial modality-specific representations for each irrigation zone. Subsequently, a sensor-quality gating mechanism dynamically assigns reliability weights to individual observations according to missingness, anomalies, drift, and neighborhood consistency, while historical sequences and neighboring-zone information are jointly exploited to reconstruct missing states. On this basis, individual irrigation zones are represented as graph nodes, while spatial adjacency, hydraulic sharing, and microclimatic similarity are encoded as heterogeneous graph edges. Pumps, main pipelines, and nutrient-solution tanks are introduced as global resource nodes to participate in graph message propagation. Multi-relational graph attention and temporal fusion are then performed to obtain a reliability-aware state representation for each irrigation zone, together with a global greenhouse state.

In the second stage, the resulting zone-level and global states are fed into an action-conditioned digital twin together with candidate control actions, including irrigation volume, irrigation duration, and nutrient-solution *EC*. An action–state interaction network is employed to predict short-term responses such as future root-zone water content, *EC*, drainage, and energy consumption. These responses are further propagated to estimate marketable yield, quality grade, resource cost, and net profit, thereby producing action-specific marginal economic values and associated risk information. In the third stage, the reliability-aware zone states, neighborhood summaries, global resource states, and digital-twin outputs are jointly provided to Graph-MAPPO. Zone-specific Actors generate continuous local control actions, whereas a graph-structured centralized Critic jointly evaluates the long-term value of all zone states, actions, and shared-resource interactions during training. The resulting policy actions are subsequently processed by a capacity-, safety-, and fairness-constrained projection layer, in which pump flow, main-pipeline capacity, nutrient-solution inventory, and safe root-zone fertigation ranges are jointly enforced. The corrected feasible actions are finally transmitted to the corresponding valves and fertigation equipment. The resulting operating outcomes generate new sensor observations that are fed back into

the state-representation module, thereby establishing a closed-loop multi-zone greenhouse fertigation-control process oriented toward long-term net-profit optimization.

3.3.2. Module I: Reliability-Aware Multi-Relational Spatiotemporal Graph State Representation

The reliability-aware multi-relational spatiotemporal graph state representation module is designed to transform heterogeneous multi-source observations into zone-level latent states with both sensor-fault suppression capability and explicit spatial–hydraulic relational information. As shown in Figure 2, for each control window, environmental, root-zone, and hydraulic-network/equipment sequences are organized as input tensors of dimensions $L \times C_e$, $L \times C_r$, and $L \times C_d$, respectively, where L denotes the historical window length, and C_e, C_r, and C_d denote the corresponding modality-specific channel dimensions. Three temporal branches adopt three-stage one-dimensional temporal encoders. The first stage maps the original channels to C_1, the second stage expands the representation to C_2, and the third stage compresses the features into a unified latent space of dimension D. The convolutional kernel widths are denoted by K_1, K_2, and K_3, respectively, while residual connections are incorporated to preserve both short-term abrupt changes and long-term temporal trends. RGB canopy images with dimensions $H \times W \times 3$ are fed into a lightweight visual encoder. After four stages of hierarchical feature extraction, a high-level feature tensor of spatial resolution $H/32 \times W/32$ and channel dimension C_v is obtained, which is subsequently transformed into a D-dimensional visual representation through global average pooling (GAP) and linear projection. Economic variables are mapped into the same D-dimensional latent space through a two-layer fully connected network, ensuring that heterogeneous modalities can subsequently be fused within a unified feature scale.

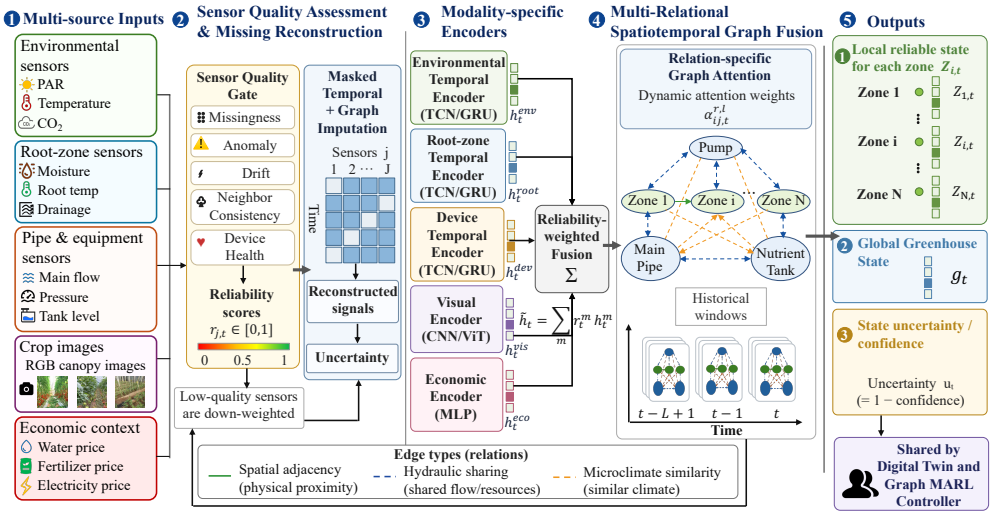


Figure 2. Architecture of the reliability-aware multi-relational spatiotemporal state representation framework. Indices i, j, t, m, r, l denote the target zone, a sensor or neighboring zone according to context, time, modality, relation type, and graph layer; N, J, and L denote the numbers of zones and sensors and the historical window length. The sensor-quality score $r_{j,t}$ lies in $[0, 1]$. Features h_t^{env}, h_t^{root}, h_t^{dev}, h_t^{vis}, and h_t^{eco} encode environmental, root-zone, equipment, visual, and economic inputs; r_t^m is a modality reliability weight and $\tilde{h}_t$ is the fused feature, with zone indices suppressed in the diagram. Relation-specific attention weights are $a_{ij,t}^{r,l}$. Outputs $z_{i,t}$ and g_t represent local zone and global shared-resource states; the scalar u_t denotes state uncertainty. Green-solid, blue-dashed, and orange-dashed edges denote spatial adjacency, hydraulic sharing, and microclimate similarity. PAR: photosynthetically active radiation; RGB: red–green–blue; TCN: temporal convolutional network; GRU: gated recurrent unit; CNN: convolutional neural network; ViT: vision transformer; MLP: multilayer perceptron; MARL: multi-agent reinforcement learning. Table 2 summarizes the principal model variables.

Table 2. Summary of primary notations, mathematical symbols, tensor dimensions, and physical descriptions.

Notation	Dimension/Type	Physical/Mathematical Description
N, T, L	Scalars	Number of greenhouse zones ($N = 8$), total time steps, and historical observation window ($L = 48$)
$x_{i,t}^m$, $h_{i,t}^m$	$\mathbb{R}^{L\times C_m}$, $\mathbb{R}^D$	Raw input sequence and temporal/visual encoded feature for modality m in zone i at time t
$q_{i,t}^m$	Quality-descriptor vector	Missingness, temporal change, residual, neighbor consistency, and equipment-state descriptors for modality m
$T_{i,t}^m$	Temporal-context representation	Local historical information supplied to the dual-path reconstruction function ψ
$\beta_{ij,t}^m$	Scalar weight	Contribution of neighboring zone j to the reconstruction of modality m in zone i
$a_{ij,t}^{r,l}$	Scalar attention weight	Contribution of neighbor j under relation r at graph layer l and time t
$\lambda_{i,t}^{(l)}$	Feature-wise gate	Residual mixing gate between the updated and preceding node representations
$W_0^{(l)}, W_r^{(l)}, W_g^{(l)}$	Linear-map matrices	Map self, relation-specific neighbor, and global-resource features to the graph-layer output space
$\phi_m(q_{i,t}^m)$	$[0,1]^D$, $D = 128$	Feature-space reliability gate matching the D-dimensional encoded and reconstructed features
$c_{i,t}^m$	$\mathbb{R}^D$	Dual-path reconstructed compensation state derived from temporal context and graph neighbors
$r_{i,t}^m$, $z_{i,t}$	$\mathbb{R}^D$, $\mathbb{R}^{D_s}$	Reliability-modulated modality feature and final fused latent node representation of zone i
g_t	$\mathbb{R}^{D_g}$	Global shared resource node state representing pump capacity, main pipeline flow, and tank volume
$\mathcal{R}$, $\mathcal{N}_i^r$	Sets	Graph edge relation types (spatial, hydraulic, microclimatic) and neighborhood of node i under relation r
$a_{i,t}$	$\mathbb{R}^{D_a}$	Candidate fertigation action: irrigation volume (V_{irr}), duration (T_{irr}), target EC (EC_{target})
$H_{i,t}^{ac}$	$\mathbb{R}^{D_h}$	Action-conditioned cross-attention state representation coupling current state with candidate action
$u_{i,t+\tau}$	$\mathbb{R}^{C_s}$	Short-term multi-step predicted trajectory (water content, pore water EC, pH, drainage volume)
h_i^{long}	$\mathbb{R}^{D_l}$	Long-term stage representation predicting marketable yield, tiered quality grade, and net return
$\mathcal{U}_{i,t}(a)$	Scalar	Risk-adjusted economic utility penalizing high-variance return via risk-sensitivity coefficient κ
C_t	Convex set	Convex feasible action set bounded by pump flow limits, pipe pressure, and root-zone safety thresholds
a_t^*	$\mathbb{R}^{N\times D_a}$	Safe projected fertigation action vector satisfying all hydraulic and agronomic hard constraints
$s_{i,t}$, $\mathcal{L}_{fair}$	Scalar, Scalar	Normalized zone service ratio and variance-based fairness regularization corresponding to Jain index

To prevent unreliable sensor measurements from directly contaminating node states, each sensor channel is simultaneously processed by a reliability-estimation network. The network consists of a two-layer multilayer perceptron (MLP) and maps the quality descriptors derived from the missing-value mask, local rate of change, historical residual, neighborhood consistency, and equipment operating state to a D-dimensional feature-space gate ($D = 128$). The raw sensor-channel dimension is C_m, while gating is applied to the encoded and reconstructed features in the common latent space. Let $h_{i,t}^m$ denote the encoded feature of modality m in zone i at time t, and let $q_{i,t}^m$ denote the corresponding quality-description vector. The reliability-modulated representation is defined as

$$r_{i,t}^m = \phi_m(q_{i,t}^m) \odot h_{i,t}^m + \left[1 - \phi_m(q_{i,t}^m)\right] \odot c_{i,t}^m, \tag{7}$$

where $\phi_m(\cdot)$ denotes the quality-gating function, and $c_{i,t}^m$ represents a compensation feature inferred from historical context and graph-neighborhood information. In plain terms, Equation (7) acts as an intelligent softswitch: when a sensor probe operates normally ($\phi_m \approx 1$), the model directly passes the genuine physical observation; conversely, when a sensor drifts, gets stuck, or drops communication ($\phi_m \approx 0$), the model seamlessly suppresses the corrupted measurement and substitutes it with the reconstructed state $c_{i,t}^m$. For miss-

ing or anomalous observations, the compensation state is obtained through a dual-path reconstruction mechanism:

$$c_{i,t}^{m} = \psi\left(T_{i,t}^{m}, \sum_{j \in \mathcal{N}_i} \beta_{ij,t}^{m} h_{j,t}^{m} \right), \tag{8}$$

where $T_{i,t}^{m}$ denotes the local temporal context, and $\beta_{ij,t}^{m}$ represents the dynamic contribution of neighboring node j to reconstruction of the current missing state. Consequently, missing-state reconstruction does not rely on global-mean imputation but instead jointly exploits the historical evolution of the same irrigation zone and synchronized information from related neighboring zones. After modality-level reliability correction, irrigation zones are represented as graph nodes, and three adjacency subgraphs are independently constructed to describe spatial relationships, hydraulic relationships, and microclimatic similarities. In the l-th multi-relational graph-attention layer, independent parameter matrices are used for different relation types, and the node update is formulated as

$$z_{i,t}^{(l+1)} = \sigma\left(W_0^{(l)} z_{i,t}^{(l)} + \sum_{r \in \mathcal{R}} \sum_{j \in \mathcal{N}_i^r} a_{ij,t}^{r,l} W_r^{(l)} z_{j,t}^{(l)} + W_g^{(l)} g_t \right), \tag{9}$$

where $\mathcal{R}$ denotes the set of relation types, and g_t represents the global resource-node state composed of the pump, main pipeline, and nutrient-solution tank. The graph network adopts a three-stage stacked architecture. Each stage contains multi-head relational attention, residual connections, Layer Normalization, and feed-forward transformation. The node latent representation is successively mapped from D to D_g, D_g, and D_s, ultimately producing a fixed-dimensional reliability-aware zone state. To mitigate over-smoothing caused by deep graph message propagation, a gated residual mechanism is adopted:

$$\hat{z}_{i,t}^{(l+1)} = \lambda_{i,t}^{(l)} \odot z_{i,t}^{(l+1)} + \left(1 - \lambda_{i,t}^{(l)}\right) \odot z_{i,t}^{(l)}. \tag{10}$$

The proposed module provides an explicit stability advantage. Suppose that an anomalous sensor introduces a perturbation Δh: the corresponding reliability-gating value is ρ, and the subsequent fusion mapping satisfies the Lipschitz continuity condition with constant L_f. The perturbation of the final state representation then satisfies

$$|\Delta z_{i,t}| \leq L_f \rho |\Delta h|. \tag{11}$$

As sensor reliability decreases, ρ is correspondingly reduced, thereby strictly suppressing the influence of anomalous observations on the final state representation. Compared with direct feature concatenation without reliability gating, the upper perturbation bound is reduced from $L_f|\Delta h|$ to $L_f \rho |\Delta h|$. In addition, multi-relational message aggregation is permutation invariant with respect to neighboring nodes; therefore, changes in zone indexing do not alter the physical semantics of the graph representation. This design enables temporal-response continuity, spatial information propagation, and shared hydraulic constraints to be represented simultaneously. Stable state representations can consequently be maintained under sensor drift, localized missing observations, and pronounced inter-zone environmental heterogeneity, thereby providing more reliable decision information for the subsequent action-conditioned digital twin and multi-agent controller.

3.3.3. Module II: Action-Conditioned Fertigation–Yield–Quality–Profit Digital Twin

The action-conditioned fertigation–yield–quality–profit digital twin is designed to learn the nonlinear mapping among current zone states, candidate control actions, and future agronomic and economic outcomes.

As shown in Figure 3, its inputs consist of the reliability-aware zone state $z_{i,t} \in \mathbb{R}^{D_s}$ generated by the preceding module, the global greenhouse state $g_t \in \mathbb{R}^{D_g}$, and the candidate

fertigation action $a_{i,t} \in \mathbb{R}^{D_a}$. The state branch adopts a three-layer fully connected encoder, in which the feature width is sequentially mapped from D_s to D_1, D_2, and a unified latent dimension D_h. The action branch employs a two-layer multilayer perceptron to map irrigation volume, irrigation duration, target nutrient solution *EC* into a D_h-dimensional action embedding. To overcome the limited representational capability of simple concatenation for characterizing different responses induced by alternative actions under the same system state, state and action features are first coupled through action-conditioned cross-attention. Let the state-query, action-key, and action-value matrices be denoted by Q_s, K_a, and V_a, respectively. The resulting interaction representation is formulated as

$$H_{i,t}^{ac} = \mathrm{LN}\left[H_{i,t}^{s} + \mathrm{softmax}\left(\frac{Q_s K_a^{\top}}{\sqrt{D_h}}\right) V_a\right], \tag{12}$$

where $\mathrm{LN}(\cdot)$ denotes Layer Normalization. The interaction module consists of L_{ac} stacked layers, each containing multi-head cross-attention, residual connections, and a feed-forward network of width D_f. This architecture enables candidate control actions to dynamically modulate latent physiological-state dimensions. The fused representation is subsequently passed to a dual-timescale digital-twin architecture. The short-term response branch contains an L_s-layer temporal decoder, which expands the latent state into a response tensor with horizon length H_s and channel number C_s. This branch jointly predicts root-zone water content, *EC*, pH, drainage, hydraulic flow, and energy consumption over multiple future control windows. Its recursive transition is expressed as

$$u_{i,t+\tau} = \mathcal{F}_s\left(u_{i,t+\tau-1}, H_{i,t}^{ac}, a_{i,t}\right), \qquad \tau = 1, \ldots, H_s. \tag{13}$$

To connect short-term fertigation responses with long-term production outcomes, attention pooling is applied to the predicted short-term trajectory to obtain a stage-level physiological representation, which is subsequently passed to the long-term outcome branch. This branch consists of an L_l-layer shared backbone followed by multiple task-specific heads for predicting marketable yield, continuous quality attributes, grade probabilities, and economic outcomes:

$$h_i^{long} = \mathcal{F}_l\left(\sum_{\tau=1}^{H_s} \omega_\tau u_{i,t+\tau}, z_{i,t}, g_t\right), \tag{14}$$

where the different prediction heads share the lower-level latent representation but employ independent parameters for regression or classification, thereby preventing any individual task from excessively dominating the shared representation space.

The interaction and feature transfer between the short-term response branch ($u_{i,t+\tau}$) and the long-term production outcome branch (h_i^{long}) follow a physiologically grounded hierarchical information flow. Specifically, the short-term branch recursively predicts transient state transitions over horizon H_s (e.g., immediate substrate moisture surge, slow drainage infiltration, and localized EC fluctuations). An attention-pooling operator dynamically weights these short-term trajectories based on their physiological relevance during specific phenological stages (e.g., prioritizing root-zone moisture deficits during flowering over vegetative stages), condensing them into a stage-level crop physiological summary vector that is injected into the long-term branch. During end-to-end backpropagation, gradients originating from long-term yield and quality classification losses ($\mathcal{L}_{long}$) directly propagate back through the attention weights into the short-term transition encoder $\mathcal{F}_s$. Simultaneously, the physical process consistency loss $\mathcal{L}_{cons}$ penalizes unphysical conservation violations (such as water accumulation exceeding applied irrigation minus drainage and evapotranspiration). This dual gradient propagation ensures that short-term representations are regularized not only by transient sensor curve fitting but also by their ultimate agronomic consequences on fruit grade and harvest volume.

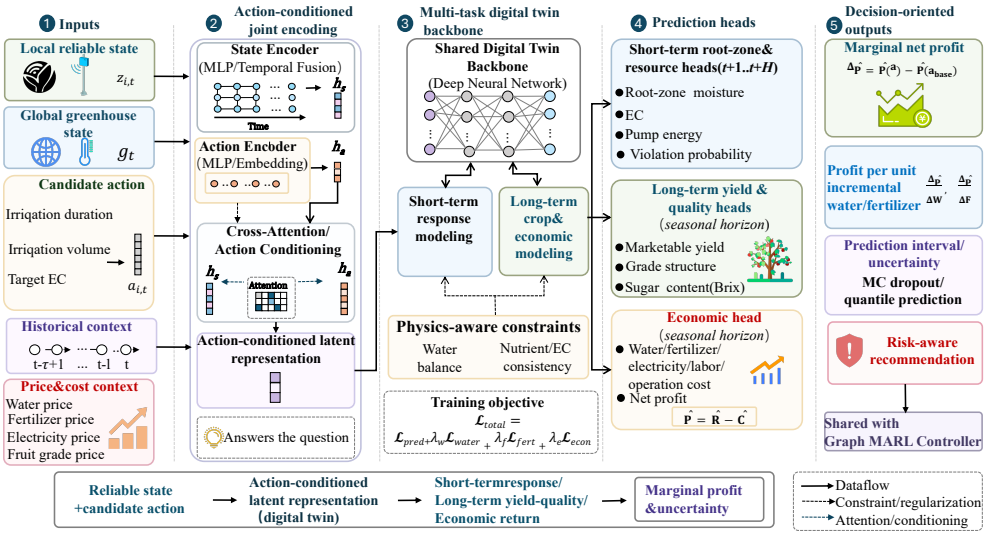


Figure 3. Architecture of the action-conditioned fertigation–yield–quality–profit digital twin. Inputs $z_{i,t}$ and g_t are local zone and global shared-resource states. The candidate action $a_{i,t}$ comprises irrigation volume, duration, and target electrical conductivity (EC). The state and action embeddings h_s and h_a form the action-conditioned representation $H^{ac}_{i,t}$. The historical span τ corresponds to the observation window L, and H denotes the short-term prediction horizon H_s. Predicted net profit $\hat{P} = \hat{R} - \hat{C}$ equals predicted revenue minus cost and corresponds to the economic return Π in the text. The difference $\Delta\hat{P} = \hat{P}(a) - \hat{P}(a_{base})$ is measured against reference action a_{base}; ΔW and ΔF are incremental water and fertilizer inputs. The ratios $\Delta\hat{P}/\Delta W$ and $\Delta\hat{P}/\Delta F$ express incremental resource returns. In the schematic training objective, $\mathcal{L}_{pred}$ denotes prediction loss, and $\mathcal{L}_{water}$, $\mathcal{L}_{fert}$, and $\mathcal{L}_{econ}$ denote water, nutrient, and economic consistency losses with weights $\lambda_w, \lambda_f, \lambda_e$. MLP: multilayer perceptron; MC: Monte Carlo; MARL: multi-agent reinforcement learning. Solid arrows indicate data flow, black dashed arrows indicate constraints or regularization, and blue dashed arrows indicate attention or conditioning.

In plain terms, Equation (12) evaluates what would happen if a specific fertigation dose were applied under current microclimatic and substrate conditions, allowing the digital twin to simulate counterfactual crop responses without physically risking the plants. Equation (15) then translates these physical predictions into a risk-adjusted economic value. In practical commercial operations, high-reward fertigation strategies often carry severe downside risks (e.g., extreme electrical conductivity that might elevate fruit sugar content but risks salt-induced root necrosis under sudden hot weather). The risk penalty term $-\kappa\sqrt{\mathrm{Var}[\Pi_{i,t}]}$ explicitly discounts volatile actions, ensuring that the controller selects robustly beneficial actions. The economic branch does not estimate profit independently. Instead, a differentiable economic mapping is constructed from predicted yield, grade probabilities, and resource consumption. The risk-adjusted utility of a candidate control action is defined as

$$\mathcal{U}_{i,t}(a) = \mathbb{E}\big[\Pi_{i,t}(a)\big] - \kappa\sqrt{\mathrm{Var}\big[\Pi_{i,t}(a)\big]}, \tag{15}$$

where $\Pi_{i,t}$ denotes the economic return jointly determined by product-sales revenue and water, fertilizer, and electricity costs, while κ denotes the risk-sensitivity coefficient. To further incorporate modern environmental sustainability criteria and time-of-use tariffs, the net economic return $\Pi_{i,t}(a)$ is formulated as:

$$\begin{aligned}\Pi_{i,t}(a) = &\sum_{c\in\{A,B,C\}} P_c \cdot Y_{i,c}(a) \\ &- \left[C_{water}V^w_{i,t} + C_{fert}M^f_{i,t} + p_{elec}(t)E_{i,t}(a)\right] \\ &- \lambda_{carb}\left(c^{GHG}_f M^f_{i,t} + c^{GHG}_e E_{i,t}\right),\end{aligned} \tag{16}$$

where P_c and $Y_{i,c}$ denote the market price and predicted yield for commercial grade c; $V_{i,t}^{w}$ and $M_{i,t}^{f}$ represent water volume (m^3) and fertilizer mass (kg); $p_{elec}(t)$ denotes the dynamic time-of-use electricity price at hour t; $E_{i,t}$ is the pumping electricity consumption (kWh); and c_f^{GHG} and c_e^{GHG} denote the carbon emission factors for synthetic fertilizer manufacturing (2.15 kg CO_2e/kg) and grid power (0.68 kg CO_2e/kWh), monetized by carbon price coefficient λ_{carb}. This enables the digital twin to penalize high-carbon inputs and shift high-energy pumping away from expensive peak-tariff periods. Therefore, candidate actions are evaluated not only according to their expected economic return but also according to the associated uncertainty, allowing aggressive actions with potentially high returns but excessive uncertainty to be suppressed. To enforce consistency among different prediction tasks, the overall training objective combines supervised prediction loss, process-consistency loss, economic-consistency loss, and uncertainty-calibration loss:

$$\mathcal{L}_{DT} = \lambda_s \mathcal{L}_{short} + \lambda_l \mathcal{L}_{long} + \lambda_c \mathcal{L}_{cons} + \lambda_u \mathcal{L}_{unc}, \tag{17}$$

where $\mathcal{L}_{short}$ and $\mathcal{L}_{long}$ constrain short-term and long-term predictions, respectively; $\mathcal{L}_{cons}$ enforces logical consistency among fertigation input, state variation, resource consumption, and economic output; and $\mathcal{L}_{unc}$ calibrates predictive uncertainty and interval estimates. The proposed structure also provides an interpretable bound on economic-risk propagation. Let ε_{Π} denote the prediction error of the action value, and assume that the economic mapping satisfies Lipschitz continuity with respect to short-term system responses and long-term yield-related outcomes, with corresponding constants L_s and L_y. The resulting value-prediction error satisfies

$$|\varepsilon_{\Pi}| \leq L_s |\varepsilon_{short}| + L_y \left|\varepsilon_{long}\right| + \varepsilon_c, \tag{18}$$

where ε_c represents the error associated with cost and price mapping. Therefore, when short-term root-zone response errors and long-term yield/quality prediction errors are jointly constrained, the resulting economic-value prediction error also remains bounded. Compared with models that only estimate water demand or crop yield, the proposed digital twin directly evaluates whether a candidate fertigation action is economically worthwhile and jointly transforms state variation, agricultural output, resource cost, and uncertainty into action-value information that can be directly exploited by the subsequent Graph-MAPPO controller. This design improves both the economic efficiency and operational safety of fertigation decisions under inter-zone resource competition.

3.3.4. Module III: Capacity-, Safety-, and Fairness-Constrained Graph Multi-Agent Fertigation Control

The capacity-, safety-, and fairness-constrained graph multi-agent fertigation-control module jointly receives the reliability-aware zone states produced by Module I, the action-value and risk estimates generated by Module II, and the current shared-resource state. Each irrigation zone is modeled as an independent agent, and coordinated decision-making is achieved using Graph-MAPPO under a centralized-training and decentralized-execution paradigm.

As shown in Figure 4, for zone i, the input feature is composed of the local state $z_{i,t} \in \mathbb{R}^{D_s}$, neighborhood summary $n_{i,t} \in \mathbb{R}^{D_n}$, digital-twin value feature $d_{i,t} \in \mathbb{R}^{D_d}$, and global resource broadcast $g_t \in \mathbb{R}^{D_g}$. The Actor adopts a three-layer fully connected policy network with an input width of $D_s + D_n + D_d + D_g$, which is sequentially projected into hidden spaces of dimensions D_{a1}, D_{a2}, and D_{a3}. Layer Normalization, nonlinear activation, and residual connections are used in each layer. The final mean and variance heads parameterize a continuous action distribution over C_a control channels, including

irrigation volume, irrigation duration, and nutrient-solution *EC*. The policy distribution is expressed as

$$\pi_{\theta_i}(a_{i,t}|o_{i,t}) = \mathcal{N}\left(\mu_{\theta_i}(o_{i,t}), \mathrm{diag}\left[\sigma^2_{\theta_i}(o_{i,t})\right]\right). \tag{19}$$

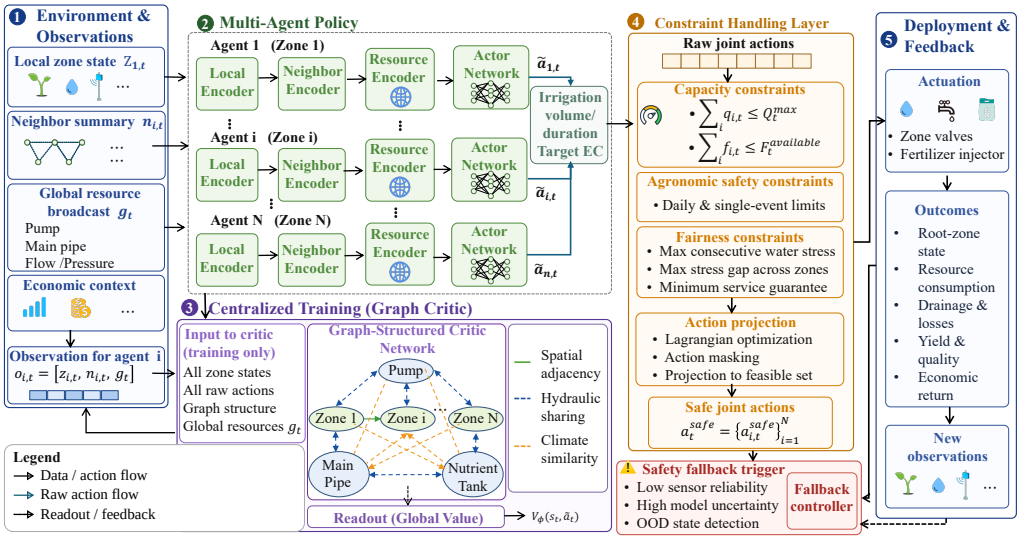


Figure 4. Architecture of the capacity-, safety-, and fairness-constrained graph multi-agent fertigation controller. For zone i at time t, $z_{i,t}$ is the local state, $n_{i,t}$ the neighborhood summary, g_t the global resource broadcast, and $o_{i,t}$ the agent observation. The Actor outputs raw action $\tilde{a}_{i,t}$; N is the number of zones, and $\tilde{a}_t$ collects their joint actions. Water-flow demand $q_{i,t}$ and fertilizer demand $f_{i,t}$ are bounded in aggregate by pump capacity Q_t^{max} and available fertilizer supply $F_t^{available}$. The projected joint action $a_t^{safe} = \{a_{i,t}^{safe}\}_{i=1}^N$ corresponds to a_t^* in the text. The Critic readout $V_\phi(s_t, \tilde{a}_t)$ denotes the global value estimate with parameters ϕ, centralized state s_t, and joint-action input. Within the graph, green-solid, blue-dashed, and orange-dashed edges denote spatial, hydraulic, and microclimate relations. Solid black arrows indicate data/action flow, solid blue arrows carry raw actor actions, and dashed black arrows indicate readout or feedback links. EC: electrical conductivity; OOD: out of distribution. Table 2 provides the principal variable definitions.

The centralized Critic consists of an L_c-layer multi-relational graph-attention network followed by a two-layer value-mapping network. Zone-node features are first projected from D_s to D_c, after which relational message passing is performed over spatial, hydraulic, and resource edges. Graph-level attention pooling is then applied to obtain a global representation $h_t^G \in \mathbb{R}^{D_G}$, which is fused with the joint-action embedding and shared-equipment state to estimate the global value:

$$V_\phi(s_t) = f_v\left(\mathrm{Pool}\left\{h_{1,t}^{(L_c)}, \ldots, h_{N,t}^{(L_c)}\right\}, e_{a,t}, g_t\right). \tag{20}$$

To simultaneously optimize economic returns, resource feasibility, and agronomic safety, policy learning is formulated as a constrained expected-return maximization problem. Let $J_R(\theta)$ denote the long-term net-return objective and $J_k(\theta)$ denote the cumulative risk associated with the k-th constraint. A Lagrangian formulation is adopted:

$$\mathcal{J}(\theta, \lambda) = J_R(\theta) - \sum_{k=1}^{K} \lambda_k [J_k(\theta) - d_k], \qquad \lambda_k \geq 0, \tag{21}$$

where d_k denotes the maximum allowable risk level associated with pump capacity, main-pipeline flow, root-zone water status, *EC*, nutrient-solution inventory, or other safety-related indicators. The Actor is optimized using a constrained PPO clipped objective. The marginal economic value and risk confidence predicted by the digital twin are jointly incorporated into

the advantage estimation, preventing actions with high expected returns but excessive risk from receiving disproportionately large policy-update weights. To regulate competition for shared resources, a normalized service ratio $s_{i,t}$ is introduced to characterize the relationship between the actual water and fertilizer resources allocated to each zone and its minimum agronomic requirement. A variance-based fairness regularization term is defined as

$$\mathcal{L}_{fair} = \frac{1}{N}\sum_{i=1}^{N}\left(s_{i,t} - \frac{1}{N}\sum_{j=1}^{N} s_{j,t}\right)^2, \tag{22}$$

thereby penalizing disparities in normalized service levels among irrigation zones. The variance-based regularization in Equation (22) is related to Jain's Fairness Index, $J(s) = (\sum_{i=1}^{N} s_{i,t})^2/(N\sum_{i=1}^{N} s_{i,t}^2)$. For a positive mean service ratio $\bar{s} = N^{-1}\sum_{i=1}^{N} s_{i,t}$ and across-zone variance $\sigma_s^2 = \mathcal{L}_{fair}$, the identity is $J(s) = \bar{s}^2/(\bar{s}^2 + \mathcal{L}_{fair})$. At fixed $\bar{s} > 0$, reducing $\mathcal{L}_{fair}$ increases Jain's index, which reaches 1 when all service ratios are equal. When both the mean and variance change during policy optimization, Jain's index depends on their ratio $\mathcal{L}_{fair}/\bar{s}^2$. Fairness regularization therefore promotes balanced relative service. With service ratios normalized by minimum agronomic requirements, minimum service corresponds to $s_{i,t} \geq 1$ for each zone and is a separate agronomic safety requirement. Both the Jain index and stress-violation rate are reported to characterize allocation equity and crop-state safety.

In plain terms, Equations (21)–(23) establish a fail-safe hierarchy of control: Equation (21) uses adaptive penalty multipliers to balance economic yield against risk boundaries; Equation (22) penalizes disparities in normalized zone service levels; and Equation (23) acts as an instantaneous physical clamp. Even if the neural network proposes an overly aggressive valve actuation that would exceed main-pipeline pressure or pump power, Equation (23) projects the action onto the nearest physically executable, safe boundary before commands reach the field hardware. Once the minimum safety requirements of all zones have been satisfied, the remaining water and fertilizer resources can still be preferentially allocated to zones with higher marginal economic returns estimated by Module II. The joint actions generated by the policy are further processed through a feasible-set projection mechanism. Each zone contributes three action components, so the eight-zone joint action has shape 8*imes*3 (24 scalar variables); the projection preserves this joint shape. The current feasible action set C_t is first constructed according to equipment-capacity limits and agronomic safety boundaries, after which the safe action closest to the original policy output is obtained as

$$a_t^* = \arg\min_{u \in C_t} |u - \tilde{a}_t|_2^2. \tag{23}$$

Because C_t is defined by linear capacity constraints and bounded agronomic intervals, it can be treated as a nonempty closed convex set, for which the Euclidean projection is unique. According to the projection theorem for closed convex sets, for any unconstrained action $\tilde{a}_t$, there exists a unique $a_t^* \in C_t$ satisfying

$$\langle \tilde{a}_t - a_t^*, u - a_t^* \rangle \leq 0, \qquad \forall u \in C_t. \tag{24}$$

Accordingly, the safety-projection mechanism minimizes the deviation from the original policy decision while satisfying all hard constraints and guarantees that the final actions transmitted to pumps, valves, and fertigation equipment remain within the executable region. Compared with independent reinforcement learning or centralized control driven solely by total economic return, the proposed module explicitly models hydraulic competition and long-term coordination among multiple irrigation zones through the graph-structured Critic. Equipment overload is avoided through capacity constraints, root-zone stress is suppressed through agronomic safety constraints, and persistent sacrifice of

economically lower-value zones is prevented through the minimum-service fairness mechanism. Consequently, limited water and fertilizer resources can be allocated more stably across multiple zones according to the principle of agronomic safety first and marginal economic value driven optimization.

4. Results and Discussion

4.1. Experimental Configuration

4.1.1. Hardware and Software Platform

All experiments were conducted on a high-performance deep learning workstation equipped with an Intel Xeon Gold 6430 processor operating at 2.10 GHz, 128 GB of DDR5 memory, and an NVIDIA RTX 4090 graphics processing unit (GPU) with 24 GB of video memory. A 2 TB NVMe solid-state drive was used to store multi-source sensor observations, crop images, and model-training outputs. The GPU was primarily used for parallel training of graph neural networks, temporal models, digital-twin components, and multi-agent reinforcement learning algorithms, whereas the central processing unit (CPU) was responsible for data preprocessing, environment simulation, reinforcement-learning interactions, and evaluation-metric computation. To improve training efficiency on large-scale multi-zone temporal data and dynamic graph structures, GPU acceleration and multithreaded data loading were enabled throughout the experiments. All deep-learning and reinforcement-learning baseline methods were executed under identical hardware conditions to minimize potential performance variations caused by differences in computing platforms.

The software environment was based on the 64-bit Ubuntu 22.04 LTS operating system, with Python 3.11 used as the primary programming language. PyTorch 2.4 was adopted as the deep-learning framework, while PyTorch Geometric 2.6 was used to implement GCNs, GATs, and multi-relational spatiotemporal graph architectures. CUDA 12.4 and cuDNN 9.1 were employed for GPU acceleration. For reproducible implementation, a compatible auxiliary software configuration is specified as follows: NumPy 1.26.4, Pandas 2.2.2, and SciPy 1.14.1 are specified for sensor-data processing and statistical computation, while Scikit-learn 1.5.2 is specified for data normalization, conventional machine-learning models, and evaluation-metric calculation. OpenCV 4.10.0 and torchvision 0.19.0 are specified for crop-canopy image preprocessing and visual feature extraction. The reference configuration for the multi-agent reinforcement-learning environment uses Gymnasium 0.29.1 and PettingZoo 1.24.3, whereas MAPPO, SAC, and PPO were implemented in PyTorch. Matplotlib 3.9.2 is specified for experimental-result visualization. In this reference configuration, torchvision 0.19.0 is paired with PyTorch 2.4.0. All models were trained and evaluated under unified software dependencies, random-seed settings, and dataset partitions to ensure reproducibility and fairness across different methods.

For dataset partitioning, samples were divided according to continuous production time and production cycles rather than by randomly shuffling adjacent temporal windows, thereby preventing future-information leakage. All samples were partitioned into training, validation, and independent test sets at a ratio of 70%:15%:15%. The training set was used for model-parameter optimization, the validation set was used for hyperparameter selection and early-stopping decisions, and the test set was used exclusively for final performance evaluation. To further reduce the randomness associated with a single data partition, five-fold cross-validation was performed on the training data. In each fold, four subsets were used for training and one subset was used for validation, and the mean and standard deviation across five experimental runs were reported. Continuous sensor observations were aggregated using a basic control interval of 30 min. The historical input sequence length was set to 48 windows, corresponding to the previous 24 h of observations used to predict subsequent responses. The image-feature dimensionality was set to 256, and the

hidden-feature dimension of each zone node was set to 128. The multi-relational GAT contained three layers, with four attention heads in each layer, and the dropout rate was set to 0.2. Model parameters were optimized using AdamW with an initial learning rate of 1×10^{-4}, a weight-decay coefficient of 1×10^{-5}, and a batch size of 64. The maximum number of training epochs was set to 300, and early stopping was activated when the validation loss failed to decrease for 20 consecutive epochs.

For reinforcement learning, SensorGraph-FertiMARL employed Graph-MAPPO as the core policy-optimization algorithm. The three Actor hidden layers have widths of 256, 256, and 128, respectively, and the Critic hidden width is 256. The discount factor γ was set to 0.99, the generalized advantage estimation parameter λ was set to 0.95, the PPO policy clipping coefficient ϵ was set to 0.2, the entropy regularization coefficient was set to 0.01, and the value-function loss coefficient was set to 0.5. The initial learning rates of the Actor and Critic were set to 3×10^{-4} and 1×10^{-3}, respectively, and the number of policy-update epochs per iteration was set to 10. The maximum gradient norm was clipped at 0.5 to improve training stability, and the rollout trajectory length was set to 2048 interaction steps. The initial threshold of the sensor-quality gate was set to 0.6, and observations with reliability scores below this threshold were assigned reduced weights during state fusion. The random dropping probability for non-critical graph edges during graph augmentation was set to 0.1, while the randomly simulated sensor-missing ratio ranged from 5% to 20%. All deep-learning and reinforcement-learning experiments were independently repeated using five different random seeds, and the mean and standard deviation were reported as the final results to improve statistical reliability and result stability. To maximize reproducibility, the complete software pipeline, sensor preprocessing scripts, and PettingZoo multi-zone greenhouse gym environment will be made publicly available upon acceptance. Furthermore, the exact structural specifications, layer dimensions, activation functions, and parameter counts for each module of the proposed framework are comprehensively detailed in Table 3.

Table 3. Detailed architectural parameters, layer dimensions, and tensor shapes of SensorGraph-FertiMARL.

Module	Sub-Component	Layer Type & Operations	Input → Output Dimensions	Parameters
Module I	Temporal Encoders	3-stage 1D-Conv, ResNet, LayerNorm, ReLU	$(L, C_m) \rightarrow (L, 64) \rightarrow (L, 128) \rightarrow 128$	148.6 K
	Visual Backbone	4-stage Hierarchical CNN, GAP, Linear	$(256, 256, 3) \rightarrow (8, 8, 256) \rightarrow 128$	524.3 K
	Reliability Gating	2-layer MLP, Sigmoid Gate, Dual-Path Imputer	$128 \rightarrow 64 \rightarrow 128$, element-wise $\odot$	49.8 K
	Multi-Relational GAT	3-layer Relational GAT (4 heads), Dropout 0.2	$N \times 128 \rightarrow N \times 128 \rightarrow N \times 128$	296.4 K
Module II	Action Embedding	2-layer MLP, LayerNorm, GELU	$3 \rightarrow 64 \rightarrow 128$	8.7 K
	State–Action Cross-Attn	2-layer Cross-Attention, feed-forward network (FFN; $D_f = 256$)	$(128, 128) \rightarrow 128$	132.1 K
	Short-Term Decoder	3-layer gated recurrent unit (GRU) decoder ($H_s = 12$)	$128 \rightarrow 128 \rightarrow 6$ channels ($H_s \times 6$)	198.5 K
	Long-Term Multi-Heads	Attn-Pooling + 4 Task-Specific Regr./Class. MLPs	$128 \rightarrow 64 \rightarrow \{1, 3, 1, 1\}$	86.2 K
Module III	Decentralized Actor	3-layer MLP $(256, 256, 128)$, Gaussian Head	$384 \rightarrow 256 \rightarrow 256 \rightarrow 128 \rightarrow 2 \times 3$	231.8 K
	Centralized Graph Critic	2-layer Multi-Relational GAT + Global MLP	$N \times 128 + 3 \rightarrow 256 \rightarrow 1$	345.2 K
	Safety Projection Layer	Differentiable convex quadratic-programming (QP) solver	$(8, 3) \rightarrow (8, 3)$ Joint projection; 24 scalars	Non-param.

4.1.2. Baseline Models and Evaluation Metrics

Several representative fertigation-control methods were selected as baseline models, including Uniform-Timer [57], Uniform-Threshold [58], Zone-Threshold [59], Proportional Allocation [60], Independent MPC [61], Centralized Economic MPC [62], LSTM-Demand + Greedy [63], GNN-Demand + Proportional [64], Independent SAC/PPO [65,66], Centralized SAC [65], MAPPO/MADDPG [67,68], and Graph-MAPPO-Yield [67]. Uniform-Timer applies identical timer-based and fixed-volume irrigation schedules together with

a fixed nutrient solution *EC* throughout the greenhouse. Its principal advantage lies in its simple control logic and ease of deployment. Uniform-Threshold activates fertigation uniformly according to water-content and *EC* thresholds obtained from representative sensors, allowing basic state-dependent feedback control. Zone-Threshold independently defines water-content and *EC* thresholds for each irrigation zone and can therefore partially accommodate differences in water and nutrient demand among zones. Proportional Allocation distributes limited resources proportionally according to zone area, crop population, or water-deficit level, providing high interpretability and straightforward adaptation to resource constraints. To resolve ambiguities regarding baseline reproduction and ensure fair comparison, the mathematical formulations and operational parameters of the model predictive control baselines are explicitly defined:

$$\min_{a_{i,t:t+N_c-1}} \sum_{\tau=1}^{N_p} \left\| \hat{x}_{i,t+\tau} - x_{ref} \right\|_{W_x}^2 + \sum_{k=0}^{N_c-1} \left\| \Delta a_{i,t+k} \right\|_{W_u}^2, \tag{25}$$

where $N_p = 12$ steps (6.0 h) is the prediction horizon, $N_c = 4$ steps (2.0 h) is the control horizon, $\hat{x}$ tracks substrate water content and EC, and W_x, W_u are diagonal weighting matrices tuned via grid search. Independent MPC solves localized quadratic programs independently for each zone via the OSQP solver without exchanging flow constraints. Centralized Economic MPC incorporates all eight zones into a unified optimization program with an economic objective that directly maximizes projected crop revenue minus water, fertilizer, and time-of-use electricity costs subject to instantaneous main-pipeline flow capacity $\sum_{i=1}^{N} Q_{i,t} \leq Q_{max} = 15.0\ \mathrm{m}^3/\mathrm{h}$, solved via IPOPT.

Furthermore, all experimental baselines were standardized under identical operational protocols: (i) an identical historical observation window of $L = 48$ steps (24.0 h); (ii) a unified continuous action space $a_{i,t} = [V_{irr}, T_{irr}, EC_{target}]$ bounded respectively within $[0, 50.0\ \mathrm{L/m^2}]$, $[0, 60.0\ \mathrm{min}]$, and $[1.0, 3.5\ \mathrm{dS/m}]$; (iii) a standardized training budget of 1.0×10^6 environment interaction steps across 300 epochs; and (iv) exhaustive hyperparameter tuning via grid search across learning rates $\{1 \times 10^{-4}, 3 \times 10^{-4}, 1 \times 10^{-3}\}$ and batch sizes $\{32, 64, 128\}$. To rigorously evaluate statistical superiority beyond ranking by sample means, two-sided paired Student's *t*-tests and non-parametric Wilcoxon signed-rank tests were performed across all five random seeds, with significance designated at $p < 0.01$. Independent MPC establishes an individual model predictive controller for each irrigation zone and optimizes local fertigation actions through state tracking, thereby exploiting future-state predictions for zone-level decision making. Centralized Economic MPC uses the global greenhouse state to jointly optimize fertigation costs and economic returns and therefore directly considers the overall economic objective of the system. LSTM-Demand + Greedy first predicts zone-specific fertigation demand using an LSTM and subsequently allocates resources through a greedy strategy, enabling nonlinear temporal demand dynamics to be learned. GNN-Demand + Proportional employs a graph neural network to model inter-zone relationships and predict fertigation demand, after which resources are allocated according to a predefined proportional rule, thereby incorporating spatial and relational information among irrigation zones. Independent SAC/PPO independently trains a reinforcement-learning policy for each zone and is capable of learning nonlinear long-term control behavior. Centralized SAC combines all zone states and actions into a single-agent formulation and directly optimizes the long-term return of the entire greenhouse. MAPPO/MADDPG adopts a centralized-training and decentralized-execution mechanism to coordinate multiple zones and is therefore capable of modeling inter-zone interactions and competition for shared resources. Graph-MAPPO-Yield further introduces graph-structured representations into MAPPO while primarily optimizing crop yield and resource-use efficiency, thereby integrating inter-zone relational modeling with collaborative decision making.

Fault-detection performance was evaluated using the area under the receiver operating characteristic curve (AUROC), area under the precision–recall curve (AUPRC), F1 score (F_1), and detection delay. Missing-sensor reconstruction was assessed using MAE, RMSE, R^2, and prediction-interval coverage probability. Main-pipeline–branch flow-balance error and physical-consistency error were used to quantify satisfaction of system-level constraints. Prediction performance for water content, *EC*, pH, drainage, water use, fertilizer use, energy consumption, marketable yield, product quality, and net profit was evaluated using MAE, RMSE, MAPE, and R^2, while Macro-F1 or weighted-F1 was additionally employed for quality-grade classification. Probabilistic predictions were evaluated using the continuous ranked probability score (CRPS), PICP, MPIW, or expected calibration error (ECE). The corresponding metrics are defined as

$$\begin{aligned}
\mathrm{AUROC} &= \int_0^1 \mathrm{TPR}(\mathrm{FPR})\, d\mathrm{FPR}, \quad \mathrm{AUPRC} = \int_0^1 P(R)\, dR, \\
F_1 &= \frac{2PR}{P+R}, \quad D = \frac{1}{N_f}\sum_{i=1}^{N_f}(\hat{t}_i - t_i), \\
\mathrm{MAE} &= \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|, \quad \mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}, \\
\mathrm{MAPE} &= \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right|, \quad R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}, \\
E_{\mathrm{bal}} &= \frac{1}{T}\sum_t \left| Q_{0,t} - \sum_j Q_{j,t} \right|, \quad E_{\mathrm{phy}} = \frac{1}{T}\sum_t |g(\hat{x}_t)|, \\
\mathrm{PICP} &= \frac{1}{N}\sum_i \mathbb{I}(L_i \le y_i \le U_i), \quad \mathrm{MPIW} = \frac{1}{N}\sum_i (U_i - L_i), \\
\mathrm{CRPS} &= \frac{1}{N}\sum_i \int_{-\infty}^{\infty} [F_i(z) - \mathbb{I}(z \ge y_i)]^2 dz, \\
\mathrm{MacroF1} &= \frac{1}{C}\sum_c F_{1,c}, \quad \mathrm{WeightedF1} = \sum_c \frac{n_c}{N} F_{1,c}, \\
\mathrm{ECE} &= \sum_{m=1}^{M} \frac{|B_m|}{N} |\mathrm{acc}(B_m) - \mathrm{conf}(B_m)|,
\end{aligned} \tag{26}$$

where y_i and $\hat{y}_i$ denote the ground-truth and predicted values, respectively, $\bar{y}$ denotes the mean of the ground-truth observations, and N is the number of samples. P and R denote Precision and Recall, respectively, while t_i and $\hat{t}_i$ represent the actual fault-occurrence time and detected fault time. $Q_{0,t}$ and $Q_{j,t}$ denote the main-pipeline flow rate and the flow rate of the j-th branch at time t, respectively, and $g(\cdot)$ represents the residual associated with the physical constraint. L_i and U_i denote the lower and upper bounds of the prediction interval, respectively, and F_i denotes the predictive cumulative distribution function. C represents the number of classes, n_c denotes the number of samples belonging to class c, and B_m represents the m-th confidence bin.

4.2. Overall Fertigation Control and Economic Performance

This experiment was designed to comprehensively evaluate the overall performance of different fertigation-control strategies in multi-zone greenhouses from four perspectives: resource consumption, production performance, economic return, and agronomic safety. Particular emphasis was placed on determining whether SensorGraph-FertiMARL could reduce water and fertilizer inputs and drainage losses while maintaining marketable yield and simultaneously improving the proportion of Grade-A products and net profit. The strategy comparisons were evaluated through offline digital-twin simulation calibrated using historical

sensor, management, and harvest records. Yield, quality, resource-use, and economic metrics in this comparison describe simulated outcomes under the evaluated control policies.

As shown in Table 4 and Figure 5, Uniform-Timer served as the basic control strategy and exhibited the highest water and fertilizer inputs, a drainage rate of 18.7%, and the lowest net profit and ROI. Uniform-Threshold reduced water and fertilizer consumption, drainage, and stress violations through state-based feedback control, whereas Zone-Threshold further improved yield and economic return by employing zone-specific thresholds. Proportional Allocation demonstrated relatively good adaptability under resource constraints; however, because the allocation rule remained fixed, its yield and profitability were slightly lower than those achieved by Zone-Threshold. Independent MPC further reduced resource inputs and improved net profit through local predictive optimization, while Centralized Economic MPC benefited from a greenhouse-level economic objective and achieved a higher net profit of 118.5 and an ROI of 1.67. LSTM-Demand + Greedy was able to learn temporal demand dependencies, although the greedy allocation mechanism limited global coordination. GNN-Demand + Proportional further improved performance by incorporating inter-zone graph relationships. Independent SAC was capable of learning long-term nonlinear control policies, but it was insufficient coordination among zones resulted in a relatively high stress-violation rate. Centralized SAC achieved improved resource utilization and economic performance through global joint optimization, while MAPPO further enhanced collaborative decision-making through centralized training and decentralized execution. Graph-MAPPO-Yield achieved the highest marketable yield of 109.2, indicating that graph-based spatial relational modeling was beneficial for production-oriented optimization. SensorGraph-FertiMARL achieved the lowest water use, fertilizer use, drainage rate, and stress-violation rate while simultaneously obtaining the highest Grade-A rate, net profit, and ROI, thereby demonstrating the most balanced overall performance. The simulated marketable yields, expressed per unit greenhouse area, were 18.42 ± 0.28 kg/m^2 for Uniform-Timer, 20.11 ± 0.18 kg/m^2 for Graph-MAPPO-Yield, and 20.02 ± 0.17 kg/m^2 for SensorGraph-FertiMARL. The corresponding simulated Grade-A yields were 13.34 ± 0.22, 16.37 ± 0.16, and 16.94 ± 0.14 kg/m^2, respectively. Each area-normalized yield represents the simulated harvest mass over the evaluation period divided by the corresponding greenhouse area; all values are reported as mean $\pm$ standard deviation across five runs. Historical harvest measurements provide the calibration and prediction-evaluation data for the twin. Graph-MAPPO-Yield achieved a 0.45% higher simulated marketable yield than SensorGraph-FertiMARL, while its net-profit index was lower (129.8% vs. 138.9%). SensorGraph-FertiMARL combined a higher Grade-A share (84.6% vs. 81.4%) with lower resource inputs, improving the economic value of the simulated harvest. A plausible agronomic explanation is that additional water and fertilizer can increase fresh fruit mass while raising input costs and drainage losses; changes in soluble-solids concentration can also affect fruit grade. The reported comparison supports the economic benefit of jointly accounting for yield, grade, and production costs.

These results are consistent with the mathematical characteristics of the corresponding control paradigms. Rule-based methods essentially rely on fixed or piecewise decision mappings and are therefore unable to adequately represent continuous state transitions, long-term feedback, and cross-zone coupling, which limits their achievable performance. MPC explicitly exploits future-state predictions through receding-horizon optimization and consequently outperforms rule-based approaches, although its effectiveness can still be constrained by model mismatch and strong system nonlinearities. LSTM-based methods capture temporal dependencies but lack explicit spatial–topological representations, whereas GNN-based approaches encode zone adjacency and hydraulic-sharing relationships through neighborhood aggregation, which explains the improved performance of GNN-Demand compared with purely temporal models. Independent reinforcement-

learning strategies directly optimize long-term cumulative returns but lack joint value estimation across agents and can therefore generate conflicts under shared-resource constraints. MAPPO alleviates this problem through a centralized Critic, which improves both credit assignment and inter-zone coordination. Although Graph-MAPPO-Yield benefits from both graph modeling and multi-agent coordination, its optimization objective remains more strongly biased toward yield maximization. It therefore achieves slightly higher marketable yield than the complete model but sacrifices part of the resource efficiency and economic benefit. SensorGraph-FertiMARL further integrates reliability-aware states, an action-conditioned digital twin, marginal economic value, and capacity, safety, and fairness constraints. Consequently, policy optimization is driven not only by trusted system states but also by the predicted long-term consequences of candidate actions and the availability of shared resources. This prevents the policy from pursuing yield alone and enables a superior balance among production, quality, resource inputs, and operational risk, ultimately resulting in the highest long-term economic return.

Table 4. Overall fertigation control and economic performance of different methods. Results are reported as mean ± standard deviation over five independent runs. † denotes statistically significant improvement over the best performing baseline at $p < 0.01$ via paired t-test. Table Note: Water/Fertilizer Use and Net Profit are reported relative to the Uniform-Timer baseline (100.0%). Marketable yield represents simulated harvest meeting commercial standards.

Method	Water Use ↓	Fertilizer Use ↓	Drainage Rate ↓	Marketable Yield ↑
Uniform-Timer	100.0 ± 1.8	100.0 ± 1.6	18.7 ± 0.9	100.0 ± 1.5
Uniform-Threshold	94.6 ± 1.7	96.2 ± 1.5	16.1 ± 0.8	101.7 ± 1.4
Zone-Threshold	90.8 ± 1.5	92.7 ± 1.4	14.3 ± 0.7	103.2 ± 1.3
Proportional Allocation	89.7 ± 1.6	91.9 ± 1.3	13.8 ± 0.7	102.8 ± 1.4
Independent MPC	86.9 ± 1.4	89.1 ± 1.2	12.5 ± 0.6	104.6 ± 1.2
Centralized Economic MPC	84.8 ± 1.3	86.7 ± 1.2	11.4 ± 0.5	105.1 ± 1.2
LSTM-Demand + Greedy	85.9 ± 1.4	88.0 ± 1.3	11.9 ± 0.6	104.9 ± 1.3
GNN-Demand + Proportional	83.6 ± 1.3	85.5 ± 1.1	10.8 ± 0.5	105.8 ± 1.1
Independent SAC	84.1 ± 1.5	86.0 ± 1.4	11.2 ± 0.6	105.4 ± 1.4
Centralized SAC	82.3 ± 1.3	84.6 ± 1.2	10.1 ± 0.5	106.3 ± 1.2
MAPPO	80.7 ± 1.2	82.9 ± 1.1	9.5 ± 0.5	107.1 ± 1.1
Graph-MAPPO-Yield	79.4 ± 1.2	81.6 ± 1.0	9.1 ± 0.4	**109.2 ± 1.0**
SensorGraph-FertiMARL	**75.8 ± 1.0** †	**78.3 ± 0.9** †	**7.2 ± 0.4** †	108.7 ± 0.9
Method	**Grade-A Rate ↑**	**Net Profit ↑**	**ROI ↑**	**Stress Violation ↓**
Uniform-Timer	72.4 ± 1.2	100.0 ± 2.1	1.42 ± 0.04	11.8 ± 0.8
Uniform-Threshold	73.6 ± 1.1	104.8 ± 2.0	1.48 ± 0.04	9.6 ± 0.7
Zone-Threshold	75.1 ± 1.0	108.9 ± 1.9	1.54 ± 0.03	8.1 ± 0.6
Proportional Allocation	74.9 ± 1.1	108.1 ± 1.8	1.53 ± 0.04	7.9 ± 0.6
Independent MPC	76.8 ± 0.9	113.7 ± 1.7	1.61 ± 0.03	6.7 ± 0.5
Centralized Economic MPC	77.6 ± 0.9	118.5 ± 1.6	1.67 ± 0.03	6.1 ± 0.5
LSTM-Demand + Greedy	77.1 ± 0.9	115.9 ± 1.7	1.64 ± 0.03	6.4 ± 0.5
GNN-Demand + Proportional	78.3 ± 0.8	120.6 ± 1.5	1.70 ± 0.03	5.8 ± 0.4
Independent SAC	77.9 ± 1.0	119.3 ± 1.9	1.68 ± 0.04	7.2 ± 0.7
Centralized SAC	79.1 ± 0.8	123.8 ± 1.6	1.74 ± 0.03	5.9 ± 0.5
MAPPO	80.2 ± 0.8	127.4 ± 1.5	1.79 ± 0.03	5.1 ± 0.4
Graph-MAPPO-Yield	81.4 ± 0.7	129.8 ± 1.4	1.82 ± 0.03	4.8 ± 0.4
SensorGraph-FertiMARL	**84.6 ± 0.7** †	**138.9 ± 1.3** †	**1.94 ± 0.02** †	**2.9 ± 0.3** †

Bold values indicate the best mean result for each metric.

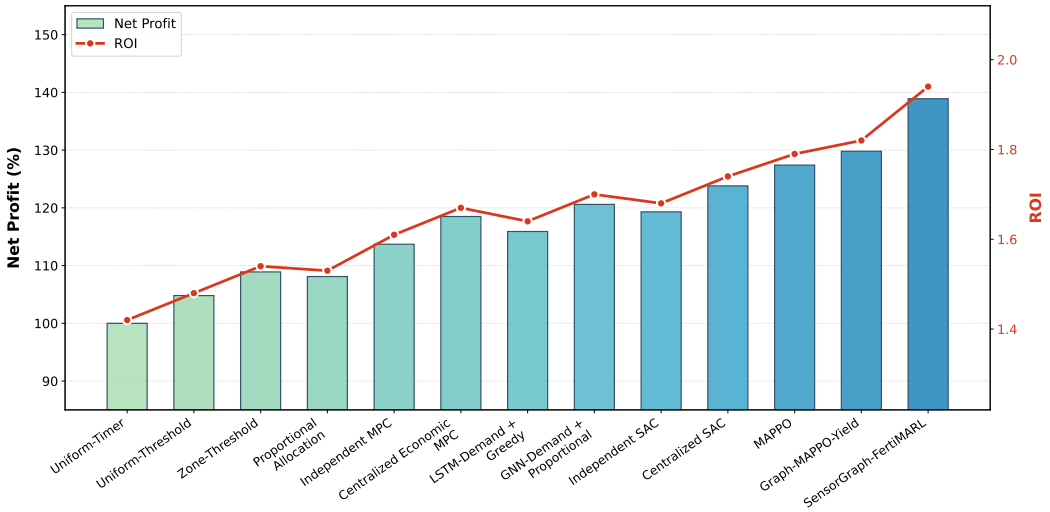


Figure 5. Comparison of net profit and return on investment across fertigation-control methods. The horizontal axis lists the control methods. Bars use the left vertical axis and report net profit as a percentage of the Uniform-Timer baseline (100%); a value of 138.9% corresponds to a 38.9% net-profit gain. The red line uses the right vertical axis and reports dimensionless return on investment (ROI). Method abbreviations are MPC, model predictive control; LSTM, long short-term memory; GNN, graph neural network; SAC, soft actor–critic; and MAPPO, multi-agent proximal policy optimization. Graph-MAPPO-Yield denotes the graph-based policy with a yield-oriented reward, and SensorGraph-FertiMARL denotes the proposed framework.

4.3. Sensor-State Reconstruction and Action-Conditioned Digital-Twin Prediction

This experiment was designed to simultaneously evaluate the ability of different methods to reconstruct missing sensor states and to predict future fertigation responses, crop yield, product quality, and economic return. Particular attention was paid to whether improvements in sensor-state reconstruction could further support the accuracy and reliability of the action-conditioned digital twin.

As shown in Table 5 and Figure 6, Mean Imputation relies exclusively on global statistical information to replace missing values and therefore produced the highest reconstruction MAE and the lowest R^2, together with the largest errors in subsequent water, *EC*, drainage, and profit prediction. Linear Interpolation improved performance by exploiting local temporal continuity, but remained unable to represent complex nonlinear dynamics. LSTM captured long-term temporal dependencies and substantially improved both reconstruction and prediction performance, while Transformer further reduced errors by modeling long-range dependencies through self-attention. After graph relationships among irrigation zones were introduced, GNN-Demand outperformed purely temporal models in both state reconstruction and yield prediction. Graph Transformer further improved multiple metrics by jointly modeling graph topology and attention-based interactions. SensorGraph without Action Conditioning produced more stable state representations through sensor-reliability modeling and multi-relational graph fusion, achieving a reconstruction R^2 of 0.891. However, profit prediction remained limited because candidate actions were not explicitly incorporated into the predictive process. The complete SensorGraph-FertiTwin achieved the best performance across all metrics, reducing reconstruction MAE to 0.036, increasing profit prediction R^2 to 0.904, and improving PICP to 0.931. These results indicate that the proposed method improved not only point-prediction accuracy but also the reliability of predictive uncertainty intervals.

From a modeling perspective, mean imputation and linear interpolation are fundamentally based on low-order statistical assumptions and are therefore inadequate for representing the nonlinear, asynchronous, and cross-zone coupled characteristics of greenhouse sensor observations. LSTM and Transformer can effectively capture temporal dependencies, but their input spaces are generally treated as regular sequences and cannot explicitly encode hydraulic sharing or spatial adjacency. GNN-based approaches exploit related irrigation-zone information through neighborhood aggregation and can therefore obtain richer complementary information when sensor observations are missing. Graph Transformer further employs attention mechanisms to adaptively determine the contribution of different neighboring nodes, resulting in more flexible spatial-relation modeling. The principal advantage of SensorGraph arises from the joint use of reliability gating and multi-relational graph modeling: anomalous or low-confidence observations are suppressed, while spatially, hydraulically, and microclimatically relevant nodes are dynamically exploited, thereby reducing the propagation of state-estimation errors into subsequent predictions. The complete FertiTwin further embeds candidate fertigation actions directly into the state-interaction process, enabling the model to learn action-conditioned state transitions rather than merely extrapolating from historical trajectories. In addition, short-term root-zone responses, long-term yield and quality outcomes, and economic returns are modeled through shared latent representations and multi-task constraints, allowing a consistent decision-oriented relationship to be established among these prediction targets. As a result, more accurate and stable estimates of fertigation responses, economic returns, and predictive uncertainty can be obtained simultaneously.

Table 5. Performance comparison for sensor-state reconstruction and action-conditioned digital-twin prediction. Results are reported as mean ± standard deviation over five independent runs. Table Note: PICP and MPIW quantify the coverage probability and mean prediction interval width of the 95% confidence bound. SensorGraph-FertiTwin simultaneously achieves the highest coverage (0.931) and the narrowest uncertainty band (0.142), demonstrating superior risk calibration.

Method	Recon. MAE ↓	Recon. R^2 ↑	Water MAE ↓	EC RMSE ↓	Dra. MAPE ↓
Mean Imputation	0.091 ± 0.004	0.721 ± 0.016	0.086 ± 0.004	0.214 ± 0.009	18.7 ± 1.1
Linear Interpolation	0.078 ± 0.004	0.754 ± 0.015	0.079 ± 0.004	0.198 ± 0.008	17.1 ± 1.0
LSTM	0.064 ± 0.003	0.812 ± 0.012	0.067 ± 0.003	0.173 ± 0.007	14.9 ± 0.9
Transformer	0.058 ± 0.003	0.836 ± 0.011	0.061 ± 0.003	0.161 ± 0.006	13.7 ± 0.8
GNN-Demand	0.054 ± 0.003	0.851 ± 0.010	0.058 ± 0.003	0.154 ± 0.006	12.8 ± 0.8
Graph Transformer	0.049 ± 0.002	0.872 ± 0.009	0.052 ± 0.002	0.143 ± 0.005	11.6 ± 0.7
SensorGraph w/o Action Cond.	0.043 ± 0.002	0.891 ± 0.008	0.048 ± 0.002	0.136 ± 0.005	10.4 ± 0.6
SensorGraph-FertiTwin	**0.036 ± 0.002**	**0.918 ± 0.007**	**0.041 ± 0.002**	**0.119 ± 0.004**	**8.7 ± 0.5**
Method	**Yield R^2 ↑**	**Quality Macro-F1 ↑**	**Profit R^2 ↑**	**PICP ↑**	**MPIW ↓**
Mean Imputation	0.771 ± 0.014	0.724 ± 0.013	0.742 ± 0.016	0.781 ± 0.015	0.328 ± 0.015
Linear Interpolation	0.789 ± 0.013	0.739 ± 0.012	0.761 ± 0.015	0.798 ± 0.014	0.294 ± 0.014
LSTM	0.821 ± 0.012	0.768 ± 0.011	0.801 ± 0.013	0.827 ± 0.012	0.231 ± 0.011
Transformer	0.842 ± 0.011	0.786 ± 0.010	0.826 ± 0.012	0.846 ± 0.011	0.208 ± 0.009
GNN-Demand	0.854 ± 0.010	0.797 ± 0.010	0.841 ± 0.011	0.858 ± 0.010	0.194 ± 0.008
Graph Transformer	0.871 ± 0.009	0.812 ± 0.009	0.858 ± 0.010	0.877 ± 0.009	0.176 ± 0.007
SensorGraph w/o Action Cond.	0.884 ± 0.008	0.826 ± 0.008	0.872 ± 0.009	0.892 ± 0.008	0.161 ± 0.007
SensorGraph-FertiTwin	**0.912 ± 0.007**	**0.851 ± 0.007**	**0.904 ± 0.008**	**0.931 ± 0.007**	**0.142 ± 0.006**

Bold values indicate the best mean result for each metric.

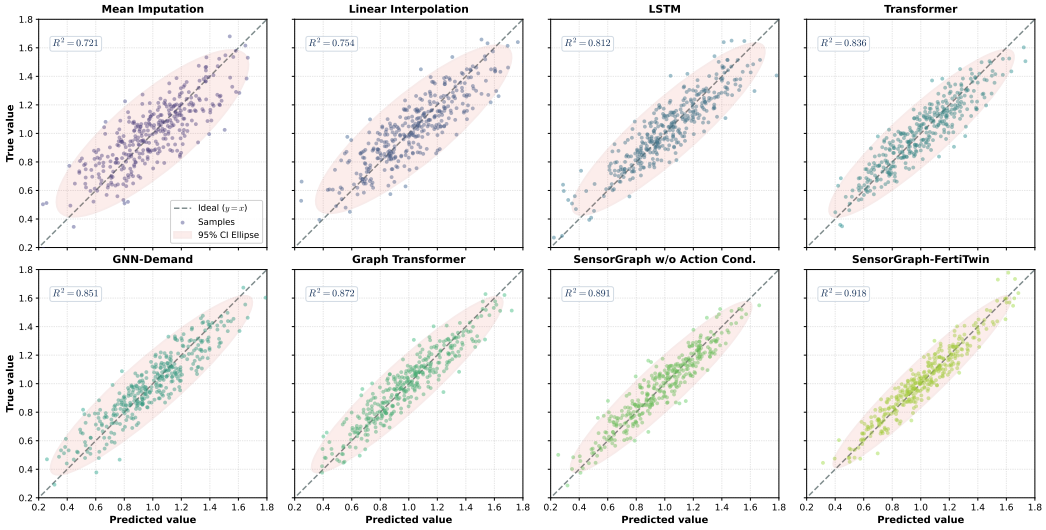


Figure 6. Comparison of agreement between predicted and ground-truth values with 95% confidence ellipses for different sensor-state reconstruction and prediction methods, where SensorGraph-FertiTwin achieves the highest R^2.

4.4. Ablation Study

This ablation experiment was designed to quantify the independent contributions of the major components of SensorGraph-FertiMARL to state reconstruction, economic prediction, resource conservation, operational safety, and inter-zone fairness.

As shown in Table 6 and Figure 7, after the Sensor Reliability Gate was removed, the reconstruction RMSE increased to 0.061, and the profit prediction R^2 decreased to 0.861, indicating that anomalous sensor observations substantially degraded the quality of state representation. Removing Graph Imputation further increased the reconstruction RMSE to 0.068, demonstrating that local temporal information alone was insufficient for recovering missing states across multiple irrigation zones. When Hydraulic Edges or the Global Resource Node were removed, the capacity-violation rate increased to 4.7 and 5.2, respectively, confirming that hydraulic topology and shared-resource nodes are essential for representing equipment competition. To elucidate the underlying mechanism of why removing hydraulic edges and global resource nodes causes the capacity-violation rate to surge from 0.6% to 4.7% and 5.2%, Figure 8 provides a visual analysis of the learned multi-relational attention weights $a^{r,l}_{ij,t}$ across the eight greenhouse zones. As visualized in the relation-specific matrices (Figure 8a–c), spatial adjacency attention is concentrated among direct topological neighbors (e.g., (Z_1, Z_2) and (Z_1, Z_5)), capturing local conductive heat and moisture diffusion. In contrast, hydraulic sharing attention (Figure 8b) exhibits dense off-diagonal couplings among zones sharing common sub-manifold branches (Zones 1–4 on Sub-manifold A and Zones 5–8 on Sub-manifold B), as well as global competition for the centralized 15 m^3/h booster pump. Crucially, the diurnal trajectory (Figure 8d) reveals dynamic relational switching: while spatial adjacency provides a stable baseline (≈25–30%), hydraulic attention dynamically surges from 20% to over 46% during morning (08:00–11:00) and afternoon (14:00–16:00) fertigation events. This enables individual zone agents to sense concurrent water demands and pipeline pressure drops in paired zones, facilitating implicit time-multiplexing and flow coordination. When hydraulic edges are ablated, zone agents operate blind to shared hydraulic load, precipitating simultaneous valve openings that overload pump capacity. Similarly, microclimatic similarity attention peaks at 42% during midday solar maximum (11:00–14:00), effectively sharing evapotranspiration representations among perimeter zones (Z_1, Z_4, Z_5, Z_8) exposed to intense solar radiation and interior zones (Z_2, Z_3, Z_6, Z_7) experiencing buffered microclimates. Removing Action

Conditioning reduced the profit prediction R^2 to 0.872, indicating that failure to explicitly model candidate actions weakened the ability to estimate downstream economic outcomes. Removing the physical and economic consistency constraints also resulted in degradation across both prediction and control metrics. Although the Yield-Only Reward preserved a strong yield-oriented objective, water-saving, fertilizer-saving, and net-profit improvement were substantially reduced. After the fairness constraint was removed, net profit remained relatively high, but the Jain Index decreased to 0.841, and the stress-violation rate increased to 6.8. Removing fairness regularization reduced water saving from 24.2% to 23.7% and fertilizer saving from 21.7% to 20.9%, while net-profit gain decreased from 38.9% to 37.9%. The larger changes in the Jain index (0.947 to 0.841) and stress-violation rate (2.9% to 6.8%) indicate a less balanced allocation with greater simulated agronomic stress. This pattern is consistent with a policy favoring zones with higher immediate marginal returns at the expense of balanced service across the greenhouse. Persistent service deficits could increase root-zone stress and compromise later crop performance. Similarly, removing the safety projection resulted in a net-profit gain close to that of the full model, but the capacity-violation rate increased sharply to 7.6. In simulation, exceeding pump or pipeline capacity does not trigger physical hardware failure, so unconstrained agents simply 'borrow' unphysical flow capacity to quench crop demand, preserving high simulated profit (38.1%). In real-world greenhouse deployments, however, this 7.6% violation rate represents acute pressure drops, pump motor thermal trips, and severe drip line cavitation, which physically cripples the irrigation system. The safety projection reduces the capacity-violation rate from 7.6% to 0.6 ± 0.1%, while net-profit gain is 38.9% for the full model and 38.1% without projection.

Table 6. Ablation study of the proposed SensorGraph-FertiMARL framework. Results are reported as mean ± standard deviation over five independent runs. Table Note: Water/Fertilizer Saving and Net Profit Gain represent percentage improvements relative to Uniform-Timer baseline. Stress Violation measures percentage of control windows exceeding safe moisture/EC thresholds, and Capacity Violation measures windows where joint flow demand exceeded booster pump rating (15 m^3/h).

Variant	Reconstruction RMSE ↓	Profit R^2 ↑	Water Saving ↑	Fertilizer Saving ↑	Net Profit Gain ↑
w/o Sensor Reliability Gate	0.061 ± 0.003	0.861 ± 0.010	18.1 ± 1.0	15.8 ± 0.9	25.7 ± 1.5
w/o Graph Imputation	0.068 ± 0.003	0.854 ± 0.011	17.4 ± 1.0	15.2 ± 0.9	24.8 ± 1.5
w/o Hydraulic Edges	0.049 ± 0.002	0.879 ± 0.009	19.3 ± 0.9	17.1 ± 0.8	27.6 ± 1.4
w/o Global Resource Node	0.047 ± 0.002	0.881 ± 0.009	18.8 ± 0.9	16.6 ± 0.8	26.9 ± 1.4
w/o Action Conditioning	0.039 ± 0.002	0.872 ± 0.010	20.2 ± 0.9	18.0 ± 0.8	28.4 ± 1.3
w/o Physics/Economic Constraints	0.040 ± 0.002	0.884 ± 0.009	20.7 ± 0.8	18.5 ± 0.8	29.3 ± 1.3
Yield-Only Reward	0.037 ± 0.002	0.901 ± 0.008	20.6 ± 1.2	18.4 ± 1.0	29.8 ± 1.4
w/o Fairness Constraint	0.036 ± 0.002	0.903 ± 0.008	23.7 ± 0.8	20.9 ± 0.7	37.9 ± 1.2
w/o Safety Projection	0.036 ± 0.002	0.904 ± 0.008	24.0 ± 0.8	21.2 ± 0.7	38.1 ± 1.2
Full SensorGraph-FertiMARL	**0.038 ± 0.001**	**0.904 ± 0.008**	**24.2 ± 1.0**	**21.7 ± 0.9**	**38.9 ± 1.3**

Variant	Stress Violation ↓	Capacity Violation ↓	Jain Index ↑
w/o Sensor Reliability Gate	5.7 ± 0.5	2.8 ± 0.3	0.891 ± 0.009
w/o Graph Imputation	6.1 ± 0.5	3.1 ± 0.3	0.887 ± 0.010
w/o Hydraulic Edges	4.9 ± 0.4	4.7 ± 0.4	0.902 ± 0.008
w/o Global Resource Node	5.1 ± 0.4	5.2 ± 0.4	0.898 ± 0.008
w/o Action Conditioning	4.6 ± 0.4	2.5 ± 0.3	0.906 ± 0.008
w/o Physics/Economic Constraints	4.3 ± 0.4	2.2 ± 0.3	0.909 ± 0.007
Yield-Only Reward	4.1 ± 0.4	1.9 ± 0.2	0.913 ± 0.007
w/o Fairness Constraint	6.8 ± 0.5	1.4 ± 0.2	0.841 ± 0.011
w/o Safety Projection	8.4 ± 0.6	7.6 ± 0.5	0.918 ± 0.007
Full SensorGraph-FertiMARL	**2.9 ± 0.3**	**0.6 ± 0.1**	**0.947 ± 0.006**

Bold values indicate the results of the full SensorGraph-FertiMARL framework.

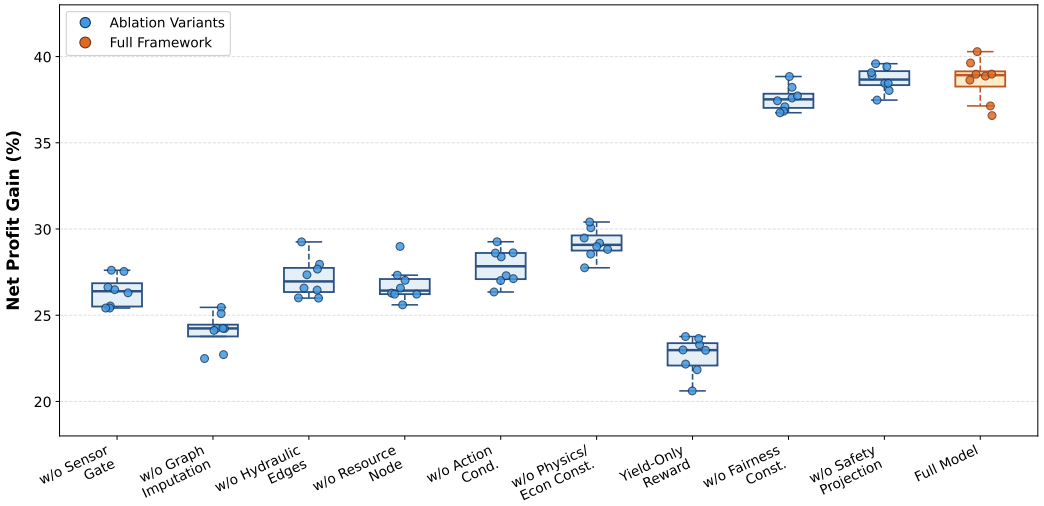


Figure 7. Distribution of net-profit gains for the SensorGraph-FertiMARL ablation variants and full framework. The horizontal axis identifies the model variants; the vertical axis reports net-profit gain relative to Uniform-Timer, calculated as $100(\Pi_{\text{variant}}/\Pi_{\text{Uniform-Timer}} - 1)\%$, where Π denotes net profit. In the labels, w/o means without, Cond. means conditioning, Const. means constraints, and Econ means economic. Boxes span the interquartile range, horizontal lines indicate medians, and circles show individual plotted observations. Blue denotes ablation variants, and orange denotes the full model.

The complete SensorGraph-FertiMARL achieved the lowest reconstruction error, the highest profit-prediction accuracy, the greatest resource savings and net-profit gain, and simultaneously the lowest stress- and capacity-violation rates.

To systematically investigate policy robustness, sim-to-real transferability, and the sensitivity of key hyperparameters, additional sensitivity experiments were conducted as reported in Table 7. First, synthetic sim-to-real mismatch errors ($\pm 5\%$ to $\pm 20\%$) were injected into the substrate water-retention curves and drainage infiltration rates. Owing to closed-loop reliability-aware feedback and conservative safety projections, the framework demonstrated remarkable stability, maintaining a 21.8% water saving and 33.4% net-profit gain even under severe 15% sim-to-real divergence. Second, tuning the physical constraint weights (λ_c, λ_u) revealed that balanced weights $(\lambda_c = \lambda_u = 0.10)$ enforce physical mass balance without restricting representation learning. Third, increasing the risk-sensitivity coefficient from $\kappa = 0.0$ to $\kappa = 1.0$ reduced stress violations from 4.8% to 2.9%, with net-profit gain changing from 39.2% to 38.9%. At $\kappa = 2.0$, stress violations decreased further to 2.5%, while water saving fell to 20.1% and net-profit gain to 31.2%. These results describe a trade-off between stress reduction, resource efficiency, and economic return. The default $\kappa = 1.0$ retains most of the risk-neutral profit gain while substantially reducing stress violations. Finally, increasing the fairness weight from $\beta_{fair} = 0.00$ to 0.05 improved the Jain index from 0.841 to 0.947, reduced stress violations from 6.8% to 2.9%, and increased net-profit gain from 37.9% to 38.9%. At $\beta_{fair} = 0.50$, the Jain index reached 0.978, while water saving and net-profit gain declined to 19.8% and 28.4%, respectively. The default $\beta_{fair} = 0.05$ balances allocation equity with resource savings and profitability among the tested settings.

These results demonstrate that the individual components provide distinct but complementary mathematical functions. Reliability gating adaptively downweights low-confidence observations and consequently suppresses the propagation of anomalous perturbations through the latent feature space. Graph-based imputation introduces spatial redundancy through neighborhood aggregation, thereby reducing estimation bias under missing observations. Hydraulic edges and the global resource node enable graph message passing to

explicitly encode shared-equipment constraints; when these components are removed, individual agents receive insufficient information regarding system-wide capacity and are more likely to generate locally optimal but globally infeasible actions. The action-conditioned digital twin distinguishes the consequences of different actions under identical system states through conditional mappings, while physical and economic constraints reduce the feasible prediction space and improve the stability of profit estimation. A yield-only reward tends to shift the policy toward high-input regions, whereas fairness regularization penalizes inter-zone disparities in normalized service level. The safety projection restricts policy outputs to the feasible operating region and therefore substantially reduces hard-constraint violations. By jointly integrating state reliability, structural relationships, action value, economic objectives, and operational constraints, the complete model achieves a more balanced and stable overall performance.

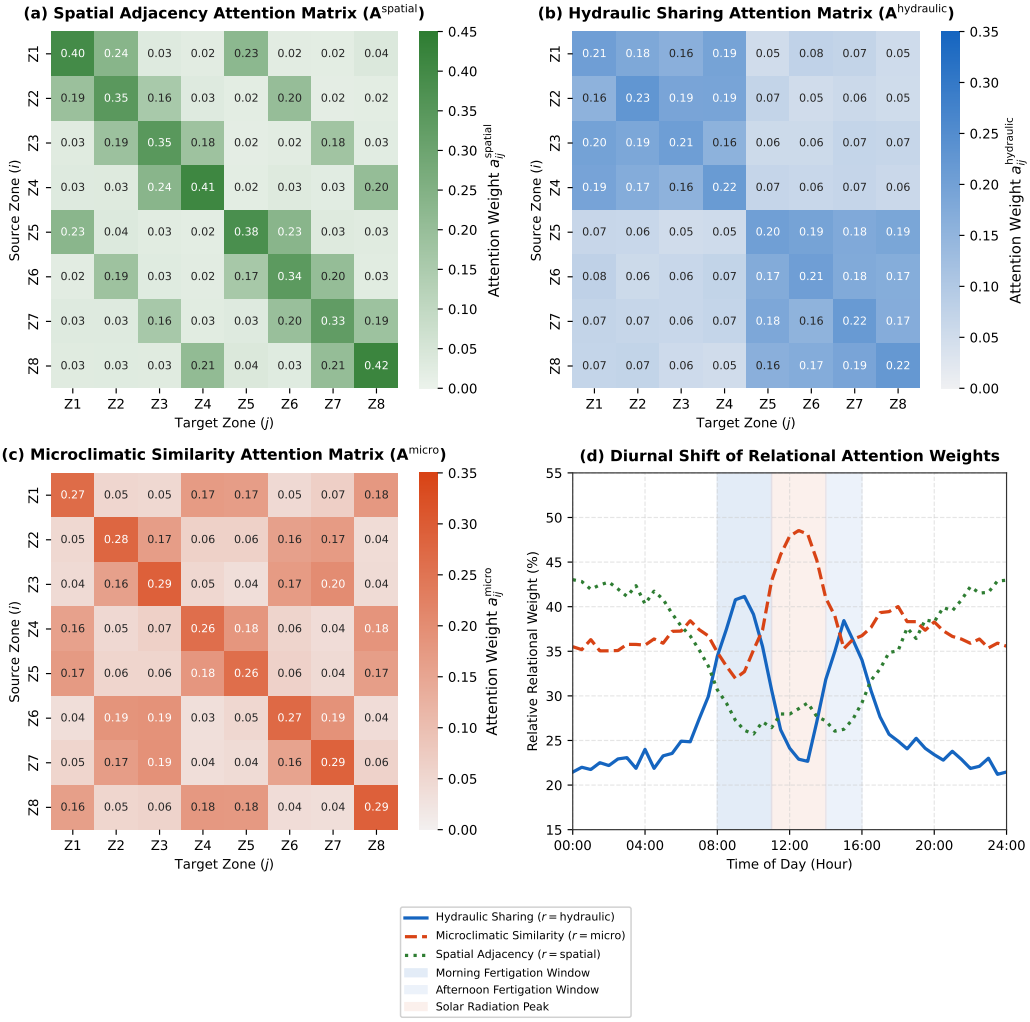


Figure 8. Visual analysis of learned multi-relational graph attention weights across eight greenhouse zones and diurnal operational cycles. (**a**) Spatial adjacency attention matrix A^{spatial}, exhibiting localized diffusion among topologically neighboring zones. (**b**) Hydraulic sharing attention matrix $A^{\text{hydraulic}}$, showing prominent cross-attention between zones sharing sub-manifold branches and booster pump capacity. (**c**) Microclimatic similarity attention matrix A^{micro}, displaying block structures between peripheral zones (Z1, Z4, Z5, Z8) and interior zones (Z2, Z3, Z6, Z7). (**d**) Diurnal variation of relational attention weights, highlighting the adaptive surge in hydraulic attention during morning (08:00–11:00) and afternoon (14:00–16:00) peak fertigation events to enforce equipment capacity constraints, while microclimatic attention peaks during solar maximum (11:00–14:00).

Table 7. Sensitivity analysis of SensorGraph-FertiMARL under sim-to-real transfer errors, physical constraint weights, risk aversion, and fairness regularization. Results are reported as mean ± standard deviation over five independent runs.

Evaluation Parameter	Tested Range/Value	Water Saving (%)	Net Profit Gain (%)	Stress Violation (%)	Jain Index
Sim-to-Real Mismatch Error	±5%	23.8 ± 0.8	38.2 ± 1.1	3.1 ± 0.3	0.945 ± 0.006
	±10%	22.9 ± 0.8	36.7 ± 1.2	3.5 ± 0.4	0.940 ± 0.007
	±15%	21.8 ± 0.9	33.4 ± 1.3	4.2 ± 0.4	0.932 ± 0.008
	±20%	19.4 ± 1.0	29.1 ± 1.5	5.6 ± 0.5	0.918 ± 0.009
Physics Loss Weight λ_c	$\lambda_c = 0.01$	21.4 ± 0.9	32.8 ± 1.3	4.9 ± 0.4	0.935 ± 0.007
	$\lambda_c = 0.10$ (Default)	**24.2 ± 1.0**	**38.9 ± 1.3**	**2.9 ± 0.3**	**0.947 ± 0.006**
	$\lambda_c = 1.00$	22.6 ± 0.8	34.2 ± 1.2	3.4 ± 0.3	0.941 ± 0.006
Risk Coefficient κ	$\kappa = 0.0$ (Risk-neutral)	24.6 ± 0.8	39.2 ± 1.2	4.8 ± 0.5	0.931 ± 0.008
	$\kappa = 1.0$ (Default)	**24.2 ± 1.0**	**38.9 ± 1.3**	**2.9 ± 0.3**	**0.947 ± 0.006**
	$\kappa = 2.0$ (Higher risk aversion)	20.1 ± 0.9	31.2 ± 1.4	2.5 ± 0.2	0.952 ± 0.005
Fairness Weight β_{fair}	$\beta_{fair} = 0.00$	23.7 ± 0.8	37.9 ± 1.2	6.8 ± 0.5	0.841 ± 0.011
	$\beta_{fair} = 0.05$ (Default)	**24.2 ± 1.0**	**38.9 ± 1.3**	**2.9 ± 0.3**	**0.947 ± 0.006**
	$\beta_{fair} = 0.50$ (Higher fairness weight)	19.8 ± 0.9	28.4 ± 1.4	2.7 ± 0.3	0.978 ± 0.004

Bold values identify the default parameter settings.

4.5. Discussion

The results indicate that the central challenge in multi-zone greenhouse fertigation management is not merely to improve irrigation accuracy within individual zones, but to resolve the conflict between heterogeneous zone-specific crop requirements and the limited capacity of shared infrastructure under realistic production conditions. A representative example occurs during periods of high midday radiation, when zones located near greenhouse edges or ventilation openings typically experience stronger evapotranspiration, causing root-zone water content to decline more rapidly than in central zones. Under a uniform timer-based irrigation strategy, all zones receive similar amounts of water and fertilizer regardless of their actual demand, which can increase drainage in low-evapotranspiration central zones while still allowing short-term water stress to occur in high-evapotranspiration edge zones. SensorGraph-FertiMARL establishes differentiated resource priorities according to zone-specific root-zone states, canopy load, and microclimatic conditions. When main-pipeline capacity becomes limiting, resources can therefore be preferentially allocated to zones with a higher stress risk while irrigation volumes are reduced in low-demand zones. This mechanism explains why the complete model was able to maintain a relatively high marketable yield while simultaneously reducing total water consumption and drainage, and further demonstrates that coordinated zone-level control provides greater practical value than improvements in predictive accuracy alone.

Constraints associated with shared pumps and main-pipeline capacity are equally important in practical greenhouse operation. During concentrated morning irrigation periods, several irrigation zones may simultaneously reach their lower water-content thresholds. If each zone is controlled independently, simultaneous valve opening can cause a reduction in main-pipeline pressure, insufficient actual flow at terminal branches, and even deviations in nutrient-solution concentration from the intended setpoint. The proposed method anticipates this form of resource competition through hydraulic relationships and global resource nodes and subsequently redistributes irrigation demand across time and flow capacity through the safety-projection mechanism. The increased capacity-violation rates observed after removing hydraulic edges or global resource nodes are consistent with this practical operating phenomenon. Accordingly, the contribution of the graph structure is not limited to abstract improvements in network representation capability; rather, it directly corresponds to the physical characteristics of real greenhouse systems, in which multiple

zones share the same hydraulic network, compete simultaneously for limited capacity, and mutually influence one another.

The economic results further demonstrate that fertigation optimization should not be driven solely by maximum crop yield. During fruit-expansion and peak-harvest stages, additional water and fertilizer supplied to high-load zones may increase gross yield, but excessive irrigation can also increase drainage, fertilizer losses, and energy consumption while potentially reducing soluble-solids content or market grade. Graph-MAPPO-Yield achieved the highest marketable yield but generated a lower net profit than the complete model, indicating that further production intensification becomes economically inefficient when marginal output is insufficient to offset additional resource inputs. SensorGraph-FertiMARL evaluates yield, product quality, and production cost simultaneously through the digital twin, allowing additional resources to be preferentially allocated to zones in which incremental fertigation remains economically beneficial rather than simply to zones with the greatest instantaneous water demand. From a production-management perspective, this control logic is more consistent with practical greenhouse operation, where crop safety and baseline production must first be maintained, and the economic value generated by each additional unit of water and fertilizer subsequently becomes the central consideration.

4.6. Limitation and Future Work

Although SensorGraph-FertiMARL demonstrated favorable overall performance in coordinated multi-zone greenhouse fertigation control, resource conservation, and economic-return optimization, several limitations should be explicitly acknowledged. First and foremost, the policy optimization and comparative evaluations in this work were conducted using historical sensor–production datasets and calibrated offline digital-twin simulations; long-term continuous online closed-loop deployment in commercial production has not yet been conducted. Real-world autonomous execution inevitably encounters unmodeled bio-physical drift, unanticipated hardware wear, and complex microbial root dynamics, demanding robust supervisory oversight during field commissioning.

Second, while our dataset encompasses an extensive 184-day period capturing all five tomato ontogenetic stages under extreme northern Chinese continental climatic fluctuations, the empirical data originate from a single eight-zone Venlo-type glass greenhouse facility. Variations in greenhouse cladding materials (e.g., polyethylene films vs. double-pane glass), substrate characteristics (perlite vs. coir slabs), and cultivar genotypes will alter thermal inertia and transpiration coefficients, necessitating domain adaptation and few-shot fine-tuning for cross-facility generalization.

Third, while we formulated preliminary mathematical structures for time-of-use dynamic electricity tariffs and carbon emission penalties in Equation (16), their empirical feedback was demonstrated within simulated environments. Future extensions will integrate real-time grid price telemetry, localized solar photovoltaic generation, on-site carbon capture metering, and automated disease vision detectors into an expanded multi-objective reward formulation, validating long-term autonomous operations across multi-year multi-site greenhouse networks.

5. Conclusions

This study addressed the tripartite conflict among sensor unreliability, multi-zone spatial heterogeneity, and shared hydraulic competition in modern protected cultivation by establishing an end-to-end closed-loop decision system, SensorGraph-FertiMARL. By embedding sensor-reliability gating, multi-relational graph attention, action-conditioned digital twinning, and constrained multi-agent reinforcement learning into an integrated pipeline, fertigation management is transformed from traditional static empirical schedules or single-

agent yield maximization toward cooperative allocation governed by marginal economic value, agronomic safety, and inter-zone equity.

Evaluated across an 184-day multi-stage tomato production dataset from an eight-zone commercial greenhouse via calibrated digital-twin simulation, the proposed framework demonstrated pronounced advantages across resource conservation, risk containment, and profitability. Compared to conventional uniform timer-based control, SensorGraph-FertiMARL achieved water and fertilizer saving rates of 24.2% and 21.7% (reducing consumption to 75.8% and 78.3%), lowered substrate drainage rates from 18.7% to 7.2%, restricted agronomic stress violations to 2.9%, and boosted Grade-A tomato ratios to 84.6%. Crucially, net economic profit improved by +38.9% with an outstanding return on investment of 1.94, significantly outperforming both independent control baselines and yield-only reinforcement learning policies that induce excessive nutrient leaching. Extensive ablation and sensitivity analyses further verified the critical role of reliability gating against sensor drift, hydraulic graph topology against equipment pressure drops, Euclidean safety projection against hardware overload, and Jain-fairness regularization against inter-zone service disparities.

In conclusion, SensorGraph-FertiMARL provides a technically rigorous, practically viable paradigm for next-generation precision horticulture. By explicitly linking upstream sensor integrity with downstream market returns and physical hardware limits, this work demonstrates that digital twins coupled with constrained multi-agent reinforcement learning can effectively reconcile economic profitability with ecological sustainability in multi-zone protected agriculture.

Author Contributions: Conceptualization, X.Z., M.L., Y.B. and Y.S.; Methodology, X.Z., M.L. and Y.B.; Software, X.Z., M.L. and Y.B.; Validation, H.Y. and J.P.; Formal analysis, H.Y. and J.P.; Investigation, H.Y. and J.P.; Resources, J.L.; Data curation, J.L.; Writing—original draft, X.Z., M.L., Y.B., H.Y., J.L., J.P. and Y.S.; Visualization, J.L.; Supervision, Y.S.; Project administration, Y.S.; X.Z., M.L., Y.B. and H.Y. contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bicamumakuba, E.; Reza, M.N.; Jin, H.; Samsuzzaman; Lee, K.H.; Chung, S.O. Multi-sensor monitoring, intelligent control, and data processing for smart greenhouse environment management. *Sensors* **2025**, *25*, 6134. [CrossRef] [PubMed]
2. Dauda, S.O.; Cortellazzi, C.; Panchal, S.K.; Yong, M.; Babla, M.; Donovan-Mak, M.; Bose, J.; Li, L.; Liang, W.; Tissue, D.; et al. Large variation of fertiliser levels does not improve physiology, nutritional quality and economic return of greenhouse capsicum. *Plant Growth Regul*. **2025**, *105*, 2215–2230. [CrossRef]
3. Li, K.; Shi, J.; Hu, C.; Xue, W. The intelligentization process of agricultural greenhouse: A review of control strategies and modeling techniques. *Agriculture* **2025**, *15*, 2135. [CrossRef]
4. Wang, Z.; Chen, J.; Zhao, H.; Wei, B. Deterministic Edge-Controlled Precision Fertigation System with Spatial Task Scheduling and Hardware–Software Safety Interlock. *Sensors* **2026**, *26*, 4289. [CrossRef] [PubMed]
5. Ghazi, N.M.; Tahir, H.T.; Othman, S.S.; Hawar, D.A.; Abdulkadir, M.A.; Korkmaz, C. Smart Irrigation Systems: A Comprehensive Review of IoT, AI, and Sustainable Agriculture Technologies. A Review Article. *J. Kirkuk Univ. Agric. Sci.* **2025**, *16*, 266. [CrossRef]
6. Adamo, T.; Caivano, D.; Colizzi, L.; Dimauro, G.; Guerriero, E. Optimization of irrigation and fertigation in smart agriculture: An IoT-based micro-services framework. *Smart Agric. Technol.* **2025**, *11*, 100885. [CrossRef]

7. Wang, C.; Gong, J. Intelligent agricultural greenhouse control system based on internet of things and machine learning. *arXiv* **2024**, arXiv:2402.09488.
8. Li, J.; Li, X.; Fan, X. Efficient Utilization and Variable Rate Regulation of Sprinkler Fertigation for Grain Crops: A Review of 15 Years of Research at the IWHR, China. *Irrig. Drain.* 2026, *early view*. [CrossRef]
9. Yang, B.; Fu, P.; Lu, J.; Ma, F.; Sun, X.; Fang, Y. Regulated deficit irrigation: An effective way to solve the shortage of agricultural water for horticulture. *Stress Biol.* **2022**, *2*, 28. [CrossRef] [PubMed]
10. Chen, W.H.; Mattson, N.S.; You, F. Intelligent control and energy optimization in controlled environment agriculture via nonlinear model predictive control of semi-closed greenhouse. *Appl. Energy* **2022**, *320*, 119334. [CrossRef]
11. Kachwaya, D.S.; Raturi, H.C.; Tiwana, A.S.; Tufail, M. Irrigation and fertigation systems for protected cultivation. In *Protected Cultivation: Structural Design, Crop Management Modeling, and Automation*; CRC Press: Boca Raton, FL, USA, 2024; pp. 161–192.
12. Gao, L.; Li, G.; Zhang, J.; Sun, L. Hierarchical Strategy Optimization for Irrigation, Ventilation, and Heating in Multi-Zone Greenhouses: A Year-Long Field Evaluation. In *Proceedings of the 2026 7th International Conference on Artificial Intelligence and Electromechanical Automation (AIEA)*; IEEE: New York, NY, USA , 2026; pp. 193–196.
13. Islam, M.M.; He, J.; Yong, M.; Babla, M.; Liang, W.; Li, L.; Bose, J.; Donovan-Mak, M.; Huda, S.; Tissue, D.; et al. Reduced fertigation input sustains yield and physiological performance for improved economic returns and cleaner production of greenhouse eggplant. *Sci. Hortic.* **2024**, *331*, 113097. [CrossRef]
14. Wang, Z.; Liu, Z.; Yuan, M.; Yin, W.; Zhang, C.; Zhang, Z.; Hu, X. A machine learning-based irrigation prediction model for cherry tomatoes in greenhouses: Leveraging optimal growth data for precision irrigation. *Comput. Electron. Agric.* **2025**, *237*, 110558. [CrossRef]
15. Bayat, P.; Bayat, P. Deep learning–driven adaptive framework for precision greenhouse management based on species-specific needs. *Greenh. Plant Prod. J.* **2025**, *2*, 28–44. [CrossRef]
16. Seri, E.; Petitta, M.; Papaioannou, C.; Katsoulas, N.; Cornaro, C. Sustainable greenhouse microclimate modeling: A comparative analysis of recurrent and Graph Neural Networks. *Build. Environ.* **2025**, *284*, 113473. [CrossRef]
17. Wang, Z.; Xiao, S.; Wang, J.; Parab, A.; Patel, S. Reinforcement learning-based agricultural fertilization and irrigation considering N_2O emissions and uncertain climate variability. *AgriEngineering* **2025**, *7*, 252. [CrossRef]
18. Mahajan, A.; Hegde, S.; Shay, E.; Wu, D.; Prins, A. Comparative Analysis of Multi-Agent Reinforcement Learning Policies for Crop Planning Decision Support. *arXiv* **2024**, arXiv:2412.02057.
19. Maraveas, C. Incorporating artificial intelligence technology in smart greenhouses: Current State of the Art. *Appl. Sci.* **2022**, *13*, 14. [CrossRef]
20. Banerjee, P.; Bhat, A. Techno-agricultural Synergies: Advancing Agricultural Economics Through Innovation and Digital Transformation. *Int. J. Food Sci. Agric.* **2025**, *9*, 150–164. [CrossRef]
21. Anas, A.; Mouna, E.; Abdellatif, K.; Hamidou, T.; Anis, L.; Cédric, A. Adaptive Control in Multi-Agent Digital Twins for Sustainable Smart Farming: Game Theory-Driven Optimization. *IEEE Internet Things J.* **2026**, *13*, 31838–31851. [CrossRef]
22. Fu, Z.; Zhang, Y.; Yu, S.; Lv, W.; Ru, S.; Wang, W.; Song, Y. Sensor-Based Cross-Modal Spatiotemporal Alignment and Causal Profit Modeling for Agricultural Input Optimization. *Sensors* **2026**, *26*, 4994. [CrossRef] [PubMed]
23. Bonelli, L.; Montesano, F.F.; D'Imperio, M.; Gonnella, M.; Boari, A.; Leoni, B.; Serio, F. Sensor-based fertigation management enhances resource utilization and crop performance in soilless strawberry cultivation. *Agronomy* **2024**, *14*, 465. [CrossRef]
24. Lei, H.; Lian, Y.; Du, J.; Pan, H.; Li, X.; Li, D.; Jin, C.; Xiao, Z.; Hou, Y. Dynamic optimization of greenhouse tomato irrigation schedule based on water, fertilizer and air coupled production function. *Agronomy* **2023**, *13*, 776. [CrossRef]
25. Sood, T.; Kapoor, S.; Kaur, J.; Hussain, N.; Sood, S. Fertigation: A paradigm shift in nutrient delivery for sustainable agriculture. In *Agricultural Nutrient Pollution and Climate Change: Challenges and Opportunities*; Springer: Cham, Switzerland, 2025; pp. 135–164.
26. Dirlik, I.; Uğurlar, F.; Kaya, C. Sensor-Guided Smart Irrigation for Tomato Production: Comparing Low and Optimum Soil Moisture in Greenhouse Environments. *Food Energy Secur.* **2025**, *14*, e70082. [CrossRef]
27. Savvas, D.; Giannothanasis, E.; Ntanasi, T.; Karavidas, I.; Ntatsi, G. State of the art and new technologies to recycle the fertigation effluents in closed soilless cropping systems aiming to maximise water and nutrient use efficiency in greenhouse crops. *Agronomy* **2023**, *14*, 61. [CrossRef]
28. Sun, W.; Chang, F.J. Empowering greenhouse cultivation: Dynamic factors and machine learning unite for advanced microclimate prediction. *Water* **2023**, *15*, 3548. [CrossRef]
29. Hasan, M.; Bhat, A.G.; S, V.K.; R, M.; Yadav, A.K.; Saji, A.; G, B.; Prakash, S.A. IoT-Based Smart Fertigation Scheduling and Wireless Microclimate Monitoring for a Greenhouse Dutch Bucket Hydroponic System. *Irrig. Drain.* **2025**, *74*, 1889–1903. [CrossRef]
30. Wang, Q.; Liu, Y.; Tang, Y. Precision fertilization and dynamic monitoring of soil nutrients based on big data: Research progress, key challenges, and future directions. *Adv. Resour. Res.* **2026**, *6*, 1333–1359.
31. Wang, S.; Jiang, K.; Li, Y.; Li, T.; An, Z.; Wang, M.; Zhang, J.; Liu, C.; Gao, H.; Yu, L.; et al. From Field Sensing to Smart Spreading: A Comprehensive Review of Variable Rate Fertilization Technologies in Precision Agriculture. *Sustainability* **2026**, *18*, 7095. [CrossRef]

32. Mohammed, S.; Malhotra, N.; Sani, J.M.; Magayaki, A.J. IoT and Data Analytics for Greenhouse Optimization: A Critical Review of Methods, Limitations, and Future Directions. *Innov. Comput. Data Sci.* **2026**, *2*, 1–28. [CrossRef]
33. El Ouaham, W.; Sadik, M.; Ennajih, A.; Mouzouna, Y.; Orchi, H.; Elouaham, S. Smart Greenhouses in the Era of IoT and AI: A Comprehensive Review of AI Applications, Spectral Sensing, Multimodal Data Fusion, and Intelligent Systems. *Agriculture* **2026**, *16*, 761. [CrossRef]
34. Imbernon-Mulero, A.; Maestre-Valero, J.F.; Martínez-Alvarez, V.; García-García, F.J.; Jódar-Conesa, F.J.; Gallego-Elvira, B. Evaluation of an autonomous smart system for optimal management of fertigation with variable sources of irrigation water. *Front. Plant Sci.* **2023**, *14*, 1149956. [CrossRef] [PubMed]
35. Tan, X.; Li, Y.; Gou, W.; Yang, S.; Wen, W.; Zuo, Q.; Xing, X.; Liang, D.; Huang, L.; Guo, X. Design and application of a multi-source sensor data fusion system based on a robot phenotype platform. *Fruit Res.* **2025**, *5*, e032. [CrossRef]
36. Xiang, J.; Ge, J.; Fu, C.; Huang, J.; Lu, M.; Zhang, Z.; Song, Y.; Yan, S. An environment-guided visual–temporal deep learning framework for early disease detection in greenhouse horticultural crops. *Front. Plant Sci.* **2026**, *17*, 1796407. [CrossRef] [PubMed]
37. Jeon, Y.J.; Park, S.J.; Jung, D.H. From Phenotyping to Supervised Agentic Decision Support: A Review of Sensing and Artificial Intelligence for Greenhouse Strawberry Cultivation. *Horticulturae* **2026**, *12*, 765. [CrossRef]
38. Zhou, W.; Li, Z.; Dong, Y.; Li, L.; Liu, S. Proactive Irrigation Timing Decision-Making for Greenhouse Tomatoes via STL-LSTM Deep Learning and Plant–Soil Dual-Threshold Sensing. *Sensors* **2026**, *26*, 2981. [CrossRef] [PubMed]
39. Tang, R.; Hu, H.; Lin, H.; Li, J.; Wang, Z.; Zhu, G.; Mei, Z.; Dai, J. Design and implementation of an autonomous intelligent fertigation system for cross-regional applications. *Actuators* **2025**, *14*, 413. [CrossRef]
40. Zhang, J.; Xiao, F.; Li, A.; Ma, T.; Xu, K.; Zhang, H.; Yan, R.; Fang, X.; Li, Y.; Wang, D. Graph neural network-based spatio-temporal indoor environment prediction and optimal control for central air-conditioning systems. *Build. Environ.* **2023**, *242*, 110600. [CrossRef]
41. Li, H.; Hou, L.; Cui, J.; Wang, Y.; Tao, Y.; Wang, H.; Savic, D.; Wang, A.; Ren, N. Spatiotemporal Graph Learning on Urban Environments. *Environ. Sci. Technol.* **2026**, *60*, 5191–5207. [CrossRef] [PubMed]
42. Cebolla-Alemany, J.; Cheng, Y.; Pintó-Espín, L.; Albano, M.; Macarulla, M.; Gassó-Domingo, S. Short-term forecasting and predictive control of rooftop greenhouse microclimate using multi-horizon machine learning models. *Smart Agric. Technol.* **2026**, *13*, 101733. [CrossRef]
43. Chen, S.; Liu, A.; Tang, F.; Hou, P.; Lu, Y.; Yuan, P. A review of environmental control strategies and models for modern agricultural greenhouses. *Sensors* **2025**, *25*, 1388. [CrossRef] [PubMed]
44. Wu, Y.; Dai, J.; Zhu, Y.; Zhang, T. Intelligent Agricultural Greenhouse Extraction Method Based on Multifeature Modeling: Fusion of Geometric, Spatial, and Spectral Characteristics. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 5568–5581. [CrossRef]
45. Yang, J.; Zhang, X.; Li, Z.; Yu, Y.; Yin, S.; Lou, X.; Song, Y. Revenue-Oriented Smart Greenhouse Regulation Using Agro-Economic State Representation and Deep Reinforcement Learning. *Sustainability* **2026**, *18*, 7332. [CrossRef]
46. Wang, Z. Reinforcement Learning-Based Agricultural Management Subject to Climate Variability. Ph.D. Thesis, The University of Iowa, Iowa City, IA, USA, 2026.
47. Zhang, Z.; Xie, D.; Cheng, K.; Liu, Z.; Ischia, G.; Zhao, Y.; Yang, F. Optimization of fertilization strategies for paddy fields based on multi-task transfer learning and reinforcement learning. *Agric. Water Manag.* **2025**, *322*, 109965. [CrossRef]
48. Padilla-Nates, J.P.; Garcia, L.D.; Lozoya, C.; Orona, L.; Cortes-Perez, A. Greenhouse irrigation control based on reinforcement learning. *Agronomy* **2025**, *15*, 2781. [CrossRef]
49. van den Bemd, W.J.G.M. Robust Deep Reinforcement Learning for Greenhouse Control and Crop Yield Optimization. Master's Thesis, Eindhoven University of Technology, Eindhoven, The Netherlands, 2022. Available online: https://research.tue.nl/en/studentTheses/robust-deep-reinforcement-learning-for-greenhouse-control-and-cro/ (accessed on 1 October 2026).
50. Vijayprasath, S.; Raj, R.M.; Raaj, R.S.; Manoharan, A. Case Study on Reinforcement Learning-Based Decentralized Approach for Precision Agriculture and Environmental Monitoring. In *Climate Smart Agriculture*; Wiley: Hoboken, NJ, USA, 2026; pp. 195–218.
51. Ajagekar, A.; Decardi-Nelson, B.; You, F. Energy management for demand response in networked greenhouses with multi-agent deep reinforcement learning. *Appl. Energy* **2024**, *355*, 122349. [CrossRef]
52. Kumar, U.N.; Narayanan, T.; Kumar, B.N. PhenoMARL-Fert: Safe Multi-Agent DRL for Real-Time Irrigation and Nutrient Control. In *Proceedings of the 2026 IEEE 6th International Conference on Computing, Power, and Communication Technologies (IC2PCT)*; IEEE: New York, NY, USA, 2026; pp. 70–76.
53. AlZu'bi, S.; Mughaid, A. The Digital Farmer Agents in Precise and Sustainable Agricultural Ecosystems. In *Proceedings of the 2025 3rd International Conference on Foundation and Large Language Models (FLLM)*; IEEE New York, NY, USA, 2025; pp. 202–209.
54. Wang, L.; He, X.; Luo, D. Deep reinforcement learning for greenhouse climate control. In *Proceedings of the 2020 IEEE International Conference on Knowledge Graph (ICKG)*; IEEE: New York, NY, USA, 2020; pp. 474–480.
55. Hertel, T.W.; Lobell, D.B. Agricultural adaptation to climate change in rich and poor countries: Current modeling practice and potential for empirical contributions. *Energy Econ.* **2014**, *46*, 562–575. [CrossRef]
56. METER Group, Inc. *ECH2O 5TE Integrator Guide*; Document 18236-02; METER Group, Inc.: Pullman, WA, USA, 2019.

57. Montesano, F.; Iersel, M.V.; Parente, A. Timer versus moisture sensor-based irrigation control of soilless lettuce: Effects on yield, quality and water use efficiency. *Hort. Sci.* **2016**, *43*, 67–75. [CrossRef]
58. Phene, C.; Hoffman, G.; Austin, R. Controlling automated irrigation with soil matric potential sensor. *Trans. ASAE* **1973**, *16*, 773–0776. [CrossRef]
59. Li, X.; Zhao, W.; Li, J.; Li, Y. Maximizing water productivity of winter wheat by managing zones of variable rate irrigation at different deficit levels. *Agric. Water Manag.* **2019**, *216*, 153–163. [CrossRef]
60. He, L.; Horbulyk, T.M.; Ali, M.K.; Le Roy, D.G.; Klein, K. Proportional water sharing vs. seniority-based allocation in the Bow River basin of Southern Alberta. *Agric. Water Manag.* **2012**, *104*, 21–31. [CrossRef]
61. Negenborn, R.R.; Van Overloop, P.J.; Keviczky, T.; De Schutter, B. Distributed model predictive control of irrigation canals. *Netw. Heterog. Media* **2009**, *4*, 359–380. [CrossRef]
62. Wang, Y.; Puig, V.; Cembrano, G. Non-linear economic model predictive control of water distribution networks. *J. Process Control* **2017**, *56*, 23–34. [CrossRef]
63. Perea, R.G.; García, I.F.; Poyato, E.C.; Díaz, J.R. New memory-based hybrid model for middle-term water demand forecasting in irrigated areas. *Agric. Water Manag.* **2023**, *284*, 108367. [CrossRef]
64. Zanfei, A.; Brentan, B.M.; Menapace, A.; Righetti, M.; Herrera, M. Graph convolutional recurrent neural networks for water demand forecasting. *Water Resour. Res.* **2022**, *58*, e2022WR032299. [CrossRef]
65. Haarnoja, T.; Zhou, A.; Abbeel, P.; Levine, S. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *Proceedings of the International Conference on Machine Learning*; PMLR: Cambridge, MA, USA, 2018; pp. 1861–1870.
66. Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal policy optimization algorithms. *arXiv* **2017**, arXiv:1707.06347.
67. Yu, C.; Velu, A.; Vinitsky, E.; Gao, J.; Wang, Y.; Bayen, A.; Wu, Y. The surprising effectiveness of ppo in cooperative multi-agent games. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 24611–24624. [CrossRef]
68. Lowe, R.; Wu, Y.; Tamar, A.; Harb, J.; Abbeel, P.; Mordatch, I. Multi-agent actor-critic for mixed cooperative-competitive environments. In *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*; Curran Associates Inc.: Red Hook, NY, USA, 2017; Volume 30.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.