12th IFAC International Symposium on
Advanced Control of Chemical Processes
July 14-17, 2024. Toronto, Canada

Semi-centralized Multi-agent RL for Irrigation Scheduling

Bernard T. Agyeman * **Jinfeng Liu** * **Sirish L. Shah** *

* *Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Canada T6G 1H9 (jinfeng@ualberta.ca).*

Abstract: This study proposes a Semi-centralized Multi-agent Reinforcement Learning (SCMARL) approach for irrigation scheduling in agricultural fields, which are characterized by spatial variability and therefore delineated into management zones. The SCMARL framework is hierarchical in nature, with a coordinator agent at the top level and local agents at the second/lower level. The coordinator agent makes daily 'yes/no' irrigation decisions based on field-wide observations from all the management zones, which are then communicated to local agents. These local agents are tasked with determining the optimal daily irrigation depths for specific management zones, utilizing both the coordinator agent's decision and local observations. A comparison between the SCMARL method and a Fully Decentralized Multi-agent Reinforcement Learning approach is presented, highlighting the superior performance of the SCMARL approach in terms of water savings and improved irrigation water-use efficiency.

Keywords: Multi-agent reinforcement learning, semi-centralized multi-agent reinforcement learning, mixed-integer optimal control, irrigation scheduling.

1. INTRODUCTION

The United Nations reports that agriculture uses approximately 70% of the world's freshwater, mainly for irrigation (UN Report (2018)). Simultaneously, the world is experiencing severe freshwater shortages due to climate change and a growing population. Consequently, precise water management, especially in irrigation, has become important to conserve freshwater resources. One effective approach to realize precision in irrigation is through closed-loop irrigation scheduling, which aims to deliver the right amount of water to crops at the right time by employing feedback from agricultural fields.

Agricultural fields are known to exhibit significant variability in soil texture, attributed to the biological, chemical, and physical processes occurring within the soil. One practical approach to address the inherent variability in agricultural fields is to delineate the field under consideration into irrigation management zones. A Management Zone (MZ) is defined as a sub-field area with uniform soil and crop conditions. Subsequently, these Management Zones (MZs) can be incorporated into the design of closed-loop irrigation schedulers.

Model Predictive Control (MPC), an optimal control method that determines control actions by iteratively solving a finite horizon optimal control problem, is extensively employed in developing closed-loop irrigation schedulers. In particular, MPC has been applied to calculate the optimal irrigation application depth that optimizes performance measures such as crop yield and overall water consumption (Delgoda et al., 2016; McCarthy et al., 2014).

⋆ Financial support from Natural Sciences and Engineering Research Council of Canada (NSERC) and Alberta Innovates is gratefully acknowledged.

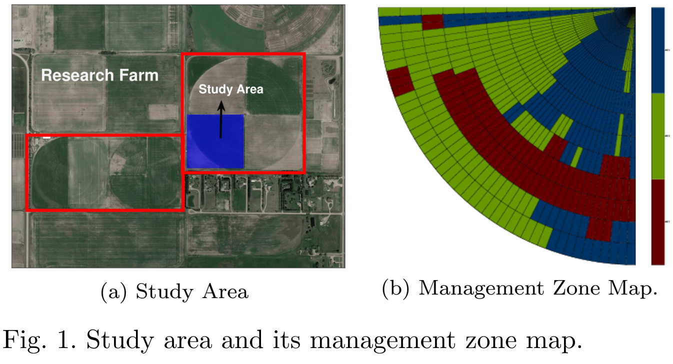
Scheduling problems inherently involve combinatorial aspects because distributing limited resources to competing tasks over time requires discrete decision-making. Presently, most MPC-based irrigation scheduling methods focus on continuous-valued controls, such as the depth of irrigation application, due to the complexities associated with handling discrete/integer-valued control actions. However, recent advancements in optimization software now enable the direct optimization of discrete-valued controls within the MPC framework. Leveraging this advancement, a mixed-integer MPC for daily irrigation scheduling was proposed in Agyeman et al. (2024).

Even in MPC-based schedulers that integrate discrete control actions, there is still a requirement for heuristic techniques to efficiently solve the optimization problem within a reasonable time-frame. Multi-agent reinforcement learning (MARL) offers a promising solution for handling the complexities of mixed-integer MPC in irrigation scheduling. In MARL, multiple agents, each with distinct observations and actions, collaborate to maximize rewards and make collective decisions. In a recent study (Agyeman et al., 2024), a combination of fully decentralized MARL, heuristic methods, and decentralized MPC with continuous variables was used to address the challenges of mixed-integer MPC-based irrigation schedulers in agricultural fields with multiple MZs. In this study, hybrid Proximal Policy Optimization (PPO) agents were trained, in a decentralized manner, for each of the MZs occurring in the spatially-variable field. These agents were trained to determine the daily irrigation decision and the corresponding daily irrigation rate for the MZs of the field. However, the need for uniform irrigation decisions across all MZs presented difficulties in the fully decentralized training framework. To address this, a heuristic approach was introduced to determine a uniform irrigation decision for all MZs. Subsequently, a decentralized MPC approach which incorporates the uniform irrigation decisions was solved to obtain the irrigation rates for each management zone. Despite notable improvements in computational efficiency and substantial advantages over traditional irrigation scheduling methods in terms of Irrigation Water Use Efficiency (IWUE), a limitation remains in the suboptimal approach that is adopted in the determination of the uniform irrigation decision. Ensuring the optimality of irrigation decisions across all zones has the potential to further enhance the water-saving and IWUE benefits of mixed-integer MPC-based irrigation scheduler.

From an RL point of view, a straightforward approach to guarantee the optimality of daily irrigation decision and the daily irrigations rates across all MZs is by training a centralized agent with both continuous and discrete action spaces. Although this approach simplifies coordination and reduces communication complexity, it poses challenges, particularly as the number of MZs increases. An alternative and potentially more effective strategy might be to combine the advantages of both centralized and fully decentralized reinforcement learning agents. This approach could benefit from insights derived from coordinator MPC concepts (Aske et al., 2008), which uses a semi-centralized control strategy to distribute control tasks between local and centralized controllers. Likewise, exploring semi-centralized MARL for irrigation scheduling in fields with multiple MZs, where the discrete irrigation decision and the irrigation rates are respectively distributed between a centralized RL agent and fully decentralized RL agents, represents a promising approach. Semi-centralized MARL has demonstrated its effectiveness in various domains. For instance, a semi-centralized deep deterministic policy gradient algorithm was proposed for cooperative tasks in StarCraft games (Xie and Zhong, 2020), and a semi-centralized multi-agent reinforcement learning algorithm was introduced to maximize energy efficiency in Internet-of-Things networks (Alajmi et al., 2022).

Despite its potential, the literature on irrigation scheduling has yet to explore the application of semi-centralized MARL to address the challenges of scheduling in spatially variable fields. Addressing this gap, this study aims to introduce a semi-centralized MARL strategy to address daily irrigation scheduling in spatially-variable fields.

2. PRELIMINARIES



(a) Study Area (b) Management Zone Map.

Fig. 1. Study area and its management zone map.

The irrigation scheduling approach was implemented in a specific section, marked by the blue rectangle in Figure 1(a), within a circular field situated at a Research Farm managed by Lethbridge College in Lethbridge, southern Alberta. Before implementing the proposed irrigation scheduler, a three-stage MZ delineation approach, previously developed in Agyeman et al. (2023a), was employed to delineate the investigated quadrant into MZs. This delineation utilized attributes such as elevation and soil hydraulic parameters, and it relied on the k-means clustering method. The soil hydraulic parameters used in the delineation were obtained through a data assimilation approach that estimated both soil moisture and hydraulic parameters within the investigated quadrant. As illustrated in Figure 1(b), the investigated quadrant is made up of three distinct MZs.

To reiterate, the MZ delineation method employed in the investigated quadrant is based on soil hydraulic parameters estimated from real-time soil moisture measurements. Consequently, the soil hydraulic parameters occurring at the center of each MZ (or cluster) offer an accurate representation of the soil properties within that specific zone. Through the calibration of a mechanistic agro-hydrological model using these centroidal soil hydraulic parameters, it is expected that the resulting model will accurately model the soil moisture dynamics in that particular MZ.

In this study, 1D Richards equation is employed to simulate the dynamics of soil moisture within each of the defined MZs. The 1D version of the Richards equation is expressed as:

$$c(\psi)\frac{\partial \psi}{\partial t}=\frac{\partial}{\partial z}\left[K(\psi)\left(\frac{\partial \psi}{\partial z}+1\right)\right]-\rho(\psi)\mathcal{R}\left(\mathrm{K_c}, \mathrm{ET_0}, \mathrm{z_r}\right) \quad (1)$$

In Equation (1), ψ (m) is the capillary pressure head, which describes the status of water in soil, t (s) represents time, $z(m)$ is the spatial coordinate, $K(\cdot)$ $(m \cdot s^{-1})$ is the unsaturated hydraulic water conductivity, $c(\cdot)$ (m^{-1}) is the capillary capacity. Note that $K(\cdot)$ and $c(\cdot)$ are parameterized functions that are relevant for solving Equation (1). $\rho(\cdot)$ $(-)$ is a dimensionless stress water factor, $\mathcal{R}(\cdot)$ is the root water uptake model which is a function of the crop coefficient $\mathrm{K_c}$ $(-)$, the reference evapotranspiration $\mathrm{ET_0}$ $(m \cdot s^{-1})$, and the rooting depth $\mathrm{z_r}$ (m).

In order to solve the 1D Richards equation, the following boundary condition is typically imposed:

$$\left.\frac{\partial \psi}{\partial z}\right|_{z=0} = -1 - \frac{u^{\mathrm{irr}} - \mathrm{EV}}{K(\psi)} \quad (2)$$

where u^{irr} $(m \cdot s^{-1})$, and EV $(m \cdot s^{-1})$ in Equation (2) represent the irrigation rate and the evaporation rate, respectively. It is important to add that the method of lines numerical approach is employed to solve Equation (1). Once a numerical value of the capillary pressure head ψ is obtained, the volumetric soil moisture content θ_v can be obtained as follows:

$$\theta_v(\psi) = \theta_r + (\theta_s - \theta_r)\left[\frac{1}{1+(-\alpha\psi)^n}\right]^{1-\frac{1}{n}} \quad (3)$$

where θ_s is the saturated moisture content, θ_r is the residual moisture content. α and n are curve fitting parameters.

The calibrated 1D Richards equation, after carrying out the temporal and spatial discretizations can be written in state-space form as:

$$x_{k+1} = \mathcal{F}(x_k, u_k, \hat{\boldsymbol{\theta}}) \tag{4}$$

$$y_k = \mathcal{H}(x_k, \hat{\boldsymbol{\theta}}) \tag{5}$$

where $x_k \in \mathbb{R}^{N_x}$ represents the state vector containing N_x capillary pressure head values for the spatial nodes in the soil column. u_k represents the input vector containing the irrigation amount, precipitation, daily reference evapotranspiration, the crop coefficient, and the rooting depth. $\hat{\boldsymbol{\theta}}$ represents the centroidal hydraulic parameters in a particular MZ. The volumetric water content θ_{v} is chosen as the output y_k. Equation (5) is thus a general representation of Equation (3) and $y_k \in \mathbb{R}^{N_y}$ represents the output vector containing $N_y = N_x$ volumetric soil moisture content values for the spatial nodes.

3. SEMI-CENTRALIZED MARL (SCMARL) DESIGN

A two-tier hierarchical SCMARL (Figure 2) strategy is proposed to address the daily irrigation scheduling problem in agricultural fields composed of multiple MZs. The framework consists of a coordinator agent which determines the daily irrigation decision and local agents tasked with determining daily irrigation rates for each MZ in the spatially variable field.

At the top level of this hierarchy is the coordinator, which serves as the root/central node. Its primary function is to make a binary daily irrigation decision, choosing either 'yes/1' or 'no/0'. The coordinator takes into account observations gathered from all MZs within the spatially-variable agricultural field, ensuring that it prescribes a daily irrigation decision that is optimal for the entire field. The coordinator's decision directly controls the irrigation across the field, serving as an on/off switch for the daily irrigation rates recommended by the local agents.

Local agents are responsible for recommending the daily irrigation rates for each of the MZs occurring in the spatially variable field, based on local observations in the MZs. The prescribed irrigation rates are then adjusted according to the coordinator's binary decision. Specifically, the irrigation rates prescribed by local agents are multiplied by the coordinator's decision, resulting in either non-zero irrigation rates for a 'yes/1' decision or zero irrigation rates for a 'no/0' decision.

This binary decision introduces a known disturbance into the environments of the local agents, leading to non-stationarity in their respective environments/MZs. For any given local agent, the future state of its environment/MZ is influenced by both its prescribed actions and the coordinator's daily irrigation decision. This non-stationarity of the learning environments affects a local agents' ability to learn stable policies.

In this study, two approaches are employed to address the non-stationarity issue. Initially, the coordinator's irrigation decision is communicated to local agents before they make their decisions. Subsequently, the coordinator's decision is combined with local observations from each MZ through state augmentation to form the input the policy of the local agents. Note that the coordinator determines its action prior to the decisions of the local agents. Consequently, the coordinator's daily irrigation decision, which is originally implemented in a switch-like manner in the SCMARL framework, can also be shared with the local agents to aid in their decision making. After sharing the coordinator's decision with the local agents, it is concatenated with local observations from each local agent's environment/MZ through a state augmentation process. This state augmentation technique ensures that local agents are aware of the global irrigation decision when making their localized irrigation rate recommendations.

Notably, the proposed architecture employs an actor-critic reinforcement learning algorithm to train all of its agents. Specifically, the PPO algorithm (Schulman et al., 2017) is utilized for this purpose.

3.1 Local Agent Design

For each MZ within the field, a dedicated local agent is assigned with the primary responsibility of determining the daily optimal irrigation rate.

In this study, a well-calibrated 1D Richards equation acts as the dynamics of the environment during the training of a local agent. This calibration is performed with the estimated hydraulic parameters that occur at the centroid of the MZ for which the local agent is designed. The transition dynamics for this calibrated 1D Richards equation are outlined in Equations (4) and (5). During the training of a local agent, the transition dynamics begin with the previously converged spatial capillary pressure head estimates within the relevant MZ. These estimates are derived from the offline simultaneous soil water and soil hydraulic parameter estimation process used during the MZ delineation in the field.

The local agent takes as input the output vector $\hat{y}$, which represents the spatial volumetric moisture contents obtained from Equation (5). Along with $\hat{y}$, the policy also receives additional key elements that are relevant for scheduling irrigation, such as the daily reference evapotranspiration ($\mathrm{ET_0}$), the daily crop coefficient ($\mathrm{K_c}$) and rain/precipitation ($\mathrm{R_n}$). Additionally, each local agent takes as input the daily irrigation decision (c) made by the coordinator. These observations, can be compactly represented as $s_{\mathrm{la}} = [\hat{y}, \mathrm{ET_0}, \mathrm{K_c}, \mathrm{R_n}, c]$. The actor network $\mathcal{A}_{\mathrm{la}}$ of a local agent takes s_{la} as input and it provides the daily irrigation rate a_{la} of its respective MZ. a_{la} is multiplied with c and the final product denoted as $u_{\mathrm{la}}^{\mathrm{irrig}}$ is applied to the environment/MZ. After applying $u_{\mathrm{la}}^{\mathrm{irrig}}$ to the environment/MZ, $\hat{y}$ transitions to $\hat{y}^+$, based on Equations (4) and (5). In determining the successor observation s_{la}^+, the one-day-ahead weather prediction is leveraged to obtain the weather conditions for the next day, which can be represented as $\mathrm{ET_0^+}$, $\mathrm{K_c^+}$, and $\mathrm{R_n^+}$. The updated actor network ($\mathcal{A}_{\mathrm{ca}}$) of the coordinator agent is leveraged to determine the prediction of the daily irrigation decision, c^+. In this approach, the successor spatial soil water contents $\hat{y}^+$ in all the n MZs are concatenated, together with the one-day-ahead weather predictions $\mathrm{ET_0^+}$, $\mathrm{K_c^+}$, and $\mathrm{R_n^+}$ to obtain $s_{\mathrm{ca}}^+ = [\hat{y}_1^+, \hat{y}_2^+, ..., \hat{y}_n^+, \mathrm{ET_0^+}, \mathrm{K_c^+}, \mathrm{R_n^+}]$. c^+ is obtained by evaluating $\mathcal{A}_{\mathrm{ca}}$ for s_{ca}^+, i.e. $c^+ = \mathcal{A}_{\mathrm{ca}}(s_{\mathrm{ca}}^+)$. Consequently, for each local agent, $s_{\mathrm{la}}^+ = [\hat{y}^+, \mathrm{ET_0^+}, \mathrm{K_c^+}, \mathrm{R_n^+}, c^+]$.

The primary goal of each local agent is to determine the daily irrigation rate, denoted as a_{la}, for a specific MZ. This daily irrigation rate should maintain the daily root zone soil moisture content θ^{RZ} within a prescribed target soil moisture range, bounded by the upper limit $\overline{\nu}$ and the lower limit $\underline{\nu}$. Note that θ^{RZ} is calculated as a weighted sum of the spatial soil moisture content $\hat{y}$, with 40% of the weight assigned to the average moisture in the top quarter of $\mathrm{z_r}$, 30% to the average moisture in the second quarter of $\mathrm{z_r}$, 20% to the average moisture in the third quarter of $\mathrm{z_r}$, and 10% to the average moisture in the bottom quarter of $\mathrm{z_r}$. Additionally, the agent considers a secondary objective, which is the minimization of the daily irrigation rate. Consequently, the reward r_{la} of a local agent consists of two parts: the target range tracking reward r^z_{la} and the irrigation rate minimization reward r^u_{la}

$$r_{\mathrm{la}} = \alpha_{\mathrm{la}} r^z_{\mathrm{la}} + \beta_{\mathrm{la}} r^u_{\mathrm{la}} \tag{6}$$

where α_{la} and β_{la} are weights for the target range tracking and irrigation rate minimization rewards, respectively. Specifically, $r^{\mathrm{z}}_{\mathrm{la}}$ is defined as:

$$r^{\mathrm{z}}_{\mathrm{la}} = \begin{cases} -\underline{Q} \times |\theta^{\mathrm{RZ}} - \underline{\nu}| & \text{if } \theta^{\mathrm{RZ}} < \underline{\nu} \\ -\overline{Q} \times |\theta^{\mathrm{RZ}} - \overline{\nu}| & \text{if } \theta^{\mathrm{RZ}} > \overline{\nu} \\ 0 & \text{if } \underline{\nu} \leq \theta^{\mathrm{RZ}} \leq \overline{\nu} \end{cases} \tag{7}$$

where $\underline{Q} > 0$, $\overline{Q} > 0$ are adjustable weights that penalize the violation of $\underline{\nu}$ and $\overline{\nu}$, respectively. r^u_{la} is defined as:

$$r^u_{\mathrm{la}} = -R_u a_{\mathrm{la}} \tag{8}$$

where $R_u > 0$ is an adjustable weight that represents the per unit cost of the water employed for irrigation.

3.2 Coordinator Agent Design

The role of the coordinator agent is to determine the daily 'yes/no' irrigation decision (c) that ensures that the root zone soil water contents in all the MZs of the field lie within some target soil moisture range. The n independently calibrated 1D Richards equations of the n MZs within the field are employed as the environment of the coordinator agent. During the simulation of the n calibrated 1D Richards equations, the initial conditions are drawn from the n local agents within the field. While the coordinator agent interacts with the n 1D Richards equations, it is important to acknowledge that these models are already defined during the training of the n local agents. Therefore, there is no need to explicitly define these n 1D Richards equations as individual components during the coordinator agent's training. Instead, the coordinator agent can rely on the same set of n 1D Richards equations that are employed in the training of the local agents.

The coordinator agent takes as input the a concatenation of the n spatial volumetric moisture contents $[\hat{y}_1, \hat{y}_2, ..., \hat{y}_n]$. Along with a concatenation of the output vectors, the coordinator agent takes as input $\mathrm{ET_0}$, $\mathrm{K_c}$, and $\mathrm{R_n}$. The observation to the coordinator agent can be represented as $s_{\mathrm{ca}} = [\hat{y}_1, ..., \hat{y}_n, \mathrm{ET_0}, \mathrm{K_c}, \mathrm{R_n}]$. The actor network $\mathcal{A}_{\mathrm{ca}}$ of coordinator agent takes s_{ca} as input and it provides the daily irrigation decision c for the entire field. c is applied to all the MZs of the fields where it turns 'on/off' the daily irrigation rates determined by the local agents. After applying c to all the MZs that make up the field, the n spatial volumetric water contents transition from $\hat{y}$ transitions to $\hat{y}^+$, based on the n independently calibrated versions of Equations (4) and (5). In determining the successor observation s^+_{ca}, $\hat{y}^+$ in the n MZs, together with the one-day-ahead

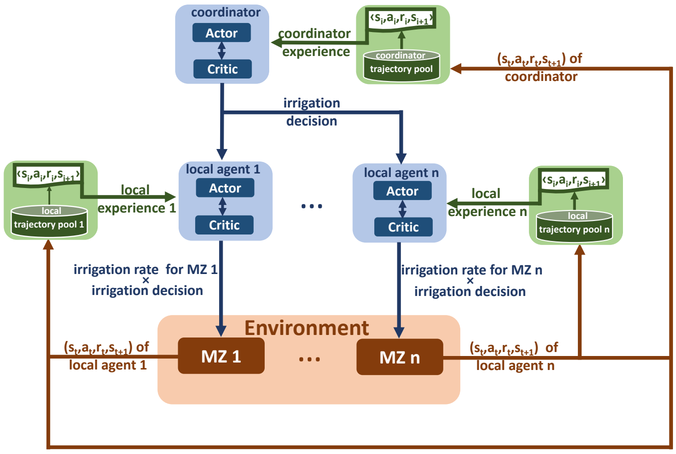


Fig. 2. Designed SCMARL for irrigation scheduling.

weather predictions $(\mathrm{ET}^+_0, \mathrm{K}^+_{\mathrm{c}}, \mathrm{R}^+_{\mathrm{n}})$ are employed. In particular, s^+_{ca} is defined as $[\hat{y}^+_1, \hat{y}^+_2, ..., \hat{y}^+_n, \mathrm{ET}^+_0, \mathrm{K}^+_{\mathrm{c}}, \mathrm{R}^+_{\mathrm{n}}]$.

Primarily, the coordinator agent seeks to determine the daily irrigation decision c that ensures that the root zone soil moisture contents in all the MZs ($\theta^{\mathrm{RZ}}_i\ \forall i \in [1, 2, ..., n]$) lie within a predetermined target range. It also considers as a secondary objective the minimization of the fixed cost associated with irrigation. Consequently, the reward r_{ca} of a local agent consists of two parts: the target range tracking reward r^z_{ca} in all the n MZs and the fixed irrigation cost minimization reward r^c_{ca}:

$$r_{\mathrm{ca}} = \alpha_{\mathrm{ca}} r^z_{\mathrm{ca}} + \beta_{\mathrm{ca}} r^c_{\mathrm{ca}} \tag{9}$$

where α_{ca} and β_{ca} are weights for the target range tracking in all the n MZs and fixed irrigation cost minimization rewards, respectively. The target range tracking reward in the n MZs is defined as:

$$r^z_{\mathrm{ca}} = \sum_{i=1}^{n} r^{\mathrm{z}}_{\mathrm{la},i} \tag{10}$$

where $r^{\mathrm{z}}_{\mathrm{la}}$ is calculated using Equation (7) for a particular MZ. The fixed irrigation cost minimization reward r^c_{ca} is defined as:

$$r^c_{\mathrm{ca}} = -R_c c \tag{11}$$

where $R_c > 0$ is an adjustable weight that represents the fixed cost of performing the irrigation event.

4. EXPERIMENTAL SETUP AND DESIGN

The study area, comprising of three irrigation MZs is employed to demonstrate the effectiveness of the proposed SCMARL approach. Additionally, this study considers the presence of wheat crop in the study area. Based on the study area, 4 agents will be trained in this experiment, specifically 3 local agents for the three MZs, and the coordinator agent. This section discusses the experimental scenario as well as the parameters employed during the training of the agents.

4.1 Environment Setup and Design

For each of the calibrated 1D Richards equations used in local agent training, a soil column with a depth of 50 cm is considered. This column is discretized into 21 equally spaced compartments. The spatial discretization follows the central difference scheme, and for temporal discretization, the backward differentiation formula is used. As previously mentioned, each Richards equation is calibrated with the estimated hydraulic parameters at the center of the MZ it describes. Furthermore, each calibrated Richards equation starts with converged estimates of spatial capillary pressure head and, consequently, the converged spatial volumetric moisture content for the specific MZ it models.

The daily reference evapotranspiration data that is employed during the training of the agents is uniformly generated between 1.04 mm/day and 9.0 mm/day, with the note that these values were chosen based on the historical daily reference evapotranspiration occurring in the study area. Similarly, historical rainfall data from the 2005 to the 2023 growing seasons were employed during the training of the agents. Additionally, crop coefficient values of wheat occurring between 0.2 and 1.25 were considered during the training of the agents. A rooting depth of 0.50 m was considered during the training of the agents, since it represents the commonly employed rooting depth that is employed for irrigation scheduling in wheat.

4.2 Training With the SCMARL Approach

In designing the rewards of the local and coordinator agents, the parameters utilized included setting α_{la} and β_{la} to 1.0 for the local agents and $\alpha_{\text{ca}} = 0.1$ and $\beta_{\text{ca}} = 1.0$ for the coordinator agent. Additionally, $\bar{Q}$ was set at 1200000, $\underline{Q}$ at 1000000, $R_c = 1000$, and $R_u = 9000$. For MZs 1 and 2, $\bar{\nu}$ and $\underline{\nu}$ were set at 0.280 and 0.200, respectively. In the case of MZ 3, the values for $\bar{\nu}$ and $\underline{\nu}$ were configured as 0.30 and 0.230, respectively. The specific values of $\bar{\nu}$ and $\underline{\nu}$, which are also referred to as Field Capacity (FC) and the threshold volumetric moisture content for the MZs, were obtained from Huffman et al. (2012).

The actor and critic networks in each agent use a learning rate of 1e-05, and input normalization is applied to these networks. In addition, the agents are trained with the following settings: a horizon (T) of 30, a minibatch size of 64, 20 epochs, a discount factor of 0.99, a generalized advantage estimation parameter of 0.97, a clipping parameter of 0.25, and an entropy coefficient of 0.01. The coordinator agent decides its daily discrete irrigation decision action by selecting it from a softmax distribution, while each local agent determines its daily irrigation rate using a Gaussian distribution. All agents' policies are represented through a fully connected multi-layer perceptron with two hidden layers, each consisting of 64 neurons and employing a hyperbolic tangent activation function. In accordance with the environment's setup, each local agent has 25 inputs, and the coordinator agent has 66 inputs. The training considers 10000 episodes (K), and the agents are configured and trained using the Tensorforce library in Python.

4.3 Evaluation of the SCMARL Approach

Following the training of the agents, the proposed method is utilized to prescribe irrigation schedules for the entire wheat growing season in the study area. This schedule relied on the weather data collected during the 2022 growing season. In particular, the proposed approach was used to determine schedules over a 123-day period. To evaluate the advantages of this approach, a comparison was made with the schedules generated by the combined fully decentralized multi-agent reinforcement learning (FDMARL) and MPC approach that was proposed in Agyeman et al. (2024). A number of approaches were adopted to facilitate a fair comparison between the SCMARL approach and the FDMARL+MPC scheduling approach. Firstly, the hyperparameters employed during the training of the agents in the SCMARL approach were adopted during the training of the decentralized hybrid PPO agents in the FDMARL approach. Additionally, the same neural network weight initialization technique was adopted during the training of the agents in both scheduling approaches. Thirdly, the SCMARL and the FDMARL approaches are used to determine the irrigation decision sequence over the scheduling horizon (14 days in this work). These sequences were employed to solve fully decentralized MPCs, proposed in Agyeman et al. (2024), to determine the daily irrigation rates in the MZs that make up the field. The comparison between the SCMARL and FDMARL scheduling approaches is made in terms of total prescribed irrigation depth and the IWUE (ratio of predicted yield to total prescribed irrigation rate).

Throughout this season-long investigation, the complete soil moisture content distribution within each MZ was not assumed to be known. Instead, the average soil moisture content in the top 25 cm of each MZ was used to estimate the overall soil moisture distribution. This estimation was realized through the application of the extended Kalman filtering technique, considering the presence of a soil moisture sensor in each MZ, which provided daily measurements. Furthermore, the state estimator utilized the calibrated 1D Richards equation for each MZ. In the design of the state estimator, the covariance matrices of the process disturbance and the measurement noise were set as $15.9\mathbb{I}_{21}$ and 19.25, respectively.

5. RESULTS AND DISCUSSION

Figure 3 illustrates the average episodic rewards obtained from running each of the coordinator agent and the local agent for MZ_1 through 10 repetitions across 10000 episodes. During the initial training phase extending up to approximately the 4000th episode, the average reward trajectories for the agents generally display an upward trend with occasional fluctuations. After the 4000th episode, the average reward trajectories remain fairly constant.

Figure 4 illustrates the trajectories of root zone soil moisture content in MZ_1 under the schedules recommended by the trained agents in the SCMARL and the FDMARL scheduling approaches. Similarly, Table 1 presents a quantitative comparative analysis between the proposed and the FDMARL scheduling approaches. This table reveals that the proposed approach prescribed a lower total irrigation depth when compared to the FDMARL approach. In particular, the proposed approach resulted in a 4.0% reduction in total irrigation depth compared to the FDMARL approach. Additionally, in terms of the IWUE, the proposed approach resulted in higher IWUE compared to the FDMARL approach. Specifically, the proposed approach increased the IWUE by 6.3%.

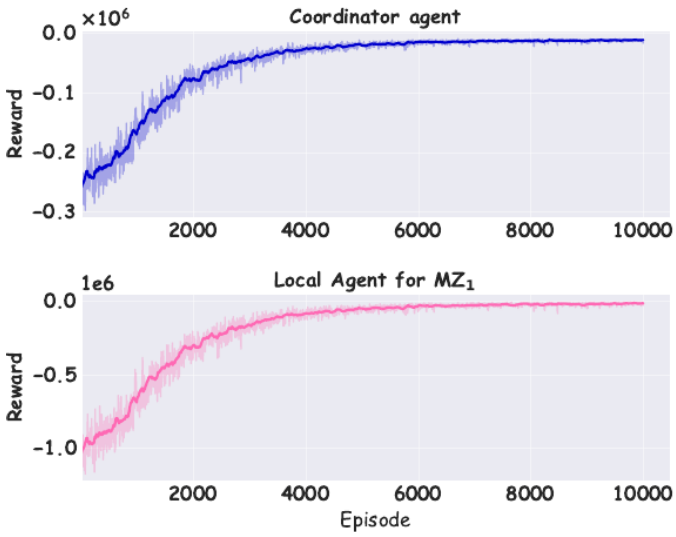


Fig. 3. Average reward trajectories over 10000 episodes.

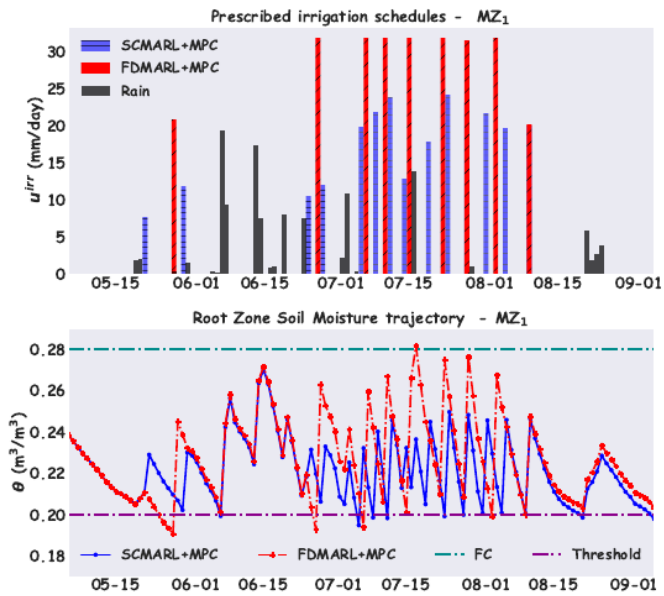


Fig. 4. Prescribed irrigation schedules and the trajectories of root zone soil moisture content under the prescribed schedules.

Table 1. Comparison between SCMARL and FDMARL scheduling approaches.

	SCMARL	**FDMARL**
Total irrigation (m)	0.774 [↓**4.0**%]	0.806
IWUE (kg/m^3)	1.134 [↑ **6.3**%]	1.067

6. CONCLUSION

A two-tier SCMARL was proposed for irrigation scheduling irrigation in spatially-variable agricultural fields. At the top level is a coordinator agent responsible for making daily irrigation decisions (yes/no) based on comprehensive soil moisture data collected from the entire field. The second level of this proposed framework consists of local agents, each assigned to specific MZs within the field. These local agents determine the daily irrigation rates for their respective zones, based on local observations and the discrete decision provided by the coordinator agent. The application of this approach to schedule irrigation demonstrated its capacity to achieve modest water savings while enhancing irrigation water-use efficiency.

REFERENCES

WWAP (United Nations World Water Assessment Programme)/UN-Water. (2018). The United Nations World Water Development Report 2018: Nature-Based Solutions for Water.

McCarthy, A. C., Hancock, N. H., Raine, S. R. (2014). Simulation of irrigation control strategies for cotton using Model Predictive Control within the VARIwise simulation framework. *Computers and Electronics in Agriculture*, 101:135–147.

Aske, E. M. B., Strand, S., & Skogestad, S.(2008). Coordinator MPC for maximizing plant throughput. *Computers & Chemical Engineering*, 32(1-2):195–204.

Delgoda, D., Malano, H., Saleem, S.K., & Halgamuge, M.N. (2016). Irrigation control based on model predictive control (MPC): Formulation of theory and validation using weather forecast data and AQUACROP model. *Environmental Modelling & Software*, 78:40–53.

Agyeman, B. T., Nouri, M. , Appels, W., Liu, J., & Shah, S. L.(2024). Learning-based multi-agent MPC for irrigation scheduling *Control Engineering Practice*, 147:105908.

Schulman, J., Wolski, F. , Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. In: *arXiv preprint arXiv:1707.06347*.

Agyeman, B. T., Nouri, M. , Appels, W., & Liu, J. (2023a). Irrigation management zone delineation and optimal irrigation scheduling for center pivot irrigation systems *In Proceedings of IFAC World Congress*, 10634–10639.

Agyeman, B. T., Sahoo, S. R., Liu, J., & Shah, S. L. (2023b). An LSTM-based mixed integer model predictive control for irrigation scheduling. *The Canadian Journal of Chemical Engineering*, 101(6):3362–3381.

Xie, D., Zhong, & X. (2020). Semicentralized deep deterministic policy gradient in cooperative StarCraft games. *IEEE transactions on neural networks and learning systems*, 33(4):1584–1593.

Alajmi, A., Fayaz, M., Ahsan, W., & Nallanathan, A. (2022). Semi-Centralized Optimization for Energy Efficiency in IoT Networks With NOMA. *IEEE Wireless Communications Letters*, 12(2):366–370.

Huffman, R. L. , Fangmeier, D. D., Elliot, W. J. & Workman, S. R.(2012). PIrrigation principles. In *Soil and Water Conservation Engineering, Seventh Edition*, 351–373, American Society of Agricultural and Biological Engineers.