WOFOSTGym: A Crop Simulator for Learning Perennial Crop Management Strategies

Anonymous authors
Paper under double-blind review

Keywords: Reinforcement Learning, Crop Simulator, Agriculture, Benchmark

Summary

Daily crop management decisions play a crucial role in maximizing yield and economic returns while minimizing environmental impact. Optimal policies are unknown, making the environment well suited for reinforcement learning (RL) application. We introduce WOFOSTGym, a novel crop simulation environment that addresses key limitations of existing crop simulators for RL by supporting perennial crop modeling and enabling learning across multiple growing seasons and multiple farms. Expanding on prior work, WOFOSTGym allows RL agents to train in multi-year, multi-crop, and multi-farm settings to learn a range of agromanagement practices. Based on real world agromanagement problems, WOFOSTGym offers a suite of challenging tasks for RL agents including partial observability, non-Markovian dynamics, and delayed feedback. Its user-friendly design ensures that non-experts can still access a wide range of agromanagement decision problems. Our experiments showcase the behaviors that RL agents can learn across various crop varieties and soil types, demonstrating WOFOSTGym's potential for advancing RL applications in agromanagement decision support.

Contribution(s)

1. We present WOFOSTGym, a simulator for RL based on the WOFOST crop growth model (van Diepen et al., 1989). WOFOSTGym enables the development of agromanagement policies for multi-season perennial crops, advancing AI-driven decision support tools for agriculture.
Context: Previous work (Overweg et al., 2021; Gautron et al., 2022b) on crop simulators for RL has not considered the problem of agromanagement for perennial crops. To maintain a wide breadth of use, WOFOSTGym still supports simulating many single-season annual crop problems. The user-friendly design of WOFOSTGym enables non-expert users to create experiments with multiple farms and multiple crops, across a range of tasks with varying observability to reflect real world sensing challenges.
2. We modify the WOFOST crop growth model (CGM) to simulate the growth of perennial crops across multiple growing seasons.
Context: Bai et al. (2019) showed that WOFOST can be used to model the growth of the perennial jujube tree across multiple seasons. Inspired by their work, we modified the WOFOST CGM to support perennial growth within WOFOSTGym, and extended their work to model continuous multi-year growth with the addition of a dormancy phase. We update WOFOST nutrient modules to be able to investigate the impact of agromanagement decisions on the surrounding environment.
3. We introduce a novel application of Bayesian Optimization to calibrate the parameters of the WOFOST crop growth model (CGM) to increase model fidelity and compare our results with those of Zapata et al. (2017), who collected phenology data for 17 grape cultivars.
Context: To enable sim-to-real transfer in open-field agriculture, high fidelity CGMs are a requirement. However, calibrating the parameters is difficult and time consuming. State of the art methods in agronomy use linear regression techniques or Monte Carlo sampling (Lara et al., 2012). Our proposed Bayesian Optimization method offers a principled way of searching the CGM parameter space, producing better or comparable results with less computation and a limited amount of field collected data.

WOFOSTGym: A Crop Simulator for Learning Perennial Crop Management Strategies

Anonymous authors
Paper under double-blind review

Abstract

We introduce WOFOSTGym, a novel crop simulation environment designed for modeling perennial crops in multi-farm settings. Effective crop management requires optimizing yield and economic returns while minimizing environmental impact, a complex sequential decision-making problem well suited for reinforcement learning (RL). However, the lack of simulators for perennial crops in multi-farm contexts has hindered RL applications in this domain. Our WOFOSTGym addresses this gap by enabling RL agents to train in multi-year, multi-crop, and multi-farm scenarios, learning diverse agromanagement strategies. Our simulator offers a suite of challenging tasks for RL agents that include partial observability, non-Markovian dynamics, and delayed feedback. Its user-friendly design allows non-experts to explore a wide range of agromanagement decision problems. Our experiments demonstrate the learned behaviors of RL agents across various crop varieties and soil types, highlighting WOFOSTGym's potential for advancing RL-driven decision support in agriculture.

1 Introduction

During a growing season, farmers face many decisions about how to optimally manage their crops to increase yield while reducing cost and environmental impact (Javaid et al., 2023). For example, irrigation planning must account for constraints on water use, and optimizing irrigation application can improve crop yield (Elliott et al., 2014). Motivated by the promising results of using reinforcement learning (RL) in other areas of precision agriculture, there is increasing interest from researchers and government agencies in applying RL to crop management decision problems in open-field agriculture, especially for specialty crops (e.g. apples, grapes, cherries) (Astill et al., 2020; Gautron et al., 2022a). While RL has been explored as a tool for optimizing open-field crop management decisions (Wu et al., 2022; Tao et al., 2023), its real-world adoption is limited to controlled settings such as greenhouses (An et al., 2021; Wang et al., 2020) and crop monitoring (Din et al., 2022; Zhang et al., 2020). This paper aims to bridge this gap by presenting a simulator for perennial crops in a multi-farm setting.

Training RL agents in the real world to optimize agromanagement decisions is infeasible because growing seasons are too long, and unconstrained exploration can cause costly errors like crop loss and soil degradation (Tevenart & Brunette, 2021). Similar challenges in robotics and autonomous driving have been addressed with high-fidelity simulators, enabling RL applications (Kober et al., 2013; Kiran et al., 2022; Dauner et al., 2024; Todorov et al., 2012). While high-fidelity crop growth models (CGMs) (Boote et al., 1996) offer an approach to testing crop management policies for open-field agriculture, they are *not* designed to interact with RL algorithms and require substantial domain expertise. Existing agriculture simulators (Overweg et al., 2021; Gautron et al., 2022b) lack the expressiveness needed for perennial crop management (see Table 1). These simulators only simulate the growth of a single crop and cannot be customized to study other crops or sites without domain expertise of the underlying CGM. Open-field agriculture problems are often modeled as a

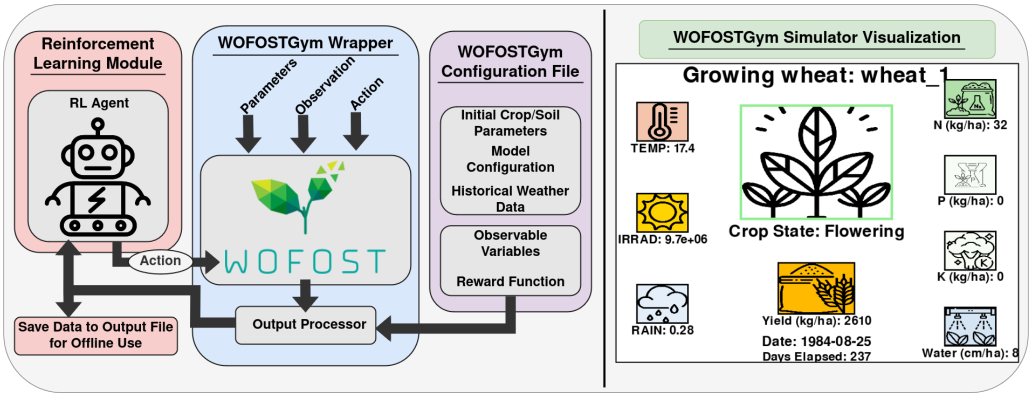


Figure 1: (Left) The structure of the WOFOSTGym simulator. WOFOSTGym provides a wrapper around the WOFOST Crop Growth Model with a variety of environments (limited nutrients and water, limited water, potential production, etc.) to train RL agents and generate data. Configuration files are well documented YAML files, and the user only needs to interact with a short YAML file for most use cases. (Right) A visualization of the growth of the wheat crop with the daily weather and cumulative agromanagement decisions rendered in WOFOSTGym.

partially observable environment but the current crop simulators do not allow for varying the hidden factors for partial observability and do not support the creation of a wide range of agromanagement tasks across crop and soil types, which limit the scenarios that can be modeled (Tao et al., 2023).

We present WOFOSTGym (see Figure 1), a crop simulator for perennial, multi-farm crop management. WOFOSTGym is built on the WOFOST CGM (van Diepen et al., 1989) to model the growth of perennial crops, and includes parameter sets for 23 annual crops and 2 perennial crops. As a step towards the high-fidelity modeling of perennial crop growth, we demonstrate the effectiveness of a Bayesian Optimization based method to calibrate CGM parameters to increase the fidelity of the phenological model for 32 grape cultivars. To make WOFOSTGym accessible to RL researchers, we prioritize usability through extensive customization, seamless integration with standard RL algorithms, and a thorough documentation. Our experiments showcase scenarios in WOFOSTGym where standard RL algorithms achieve optimal performance, alongside more complex cases that remain difficult—highlighting opportunities for developing new learning approaches. We also design agromanagement decision-making tasks in WOFOSTGym, illustrating both the potential and challenges of applying RL to this class of problems.

2 Related Work

Reinforcement Learning for Crop Management Building on RL's success in robotics, autonomous driving, and healthcare, there is growing interest in applying RL to optimize crop yield (Binas et al., 2019). While RL has proven effective in controlled greenhouse environments (An et al., 2021), its application in open-field agriculture remains limited due to reduced sensing capabilities and long growing seasons. To bridge this gap, several crop simulators have been developed. CropGym simulates winter wheat in a nitrogen-limited soil, via a Gym wrapper around a CGM (Overweg et al., 2021). Gym-DSSAT focuses on maize growth optimization through fertilization and irrigation decisions (Gautron et al., 2022b). CyclesGym, built around the Cycles CGM (Kemanian et al., 2022), focuses on learning crop rotation strategies for annual crops but is limited to soybeans and maize, lacking support for perennial crop modeling (Turchetta et al., 2022). Table 1 summarizes the capabilities of different crop simulators.

Existing crop simulators support RL training for fertilization and irrigation but *lack support for perennial crops*, a key research area (Gautron et al., 2022a). Additionally, customization is infeasible without expert knowledge of the underlying CGMs, since most CGMs are run through separate

Name	Perennial Crop Support	Multiple Crops and Farms	Easily Customizable	Models Crop Sub-processes
CyclesGym	✗	✗	✓	✓
CropGym	✗	✗	✗	✓
gym-DSSAT	✗	✗	✗	✓
FarmGym	✗	✗	✓	✗
WOFOSTGym (Ours)	✓	✓	✓	✓

Table 1: A comparison of available crop simulators based on four important desiderata for use with RL. A simulator is easily customizable if it does not require agriculture domain expertise to run different experiments. Modeling crop sub-processes (phenology, roots, stems, leaves, etc.) as it generally leads to a higher fidelity model.

executables. In contrast, WOFOSTGym offers easy domain customization for RL researchers while providing high-fidelity parameters for 23 annual and 2 perennial crops, high-fidelity model parameters for grape phenology for 32 cultivars, and access to diverse soil types and weather patterns.

Crop Growth Models Crop Growth Models (CGMs) simulate the growth of crops in varying environments subject to different agromanagement decisions (Jones et al., 2017). CGMs are typically one of three types: mechanistic, empirical, or hybrid. Mechanistic models simulate canopy or nutrient level crop processes and aim to validate scientific understanding of crop growth (Estes et al., 2013). Empirical models, in contrast, rely on observed field data, offering greater scalability with lower computational overhead (Di Paola et al., 2016). Hybrid crop models simulate crop growth using both mechanistic and empirical modeling decisions (Yang et al., 2004). Examples of widely-used CGMs include WOFOST (de Wit et al., 2019), DSSAT (Jones et al., 2003), APSIM (McCown et al., 1996), EPIC (Cabelguenne et al., 1990), CropSyst (Stockle et al., 1994), Cycles (Kemanian et al., 2022) and AquaCrop (Andarzian et al., 2011). The relevant features of these CGMs are highlighted in Appendix A. We note that none of the available CGMs support perennial crop modeling.

Our simulator is built on WOFOST, a hybrid CGM that models annual crop growth subject to nutrient (nitrogen, phosphorus, and potassium) and water-limited conditions (van Diepen et al., 1989). WOFOST is an ideal candidate to build on, since it can model the growth of perennial crops with a high fidelity (Bai et al., 2020; Shi et al., 2022). It also accounts for varying CO2 concentrations, making it valuable for climate-impacted agricultural research (Gilardelli et al., 2018). Additionally, its modular design facilitates modifications to crop process models (de Wit, 2024), and its Python implementation enables seamless integration with OpenAI Gym (Brockman et al., 2016).

3 Background

The Partially Observable Markov Decision Process (POMDP) is a popular framework to model partially observable environments (Kaelbling et al., 1998). POMDPs are well-suited for open-field agriculture problems, since many crop and soil-related features that are essential for defining the system's full state cannot be directly observed (Tao et al., 2023). Formally, a POMDP is a tuple $M = \langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \Omega, \mathcal{O}, \gamma \rangle$ where $\mathcal{S}$ is a set of states, $\mathcal{A}$ is a set of actions, $\mathcal{P} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$ is the transition kernel, and $\mathcal{R} : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ is the reward function. Ω is the set of possible observations and $\mathcal{O} : \mathcal{S} \times \mathcal{A} \times \Omega \to [0, 1]$ is the probability of obtaining observation o when taking action a in state s. Finally, γ is the discount factor that determines the importance of present versus future rewards. The RL agent computes a policy $\pi : \Omega \times \mathcal{A} \to [0, 1]$ that maximizes the expected sum of discounted rewards, $\mathbb{E}_{\rho^\pi}\left[\sum_{t=0}^{T} \gamma^t \mathcal{R}(s_t, a_t)\right]$, where ρ^π is the distribution of states and actions induced by the policy π and T is the time horizon.

4 WOFOSTGym

WOFOSTGym is built on the WOFOST CGM (van Diepen et al., 1989) and interfaces with the OpenAI Gym API to create a high-fidelity and easy-to-use crop simulator for RL. WOFOSTGym supports a variety of agromanagement decisions such as fertilizing, irrigating, planting, and harvesting. In the interest of clarity, we focus on fertilization and irrigation decisions in the rest of the paper, since they are popular tasks that are supported by all existing crop simulators. In these tasks, the agent must optimize fertilization and irrigation strategies that maximize the cumulative yield of a crop subject to a set of penalties or constraints over one or more growing seasons.

The rest of this section is organized as follows. We begin with an overview of the environment design, including modifications made to the WOFOST CGM to support perennial crop modeling and multi-farm environments, and highlight a few of the open-field agriculture tasks that can be modeled in WOFOSTGym. We then propose a model calibration method to fine-tune the model parameters of the WOFOST CGM to increase the fidelity of the simulator as a step towards sim-to-real transfer (Peng et al., 2018).

4.1 Environment Design

States and Observations The model state in WOFOSTGym is the concatenation of two features vectors, $\mathbf{x} = (x_1 \ldots x_{203})$ and $\mathbf{y} = (y_1, \ldots, y_7)$, where $\mathbf{x}$ contains the crop and soil state and $\mathbf{y}$ contains the weather state for a given day. However, it is unrealistic to have access to the vast majority of the state features because sampling is prohibitively expensive or time consuming. As a result, an RL agent receives a subset of the model state as observation $\mathbf{o} = (x_1, ..., x_n)$, with $n \ll 210$. In the multi-farm environments, the agent receives an observation for each farm and the daily weather observation is shared across farms.

Additionally, WOFOSTGym allows for domain randomization by 1) adding small amounts of random uniform noise to parameters in the WOFOST GGM, 2) allowing RL agents to train on different crops and soil types simultaneously, and 3) enabling RL agents to train on a wide breadth of historical weather data.

Action Space The action space of WOFOSTGym is a discrete space consisting of fertilization (nitrogen, phosphorus, and potassium) and irrigation actions. A single fertilization or irrigation action can be applied at each time step, at an amount and efficacy specified in the environment. The frequency of intervention can range from one day to multiple weeks, to reflect the reality that farmers cannot act on agromanagement decisions daily. WOFOSTGym simulates crop growth subject to limited Nitrogen, Phosphorus, Potassium, and water, and also includes other environments to model crop growth subject to unlimited water or fertilizer.

Reward WOFOSTGym rewards crop yield, as profitability is the primary driver of agromanagement adoption (Turchetta et al., 2022). Consequently, an agent should learn a policy that maximizes the crop yield. However, real-world agriculture requires balancing yield with constraints such as fertilizer costs, water usage limits, and surface runoff restrictions (see Section 5.1). For instance, exceeding 150 kg/ha of Nitrogen fertilizer could incur a significant penalty.

Opportunities for RL Research The agriculture domain poses a few key challenges for RL agents, making it a valuable testbed for RL research. 1) Actions with delayed feedback make it difficult for RL agents to learn the value of those actions. For example, in WOFOSTGym, the effect of applying fertilizer in March is not seen until the crop flowers in June. 2) Sparse rewards are difficult for an agent to learn the optimal policy (Vecerik et al., 2018). All reward functions in WOFOSTGym are sparse as the yield of the crop is unknown until the end of the episode. 3) As many crop and soil states are infeasible or prohibitively expensive to measure, WOFOSTGym's observations are subsets of the entire state space, which models partial observability.

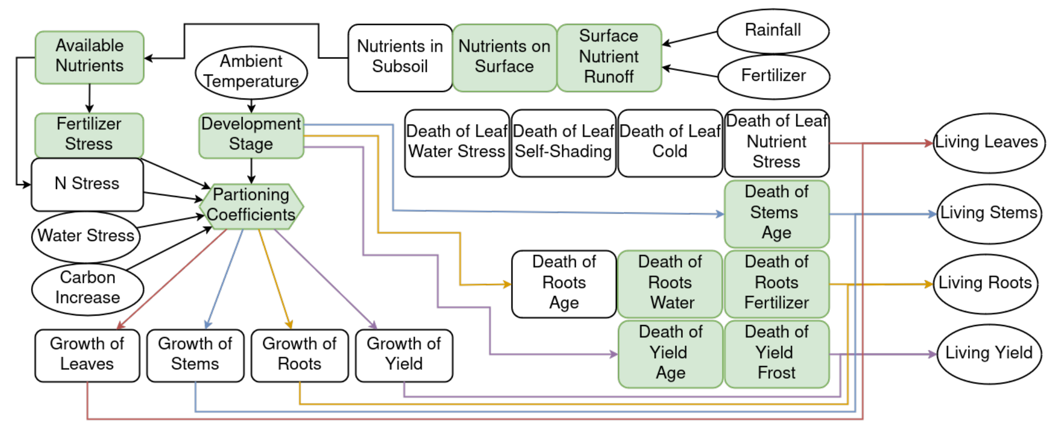


Figure 2: A simplified flowchart of the perennial crop growth in WOFOSTGym. Boxes highlighted in green denote additions or areas of substantial change to the underlying WOFOST CGM to support perennial crop growth. The nutrient balance module was changed to model a bilevel profile with nutrients on the soil surface and in the subsoil. The development stage of the crop is driven by the daily ambient temperature. The development stage determines how accumulated dry matter is partitioned to crop organs subject to available nutrients. The weight of the living organs (yield) is calculated as the accumulated difference between the growth and death rates.

Available Crops WOFOSTGym includes parameters that describe 23 annual crops: barley, cassava, chickpea, cotton, cowpea, faba bean, groundnut, maize, millet, mung bean, pigeon pea, potato, rapeseed, rice, onion, sorghum, soybean, sugar beet, sugarcane, sunflower, sweet potato, tobacco, and wheat and 2 perennial crops: pear (Wang et al., 2022) and jujube (Bai et al., 2020). Each crop contains between 1 and 10 varieties. WOFOST CGM parameters for each variety were calibrated empirically from field data (de Wit, 2025).

In addition to the high-fidelity models for 25 crops, WOFOSTGym also includes parameters for modeling the phenology of 32 grape cultivars. These cultivars are: Aligote, Alvarinho, Auxerrois, Barbera, Cabernet Franc, Cabernet Sauvignon, Chardonnay, Chenin Blanc, Concord, Dolcetto, Durif, Gewurztraminer, Green Veltliner, Grenache, Lemberger, Malbec, Melon, Merlot, Muscat Blanc, Nebbiolo, Petit Verdot, Pinot Blanc, Pinot Gris, Pinot Noir, Riesling, Sangiovese, Sauvignon Blanc, Semillon, Syrah, Tempranillo, Viognier, and Zinfandel. In Section 4.3 we introduce a novel application of Bayesian Optimization and show how historical data of grape phenology can be used to calibrate CGM parameters to increase model fidelity.

4.2 WOFOSTGym Modified Perennial Crop Growth Model

With the limited availability of standalone perennial CGMs, Bai et al. (2019) and Wang et al. (2022) showed that the WOFOST CGM (de Wit, 2024) could be adjusted to model the growth of pear and jujube crops across multiple growing seasons. Thus, to create a crop simulator with realistic dynamics for perennial crop modeling, the WOFOST CGM was a reasonable starting point. Below, we highlight the important changes that were made to WOFOST to enable perennial crop modeling.

Perennial Phenology We primarily focused on modifying the phenology submodule within the WOFOST CGM given the differences between annual and perennial phenology. Unlike annual crop growth, the phenology of perennial crops is characterized by a dormancy stage induced by day length in autumn and released by temperature in spring (Rohde & Bhalerao, 2007). We added parameters for dormancy induction based on day length, release temperature, and a minimum dormancy period. In the modified WOFOST CGM, dormancy can also be triggered by prolonged growth stagnation, indicating insufficient ambient temperature for crop growth (Jones et al., 1978).

Perennial Organ Growth In addition to differences in phenology, perennial crops exhibit differences in their visible growth organs (Thomas et al., 2000). The roots and stems of perennial crops do not die each fall while the leaves and storage organs regrow each season subject to intercepted light and nutrient uptake. The death rates of all crop organs are modeled as a function of the development stage of the crop (Lindén et al., 1996). Crucially, perennials exhibit a decrease in seasonal growth as they age (Munné-Bosch, 2007). While the exact mechanism for this decrease is difficult to quantify, the process can be modeled empirically by an increased maintenance respiration and decreased carbon conversion efficiency (Zhu et al., 2021). For a simplified explanation of how the principle observable features evolve throughout the course of a perennial crop simulation, see Figure 2.

Modified Nutrient Module A multi-layer nutrient balance is important for modeling the effects of fertilization stressors on the roots and stems and the change in nutrient partitioning (Albornoz, 2016). We change WOFOST's single-layer nutrient balance to a multi-layered nutrient balance within the soil module (He et al., 2013). When nutrients are applied via fertilization, they reside on the soil surface. As the simulation evolves, nutrients are absorbed into the subsoil and then the roots of the plant. When surface nutrient levels are too high, the partitioning of dry matter is changed to limit allocation to the storage organs in favor of stems and leaves (He et al., 2013).

WOFOSTGym Interaction with WOFOST As outlined in Figure 1, WOFOSTGym provides a Gym wrapper for RL agents to interact with the modified WOFOST crop growth model. WOFOSTGym passes a set of parameters that describe the crop and soil dynamics to WOFOST and the daily weather provided by the NASAPower API which has over 40 years of historical weather data from locations around the globe. Unlike other crop simulators such as Turchetta et al. (2022) and Gautron et al. (2022b), WOFOST's Python implementation and reduced overhead enables faster interaction with Gym, which leads to a dramatically reduced run time (see Table 3).

The WOFOST CGM simulates the growth of the crop for a length of time specified by WOFOSTGym and returns the state of the crop and soil each day. WOFOSTGym passes an observation of the state space along with the reward to an RL agent. A RL agent can use this observation and reward to choose the next best action, which is then passed through WOFOSTGym to WOFOST. With this construction, substantial modifications can be made to the underlying CGM without necessitating changes to the Gym environment or RL pipeline. As a result, a wide variety of crop configurations and tasks are able to be simulated in WOFOSTGym with minimal effort. Due to space limitations, we defer the discussion on the usability of WOFOSTGym to Appendix B.

4.3 Parameter Calibration for Crop Growth Models

Before a CGM can be used for prediction tasks it must be calibrated (Bhatia, 2014). While we have detailed the changes made to the WOFOST CGM for modeling perennial crops, we have not discussed how to obtain parameters for high-fidelity modeling of a specific crop. Parameters for CGMs are calibrated from field experiments (Berghuijs et al., 2024) and are optimized using regression techniques to find the parameters that give the best fit (Zapata et al., 2017). Parameter spaces for CGMs are high dimensional and highly non-linear (Sinclair & Seligman, 2000) and so the brute force and regression techniques that are used in the agronomy research community may be insufficient to find an optimal solution.

To address the shortcomings in current CGM calibration methods, we propose a Bayesian Optimization approach that requires minimal domain knowledge, matches or surpasses the performance of regression-based methods, and significantly reduces the number of simulations needed compared to Monte Carlo approaches (Lara et al., 2012). If historical crop data is available, Bayesian Optimization can be used with a WOFOSTGym to increase model fidelity, which is an important step towards obtaining successful sim-to-real transfer for agromanagement decision support.

Bayesian Optimization for Grape Phenology Calibration Washington State University maintains a research vineyard and collects historical phenology observations for 32 grape cultivars. Each

Cultivar	**Bud Break**				**Bloom**				**Veraison**			
	Ours	BB-T_b	BL-T_b	V-T_b	Ours	BB-T_b	BL-T_b	V-T_b	Ours	BB-T_b	BL-T_b	V-T_b
Cabernet Franc	4.0	6.1	6.2	13.8	3.5	3.1	2.9	4.7	7.7	6.7	7.1	9.6
Cabernet Sauvignon	5.0	8.7	10.5	10.7	5.2	5.8	5.7	6.4	9.8	6.6	7.0	7.9
Malbec	3.7	5.6	6.2	8.3	2.8	3.2	2.9	3.1	8.3	5.7	6.0	6.8
Pinot Noir	3.6	4.2	3.9	9.2	2.4	2.6	2.3	3.1	8.6	6.6	7.7	9.0
Zinfandel	3.7	6.8	9.0	11.8	3.8	4.3	4.0	5.3	6.0	4.1	3.8	4.9
Chardonnay	7.2	6.3	5.9	4.5	4.1	3.7	3.2	2.6	7.8	5.6	5.9	6.1
Chenin Blanc	5.0	6.1	6.2	5.4	3.8	4.8	4.6	4.3	8.5	9.2	9.4	9.9
Sauvignon Blanc	3.4	6.4	5.7	7.1	5.9	3.7	3.5	5.5	1.6	7.7	8.5	10.1
Semillon	4.7	6.0	7.0	11.7	2.7	6.0	5.8	5.5	8.8	11.2	11.6	12.2
Riesling	3.7	4.2	5.7	7.4	3.8	4.1	3.7	4.0	8.5	8.5	9.0	9.9
Average	4.4	6.0	6.6	9.9	3.8	4.1	3.9	4.4	7.4	7.2	7.6	8.6

Table 2: Root Mean Squared Error (RMSE) in days when predicting the key phenological stages (Bud Break, Bloom, and Veraison) in 10 grape cultivars using the Growing Degree Day (GDD) model (Zapata et al., 2015). The major columns are each of the three phenological stages. The minor columns are the RMSE in the GDD model with a given parameterization. Ours: Using parameter set tuned with Bayesian Optimization. BB-T_b: Parameter set tuned for Bud Break. BL-T_b: Parameter set for Bloom. V-T_b: Parameter set for Veraison. Values for BB:T_b, BL-T_b, and V-T_b correspond to columns 2, 3, and 4 in Table 6 in Zapata et al. (2017) and were all tuned using regression.

cultivar has between 6 and 15 years of data. With this data and the output of the phenology submodule in WOFOST, we crafted a loss function for our Bayesian Optimization approach as the Root Mean Squared Error (RMSE) in days between the predicted phenological stage and the true phenological stage. Our loss function also accounts for the RMSE in days of the previous phenological stage to prevent overfitting to the current stage.

Non-dormant grape phenology is subdivided into three crucial phenological stages: bud break, bloom, and veraison and is described by seven parameters in WOFOSTGym (Base Emergence Temperature, Max Effective Temperature, Temperature Sum for Bud Break, Temperature Sum for Bloom, Temperature Sum for Veraison, Temperature Sum for Ripe). Across 500 iterations, we sample parameters from the underlying Gaussian Processes using the expected improvement acquisition function that correspond to Bud Break, Bloom, and Veraison. We compute the loss for the selected parameters and update the underlying RBF kernel, as is standard in Bayesian Optimization. We maintain the best fit parameters for each stage which results in a parameter set that minimizes the RMSE across all stages.

We compare our results from Bayesian Optimization with the results of (Zapata et al., 2017) who use a subset of the same historical data from the Washington State University research site in Prosser, WA to find model parameters for grape phenology. They used linear regression to find best fit parameters for the phenological stages of Bud Break, Bloom, and Veraison. Moreover, they focus on finding parameters that best minimize the error for a specific phenological stage. In contrast, we minimize the cumulative error across all stages. Despite focusing on globally minimizing error, we find that our selected parameters generally reduce the RMSE in days for all three phenological stages across 10 different cultivars (Table 2). We highlight that the parameters for our Bayesian Optimization method were obtained using only 500 simulations, which is much less than the 10K-15K used in Monte Carlo methods for CGM calibration (Lara et al., 2012). The parameters sets that describe grape phenology for 32 cultivars are included in WOFOSTGym as a step towards effective sim-to-real transfer for crop management policies in open-field agriculture.

5 Experiments

We present RL agent experiments on a variety of different tasks to learn crop management policies for annual and perennial crops on realistic tasks with constraints. In contrast to previous work which only investigated crop management strategies for a single crop (Turchetta et al., 2022; Overweg et al.,

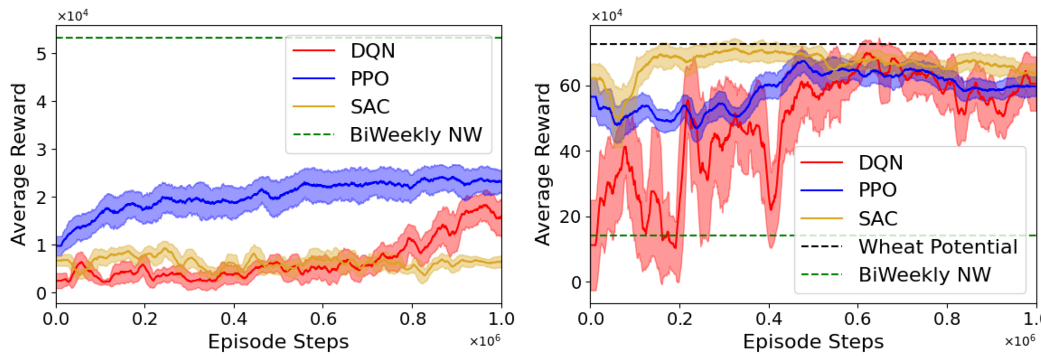


Figure 3: The average reward (in seasonal yield) of different policies: the jujube crop (left) and the wheat crop (right). The BiWeekly NW policy alternates applying nitrogen and water biweekly while the Wheat Potential (right) is the maximum growth obtainable. We omit the Jujube Potential because it assumes daily intervention, while we only allow the RL agents to intervene biweekly in the Jujube experiment.

2021), we show results from a variety of crop and soil types, demonstrating the customizability of WOFOSTGym. WOFOSTGym also allows for phosphorous and potassium application, and so our experiments consider their application as agromanagement decisions as well in contrast to previous work which only considered nitrogen and water application. We highlight the specific shortcomings of RL algorithms in our experiments, namely their inability to adhere to hard constraints and difficulty across long time horizons with delayed feedback. These difficulties naturally arise in agriculture and are reflected in WOFOSTGym, making it an appropriate simulator for both fundamental RL research and research on RL as an agromanagement decision support tool.

5.1 RL Learning Curves on Various Tasks

In this section, we demonstrate that standard RL agents (PPO, SAC, DQN) can achieve near optimal performance on some instances of WOFOSTGym environments and substantially outperform common agromanagement practices. However, on other environment instances, RL algorithms struggle to learn, demonstrating the need for further research on RL in problems with long episode horizons with delayed feedback.

The PPO, SAC, and DQN algorithms are implemented with CleanRL[1] and hyperparameters were tuned experimentally to yield best performance in the WOFOSTGym domain. The agent can choose between 16 actions corresponding to varying amounts of nitrogen, phosphorus, potassium, and Water. Unless otherwise noted, the agent observes crop states: development stage and weight of storage organs, the soil states: total nitrogen, phosphorus, potassium, and Water applied, soil moisture content, nitrogen, phosphorous, and potassium available in subsoil, and the weather states: solar irradiation, average temperature, and rainfall.

Sub-Optimal Performance of Standard RL Algorithms We train RL agents to maximize the growth of jujube over three seasons and wheat over one season in WOFOSTGym and compare them with the maximum potential yield and a common agromanagement policy that alternates between nitrogen fertilization and irrigation biweekly. Results are shown in Figure 3. In the wheat experiment, we see that RL agents dramatically outperform the baseline of a Biweekly nitrogen and water application policy, yet fall short of the potential production in an unlimited nutrient setting. Meanwhile, in the jujube experiment, we see that RL agents are unable to match the performance of a monthly nitrogen and water application policy, necessitating future research.

[1] https://github.com/vwxyzjn/cleanrl

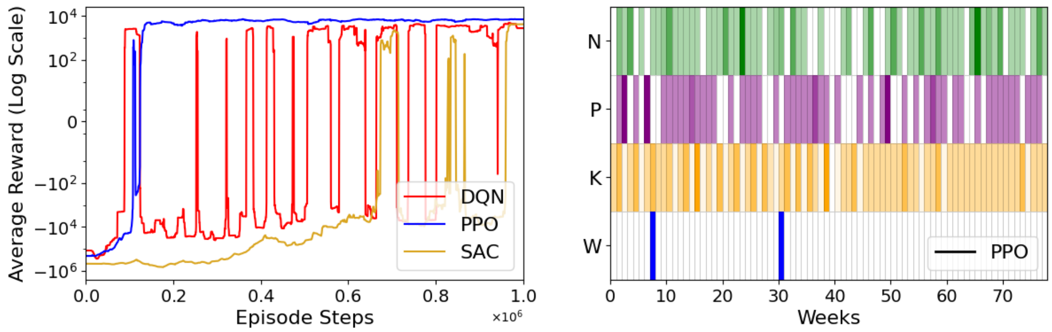


Figure 4: (Left) The reward running average of growing the pear crop over three seasons. (Right) The average day of fertilization and irrigation actions growing the pear tree using the trained PPO agent over 30 years. Darker colors signify more frequent application.

Failure of RL Algorithms to Consistently Satisfy Constraints Agriculture is modeled as a yield maximization problem subject to a set of constraints. One possible constraint is a limit on the amount of fertilizer and water that can be applied per season. To incentivize that behavior, we reward total yield and apply a large negative penalty if the fertilization and irrigation thresholds (in kg/ha and cm/ha) are exceeded. We train RL algorithms using this reward function and show the results in Figure 4. Unlike the previous experiment, there is no principled way to find the maximum total reward. We compare the RL agents to a baseline that applies nitrogen fertilizer and water until it meets the thresholds of 80 kg/ha of fertilizer or 40 cm/ha of water. While the baseline policy does not violate constraints, it does not obtain as high of a reward as the trained RL agents. However, the RL agents do not guarantee that they will never violate the fertilization constraints, which demonstrates the need for future research.

Ablation Study of RL Algorithms under Partial Observability As previously discussed, agriculture should be modeled as a partial observability problem. To demonstrate the effects of partial observability in WOFOSTGym, we design a reward function that rewards crop yield subject to a large penalty if nutrient runoff occurs. Nutrient runoff happens when large amounts of fertilizer are present on the soil surface and irrigation or rainfall occur. We consider two relevant state features, RAIN and TOTN, the daily rainfall and fertilizer on the soil surface, respectively, and create four partially observable environments based on the combinations of the two variables. We train a PPO agent on all four of these scenarios and show the results on the potato crop in Figure 5.

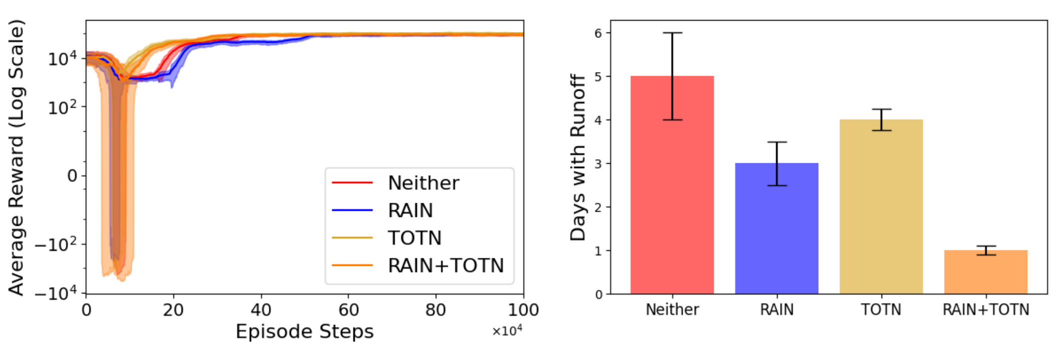
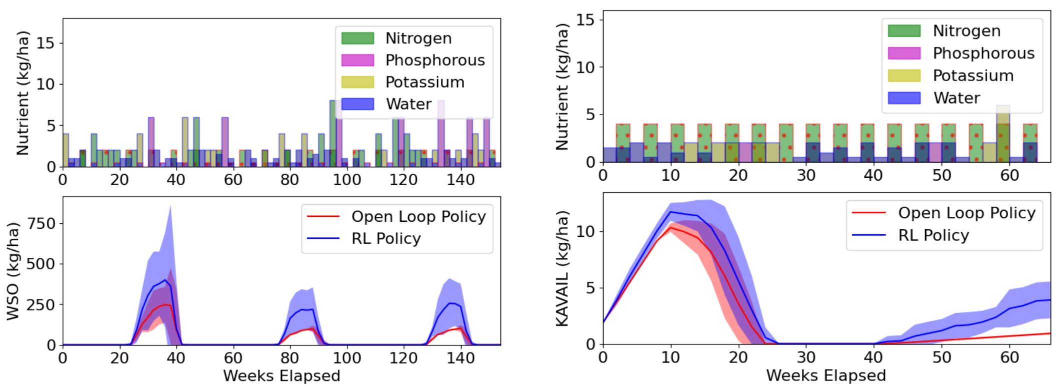


Figure 5: (Left) The learning curves of PPO when growing the potato crop subject to a reward for maximizing yield and a large penalty for incurring nitrogen surface runoff. Each PPO agent is trained on some degree of partial observability indicated by the label if it has access to TOTN or RAIN, the available surface nitrogen and daily rainfall, respectively.



(a) (Top) The biweekly action taken by two policies when growing the jujube tree across three seasons. (Bottom) The yield (WSO) of the jujube tree.

(b) (Top) The weekly action by two policies when growing the wheat plant across one season. (Bottom) The phosphorus available in the subsoil.

Figure 6: Comparing the 10 season average weight of the storage organs of the jujube tree (left) and the potassium available in the subsoil of the wheat crop (right). Bars with dots correspond to actions taken by the open loop biweekly water and nitrogen policy, solid bars are actions taken by the RL policy.

5.2 Visual Comparison of Agromanagement Decisions on Multiple Farms

Comparing yield and nutrient levels under different agromanagement policies is desirable for farmers, but unrealistic to perform in the field due to the risk of untested actions on yield. WOFOSTGym enables agromanagement policies to be evaluated in simulation which could be a useful tool for farmers to understand the impacts of their agromanagement decisions on crop and soil health.

We train two PPO agents: one to maximize the yield of jujube over three seasons, and one to maximize the growth of barley over one season. We evaluate the yield obtained (WSO) by the PPO jujube policy to the yield obtained by a monthly fertilization policy across three growing seasons. Second, we compare the phosphorus available (PAVAIL) by the PPO barley policy to a biweekly fertilization policy during one growing season. See Figure 6.

In addition to visualizing agromanagement decisions, it is desirable to be able to learn a general agromanagement policy that can be applied to multiple farms. We create a WOFOSTGym environment that simultaneously simulates the growth of 5 sunflower farms experiencing the same weather. The observation space is the growth and soil variables for each farm. The weather is shared between farms and the action selected is uniformed applied to each farm. We train a PPO, DQN, and SAC agent in this multi farm scenario and report the soil moisture content and the average yield with each agent policy on each farm in Figure 7.

5.3 Simulator Run Times

Fast simulators are central to the successful application of RL given the high sampling complexity of RL algorithms (Lechner et al., 2023). We benchmark the run times of three crop simulators: WOFOSTGym, CyclesGym, and gym-DSSAT (Gautron et al., 2022b; Turchetta et al., 2022). We compare run times for a single episode growing the maize crop (155 episode steps), and with the large potential for overhead when resetting the underlying CGM, we investigate the run times of the step and reset functions as well.

Our results in Table 4 show that WOFOSTGym outperforms CyclesGym, the only crop simulator that supports multi-year simulations, by an order of magnitude. WOFOSTGym is also faster than gym-DSSAT, with the improvement coming from the dramatic reduction in run time in the reset function. While the step function in gym-DSSAT outperforms that of WOFOSTGym, we note that

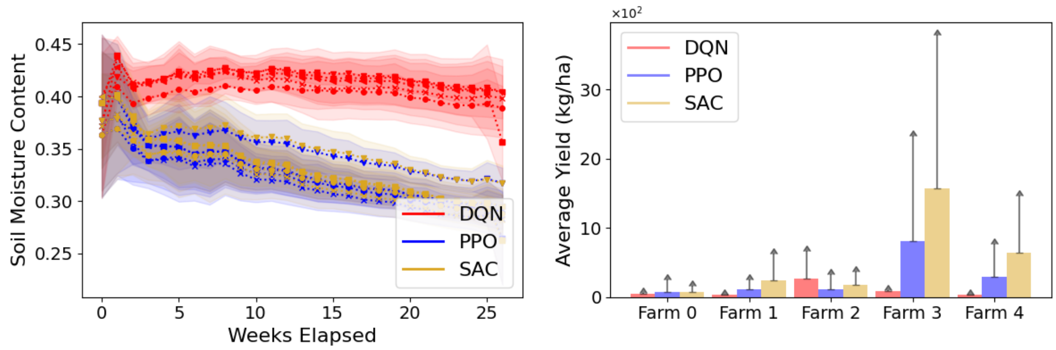


Figure 7: The soil moisture content (left) for the 5 farms, each denoted with a different marker, under 3 different RL agromanagement policies. The average yield (right) for the three agents across the 5 farms over 10 growing seasons of the sunflower plant.

WOFOSTGym performs many more computations per step by maintaining a nitrogen, phosphorus and potassium balance throughout the crop, whereas gym-DSSAT maintains only a nitrogen balance.

Run Time (s)	WOFOSTGym	CyclesGym	gym-DSSAT
1 Episode	0.34±.012	2.08±.221	0.38±0.018
Step Function	0.003±0.0005	0.04±.0020	0.001±0.0001
Reset Function	0.012±0.002	0.055±.002	0.191±0.012

Table 3: The average runtime (in seconds) of three different crop simulators on an Nvidia 3080Ti. Averages and standard deviations were computed over 100 trials.

6 Limitations

Although WOFOSTGym offers an improved run time compared to other crop simulators, we note that for multi-year simulations, run time is non-trivial. WOFOSTGym takes $\tilde{2}$ seconds to run a 3 year simulation of a perennial jujube crop. This is not an insignificant amount of time when RL algorithms require millions of episodes to learn a good policy. As the required episode horizon continues to increase for modeling perennial crop management decisions, developing a framework for a faster crop dynamics simulator will be a valuable direction for future work.

WOFOSTGym does not directly support training RL agents to optimize long term managerial crop rotation strategies and instead focuses on modeling perennial crop growth. The excellent work of Turchetta et al. (2022) handles managerial strategies for crop rotations with RL Agents. The WOFOST CGM provides support for crop rotations, and extending WOFOSTGym to support such problems could be an interesting extension.

The fidelity of the modified WOFOST CGM we presented is limited by the data available. While the parameter sets included can be considered high fidelity models as they were tuned against field data, sim-to-real transfer using WOFOSTGym should be attempted with caution. As research bridging RL to open-field agriculture continues to proliferate and the fidelity of CGMs increase with approaches like those in Section 4.3, direct sim-to-real transfer may become possible.

7 Conclusion

We present WOFOSTGym, the first RL simulator for perennial crop management decision support. The WOFOSTGym repository includes high fidelity parameters for 2 pernenial crops and 23 annual crops and a wide variety of pre-specified agromanagement policies to compare RL agents against. WOFOSTGym is easily customizable, which allows researchers to design and run their own experiments even without domain expertise in agriculture.

We propose a new method for CGM calibration with the aim of increasing the fidelity of CGMs like WOFOST as a step towards sim to real transfer in open-field agriculture. Our experiments demonstrate the difficulty that current RL algorithms have in the agriculture domain, thus necessitating further RL research to address the specific challenges that the agriculture domain presents. We suggest some realistic benchmarks to evaluate RL algorithms against to support further research on AI decision support for agriculture *before* deployment.

Broader Impact Statement

Reinforcement Learning for crop modeling promises to aid growers in achieving policies that increase crop yield while lowering costs and environmental impact. WOFOSTGym enables researchers to learn and test policies on a high-fidelity simulator. However, there is still a simulation gap, and the performance of these policies in sim should not be taken to be reflective of their potential when deployed in the field.

References

Francisco Albornoz. Crop responses to nitrogen overfertilization: A review. *Scientia Horticulturae*, 205:79–83, June 2016. ISSN 0304-4238. DOI: 10.1016/j.scienta.2016.04.026. URL https://www.sciencedirect.com/science/article/pii/S0304423816302151.

Zhicheng An, Xiaoyan Cao, Yao Yao, Wanpeng Zhang, Lanqing Li, Yue Wang, Shihui Guo, and Dijun Luo. A Simulator-based Planning Framework for Optimizing Autonomous Greenhouse Control Strategy. *Proceedings of the International Conference on Automated Planning and Scheduling*, 31:436–444, May 2021. ISSN 2334-0843. DOI: 10.1609/icaps.v31i1.15989. URL https://ojs.aaai.org/index.php/ICAPS/article/view/15989.

B. Andarzian, M. Bannayan, P. Steduto, H. Mazraeh, M. E. Barati, M. A. Barati, and A. Rahnama. Validation and testing of the AquaCrop model under full and deficit irrigated wheat production in Iran. *Agricultural Water Management*, 100(1):1–8, November 2011. ISSN 0378-3774. DOI: 10.1016/j.agwat.2011.08.023. URL https://www.sciencedirect.com/science/article/pii/S0378377411002307.

Gregory Astill, Agnes Perez, and Suzanne Thornsbury (eds.). *Developing Automation and Mechanization for Specialty Crops: A Review of U.S. Department of Agriculture Programs: A Report to Congress*. Administrative Publication Number 082. 2020. DOI: 10.22004/ag.econ.320792.

Tie-cheng Bai, Tao Wang, Nan-nan Zhang, You-qi Chen, and Benoit Mercatoris. Growth simulation and yield prediction for perennial jujube fruit tree by integrating age into the WOFOST model. *Journal of Integrative Agriculture*, 19(3):721–734, March 2020. ISSN 2095-3119. DOI: 10.1016/S2095-3119(19)62753-X. URL https://www.sciencedirect.com/science/article/pii/S209531191962753X.

Tiecheng Bai, Nannan Zhang, Youqi Chen, and Benoit Mercatoris. Assessing the Performance of the WOFOST Model in Simulating Jujube Fruit Tree Growth under Different Irrigation Regimes. *Sustainability*, 11(5):1466, January 2019. ISSN 2071-1050. DOI: 10.3390/su11051466. URL https://www.mdpi.com/2071-1050/11/5/1466. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.

Herman N. C. Berghuijs, João Vasco Silva, Pytrik Reidsma, and Allard J. W. de Wit. Expanding the WOFOST crop model to explore options for sustainable nitrogen management: A study for winter wheat in the Netherlands. *European Journal of Agronomy*, 154:127099, March 2024. ISSN 1161-0301. DOI: 10.1016/j.eja.2024.127099. URL https://www.sciencedirect.com/science/article/pii/S1161030124000200.

Avnish Kumar Bhatia. Crop Growth Simulation Modeling. In S.K. Basu and Naveen Kumar (eds.), *Modelling and Simulation of Diffusive Processes: Methods and Applications*, pp. 315–332. Springer International Publishing, Cham, 2014. ISBN 978-3-319-05657-9. DOI: 10.1007/978-3-319-05657-9_15. URL https://doi.org/10.1007/978-3-319-05657-9_15.

Jonathan Binas, Leonie Luginbuehl, and Yoshua Bengio. Reinforcement Learning for Sustainable Agriculture. In *Climate Change AI*. Climate Change AI, June 2019. URL https://www.climatechange.ai/papers/icml2019/32.

Kenneth J. Boote, James W. Jones, and Nigel B. Pickering. Potential Uses and Limitations of Crop Models | Agronomy Journal, September 1996. URL https://acsess.onlinelibrary.wiley.com/doi/abs/10.2134/agronj1996.00021962008800050005x.

Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. OpenAI Gym, June 2016. URL http://arxiv.org/abs/1606.01540. arXiv:1606.01540.

M. Cabelguenne, C. A. Jones, J. R. Marty, P. T. Dyke, and J. R. Williams. Calibration and validation of EPIC for crop rotations in southern France. *Agricultural Systems*, 33(2):153–171, January 1990. ISSN 0308-521X. DOI: 10.1016/0308-521X(90)90078-5. URL https://www.sciencedirect.com/science/article/pii/0308521X90900785.

Daniel Dauner, Marcel Hallgarten, Tianyu Li, Xinshuo Weng, Zhiyu Huang, Zetong Yang, Hongyang Li, Igor Gilitschenski, Boris Ivanovic, Marco Pavone, Andreas Geiger, and Kashyap Chitta. NAVSIM: Data-Driven Non-Reactive Autonomous Vehicle Simulation and Benchmarking, October 2024. URL http://arxiv.org/abs/2406.15349. arXiv:2406.15349 [cs].

Allard de Wit. WOFOST - WOrld FOod STudies, August 2019. URL https://www.wur.nl/en/research-results/research-institutes/environmental-research/facilities-tools/software-models-and-databases/wofost.htm.

Allard de Wit. PCSE Documentation, August 2024. URL https://pcse.readthedocs.io/en/stable/.

Allard de Wit. ajwdewit/WOFOST_crop_parameters, January 2025. URL https://github.com/ajwdewit/WOFOST_crop_parameters. original-date: 2016-11-29T20:18:52Z.

Allard de Wit, Hendrik Boogaard, Davide Fumagalli, Sander Janssen, Rob Knapen, Daniel van Kraalingen, Iwan Supit, Raymond van der Wijngaart, and Kees van Diepen. 25 years of the WOFOST cropping systems model. *Agricultural Systems*, 168:154–167, January 2019. ISSN 0308-521X. DOI: 10.1016/j.agsy.2018.06.018. URL https://www.sciencedirect.com/science/article/pii/S0308521X17310107.

Arianna Di Paola, Riccardo Valentini, and Monia Santini. An overview of available crop growth and yield models for studies and assessments in agriculture. *Journal of the Science of Food and Agriculture*, 96(3):709–714, 2016. ISSN 1097-0010. DOI: 10.1002/jsfa.7359. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jsfa.7359. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jsfa.7359.

Ahmad Din, Muhammed Yousoof Ismail, Babar Shah, Mohammad Babar, Farman Ali, and Siddique Ullah Baig. A deep reinforcement learning-based multi-agent area coverage control for smart agriculture. *Computers and Electrical Engineering*, 101:108089, July 2022. ISSN 0045-7906. DOI: 10.1016/j.compeleceng.2022.108089. URL https://www.sciencedirect.com/science/article/pii/S0045790622003445.

Joshua Elliott, Delphine Deryng, Christoph Müller, Katja Frieler, Markus Konzmann, Dieter Gerten, Michael Glotter, Martina Flörke, Yoshihide Wada, Neil Best, Stephanie Eisner, Balázs M. Fekete, Christian Folberth, Ian Foster, Simon N. Gosling, Ingjerd Haddeland, Nikolay Khabarov, Fulco Ludwig, Yoshimitsu Masaki, Stefan Olin, Cynthia Rosenzweig, Alex C. Ruane, Yusuke Satoh, Erwin Schmid, Tobias Stacke, Qiuhong Tang, and Dominik Wisser. Constraints and potentials of future irrigation water availability on agricultural production under climate change. *Proceedings of the National Academy of Sciences*, 111(9):3239–3244, March 2014. DOI: 10.1073/pnas.1222474110. URL https://www.pnas.org/doi/abs/10.1073/pnas.1222474110. Publisher: Proceedings of the National Academy of Sciences.

L. D. Estes, B. A. Bradley, H. Beukes, D. G. Hole, M. Lau, M. G. Oppenheimer, R. Schulze, M. A. Tadross, and W. R. Turner. Comparing mechanistic and empirical model projections of crop suitability and productivity: implications for ecological forecasting. *Global Ecology and Biogeography*, 22(8):1007–1018, 2013. ISSN 1466-8238. DOI: 10.1111/geb.12034. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/geb.12034. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/geb.12034.

Romain Gautron, Odalric-Ambrym Maillard, Philippe Preux, Marc Corbeels, and Régis Sabbadin. Reinforcement learning for crop management support: Review, prospects and challenges. *Computers and Electronics in Agriculture*, 200:107182, September 2022a. ISSN 0168-1699. DOI:

10.1016/j.compag.2022.107182. URL https://www.sciencedirect.com/science/article/pii/S0168169922004999.

Romain Gautron, Emilio Padron, Philippe Preux, Julien Bigot, Odalric-Ambrym Maillard, and David Emukpere. gym-DSSAT: a crop model turned into a Reinforcement Learning Environment, September 2022b. URL https://arxiv.org/pdf/2207.03270.

Carlo Gilardelli, Roberto Confalonieri, Giovanni Alessandro Cappelli, and Gianni Bellocchi. Sensitivity of WOFOST-based modelling solutions to crop parameters under climate change. *Ecological Modelling*, 368:1–14, January 2018. ISSN 0304-3800. DOI: 10.1016/j.ecolmodel.2017.11.003. URL https://www.sciencedirect.com/science/article/pii/S0304380017304453.

Y. He, K. L. Hu, H. Wang, Y. F. Huang, D. L. Chen, B. G. Li, and Y. Li. Modeling of water and nitrogen utilization of layered soil profiles under a wheat–maize cropping system. *Mathematical and Computer Modelling*, 58(3):596–605, August 2013. ISSN 0895-7177. DOI: 10.1016/j.mcm.2011.10.060. URL https://www.sciencedirect.com/science/article/pii/S0895717711006650.

Mohd Javaid, Abid Haleem, Ibrahim Haleem Khan, and Rajiv Suman. Understanding the potential applications of Artificial Intelligence in Agriculture Sector. *Advanced Agrochem*, 2 (1):15–30, March 2023. ISSN 2773-2371. DOI: 10.1016/j.aac.2022.10.001. URL https://www.sciencedirect.com/science/article/pii/S277323712200020X.

J. W Jones, G Hoogenboom, C. H Porter, K. J Boote, W. D Batchelor, L. A Hunt, P. W Wilkens, U Singh, A. J Gijsman, and J. T Ritchie. The DSSAT cropping system model. *European Journal of Agronomy*, 18(3):235–265, January 2003. ISSN 1161-0301. DOI: 10.1016/S1161-0301(02)00107-7. URL https://www.sciencedirect.com/science/article/pii/S1161030102001077.

James W. Jones, John M. Antle, Bruno Basso, Kenneth J. Boote, Richard T. Conant, Ian Foster, H. Charles J. Godfray, Mario Herrero, Richard E. Howitt, Sander Janssen, Brian A. Keating, Rafael Munoz-Carpena, Cheryl H. Porter, Cynthia Rosenzweig, and Tim R. Wheeler. Brief history of agricultural systems modeling. *Agricultural Systems*, 155:240–254, July 2017. ISSN 0308-521X. DOI: 10.1016/j.agsy.2016.05.014. URL https://www.sciencedirect.com/science/article/pii/S0308521X16301585.

M. B. Jones, E. L. Leafe, W. Stiles, and B. Collett. Pattern of Respiration of a Perennial Ryegrass Crop in the Field. *Annals of Botany*, 42(3):693–703, May 1978. ISSN 0305-7364. DOI: 10.1093/oxfordjournals.aob.a085504. URL https://doi.org/10.1093/oxfordjournals.aob.a085504.

Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1-2):99–134, May 1998. ISSN 00043702. DOI: 10.1016/S0004-3702(98)00023-X. URL https://linkinghub.elsevier.com/retrieve/pii/S000437029800023X.

Armen Kemanian, Yuning Shi, Charles White, Felipe Pena, Claudio Stockle, David Huggins, Maria Cangiano, Giovani Stefani-Faé, and Rachel Rozum. The Cycles Agroecosystem Model: Fundamentals, Testing, and Applications. *SSRN Electronic Journal*, January 2022. DOI: 10.2139/ssrn.4188402.

B Ravi Kiran, Ibrahim Sobh, Victor Talpaert, Patrick Mannion, Ahmad A. Al Sallab, Senthil Yogamani, and Patrick Pérez. Deep Reinforcement Learning for Autonomous Driving: A Survey. *IEEE Transactions on Intelligent Transportation Systems*, 23(6):4909–4926, June 2022. ISSN 1558-0016. DOI: 10.1109/TITS.2021.3054625. URL https://ieeexplore.ieee.org/document/9351818/?arnumber=9351818. Conference Name: IEEE Transactions on Intelligent Transportation Systems.

Jens Kober, J. Andrew Bagnell, and Jan Peters. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, 32(11), August 2013. DOI: https://doi.org/10.1177/027836491349572. URL https://journals.sagepub.com/doi/full/10.1177/0278364913495721?casa_token=8R-qBicIj_gAAAAA%3A5nkkUMu2YGc1jYvehc17X5CzVm-sK2YscqCWtTQXIqYEn01uReYCZ8_cfXjv_JYdMeBVLUCp3pPD.

Márcio A.S. Lara, Carlos G.S. Pedreira, Kenneth J. Boote, Bruno C. Pedreira, Leonardo S.B. Moreno, and Phillip D. Alderman. Predicting Growth of Panicum maximum: An Adaptation of the CROPGRO–Perennial Forage Model. *Agronomy Journal*, 104(3): 600–611, 2012. ISSN 1435-0645. DOI: 10.2134/agronj2011.0272. URL https://onlinelibrary.wiley.com/doi/abs/10.2134/agronj2011.0272. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.2134/agronj2011.0272.

Mathias Lechner, Lianhao Yin, Tim Seyde, Tsun-Hsuan Johnson Wang, Wei Xiao, Ramin Hasani, Joshua Rountree, and Daniela Rus. Gigastep - One Billion Steps per Second Multi-agent Reinforcement Learning. *Advances in Neural Information Processing Systems*, 36:155–170, December 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/00ba06ba5c324efdfb068865ca44cf0b-Abstract-Datasets_and_Benchmarks.html.

Sergey Levine, Aviral Kumar, George Tucker, and Justin Fu. Offline Reinforcement Learning: Tutorial, Review, and Perspectives on Open Problems, November 2020. URL http://arxiv.org/abs/2005.01643. arXiv:2005.01643 [cs].

Leena Lindén, Hannu Rita, and Terhi Suojala. Logit Models for Estimating Lethal Temperatures in Apple. February 1996. DOI: 10.21273/HORTSCI.31.1.91. URL https://journals.ashs.org/hortsci/view/journals/hortsci/31/1/article-p91.xml. Section: HortScience.

R. L. McCown, G. L. Hammer, J. N. G. Hargreaves, D. P. Holzworth, and D. M. Freebairn. APSIM: a novel software system for model development, model testing and simulation in agricultural systems research. *Agricultural Systems*, 50(3):255–271, January 1996. ISSN 0308-521X. DOI: 10.1016/0308-521X(94)00055-V. URL https://www.sciencedirect.com/science/article/pii/0308521X9400055V.

Sergi Munné-Bosch. Aging in Perennials. *Critical Reviews in Plant Sciences*, 26(3):123–138, June 2007. ISSN 0735-2689. DOI: 10.1080/07352680701402487. URL https://doi.org/10.1080/07352680701402487. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07352680701402487.

Hiske Overweg, Herman N. C. Berghuijs, and Ioannis N. Athanasiadis. CropGym: a Reinforcement Learning Environment for Crop Management, April 2021. URL http://arxiv.org/abs/2104.04326. arXiv:2104.04326 [cs].

Xue Bin Peng, Marcin Andrychowicz, Wojciech Zaremba, and Pieter Abbeel. Sim-to-Real Transfer of Robotic Control with Dynamics Randomization, March 2018. URL http://arxiv.org/abs/1710.06537. arXiv:1710.06537.

Antje Rohde and Rishikesh P. Bhalerao. Plant dormancy in the perennial context. *Trends in Plant Science*, 12(5):217–223, May 2007. ISSN 1360-1385. DOI: 10.1016/j.tplants.2007.03.012. URL https://www.cell.com/trends/plant-science/abstract/S1360-1385(07)00083-0. Publisher: Elsevier.

Yinfang Shi, Zhaoyang Wang, Cheng Hou, and Puhan Zhang. Yield estimation of *Lycium barbarum* L. based on the WOFOST model. *Ecological Modelling*, 473:110146, November 2022. ISSN 0304-3800. DOI: 10.1016/j.ecolmodel.2022.110146. URL https://www.sciencedirect.com/science/article/pii/S0304380022002472.

Thomas R Sinclair and No'am Seligman. Criteria for publishing papers on crop modeling. *Field Crops Research*, 68(3):165–172, November 2000. ISSN 0378-4290. DOI: 10.1016/S0378-4290(00)00105-2. URL https://www.sciencedirect.com/science/article/pii/S0378429000001052.

Claudio O. Stockle, Steve A. Martin, and Gaylon S. Campbell. CropSyst, a cropping systems simulation model: Water/nitrogen budgets and crop yield. *Agricultural Systems*, 46:335–359, January 1994. DOI: 10.1016/0308-521X(94)90006-2. URL https://ui.adsabs.harvard.edu/abs/1994AgSys..46..335S. ADS Bibcode: 1994AgSys..46..335S.

Ran Tao, Pan Zhao, Jing Wu, Nicolas F. Martin, Matthew T. Harrison, Carla Ferreira, Zahra Kalantari, and Naira Hovakimyan. Optimizing Crop Management with Reinforcement Learning and Imitation Learning, February 2023. URL http://arxiv.org/abs/2209.09991. arXiv:2209.09991 [cs].

Camille Tevenart and Marielle Brunette. Role of Farmers' Risk and Ambiguity Preferences on Fertilization Decisions: An Experiment. *Sustainability*, 13(17):9802, January 2021. ISSN 2071-1050. DOI: 10.3390/su13179802. URL https://www.mdpi.com/2071-1050/13/17/9802. Number: 17 Publisher: Multidisciplinary Digital Publishing Institute.

Howard Thomas, Huw Martin Thomas, and Helen Ougham. Annuality, perenniality and cell death. *Journal of Experimental Botany*, 51(352):1781–1788, November 2000. ISSN 0022-0957. DOI: 10.1093/jexbot/51.352.1781. URL https://doi.org/10.1093/jexbot/51.352.1781.

Philip Thomas and Emma Brunskill. Data-Efficient Off-Policy Policy Evaluation for Reinforcement Learning. In *Proceedings of The 33rd International Conference on Machine Learning*, pp. 2139–2148. PMLR, June 2016. URL https://proceedings.mlr.press/v48/thomasa16.html. ISSN: 1938-7228.

Emanuel Todorov, Tom Erez, and Yuval Tassa. MuJoCo: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5026–5033, October 2012. DOI: 10.1109/IROS.2012.6386109. URL https://ieeexplore.ieee.org/document/6386109/?arnumber=6386109. ISSN: 2153-0866.

Matteo Turchetta, Luca Corinzia, Scott Sussex, Amanda Burton, Juan Herrera, Ioannis Athanasiadis, Joachim Buhmann, and Andreas Krause. Learning Long-Term Crop Management Strategies with CyclesGym, 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/4a22ceafe2dd6e0d32df1f7c0a69ab68-Paper-Datasets_and_Benchmarks.pdf.

C.a. van Diepen, J. Wolf, H. van Keulen, and C. Rappoldt. WOFOST: a simulation model of crop production. *Soil Use and Management*, 5(1):16–24, 1989. ISSN 1475-2743. DOI: 10.1111/j.1475-2743.1989.tb00755.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-2743.1989.tb00755.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-2743.1989.tb00755.x.

Mel Vecerik, Todd Hester, Jonathan Scholz, Fumin Wang, Olivier Pietquin, Bilal Piot, Nicolas Heess, Thomas Rothörl, Thomas Lampe, and Martin Riedmiller. Leveraging Demonstrations for Deep Reinforcement Learning on Robotics Problems with Sparse Rewards, October 2018. URL http://arxiv.org/abs/1707.08817. arXiv:1707.08817 [cs].

Chengkun Wang, Nannan Zhang, Mingzhe Li, Li Li, and Tiecheng Bai. Pear Tree Growth Simulation and Soil Moisture Assessment Considering Pruning. *Agriculture*, 12(10):1653, October 2022. ISSN 2077-0472. DOI: 10.3390/agriculture12101653. URL https://www.mdpi.com/2077-0472/12/10/1653.

Lu Wang, Xiaofeng He, and Dijun Luo. Deep Reinforcement Learning for Greenhouse Climate Control. In *2020 IEEE International Conference on Knowledge Graph (ICKG)*, pp. 474–480, August 2020. DOI: 10.1109/ICBK50248.2020.00073. URL https://ieeexplore.ieee.org/abstract/document/9194467.

Jing Wu, Ran Tao, Pan Zhao, Nicolas F. Martin, and Naira Hovakimyan. Optimizing Nitrogen Management with Deep Reinforcement Learning and Crop Simulations. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1711–1719, New Orleans, LA, USA, June 2022. IEEE. ISBN 978-1-6654-8739-9. DOI: 10.1109/CVPRW56347.2022.00178. URL https://ieeexplore.ieee.org/document/9857484/.

H. S Yang, A Dobermann, J. L Lindquist, D. T Walters, T. J Arkebauer, and K. G Cassman. Hybrid-maize—a maize simulation model that combines two crop modeling approaches. *Field Crops Research*, 87(2):131–154, May 2004. ISSN 0378-4290. DOI: 10.1016/j.fcr.2003.10.003. URL https://www.sciencedirect.com/science/article/pii/S0378429003002272.

D. Zapata, M. Salazar, B. Chaves, M. Keller, and G. Hoogenboom. Estimation of the base temperature and growth phase duration in terms of thermal time for four grapevine cultivars. *International Journal of Biometeorology*, 59(12):1771–1781, December 2015. ISSN 1432-1254. DOI: 10.1007/s00484-015-0985-y. URL https://doi.org/10.1007/s00484-015-0985-y.

Diana Zapata, Melba Salazar-Gutierrez, Bernardo Chaves, Markus Keller, and Gerrit Hoogenboom. Predicting Key Phenological Stages for 17 Grapevine Cultivars (Vitis vinifera L.). *American Journal of Enology and Viticulture*, 68(1):60–72, January 2017. ISSN 0002-9254. DOI: 10.5344/ajev.2016.15077. URL https://www.ajevonline.org/content/68/1/60. Publisher: American Journal of Enology and Viticulture Section: Research Article.

Zichen Zhang, Jayson Boubin, Christopher Stewart, and Sami Khanal. Whole-Field Reinforcement Learning: A Fully Autonomous Aerial Scouting Method for Precision Agriculture. *Sensors*, 20 (22):6585, January 2020. ISSN 1424-8220. DOI: 10.3390/s20226585. URL https://www.mdpi.com/1424-8220/20/22/6585. Number: 22 Publisher: Multidisciplinary Digital Publishing Institute.

Junqi Zhu, Amber Parker, Fang Gou, Rob Agnew, Linlin Yang, Marc Greven, Victoria Raw, Sue Neal, Damian Martin, Michael C T Trought, Neil Huth, and Hamish Edward Brown. Developing perennial fruit crop models in APSIM Next Generation using grapevine as an example. *in silico Plants*, 3(2):diab021, July 2021. ISSN 2517-5025. DOI: 10.1093/insilicoplants/diab021. URL https://doi.org/10.1093/insilicoplants/diab021.

Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He. A Comprehensive Survey on Transfer Learning. *Proceedings of the IEEE*, 109 (1):43–76, January 2021. ISSN 1558-2256. DOI: 10.1109/JPROC.2020.3004555. URL https://ieeexplore.ieee.org/document/9134370/?arnumber=9134370. Conference Name: Proceedings of the IEEE.

A Crop Growth Models

WOFOST is a single-year and multi-crop agroecosystem model (Jones et al., 2017). It is built as a hybrid crop model, meaning it relies both on mechanistic and empirical processes to simulate crop growth (Di Paola et al., 2016). Crop growth in WOFOST is determined by the atmospheric CO2 concentration, irradiation, daily temperature, subject to limited water, nitrogen, phosphorus, and potassium. While WOFOST was designed for simulating the yield of annual crops (van Diepen et al., 1989), field studies have recently shown that it can be used to accurately predict yield in perennial fruit trees with some small modifications to the base model (Bai et al. (2020) Wang et al. (2022)).

Given WOFOST's wide range of customizability with over 100 crop and soil parameters, its ability to simulate a wide variety of management practices, and its modular implementation in Python, WOFOST is an ideal CGM candidate to be used to simulate perennial crop growth to address that lack of perennial CGMs available, and the lack of perennial crop benchmarks available for RL research (Gautron et al., 2022a). For an introduction to the WOFOST see de Wit (2024) and de Wit (2019).

The choice of the WOFOST Crop Growth Model (CGM) was an integral design decision in the creation of WOFOSTGym. There are a wide variety of CGMs available for use. Table 4 outlines the desiderata that we used when selecting WOFOST to serve as the base CGM for the WOFOSTGym simulator.

Crop Model	Model Type	Nutrient Balance	Water Balance	Crop Type	Language
WOFOST	Hybrid	nitrogen, phosphorous, potassium	Single Layer, Multi Layer	Annual	Python, FORTRAN
APSIM	Mechanistic	nitrogen, phosphorous, potassium	Multi Layer	Annual	FORTRAN, C++
DSSAT	Hybrid	nitrogen, phosphorous, potassium	Multi Layer	Annual	FORTRAN
CropSyst	Mechanistic	nitrogen	Single Layer	Annual	C++
EPIC	Hybrid	nitrogen, phosphorous	Multi Layer	Annual, Rotations	FORTRAN
STICS	Empirical	nitrogen	Multi Layer	Annual, Perennial	Executable
Cycles	Mechanistic	nitrogen	Multi Layer	Annual, Perennial	Executable
AquaCrop	Empirical	Abundant	Multi Layer	Annual	Python, Executable
LINTUL3	Empirical	nitrogen	Abundant	Annual	Python

Table 4: Different CGMs and their strengths and weaknesses for modeling high fidelity crop growth, interfacing with RL algorithms, and supporting perennial crop decision evaluation. We note that CGMs with Python implementations are inherently easier to create OpenAI Gym environments for while also enabling the underlying environment to be modified with additional features such as perennial growth models.

B Usability

A shortcoming in other crop simulators is the difficulty in using them by RL researchers and users unfamiliar with CGMs. WOFOSTGym aims to relieve the burden of domain knowledge required to use other crop simulators by streamlining configuration into the readable YAML file format. In this section, we will highlight the features that make WOFOSTGym easy to use for RL researchers interested in agriculture.

Simulation Configuration WOFOST CGM configuration is divided into three configuration files: the crop YAML file, site YAML file, and agromanagement YAML file. WOFOSTGym provides 25 crop YAML files and 3 site YAML files. In the agromanagement YAML file, the specific crop and site configuration to load is specified, along with the length of the simulation, year, and geographic location. The agromanagement YAML file contains 14 entries, while still enabling a user to simulate 25 different crops and their associated varieties. This feature is a dramatic improvement over other crop simulators, which either only support 1 crop, or have an opaque configuration that requires agricultural knowledge that many RL researchers will not possess.

In addition, every parameter can be modified from the command line, and each simulation automatically saves a configuration file, meaning that simulations can be easily modified and reloaded for further use. For an example, see Figure 8,

```
# Test Simulation of the Jujube crop
python3 test_wofost.py --save-folder test/ --data-file test --npk.ag.crop-name
    jujube --npk.ag.crop-variety jujube_1 --env-id perennial-lnpkw-v0

# Generate data of the default crop (wheat) and modify a few crop parameters
python3 gen_data.py --save-folder data/ --data-file wheat_data --file-type npz --
    npk.wf.TBASEM 2.0 --npk.wf.SMFCF 0.51

# Train a SAC agent to irrigate the wheat crop in an environment where nitrogen,
    phosphorus and potassium nutrients are abundant and modify the SAC algorithm
    parameters
python3 train_agent.py --save-folder RL/ --agent-type SAC --env-id lnw-v0 --SAC.
    gamma 0.95

# Train a PPO agent given a WOFOSTGym Configuration file
python3 train_agent.py --save-folder RL --agent-type PPO --config-fpath RL/
    ppo_test.yaml
```

Figure 8: Example for how to configure agent training and data generation in WOFOSTGym. Specific parameters can be modified by the command line. In addition, configuration YAML files can also be loaded for reproducibility (and are automatically saved each time a simulation is run).

Pre-Specified Agricultural Policies The goal of using RL for agriculture is to improve upon crop management strategies. However, to measure this improvement, common crop management policies must be available in a crop simulator to compare against. While other crop simulators do not include baselines, WOFOSTGym comes with 10 pre-specified agromanagement policies that are commonly used in agriculture. These policies include: "fertilize X amount every week," or "irrigate X amount if the soil moisture content is below Y." In WOFOSTGym, X and Y are easily modifiable via command line or YAML file.

Data Generation Offline data is used to for problems in Offline RL (Levine et al., 2020), Off Policy Evaluation (Thomas & Brunskill, 2016), and Transfer Learning (Zhuang et al., 2021). To facilitate research on these topics in the agriculture domain, data must be widely available. However, real world agricultural data is plagued by inaccessibility and missing data. To address this problem, crop simulators are a natural solution, yet no other available crop simulator provides an efficient pipeline for data generation across a wide range of scenarios that would be of interest in the aforementioned problems.

WOFOSTGym addresses this shortcoming by providing a pipeline that can generate data from a variety of different crops and sites and supports generating data from both RL Agent policies and the Pre-Specified policies described above. By including this data generation in WOFOSTGym, we hope to both facilitate research into these interesting RL related problems in agriculture, and set a standard of usability for crop simulators that follow WOFOSTGym.