SAND2022-11021C

A SWAT-based Reinforcement Learning Framework for Crop Management

Anonymous submission

Abstract

Crop management involves a series of critical, interdependent decisions or actions in a complex and highly uncertain environment, with distinct spatial and temporal variations. Managing resource inputs such as fertilizer and irrigation in the face of climate change, dwindling supply and soaring prices is nothing short of a Herculean task. The ability of machine learning to efficiently interrogate complex, nonlinear, and high-dimensional datasets can revolutionize decision making in agriculture. In this paper we introduce a reinforcement learning (RL) environment that leverages the dynamics in the Soil and Water Assessment Tool (SWAT) and enables management practices to be assessed and evaluated on a watershed-level. This drastically saves time and resources that would have been otherwise deployed during a full-growing season. We consider crop management as an optimization problem where the objective is to produce higher crop yield while minimizing the use of external farming inputs (specifically, fertilizer and irrigation amounts). The model is naturally subject to environmental factors such as precipitation, solar radiation, temperature, and soil water content. By using a case study of corn production in central Texas, we demonstrate the utility of our framework by developing and benchmarking various decision-making agents following management strategies informed by standard farming practice and state-of-the-art RL algorithms.

Introduction

As global demand for agricultural products increases and resources such as land, water, and nitrogen fertilizers become scarce and costly, farming practices that produce higher yields on less acreage and input resources have become more appealing (Alexandratos and Bruinsma 2012; Muller et al. 2017; Ranganathan et al. 2018). The advent of precision agriculture, an innovative approach to modern farming that has been spurred by the rapid development of agricultural technology, has enabled many farmers to maximize crop yields without having to extend arable land or increase farming inputs (Dutia 2014). Each growing season, farmers all over the world make decisions critical to maximizing yields such as crop selection, planting and harvesting scheduling, and application of fertilizers and irrigation. Complexities are exacerbated by a changing climate and the need to minimize environmental impacts, while achieving global food security in the face of increasing populations.

The wide adoption of agricultural technology has provided many researchers and practitioners in machine learning with opportunities to aid farmers with crop and resource management. The creation and availability of large datasets containing weather, soil, and crop data has made it possible for machine learning to be applied to tasks ranging from moisture and crop yield prediction to crop disease detection using satellite image data ((Chlingaryan, Sukkarieh, and Whelan 2018; Gandhi 2022; Liakos et al. 2018)). Various approaches to meet the growing demand of agricultural products have been proposed. These include increasing crop production on current arable land, adopting greener farming practices, and reducing food consumption and wastage (Muller et al. 2017). We consider the first approach and make contributions towards optimizing crop yields using machine learning.

Reinforcement learning (RL) has proven its utility in decision making across multiple domains such as healthcare, engineering, and games. Agriculture use cases have focused on optimizing for crop production subject to various resource constraints such as fertilizer and water usage ((Binas, Luginbuehl, and Bengio 2019; Elavarasan and Vincent 2020; Overweg, Berghuijs, and Athanasiadis 2021; Wu et al. 2021)). This paper employs RL to autonomously learn the optimal set and distribution of actions to direct crop growth cognizant of spatial and temporal variations. The Soil and Water Assessment Tool (SWAT) (Arnold et al. 1998; Gassman et al. 2007) numerical simulator was used to generate representative soil-water-plant synthetic data. Four decision-making agents utilizing different strategies to maximize crop yields were evaluated on the data. The aforementioned model incorporates dynamical representations of economic costs, environmental impacts, and crop yields, and produces a reward that characterizes the effects of different agent actions on crop production. We evaluated the agents and assessed the impact and practicality of the strategies they pursued. By favoring less frequent and minimal amounts of agricultural inputs, it generates a holistic framework to minimize environmental impact and operational costs while recovering optimal yields.

Contributions In this paper, we propose the SWATGym environment, an open-source reinforcement learning environment modeled after the widely used Texas A&M SWAT

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

model (Arnold et al. 2012). As our key contributions, we

1. formulate crop management as a decision-making problem characterized by an episodic Markov Decision Process
2. provide SWATGym, an OpenAI Gym environment that simulates crop growth and incorporates geographic and weather data to model the complex soil-water-plant-atmosphere system as described by the Texas A&M SWAT model
3. evaluate state of the art RL-based strategies and provide a simple API to apply and evaluate custom crop management strategies
4. highlight some promising research directions in this environment, e.g. fine-tuning it to capture the spatial-temporal variability of the seasons.

Background and Related Work

Nitrogen and water are the most limiting factors for crop production. Small amounts of each may lead to less production and excess amounts are not only costly but can lead to environmental harms downstream. Applying correct levels of nitrogen fertilizer and irrigation amounts can optimize yield and profit while reducing the amount of excess nutrients washing away to local groundwater sources, rivers, and lakes. However, determining the input application rates and schedules that meet environmental and economic goals of farming is not easy. It is complicated by uncertainties such as weather, resource availability and costs, among other factors. As a result, techniques that can inform crop management under such conditions are pivotal. Moreover, controlling for the environmental and financial impacts of crop management is central to the tenets of sustainable agriculture.

Excessive nitrogen fertilizer application elevates nitrate levels in soil, which aided by percolating irrigation water may leach down to groundwater. High nitrate levels are considered harmful to humans and livestock. has been found to affect water quality and excess irrigation washes away fertilizer and other potentially harmful substances into rivers and streams. For sustainability efforts, models that consider long-term impact of management practices on water quality are pivotal.

Reinforcement Learning for Crop Management Reinforcement learning (RL) involves a decision making agent interacting with an environment in order to learn a reward-maximizing strategy. At each discrete time step t and given a state $s \in \mathcal{S}$, the agent selects an action $a \in \mathcal{A}$ informed by its policy $\pi : \mathcal{S} \rightarrow \mathcal{A}$. The agent then receives a reward r and a new observation of the environment s'. The reinforcement learning objective is to find the optimal policy π^* that maximizes the cumulative reward.

RL has only recently been applied to crop management with notable works including (Binas, Luginbuehl, and Bengio 2019) which explores applying RL to sustainable agriculture as well as (Ashcraft and Karra 2021) and (Overweg, Berghuijs, and Athanasiadis 2021) which optimizes for crop yield subject to irrigation and fertilizer management actions.

Crop Growth models Simulated environments are fast becoming the key to the creation of state-of-the-art algorithms for various learning tasks as evidenced by the success of RL algorithms in games such as Go (Silver et al. 2016) and in robotics. Most existing environments are designed for managing either irrigation operations or fertilizer operations. The most notable environment is the Python Crop Simulation Environment (de Wit 2018) which houses various crop models such as WOFOST (Van Diepen et al. 1989) and LINTUL3 (Shibu et al. 2010) and has inspired environments such as CropGYM model (Overweg, Berghuijs, and Athanasiadis 2021). One drawback of PCSE is that it requires one to provide parameters for the various model components (soil, crop, weather) as well as specify the agromanagement activities that will take place on the field to be simulated.

The CropGym environment simulates winter wheat growth in the Netherlands, with a particular focus on fertilizer management. Other seminal models only focus on irrigation management and include the SIMPLE model (Zhao et al. 2019), implemented as an OpenAI environment in (Ashcraft and Karra 2021) and applied to potato growth simulation in Washington State, as well as the paddy rice simulation environment described in (Chen et al. 2021).

SWAT The Texas A&M Soil and Water Assessment Tool was developed to simulate physical processes such as crop growth, soil water balance, and nutrient cycling in a watershed (Arnold et al. 1998; Gassman et al. 2007). SWAT primarily considers two production levels: potential production, which represents estimated growth under optimal conditions, and actual production, which is limited by factors such as temperature stress, nutrient and water availability. SWAT uses a simplified version of the EPIC crop model (Williams et al. 1989) to compute most of the crop-related variables.

The SWATGym Environment

The main contribution of our paper is SWATGym, a reinforcement learning environment based on SWAT that simulates a crop's phenological development, growth, and yield on a daily basis by taking into account the effects of factors such as nutrient cycling, water availability, and temperature. We simulate crop growth and yield as a function of soil conditions, climatic data, and agricultural input management strategies.

SWATGym is the first Python-based implementation of SWAT, which is primarily written in FORTRAN and is not readily available for reinforcement learning applications. Our environment is built on top of OpenAI Gym framework, the gold standard framework for developing reinforcement learning environments (Brockman et al. 2016). The environment has a continuous state space comprising of 14 state variables describing various processes related to weather, soil, crop, and hydrology dynamics (see Table 2). It also has a continuous multidimensional action space. At time t, the action is given by $a_t = [F_t, I_t]$, where F and I represent fertilizer and irrigation amounts applied at that time.

Environment	Fertilizer	Irrigation
(Ashcraft and Karra 2021)	✗	✓
CropGym (Overweg, Berghuijs, and Athanasiadis 2021)	✓	✗
(Chen et al. 2021)	✓	✗
SWATGym	✓	✓

Table 1: Comparison of crop growth reinforcement learning environments.

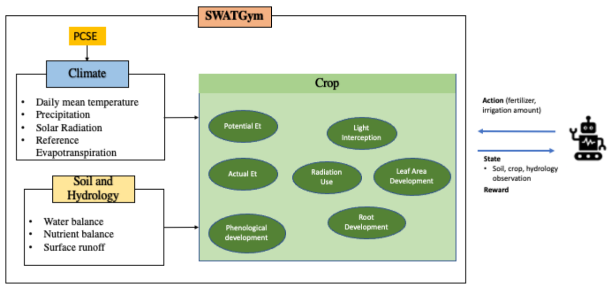

Figure 1: The SWATGym environment is the first SWAT-based reinforcement learning environment. Agents interact with the environment to learn crop management strategies in simulation

.

Reward Function

SWATGym incorporates dynamical representations of crop yields and corresponding economic costs of input resources as well as their environmental impact. It produces a reward that characterizes the effect of different choices of actions on crop production.

The reward at each time step is then computed as

$$r_t = yld_t - \alpha F_t - \beta I_t, \quad (1)$$

where yld is the estimated crop yield on a particular day and α and β are penalty terms associated with the cost of applying fertilizer F and irrigation I. In our case, $\alpha = 2.43$ and $\beta = 0.16$.

Crop Growth Simulation

SWATGym simulates crop growth for a full growing season (120 days). Users have the option to specify the location and simulation start date, otherwise default locale values for Temple, Texas, and a season start date of April 15, 2021, will be used. The simulation starts with crop emergence and ends with harvest.

The environment operates on a daily time step. An episode begins on the specified/default simulation date or when the environment is reset, and ends after the specified/default duration of the growing season (harvest day). The environment also has the option to save all relevant data about the current growing season, including weather observations, crop states, soil and hydrology balances, as well as actions taken and yield achieved to date. This feature enables the collection of expert data, which can be used for other tasks including offline learning. Below we show a plot of precipitation observed during one simulation run.

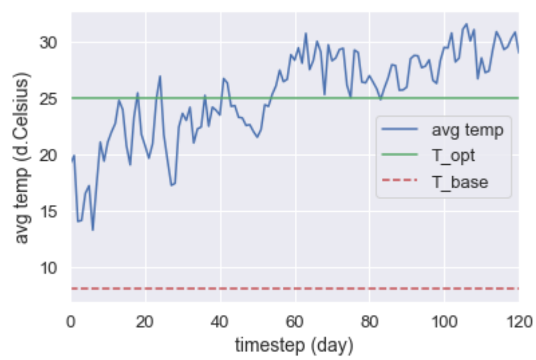

Figure 2: Temperature variation during one growing season, along with the crop-specific base temperature i.e., temperature at or below which crop growth ceases, and optimal temperatures for corn production. $T_{base} = 8°C$ and $T_{opt} = 25°C$

State-space dynamics

SWATGym has several major modules such as hydrology, weather, soil temperature, crop growth, and agricultural management (Arnold et al. 1998). It includes state variables describing time evolution of hydrological, soil, and crop variables. At time (day) t, the plant state variables are described by

$$z_p(t) = [LAI, BIO, E_a, N_{strs}, W_{strs}, T_{strs}], \quad (2)$$

and the soil state variables by

$$z_s(t) = [SW, RCN, DN, N_{up}]. \quad (3)$$

LAI is the leaf area index, BIO is the cumulative biomass, E_a is the actual evapotranspiration rate, and $N_{strs}, W_{strs},$ and T_{strs} are growth factors related to stress caused by nitrogen, water, and temperature on the plant. For soil state z_s, SW is the available soil water content, RCN is the daily surface runoff curve number, DN is the denitrification rate, and N_{up} is the nitrogen uptake.

We also define a vector of climatic inputs,

$$\xi(t) = [P, Et, Ta, Rd], \quad (4)$$

where P is the precipitation received on that day, Et is the reference evapotranspiration rate, Ta is average daily air temperature, Rd is daily solar radiation. Therefore, the plant state at time $t+1$ is given explicitly as a nonlinear function of the state and the input climatic variables at time t:

$$z_p(t+1) = f_p\left(z_p(t), z_s(t), \xi(t)\right),\ t \in [0, T],\ z_p(0) = x \tag{5}$$

where, for a given initial state x, $z_p(t)$ is the vector with the plant state variables at time t, $z_s(t)$ is the vector with the soil state variables; and $\xi(t)$ is the vector of the climatic inputs provided by PCSE (de Wit 2018) for a specific location and day.

Excluding daily weather data, most of the state variables are propagated by equations provided in the SWAT2009 Theory Documentation (Neitsch et al. 2011). Below we highlight a few of the state variables:

1. **Phenology:** Similar to the original SWAT model, we express crop growth/phenological development in term of heat units, which are driven by daily mean temperature. Growth is accelerated at or above the optimal temperature for the crop, T_{opt}, and is slowed or stopped at or below the base temperature T_{base}. A crop's phenological development is based on daily heat unit accumulation, given by

$$HU_i = \bar{T}_i - T_{base}, \quad \forall i \in 1, \ldots, T,\ T > 1 \tag{6}$$

where HU is the value of heat units and $\bar{T}_i$ is the average air temperature in °C on day i.

The fraction of potential heat units accumulated for a given day d is given by

$$fr_{PHU} = \frac{\sum_{i=1}^{d} HU}{PHU} \tag{7}$$

where PHU is the total number of heat units required for a plant to reach maturity. This is often calculated from planting date to harvest date (on the last day T) if not known beforehand.

2. **Potential Growth:** Other factors related to plant growth that are modeled by SWATGym include leaf area development, light interception, and a plant-specific radiation use efficiency metric (which measures conversion of intercepted light into biomass).

(a) *Leaf Area Development:* The leaf area index (LAI) is the area of green leaf per unit area of land. It is computed as a function of crop canopy height by

$$\Delta LAI_i = K_f\left(1 - e^{5*(LAI_{i-1} - LAI_{max})}\right) \tag{8}$$

where $K_f = LAI_{max}(fr_{LAImax,i} - fr_{LAImax,i-1})$ and $LAI_0 = 0$.

$$LAI_i = LAI_{i-1} + \Delta LAI_i \tag{9}$$

where h_c is the canopy height. For corn, $LAI_{max} = 3$ and $fr_{\text{PHU, sen}} = 0.9$ (Arnold et al. 2012).

(b) *Light Interception:* Using Beer's law, the amount of daily solar radiation intercepted by the plant is computed as

$$H_{\text{phosyn}} = 0.5 H_{\text{day}}\left(1 - exp(-k_\ell LAI)\right) \tag{10}$$

where H_{phosyn} is the amount of intercepted photosynthetically active solar radiation (MJ/m^2), H_{day} is incident total solar, k_ℓ is light extinction coefficient and LAI is leaf area index.

(c) *Biomass Production:* Radiation Use Efficiency (RUE) is defined for each plant species and is independent of the plant's growth stage. The potential increase in total plant biomass on a given day is given by

$$\Delta bio = RUE \cdot H_{\text{phosyn}} \tag{11}$$

The total plant biomass on a given day d is subsequently given by

$$bio = \sum_{i=1}^{d} \Delta bio_i, \quad d \le T \tag{12}$$

3. **Crop Yield:** SWATGym computes crop yield as the product of a plant's above ground biomass (and its roots, if they a harvest-able product) and harvest index, which is defined as the fraction of above-ground plant dry biomass removed as dry economic yield (with values typically between 0 and 1). HI, the potential harvest index for a given day in the plant's growing season is computed using the following relationship:

$$HI = HI_{\text{opt}} \frac{100 fr_{\text{PHU}}}{100 fr_{\text{PHU}} + exp(11.1 - 10 fr_{\text{PHU}})}, \tag{13}$$

where HI_{opt} is the potential harvest index at time of maturity, and fr_{PHU} is the fraction of potential heat units accumulated for the plant on a given day in the growing season. If needed, the actual harvest index can be derived from Equation 13 by taking water deficiency into account.

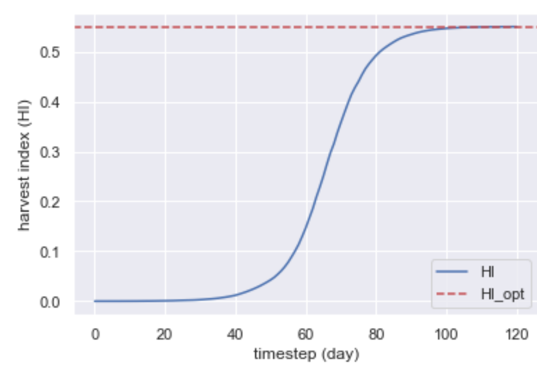

Figure 3: Potential harvest index for corn production during one growing season

Overall, estimated yield is given by

$$yld = \begin{cases} bio_{\text{ag}} HI, & \text{for } HI \le 1 \\ bio\left(\frac{HI}{HI+1}\right), & \text{otherwise,} \end{cases} \tag{14}$$

where yld is the crop yield (kg/ha), bio_{ag} is the aboveground biomass (kg/ha), HI is the harvest index, and bio

is the total plant biomass on the day of harvest. bio_{ag} is computed as follows,

$$bio_{\text{ag}} = (1 - fr_{\text{root}})\, bio, \tag{15}$$

where $fr_{\text{root}} = 0.4 - 0.2 fr_{\text{PHU}}$ is the fraction of total biomass in the roots on harvest day and fr_{PHU} is the fraction of potential heat units accumulated for the plant on a given day in the growing season.

4. **Growth Constraints:** Plant growth may be affected by insufficient or excess water, nutrients, and extreme temperatures. Stress factors are typically 0 under normal conditions and approach 1 as growth conditions veer from the optimal. Equations propagating these factors are provided in (Neitsch et al. 2011).

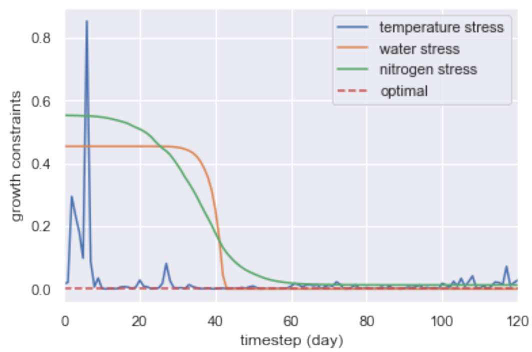

Figure 4: Growth related stress factors during one growing season

Sample Usage

The SWATGym complies with the OpenAI Gym framework. Below we show a sample code that runs a random agent on SWATGym but can be configured for any custom crop management strategy.

```
from swat_env import SWATGym

env = SWATGym()
# initialize the env
state, reward, done, info = env.reset()
while not done:
    action = env.action_space.sample()
    observation, reward, done, info = \
    env.step(action)
```

Source Code

SWATGym is released as a free and open-source environment under the terms of the MIT License. The latest version of the code is publicly available at https://github.com/ and includes a detailed setup process as well as reproducibility steps.

While SWAT accounts for a variety of crop species and environmental processes, we primarily consider corn production (in the central Texas region) as proof of concept and in order to utilize existing crop parameters in our formulation. We further limit ourselves to key processes such as surface runoff, water balance, and denitrification. For a detailed description of the physiological processes, refer to (Neitsch et al. 2011).

We provide a simplified implementation of SWAT that makes the following assumptions:

- all climatic and agromanagement inputs are applied uniformly and daily, over a single growing season (typically 120 days for corn).
- all soil layers (except the surface layer, top 10mm) have largely the same characteristics i.e. homogeneous soil profile.
- no/negligible percolation and bypass flow exiting the soil profile at the bottom; no lateral and base flow (which means soil water content is computed as $SW_t = SW_0 + \sum_{i=1}^{t}(R_{day} - Q_{surf} - Ea)$, where R_{day} is precipitation in mm, Q_{surf} is surface runoff, and Ea is actual crop evapotranspiration).
- no growth-reducing factors such as weeds and pests and negligible growth impact of all other nutrients besides nitrogen.

Experiments

In addition to the SWATGym environment, we also evaluate a selection of crop management strategies on the environment to demonstrate that reinforcement learning agents can learn useful crop management strategies. We benchmark the following agents:

- Deep Deterministic Policy Gradient (DDPG) - is a state of the art reinforcement learning algorithm for continuous control tasks (Lillicrap et al. 2015). DDPG has been widely successful in data rich applications.
- Twin Delayed Deep Deterministic policy gradient algorithm (TD3) (Fujimoto, Hoof, and Meger 2018) builds on DDPG and applies various modification to improve its stability and learning performance.

We also provide performance measurements of three baseline agents

- Random Agent - this is a dynamics-agnostic agent which selects random amounts of fertilizer and irrigation to apply at each time step.
- Standard Practice Agent - applies predetermined amounts of inputs on scheduled days during the early, mid, and late stage of the growing season. This corresponds to traditional farming methods of applying inputs during different phases of the crop's growth.
- Reactive Agent - applies high concentrations of fertilizer and irrigation whenever soil water content is below a certain threshold and nitrogen levels are depleted.

Evaluation

We train each agent for a full corn growing season, equivalent to 120 days, with evaluations done every 7 days. Each evaluation consists of 10 episodes. We provide for comparison the means and standard deviation values of all algorithms across 5 repetitions of the experiment.

Each method was evaluated as follows. An episode starts when the environment resets to an initial starting state and

Observation	Unit
Mean air temperature	°C
Precipitation	mm
Reference Evapotranspiration	mm
Solar Radiation	MJ/mm^2
Mean Vapor Pressure	hPa
Actual Evapotranspiration	mm
Water balance	mm
Daily runoff curve number	-
Leaf area index	-
Nitrogen Uptake	kg/ha
Denitrification	kg/ha
Nitrogen stress factor	-
Temperature stress factor	-
Water stress factor	-

Table 2: Environment's observations include weather inputs, plant state variables and soil state variables.

Method	Performance
Random	
Standard	2
Reactive	3
DDPG	4
TD3	5

Table 3: Average return $\pm$ one standard deviation across 5 repetitions of the experiment for one full growing season

Method	Performance
Random	-4547.152 ± 146.386
Standard	4396.402 ± 13.116
Reactive	279.121 ± 32.234
DDPG	22217.155 ± 4628.495
TD3	3263.175 ± 2295.258

Table 4: Average return $\pm$ one standard deviation over 17 policy evaluations ran across 5 repetitions of the experiment. Each evaluation has 10 episodes.

ends when the environment terminates, which happens when the growing season completes.

Results

First, we show the performance of the baseline methods over one growing season. Figure 5 shows that the standard

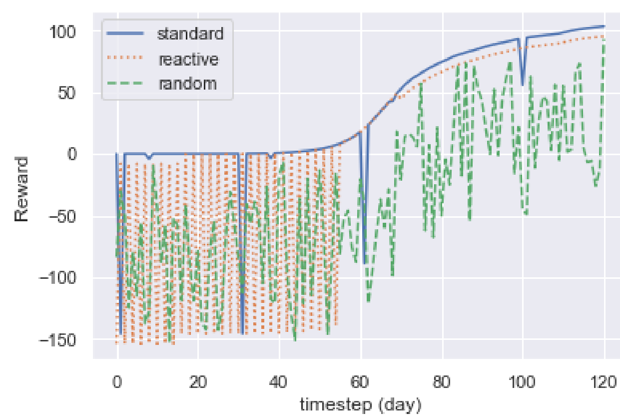

Figure 5: Performance of baseline methods on one full growing season

.

approach obtains the best performance overall compared to that of the Random agent and Reactive agent. We then benchmark this against the RL strategies and observe the following:

Beyond highlighting the potential of reinforcement learning strategies in facilitating sustainable crop management, our preliminary results demonstrate the value of SWATGym as a benchmark framework.

Our results are presented in Table and learning curves in Figure . Mean and 95% confidence interval computed over 5 seeds are reported.

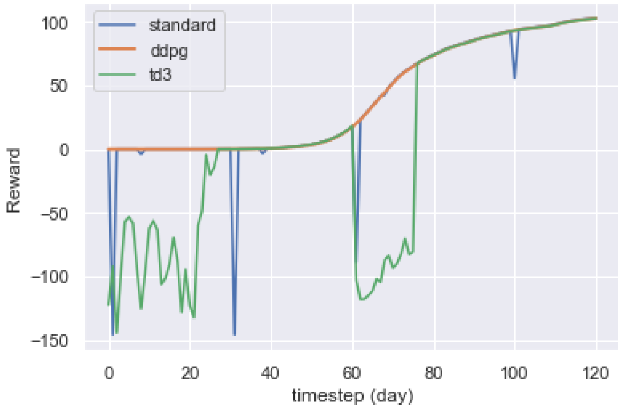

Figure 6: Performance of best baseline method against RL methods on one full growing season

.

Discussion and Research Directions

In this section we outline key challenges in modeling crop growth and briefly describe social challenges such as adoption of RL-based crop management strategies.

Challenges in crop modelling Accurately modeling crop growth is a key and active research problem. SWATGym has lots of room for improvement to increase its fidelity. Yet, one also has to balance simplicity. Future releases of the environment will extend the dynamics to include more processes such as groundwater seepage and to other applications such as offline reinforcement learning.

Seasonal spatial-temporal variability Most crops that are simulated are short-season crops. Regardless of the season duration, it is important to capture the spatial-temporal variability of the season. For example, evapotranspiration is higher in the summer than winter, so the crop management algorithm should apply more irrigation to match water loss during that period. Likewise, surface runoff could be higher in the spring or rain season (e.g. when snow melts/after it rains). So, the best strategy might be to apply less fertilizer to avoid leaching and waste or to apply it early in the growing season before runoff affects it.

Reality Gap SWATGym simulates crop growth daily and allows agents to select inputs on any given day. In practice, such operations are done over a week or more, depending on the size of the field and type of equipment used to irrigate or fertilize. Furthermore, the environment can be further constrained to limit the total amounts of inputs applied throughout the growing season and terminate whenever this threshold is reached.

Social Impact Through a pro bono social impact program which leverages technologies such as hybrid cloud and AI to enhance and scale non-profit and government organizations, plans are underway to deploy this framework in smallholding farms in central Texas. The goal is to help farmers make better decisions with regards to crop management and offer a platform where they can easily evaluate different strategies in order to optimize crop production. This highlights a challenging aspect of this work, reinforcement learning, and digital solutions to real-world problems: adoption. RL has widely been used in simulated environments but ever so rarely in real-world applications. Working with smallholding farmers will present unique opportunities to learn how to translate research ideas into useful real-world products.

Conclusion

Evaluating and comparing different crop management strategies in the real world is a costly and time-consuming task. Simulated environments offer a compelling solution and enable multiple strategies to be benchmarked simultaneously and without much of the cost overhead of the former approach. To this end, we introduced SWATGym, a reinforcement learning environment designed to make it easy to simulate crop growth and evaluate crop management strategies. SWATGym models crop growth processes from emergence to harvest and can be used to benchmark crop management strategies, which in turn can inform decision-makers towards sustainable agriculture practices. We hope that the framework introduced in this paper facilitates follow-up work and encourages researchers and practitioners in both reinforcement learning and agriculture to collaborate on developing better crop management strategies and contribute towards promoting sustainable agriculture.

References

Alexandratos, N.; and Bruinsma, J. 2012. World agriculture towards 2030/2050: the 2012 revision. *ESA Working Papers 12-03, Food and Agriculture Organization of the United Nations (FAO).*

Arnold, J.; Kiniry, J.; Srinivasan, R.; Williams, J.; Haney, E.; and Neitsch, S. 2012. Soil and Water Assessment Tool Input/Output file documentation version 2012. Technical report, Texas Water Resources Institute.

Arnold, J. G.; Srinivasan, R.; Muttiah, R. S.; and Williams, J. R. 1998. Large area hydrologic modeling and assessment part I: model development 1. *JAWRA Journal of the American Water Resources Association*, 34(1): 73–89.

Ashcraft, C.; and Karra, K. 2021. Machine Learning aided Crop Yield Optimization. *arXiv preprint arXiv:2111.00963*.

Binas, J.; Luginbuehl, L.; and Bengio, Y. 2019. Reinforcement learning for sustainable agriculture. In *ICML 2019 Workshop Climate Change: How Can AI Help*.

Brockman, G.; Cheung, V.; Pettersson, L.; Schneider, J.; Schulman, J.; Tang, J.; and Zaremba, W. 2016. Openai gym. *arXiv preprint arXiv:1606.01540*.

Chen, M.; Cui, Y.; Wang, X.; Xie, H.; Liu, F.; Luo, T.; Zheng, S.; and Luo, Y. 2021. A reinforcement learning approach to irrigation decision-making for rice using weather forecasts. *Agricultural Water Management*, 250: 106838.

Chlingaryan, A.; Sukkarieh, S.; and Whelan, B. 2018. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and electronics in agriculture*, 151: 61–69.

de Wit, A. 2018. PCSE: Python Crop Simulation Environment, http://pcse.readthedocs.io.

Dutia, S. 2014. Agtech: Challenges and opportunities for sustainable growth. *Available at SSRN 2431316*.

Elavarasan, D.; and Vincent, P. D. 2020. Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications. *IEEE access*, 8: 86886–86901.

Fujimoto, S.; Hoof, H.; and Meger, D. 2018. Addressing function approximation error in actor-critic methods. In *International conference on machine learning*, 1587–1596. PMLR.

Gandhi, R. 2022. Deep Reinforcement Learning for Agriculture: Principles and Use Cases. In *Data Science in Agriculture and Natural Resource Management*, 75–94. Springer.

Gassman, P. W.; Reyes, M. R.; Green, C. H.; and Arnold, J. G. 2007. The soil and water assessment tool: historical development, applications, and future research directions. *Transactions of the ASABE*, 50(4): 1211–1250.

Liakos, K. G.; Busato, P.; Moshou, D.; Pearson, S.; and Bochtis, D. 2018. Machine learning in agriculture: A review. *Sensors*, 18(8): 2674.

Lillicrap, T. P.; Hunt, J. J.; Pritzel, A.; Heess, N.; Erez, T.; Tassa, Y.; Silver, D.; and Wierstra, D. 2015. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*.

Muller, A.; Schader, C.; El-Hage Scialabba, N.; Brüggemann, J.; Isensee, A.; Erb, K.-H.; Smith, P.; Klocke, P.; Leiber, F.; Stolze, M.; et al. 2017. Strategies for feeding the world more sustainably with organic agriculture. *Nature communications*, 8(1): 1–13.

Neitsch, S. L.; Arnold, J. G.; Kiniry, J. R.; and Williams, J. R. 2011. Soil and Water Assessment Tool theoretical documentation version 2009. Technical report, Texas Water Resources Institute.

Overweg, H.; Berghuijs, H. N.; and Athanasiadis, I. N. 2021. CropGym: a Reinforcement Learning Environment for Crop Management. *arXiv preprint arXiv:2104.04326*.

Ranganathan, J.; Waite, R.; Searchinger, T.; and Hanson, C. 2018. How to sustainably feed 10 billion people by 2050, in 21 charts. *Insights, World Resources Institute*.

Shibu, M.; Leffelaar, P.; Van Keulen, H.; and Aggarwal, P. 2010. LINTUL3, a simulation model for nitrogen-limited situations: Application to rice. *European Journal of Agronomy*, 32(4): 255–271.

Silver, D.; Huang, A.; Maddison, C. J.; Guez, A.; Sifre, L.; Van Den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M.; et al. 2016. Mastering the game of Go with deep neural networks and tree search. *nature*, 529(7587): 484–489.

Van Diepen, C. v.; Wolf, J. v.; Van Keulen, H.; and Rappoldt, C. 1989. WOFOST: a simulation model of crop production. *Soil use and management*, 5(1): 16–24.

Williams, J.; Jones, C.; Kiniry, J.; and Spanel, D. A. 1989. The EPIC crop growth model. *Transactions of the ASAE*, 32(2): 497–0511.

Wu, J.; Zhao, P.; Tao, R.; Hovakimyan, N.; Marcillo, G.; Martin, N.; Ferreira, C.; Kalantari, Z.; and Hobbs, J. 2021. Optimization of Agricultural Management for Soil Carbon Sequestration based on Deep Reinforcement Learning and Large-Scale Simulations. In *NeurIPS 2021 Workshop on Tackling Climate Change with Machine Learning*.

Zhao, C.; Liu, B.; Xiao, L.; Hoogenboom, G.; Boote, K. J.; Kassie, B. T.; Pavan, W.; Shelia, V.; Kim, K. S.; Hernandez-Ochoa, I. M.; et al. 2019. A SIMPLE crop model. *European Journal of Agronomy*, 104: 97–106.