Optimizing Crop Management with Reinforcement Learning and Imitation Learning

Extended Abstract

Ran Tao
University of Illinois at Urbana-Champaign
Champaign, USA

Pan Zhao
University of Illinois at Urbana-Champaign
Champaign, USA

Jing Wu
University of Illinois at Urbana-Champaign
Champaign, USA

Nicolas F. Martin
University of Illinois at Urbana-Champaign
Champaign, USA

Matthew T. Harrison
University of Tasmania
Hobart, Australia

Carla Ferreira
Stockholm University
Stockholm, Sweden

Zahra Kalantari
KTH Royal Institute of Technology
Stockholm, Sweden

Naira Hovakimyan
University of Illinois at Urbana-Champaign
Champaign, USA

ABSTRACT

To increase crop yield while minimizing environmental impact, we present an intelligent crop management system that optimizes nitrogen fertilization and irrigation simultaneously via reinforcement learning (RL), imitation learning (IL), and crop simulations using DSSAT. We first use deep RL to train management policies that require a large number of state variables from the simulator as observations (denoted as full observation). We then invoke IL to train management policies that only need a limited number of variables that are measurable in the real world (denoted as partial observation) by mimicking the actions of the RL-trained policies under full observation. Simulation experiments using maize in Florida demonstrate that our trained policies under both full and partial observations achieve better outcomes than a baseline policy. Most importantly, the IL-trained management policies are directly deployable in the real world as they use readily available information.

KEYWORDS

Reinforcement Learning; Imitation Learning; Intelligent Crop Management; Sustainable Agriculture

ACM Reference Format:
Ran Tao, Pan Zhao, Jing Wu, Nicolas F. Martin, Matthew T. Harrison, Carla Ferreira, Zahra Kalantari, and Naira Hovakimyan. 2023. Optimizing Crop Management with Reinforcement Learning and Imitation Learning: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

Nitrogen (N) fertilization and irrigation are two crop management practices that affect the crop yield and environment most [4]. Based on empirical experience and existing agricultural studies, local best management practices exist among farmers. However, it remains to be seen whether the current management practices are optimal and whether these strategies perform well in the presence of changes in climate and market conditions.

Here, we present an intelligent crop management system, depicted in Figure 1, that generates *deployable* and *adaptable* management policies based on reinforcement learning (RL), imitation learning (IL), and crop simulations via Decision Support System for Agrotechnology Transfer (DSSAT). Compared with the previous studies [9], [6], and [1], we advance the state of the art by increasing the action space to include both N fertilization and irrigation and testing the RL-based crop management architecture with different reward functions. More importantly, we leverage IL as a new tool to train policies that require only state variables measurable in the real world for decision-making, which paves the way for real-world deployment of our framework. More details of this paper can be found in [7].

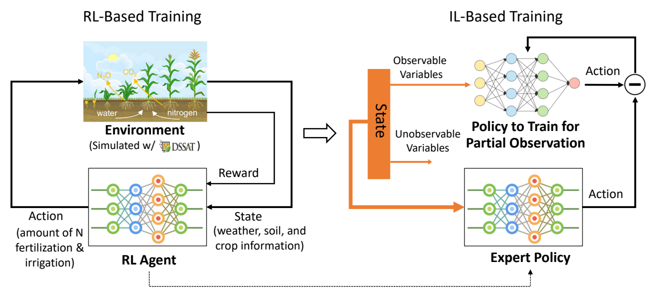

Figure 1: Framework of the intelligent crop management system using RL and IL

2 METHODS

The N fertilization and irrigation management is formulated as a finite Markov decision processes (MDP) here. On each day t, the

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

Table 1: Evaluation results of trained policies under full observation and the baseline policy. Trained Policy x indicates the training result of the RL agent using reward function (RF) x. For each RF, the largest cumulative reward value is shown in bold.

	N Input (kg/ha)	Irrigation (L/m^2)	N_l (kg/ha)	Yield (kg/ha)	RF 1	RF 2	RF 3	RF 4	RF 5
Baseline Policy	360	394	213	10772	984	1417	1269	700	338
Trained Policy 1	200	120	36	10852	**1425**	1557	1538	1267	**1673**
Trained Policy 2	200	732	59	11244	813	**1619**	971	655	1020
Trained Policy 3	19920	108	6205	10865	-1.4e4	-1.4e4	**1598**	-3.0e4	-4.9e4
Trained Policy 4	160	102	35	10358	1398	1510	1524	**1272**	1635
Trained Policy 5	200	138	39	10926	1417	1568	1575	1259	1651

Table 2: Performance comparison between the RL-trained policies (experts) and their corresponding IL-trained policies

	N Input (kg/ha)	Irrigation (L/m^2)	N_l (kg/ha)	Yield (kg/ha)	RF 1	RF 5
Baseline Policy	360	393.7	212.6	10771.5	984.4	337.6
RL-Trained Policy 1 (Full)	200	120	35.5	10852.4	1424.7	N/A
IL-Trained Policy 1 (Partial)	200	138	37	10870.0	1407.7	N/A
RL-Trained Policy 5 (Full)	200	138	39.2	10926.1	N/A	1651.0
IL-Trained Policy 5 (Partial)	200	138	39.2	10926.1	N/A	1651.0

agent receives the state of the environment, s_t, and chooses the action a_t, which consists of N fertilization amount N_t and irrigation amount W_t. Given s_t and a_t, the reward $r_t(s_t, a_t)$ is defined as:

$$r_t(s_t, a_t) = \begin{cases} w_1 Y - w_2 N_t - w_3 W_t - w_4 N_{l,t} & \text{if harvest at } t, \\ -w_2 N_t - w_3 W_t - w_4 N_{l,t} & \text{otherwise,} \end{cases} \tag{1}$$

where w_i $(i = 1, 2, 3, 4)$ are weights to balance the input, yield Y, and nitrate leaching $N_{l,t}$. The goal of the RL agent is to find the optimal policy $\pi(s_t, \theta_t)$ that maximizes the future discounted return, which is defined as $R_t = \sum_{\tau=t}^{T} \gamma^{\tau-t} r_\tau$ with $\gamma \in (0, 1]$ being a discount factor.

With recently developed Gym-DSSAT [5], we achieve daily interaction between the crop environment simulated via DSSAT and the agent, which enables the RL-based management policy training.

Imitation learning (IL) trains the agent to perform a task by mimicking the behavior of an expert [3]. For the crop management problem, not all state variables from the simulator can be observed or measured by farmers. Management policies should only utilize state variables accessible to farmers for real-world deployment. Given any state s, denote s^o as the observable state which contains variables from s that are measurable in the real world. For the IL-based training under partial observation, on each day t, the agent receives s_t^o and aims to learn an optimal policy $\pi(s_t^o, \theta)$ that generates an action a_t^o that is the same as a_t, where a_t is the action determined by the expert given an observation of s_t. The RL-trained policies under full observation are used as the expert during IL training.

3 EXPERIMENTS AND RESULTS

Experiments were conducted using the simulation of the maize crop in Florida in 1982. For comparison with the trained policies, we used a baseline policy that follows a Florida corn production guide written by domain experts [8]. For RL-based training, we used Deep Q-network [2] and tested with five different reward functions (RFs) to demonstrate the adaptability of our framework to different targets. The details of the RF design are given in Table 3, where both economic and environmental factors were considered in the RF design. The training results can be found in Table 1, and we achieved five trained policies using five RFs. The results show that given an RF to compute the cumulative rewards of different trained policies, the largest reward is always achieved by the policy trained with this particular RF (e.g., Trained Policy 1 achieves the highest cumulative reward with RF 1), except for the case of RF 5, where Trained Policy 5 still achieves a much larger reward than the baseline policy. Thus, given a specific target represented by a corresponding RF, we can always apply RL-based training to find an optimal management policy.

Table 3: Weights used in each reward function defined by (1)

	w_1 (Y)	w_2 (N_t)	w_3 (W_t)	w_4 ($N_{l,t}$)	Note
RF 1	0.158	0.79	1.1	0	Economic profit
RF 2	0.158	0.79	0	0	Free water
RF 3	0.158	0	1.1	0	Free N fertilizer
RF 4	0.158	1.58	1.1	0	Doubled N price
RF 5	0.2	1	1	5	With N Leaching

For the IL-based training, we used RL-Trained Policy 1 and RL-Trained Policy 5 as experts. The results are shown in Table 2. The IL-Trained Policy 1 under partial observation achieves a cumulative reward of 1407.7 with only a negligible decrease of 1.4% in the cumulative reward compared with the RL-Trained Policy 1. In addition, the IL-Trained Policy 5 achieves exactly the same results as the RL-Trained Policy 5. In conclusion, IL can help find crop management policies that behave very closely to RL-trained policies but require much fewer state variables.

ACKNOWLEDGMENTS

This work was supported by the C3.ai Digital Transformation Institute and NSF under the RI grant #2133656.

REFERENCES

[1] Chace Ashcraft and Kiran Karra. 2021. Machine learning aided crop yield optimization. *arXiv preprint arXiv:2111.00963* (2021).

[2] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. 2015. Human-level control through deep reinforcement learning. *Nature* 518, 7540 (2015), 529–533.

[3] Takayuki Osa, Joni Pajarinen, Gerhard Neumann, J Andrew Bagnell, Pieter Abbeel, and Jan Peters. 2018. An algorithmic perspective on imitation learning. *Foundations and Trends® in Robotics* 7, 1-2 (2018), 1–179.

[4] B.V.S. Reddy, P Sanjana Reddy, F Bidinger, and Michael Blümmel. 2003. Crop management factors influencing yield and quality of crop residues. *Field Crops Research* 84, 1-2 (2003), 57–77.

[5] Gautron Romain, Preux Philippe, Bigot Julien, Maillard Odalric-Ambrym, Emukpere David, et al. 2022. gym-DSSAT: a crop model turned into a Reinforcement Learning environment. *arXiv preprint arXiv:2207.03270* (2022).

[6] Lijia Sun, Yanxiang Yang, Jiang Hu, Dana Porter, Thomas Marek, and Charles Hillyer. 2017. Reinforcement learning control for water-efficient agricultural irrigation. In *2017 IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC)*. 1334–1341.

[7] Ran Tao, Pan Zhao, Jing Wu, Nicolas F Martin, Matthew T Harrison, Carla Ferreira, Zahra Kalantari, and Naira Hovakimyan. 2022. Optimizing Crop Management with Reinforcement Learning and Imitation Learning. *arXiv preprint arXiv:2209.09991* (2022).

[8] David Wright, Ian Small, Cheryl Mackowiak, Zane Grabau, Pratap Devkota, and Silvana Paula-Moraes. 2022. Field corn production guide: SS-AGR-85/AG202, rev. 8/2022. *EDIS* 2022, 4 (2022).

[9] Jing Wu, Ran Tao, Pan Zhao, Nicolas F Martin, and Naira Hovakimyan. 2022. Optimizing nitrogen management with deep reinforcement learning and crop simulations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 1712–1720.