WOFOSTGym: A Crop Simulator for Learning Annual and Perennial Crop Management Strategies

William Solow, Sandhya Saisubramanian, Alan Fern

Keywords: Crop Simulator, Reinforcement Learning for Agriculture

Summary

We introduce WOFOSTGym, a novel crop simulation environment designed to train reinforcement learning (RL) agents to optimize agromanagement decisions for annual and perennial crops in multi-farm settings. Effective crop management requires optimizing yield and economic returns while minimizing environmental impact, which is a complex sequential decision-making problem well-suited for RL. However, the lack of simulators for perennial crops in multi-farm contexts has hindered RL applications in this domain. Existing crop simulators also do not support multiple annual crops. WOFOSTGym addresses the shortcomings of available crop simulators by supporting 23 annual crops and two perennial crops, enabling RL agents to learn diverse agromanagement strategies in multi-year, multi-crop, and multi-farm settings. Our simulator offers a suite of challenging tasks for learning under partial observability, non-Markovian dynamics, and delayed feedback. Our extensive experiments across a wide variety of crops in single and multi-farm settings, including the constrained optimization tasks that arise in agriculture, demonstrate the learning capabilities and challenges of RL and imitation learning agents. The experiments highlight WOFOSTGym's potential for advancing core RL research and RL-driven decision support in agriculture.

Contribution(s)

1. We introduce WOFOSTGym, an RL simulator built on the WOFOST crop growth model, designed for developing agromanagement policies across multiple annual and multi-season perennial crops, advancing AI-driven decision support in agriculture.
   Context: Existing crop simulators do not support perennial crops or multiple annual crops. WOFOSTGym addresses this gap, enabling users without agricultural expertise 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, and update WOFOST nutrient modules to be able to investigate the impact of agromanagement decisions on the surrounding environment.
   Context: Bai et al. (2019) used the WOFOST crop growth model (CGM) 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 to model continuous multi-year growth with the addition of a dormancy phase.
3. We apply Bayesian Optimization to calibrate the parameters of the WOFOST CGM to increase model fidelity and compare our results with those of an existing work that collected phenology data for 10 grape cultivars.
   Context: High-fidelity CGMs are essential for sim-to-real transfer in open-field agriculture, but parameter calibration is challenging and time-consuming. Traditional agronomic methods rely on linear regression or Monte Carlo sampling. In contrast, our Bayesian Optimization approach provides a more efficient, principled search of the CGM parameter space, achieving comparable or superior results with fewer computations and limited field data.

WOFOSTGym: A Crop Simulator for Learning Annual and Perennial Crop Management Strategies

William Solow[1], **Sandhya Saisubramanian**[1], **Alan Fern**[1]
{soloww,sandhya.sai,alan.fern}@oregonstate.edu

[1]**Oregon State University, Corvallis, Oregon**

Abstract

We introduce WOFOSTGym, a novel crop simulation environment designed to train reinforcement learning (RL) agents to optimize agromanagement decisions for annual and perennial crops in single and 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 RL. However, the lack of simulators for perennial crops in multi-farm contexts has hindered RL applications in this domain. Existing crop simulators also do not support multiple annual crops. WOFOSTGym addresses these gaps by supporting 23 annual crops and two perennial crops, enabling RL agents to learn diverse agromanagement strategies in multi-year, multi-crop, and multi-farm settings. Our simulator offers a suite of challenging tasks for learning under partial observability, non-Markovian dynamics, and delayed feedback. WOFOSTGym's standard RL interface allows researchers without agricultural expertise to explore a wide range of agromanagement problems. Our experiments demonstrate the learned behaviors 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 optimal irrigation scheduling can improve crop yield (Elliott et al., 2014). Motivated by promising results in other areas of precision agriculture, researchers and government agencies are increasingly interested in applying reinforcement learning (RL) to crop management decision problems in open-field agriculture, particularly for perennial crops (e.g., pears, grapes) (Astill, 2020; Gautron et al., 2022a).

Agriculture presents key challenges for RL, making it a valuable testbed for research: (1) *delayed feedback*—actions like fertilization affect yield only months later, complicating credit assignment; (2) *sparse rewards*—since yield is only known at the episode's end, learning an optimal policy is difficult (Vecerik et al., 2018); and (3) *partial observability*—many crop and soil states are unmeasurable or costly to obtain. 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). We bridge this gap by presenting a simulator for annual and perennial crops in single and multi-farm settings.

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 death and soil degradation (Tevenart & Brunette, 2021). Similar challenges in robotics and autonomous

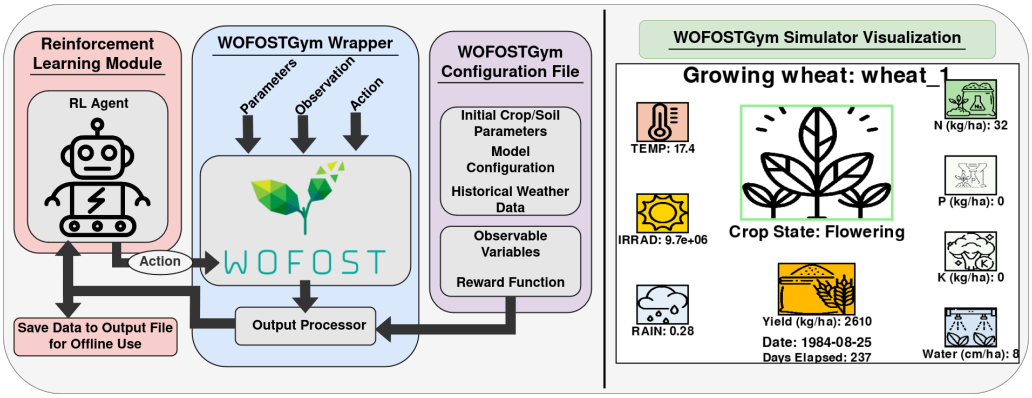

Figure 1: The structure and visualization of the WOFOSTGym simulator. WOFOSTGym provides an API around the WOFOST Crop Growth Model with a variety of environments to train RL agents and generate data. Well documented configuration files control crop and soil dynamics.

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, they are *not* designed to interact with RL algorithms and require substantial domain expertise to use.

Existing agriculture simulators (see Table 1) only simulate the growth of a *single* annual crop. They lack the functionality needed for perennial crop management as they do not capture crop growth across multiple years, including the dormant season (Forcella, 1998). Moreover, these simulators cannot be customized to study other crops or sites without domain knowledge of the underlying CGM and cannot learn joint agromanagement policies for multiple farms. Open-field agriculture problems are often modeled as a partially observable environment, but the current crop simulators do not allow the user to change the observable features 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 learning annual and perennial crop management strategies across single and multiple farms. WOFOSTGym is built on the WOFOST CGM (van Diepen et al., 1989) to model the growth of perennial crops, and includes high fidelity parameter sets for 23 annual crops and two perennial crops, calibrated with real-world data. Each crop contains between one and ten varieties. As a step towards high-fidelity modeling of perennial crop growth, we employ 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: `https://tinyurl.com/WOFOSTGym-Docs`.

Our experiments highlight scenarios in WOFOSTGym where standard RL algorithms and imitation learning (IL) agents achieve optimal performance, alongside more complex cases that remain difficult, underscoring opportunities for advancing learning approaches in agromanagement for both annual and perennial crops. We also design agromanagement decision-making tasks in WOFOSTGym that illustrate both the potential and challenges of applying RL to agriculture, positioning WOFOSTGym as a rigorous testbed for developing and evaluating new algorithms.

2 Background and Related Work

Partially Observable Markov Decision Process We formulate our agromanagement problems using the framework of partially observable Markov decision process (POMDP) (Kaelbling et al.,

Name	Perennial Crop Support	Multiple Crops and Farms	Easily Customizable	Models Crop Sub-processes
CyclesGym	✗	✗	✓	✓
CropGym	✗	✗	✗	✓
gym-DSSAT	✗	✗	✗	✓
SWATGym	✗	✗	✗	✓
Chen et al. (2021)	✗	✗	✗	✓
FarmGym	✗	✗	✓	✗
WOFOSTGym (Ours)	✓	✓	✓	✓

Table 1: 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.

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 $M = \langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \Omega, \mathcal{O} \rangle$ 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. A reward discount factor γ determines the importance of immediate 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.

RL 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. Tao et al. (2023) proposed an imitation learning approach to learn expert actions under partial observability, but it has not been tested in the real-world. 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). Chen et al. (2021) proposed a similar environment to CropGym, with a finer-grained time step based on the SIMPLE crop model. 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). SWATGym uses a simplified EPIC crop model and focuses primarily on soil-level properties (Madondo et al., 2023). 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* (Gautron et al., 2022a). Additionally, customization is infeasible without expert knowledge of the underlying CGMs, since most CGMs are run through separate executables. In contrast, WOFOSTGym offers easy domain customization for RL researchers while providing high-fidelity parameters for 23 annual and two 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). 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). None of the available CGMs support perennial crops. The relevant features of these CGMs are highlighted in the Supplementary Materials.

Our simulator is built on WOFOST, a CGM that models annual crop growth subject to nutrient (nitrogen, phosphorus, and potassium) and water-limited conditions (van Diepen et al., 1989). We choose WOFOST because studies show that it can be modified to model the growth of perennial crops with 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 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. Agromanagement decisions supported in WOFOSTGym are: fertilizing, irrigating, planting, and harvesting. In the interest of clarity, we focus on fertilization and irrigation decisions in the rest of the paper, since these tasks 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 and across one or more farms.

The rest of this section is organized as follows. We begin with an overview of the environment design. We then propose a model calibration method to fine-tune the model parameters of the WOFOST CGM to increase the fidelity as a step towards sim-to-real transfer (Peng et al., 2018).

3.1 Environment Design

A WOFOSTGym instance is defined by its Gym environment ID, the reward wrapper, and an agromanagement configuration file. WOFOSTGym contains 54 Gym environments that relate to annual and perennial crop simulation, single and multi-farm simulations, and six combinations of nutrient-limited environments. Our documentation includes three examples on how to modify the reward function, if needed, via the Gym reward wrappers. The agromanagement YAML file defines crop and soil dynamics and specifies the weather data which is provided by the NASAPower database. Gym environments, reward wrappers, and agromanagement files are *configurable*, allowing customization to simulate agromanagement decision problems across various crops, farms, and tasks.

States and Observations The model state in WOFOSTGym is the concatenation of two feature vectors, $\mathbf{c} = (c_1 \dots c_{203})$ and $\mathbf{w} = (w_1, \dots, w_7)$, where $\mathbf{c}$ contains the crop and soil state and $\mathbf{w}$ contains the weather state for a given day. However, most of these state features are not directly observable in the real-world. Thus, the state features available to the RL agent are a subset of the model state as observation $\mathbf{o} = (o_1, ..., o_n)$, with $n \ll 210$. An observation could be: $\mathbf{o} =$ (Weight of Storage Organs, Development Stage, Leaf Area Index, Soil Moisture Content, Rainfall, Solar Irradiation, Daily Temperature). WOFOSTGym supports any combination of state features as an observation. In the multi-farm environments, the agent receives an observation for each farm and the daily weather observation is shared across farms.

Action Space WOFOSTGym's action space consists of fertilization (F): nitrogen (n), phosphorus (p), and potassium (k), and irrigation (I) actions. At each time step, an action can be chosen from $A = \{F_n, F_p, F_k, I\}$ which corresponds to applying fertilizer (F_i) or water (I) in the following amounts, where f, n, i, and m can all be modified:

$$F_i = \left\{ f \cdot k \tfrac{\text{kg}}{\text{ha}} \middle| k \in \{0 \dots n\} \right\}, I = \left\{ i \cdot k \tfrac{\text{cm}}{\text{ha}} \middle| k \in \{0 \dots m\} \right\}$$

meaning that $|A| = 3n + m$. By default, a time step represents a single day, but can be modified to denote multiple weeks to model the varying length between agromanagement decisions.

Reward Real-world agriculture requires balancing yield with constraints such as fertilizer costs, water usage limits, and surface runoff restrictions. WOFOSTGym includes reward wrappers to penalize the violation of these constraints. Profitability is the primary driver of agromanagement policy adoption (Turchetta et al., 2022), so reward is a function of crop yield in kg/ha times a variable profit coefficient (by default we use 10, in line with previous work Overweg et al. (2021)). However, to enable wide-ranging task specification, WOFOSTGym's reward wrapper design enables the reward to be any function of the *entire state space*. An example reward function in WOFOSTGym is: $R_t = \text{Yield} - C \cdot (F_t + I_t)$, where C is a constant that modifies the penalty for nutrient application.

Domain Randomization Domain randomization enables successful sim-to-real transfer (Mehta et al., 2020), a key feature missing from existing crop simulators. WOFOSTGym supports three types of domain randomization, which can be used individually, in combination, or not at all. They are: 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.

Available Crops and Modifications to WOFOST WOFOSTGym includes parameters for 23 annual crops and 2 perennial crops which were all calibrated empirically from field data (de Wit, 2025; Wang et al., 2022; Bai et al., 2019). For perennial crops, it is insufficient to model individual seasons of crop growth, as important agromanagement decisions are made during the dormant season (Forcella, 1998). It is more appropriate to model growth over multiple consecutive years, which requires modification to the phenology, crop organ, and nutrient balance modules WOFOST. We outline the modifications made to WOFOST and list all available crops in the Supplementary Materials.

3.2 Parameter Calibration for Crop Growth Models

Before a CGM can be used for sim-to-real transfer with RL, it must be calibrated (Bhatia, 2014). CGM parameters are typically derived from field experiments and optimized using regression to find the best fit (Berghuijs et al., 2024; Zapata et al., 2017). Parameter spaces for CGMs are high-dimensional and highly non-linear (Sinclair & Seligman, 2000), so brute force and regression techniques that are commonly used in agronomy research may be insufficient to find an optimal solution. To overcome the limitations in current CGM calibration methods, we propose a Bayesian optimization approach that requires minimal domain knowledge and outperforms regression-based methods. When historical crop data is available, Bayesian optimization is a more principled way of exploring the parameter space to increase the model fidelity of WOFOSTGym.

Example: Bayesian Optimization for Grape Phenology Calibration Grape phenology is divided into three key phenological stages: Bud Break, Bloom, and Veraison (Lorenz et al., 1995). Accurately predicting the onset of a phenological stage allows growers to implement effective agromanagement strategies, and the Root Mean Squared Error (RMSE) is the widely accepted measure of performance in grape phenology modeling (Parker et al., 2013). We use an *iterative* optimization process that uses Bayesian optimization in each iteration to refine parameters and minimize error across *all* phenological stages. Phenology in WOFOSTGym is a crop sub-module and is described by a set of seven parameters, θ. Each iteration aligns with minimizing RMSE for a stage k, where θ_k is a subset of θ.

Using a dataset of six to 15 years of historical weather and phenology observations per cultivar collected by Zapata et al. (2017), we define the following loss function for Bayesian Optimization:

$$L_{RMSE}(\theta_k) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(P_i^k(\theta_k) - O_i^k\right)^2 + \frac{1}{n}\sum_{i=1}^{n}\left(P_i^{k-1}(\theta_k) - O_i^{k-1}\right)^2}$$

Cultivar	Bud Break			Bloom			Veraison			Cumulative Error		
	Ours	BB-T_b	BL-T_b	Ours	BB-T_b	BL-T_b	Ours	BB-T_b	BL-T_b	Ours	BB-T_b	BL-T_b
Cabernet Franc	**4.0**	6.1	6.2	3.5	3.1	**2.9**	7.7	**6.7**	7.1	**15.2**	15.9	16.2
Cabernet Sauvignon	**5.0**	8.7	10.5	**5.2**	5.8	5.7	9.8	**6.6**	7.0	**20.0**	21.1	23.2
Malbec	**3.7**	5.6	6.2	**2.8**	3.2	2.9	8.3	**5.7**	6.0	14.8	**14.5**	15.1
Pinot Noir	**3.6**	4.2	3.9	2.4	2.6	**2.3**	8.6	**6.6**	7.7	14.6	**13.4**	13.9
Zinfandel	**3.7**	6.8	9.0	**3.8**	4.3	4.0	6.0	4.1	**3.8**	**13.5**	15.2	16.8
Chardonnay	7.2	6.3	**5.9**	4.1	3.7	**3.2**	7.8	**5.6**	5.9	19.1	15.6	**15.0**
Chenin Blanc	**5.0**	6.1	6.2	**3.8**	4.8	4.6	**8.5**	9.2	9.4	**17.3**	20.1	20.2
Sauvignon Blanc	**3.4**	6.4	5.7	5.9	3.7	**3.5**	**1.6**	7.7	8.5	**10.9**	17.8	17.7
Semillon	**4.7**	6.0	7.0	**2.7**	6.0	5.8	**8.8**	11.2	11.6	**16.2**	23.2	24.4
Riesling	**3.7**	4.2	5.7	3.8	4.1	**3.7**	**8.5**	**8.5**	9.0	**16.0**	16.8	18.4
Average	**4.4**	6.0	6.6	**3.8**	4.1	3.9	7.4	**7.2**	7.6	**15.6**	17.3	18.1

Table 2: RMSE in days when predicting the key phenological stages (Bud Break, Bloom, and Veraison) in ten grape cultivars. The columns represent the RMSE between the model's predicted phenology for a given parameterization and the observed phenology. Ours: Using parameter set tuned with Bayesian Optimization. BB-T_b: Parameter set tuned for Bud Break. BL-T_b: Parameter set for Bloom. Values for BB-T_b and BL-T_b are columns 2 and 3 in Table 6 in Zapata et al. (2017).

where $P_i^k(\theta_k)$ and O_i^k denote the predicted and observed onset day for phenological stage k for year i with parameter set θ_k. We run three iterations of Bayesian optimization (Noguiera, 2014) with a RBF kernel and the expected improvement acquisition function for 500 steps. By retaining the best-fit parameters found by each iteration, we obtain $\theta = \{\theta_{\text{Bud Break}}, \theta_{\text{Bloom}}, \theta_{\text{Veraison}}\}$, which minimizes the RMSE across all phenological stages. We compare our Bayesian Optimization results with Zapata et al. (2017) who use linear regression. They find parameter sets for grape phenology, BB-T_b and BL-T_b, that aim to minimize the error for Bud Break and Bloom, and report the RMSE for all stages. Our results in Table 2 show that our model outperforms others, providing a 10% reduction in RMSE over the next best parameter set, BB-T_b.

The 32 calibrated grape phenology parameterizations included in WOFOSTGym increase model fidelity and represent a step towards sim-to-real transfer for crop management policies in open-field agriculture. Grape growers can use the high-fidelity phenology models in WOFOSTGym as *digital twins* to examine the effects of different agromanagement policies on their grape vines without the risk of crop loss. As more crop data becomes widely available, our Bayesian optimization method can be used to calibrate WOFOSTGym parameters for all crop sub-modules to more accurately model a variety of crop processes. In the absence of publically available historical data, we rely on previously calibrated crop parameters from de Wit (2025); Wang et al. (2022); Bai et al. (2019).

4 Experiments and Results

To illustrate the use of WOFOSTGym, we ran RL and IL experiments on diverse tasks to learn crop management policies for annual and perennial crops under realistic constraints. We present results using varied crops and soil types, demonstrating WOFOSTGym's customizability. Overall, our results showed that off-the-shelf RL algorithms struggle with hard constraints, long horizons, and delayed feedback—challenges inherent to agriculture and captured by WOFOSTGym, making it a valuable platform for both core RL research and agromanagement decision support.

Our crop selection was guided by common agronomic challenges: wheat for nutrient-limited growth due to its high nitrogen and water demands, potatoes for soil nutrient runoff risks, and grapes for their sensitivity to irrigation timing. We tested with pears and jujubes for the long-horizon decision-making challenges in perennial crop management, and maize as it is the only crop supported by other simulators.

Agents in our experiments could choose from 16 actions, each corresponding to one of four discrete amounts of nitrogen, phosphorus, potassium, and water. Unless otherwise noted, the agent observed a small subset of the 210 features: the development stage; weight of storage organs; total nitro-

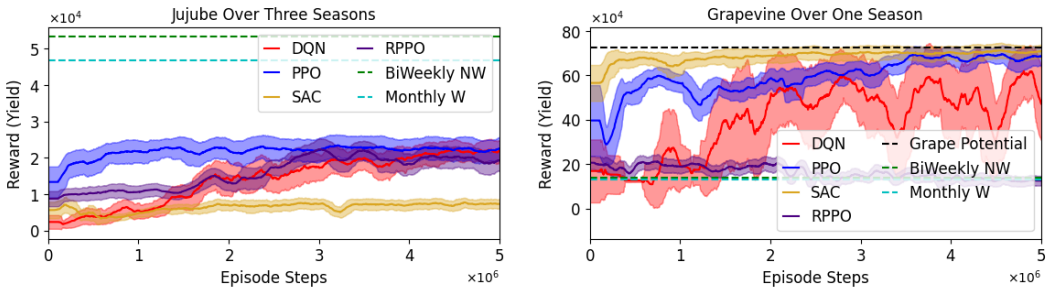

Figure 2: Unconstrained Control. The average reward, as seasonal yield, of different policies. The BiWeekly NW policy alternates applying nitrogen and water biweekly, the Monthly W policy applies water every month, and the Grape Potential is the maximum growth obtainable. Average and standard deviation reported over three seeds.

gen, phosphorus, potassium, and water applied; soil moisture content; nitrogen, phosphorus, and potassium in subsoil; solar irradiation; average temperature; and rainfall. Thus, all experiments are conducted under partial observability, reflecting the reality of decision making in agriculture. For our RL experiments, we used PPO, Recurrent-PPO, SAC, and DQN (Schulman et al., 2017; Haarnoja et al., 2018; Mnih et al., 2013), using implementations from Huang et al. (2021) and hyperparameters tuned experimentally to yield best performance in the WOFOSTGym domain. For our IL experiments, we used implementations from Gleave et al. (2022) of BC, GAIL, and AIRL (Bain & Sammut, 2000; Ho & Ermon, 2016; Fu et al., 2018). GAIL and AIRL used a PPO policy, and BC used an Actor Critic Policy, provided by Raffin et al. (2021). When reporting statistics of trained RL agents, we evaluated the trained RL agent on 15 different years of weather data and compute the average. All experiments and code can be found at: `https://tinyurl.com/WOFOSTGym-Code`.

Learning Efficiency Figure 2 presents learning curves for maximizing jujube growth over three seasons and grapevine over one season in WOFOSTGym. We compared RL performance against (1) the maximum potential yield; (2) an agromanagement policy that alternated nitrogen fertilization and irrigation biweekly; and (3) and a policy that applied irrigation monthly. In the grapevine experiment, the RL algorithms significantly outperformed the baseline of a bi-weekly nitrogen and water application policy but fell short of reaching the potential production of an unlimited nutrient setting. For jujube, we saw that RL agents were unable to match the performance of a monthly nitrogen and water application policy. We omitted the jujube potential because it assumed daily intervention while we only allowed the RL agents to perform biweekly interventions. These examples show the potential for off-the-shelf RL algorithms to achieve non-trivial performance for some crop scenarios, but also indicate that there is significant room for improvement in multi-year settings.

Learning Under Constraints In the real world, yield maximization is always subject to multiple constraints, such as a limit on the amount of fertilizer and water that can be applied per season. To model this, we rewarded total yield and applied a large negative penalty if the fertilization and irrigation thresholds (in kg/ha and cm/ha) were exceeded. Figure 3 shows the results of RL algorithms with this reward function: positive reward indicates no constraint violation, rewards less than zero indicate a constraint violation, and rewards less than -10^5 indicate more than five constraint violations. Note that unlike the previous experiment, there is no principled way to find the maximum reward obtainable in this setting. We compared the RL agents to a baseline that applies nitrogen fertilizer and water until it meets the same thresholds of 80 kg/ha of fertilizer or 40 cm/ha of water (Bushong et al., 2016). While this baseline satisfied constraints, it achieved a lower average reward than the trained RL agents. While this simple approach may be insufficient to guarantee that constraint satisfaction in the real world, the ability to construct such experiments demonstrates WOFOSTGym's potential as a testbed for constrained RL research.

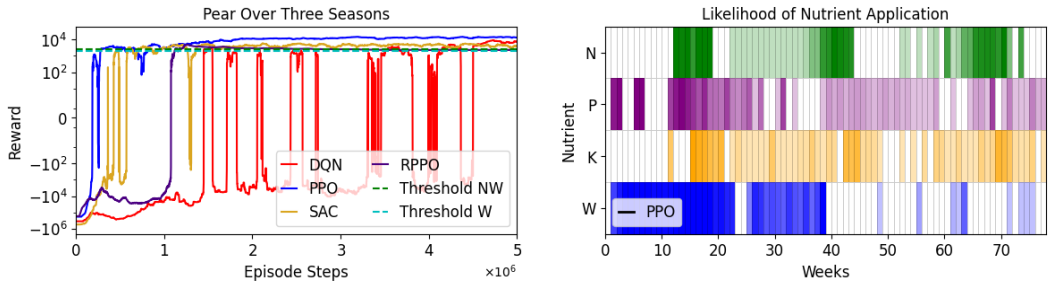

Figure 3: Constrained Control. (Left) The average over three seeds of the reward during training. (Right) The likelihood of fertilization or irrigation action each week. Likelihoods were computed over 15 seasons of weather data with darker colors signifying more likely nutrient application.

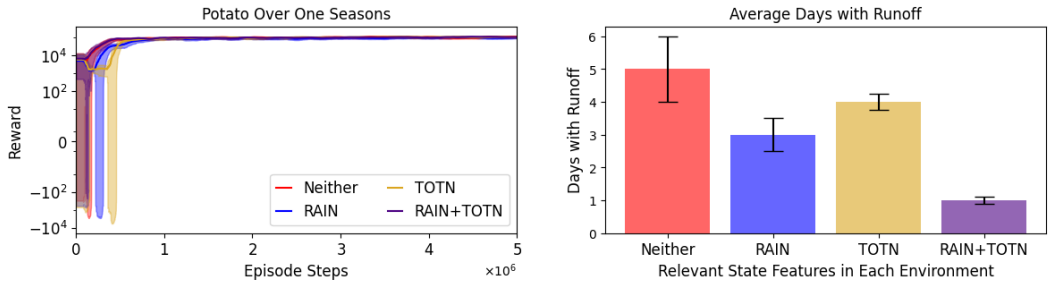

Figure 4: Constrained Control Under Partial Observability. (Left) The average reward over three seeds of RPPO agents during training. (Right) The average days of runoff after completing training. Evaluation on 15 seasons of weather data.

Effect of Reduced Observability on Constraint Adherence Limitations on sensing capabilities further restrict the observation space and are a constant source of uncertainty in agromanagement decisions. To illustrate the effects of reduced observability in WOFOSTGym (our agents already observe a subset of the 210 state features), we omitted two state features, RAIN and TOTN, the daily rainfall and fertilizer on the soil surface, respectively, and created four environments with fewer observable features based on the omission of the two variables from the observation space. We designed a reward function that rewarded crop yield subject to a -10^4 penalty if nutrient runoff occurred. Nutrient runoff happens when fertilizer amasses on the soil surface and irrigation or rainfall occur. We trained RPPO agents to grow the potato crop in each environment and show the results in Figure 4. Access to all relevant features improved constraint adherence of an agent policy. However, even in the fully observable case, constraint satisfaction was not guaranteed, exhibited by the non-zero days of runoff on average. Future research could use WOFOSTGym to inform the importance of obtaining costly field measurements, and study constraint inference in partially observable environments.

Imitation Learning for Agromanagement Decisions Given the difficulty of learning in the real world with RL, learning from past farmer demonstrations may be a viable option. We investigated the ability to learn from demonstrations using BC, GAIL, and AIRL algorithms. We provided each IL agent with 100 seasons of data generated from an expert PPO agent trained to maximize wheat yield subject to strict limits of 20 kg/ha of fertilizer and 20 cm/ha of water per season. We only included trajectories where the PPO agent successfully adhered to the underlying operational constraints. Table 3 shows that while GAIL and AIRL failed to match the expert's yield and behavior as shown by the differences in nitrogen and water application, BC most closely mimicked the expert but was unable to avoid all constraint violations. This shows that WOFOSTGym can serve as a non-trivial benchmark for IL and in particular, research on implementing constraints into IL.

Agent	Max Yield (kg/ha)	Constraints Violated	Nitrogen (kg/ha)	Phosphorus (kg/ha)	Potassium (kg/ha)	Water (cm/ha)
Expert (PPO)	4376 ± 805	0.00 ± 0.00	16.53 ± 3.3	6.13 ± 2.47	14.53 ± 3.46	3.33 ± 0.83
AIRL	2975 ± 335	4.53 ± 2.33	28.93 ± 5.56	2.67 ± 2.02	4.40 ± 2.65	1.80 ± 0.93
GAIL	2647 ± 562	0.67 ± 0.94	8.93 ± 2.72	9.47 ± 4.10	19.87 ± 4.29	4.00 ± 1.13
BC	4598 ± 790	0.33 ± 0.60	16.8 ± 3.56	6.53 ± 3.38	14.67 ± 2.98	4.30 ± 1.25

Table 3: Results of three IL agents trained to maximize wheat yield subject to constraints on nutrient application. The Constraints Violated column shows the number of days where excess nutrients were applied after the threshold was reached. Results are averaged over 15 seasons.

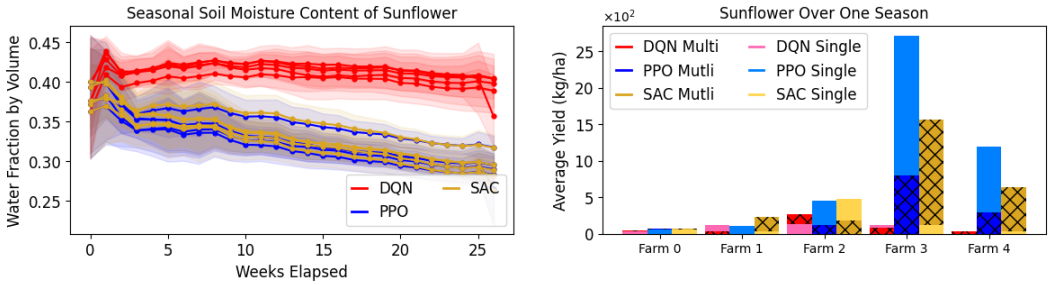

Figure 5: (Left) The soil moisture content of each field under three joint RL agromanagement policies. Averages taken over 15 seasons. (Right) The average yield obtained by trained multi-field agents. Lighter colors indicate the yield obtained by an agent trained on that specific field as a baseline for obtainable crop yield.

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 exploratory actions decreasing crop yield. WOFOSTGym enables agromanagement policies to be evaluated in simulation which could be a useful tool for farmers to understand the impacts of agromanagement decisions on crop and soil health. WOFOSTGym instances describe the dynamics of a field growing a specific crop. Each field can represent a farm. Using WOFOSTGym, we compared joint multi-field policies with field-specific policies to analyze their trade-offs in accumulated yield.

As farmers often apply the same policy to multiple fields, we created a WOFOSTGym environment that simulates the growth of five sunflower fields experiencing the same weather. The observation space was the growth and soil variables for each field. The weather was shared between fields and the action selected was uniformly applied to each field. We trained a PPO, DQN, and SAC agent in this multi-field scenario and reported the soil moisture content and the average yield with each agent policy on each field in Figure 5. We then trained the respective agents on each individual field to understand the value of using a specialized policy compared to a joint policy. The increased soil moisture content achieved under the DQN policy led to the lowest yield across all fields, providing meaningful insight into soil dynamics and the value of learning field-specific irrigation policies.

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 benchmarked the run times of three crop simulators: WOFOSTGym, CyclesGym, and gym-DSSAT (Gautron et al., 2022b; Turchetta et al., 2022). We compared run times for a single episode of growing the maize crop (155 episode steps). Given the large potential overhead when resetting the underlying CGM, we also measured the run times of the step and reset functions. 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 was also faster than gym-DSSAT, due to its significantly faster reset function while also maintaining phosphorus and potassium nutrient balances.

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

Table 4: The average runtime and standard deviation, computed over 100 trials, of three different crop simulators on an Nvidia 3080Ti.

5 Limitations and Future Work

WOFOSTGym takes around two seconds to run a three-year simulation of a perennial jujube crop. Although WOFOSTGym offers an improved run time compared to other crop simulators, the run time quickly adds up when RL algorithms require millions of episodes to learn a good policy. As episode horizon increases for modeling perennial crop management decisions, accelerating the modeling of crop dynamics will become critical.

Although WOFOSTGym provides high fidelity models for many annual and perennial crops, it was not designed for direct sim-to-real transfer. Such a simulator would also need to consider long-term crop rotation strategies and exogenous processes such as weed growth and pests, which impact farming operations seasonally. Creating a simulator for direct sim-to-real transfer is a long term goal. As research bridging RL to open-field agriculture advances and CGM fidelity improves through approaches like those in Section 3.2, direct sim-to-real transfer may become feasible.

6 Summary

We present WOFOSTGym, the first RL simulator for annual and perennial crop management decision support. The WOFOSTGym repository includes high-fidelity parameters for two perennial crops and 23 annual crops, along with diverse pre-specified agromanagement policies for benchmarking RL agents. Its customizable design enables researchers to conduct experiments without requiring agricultural domain expertise. To improve CGM fidelity and facilitate sim-to-real transfer in open-field agriculture, we propose a Bayesian optimization-based calibration method. Our results reveal the limitations of current RL and IL algorithms in this domain, emphasizing the need for further research to address the specific challenges presented in the agriculture domain. We outline realistic benchmarks to assess RL algorithms before deployment for agricultural decision support.

Broader Impact Statement

Reinforcement learning for crop modeling has the potential to help growers optimize yield while reducing costs and environmental impact. WOFOSTGym provides a high-fidelity platform for researchers to develop and evaluate agromanagement policies. However, due to the gap between simulation and real-world environments, RL policy performance in simulation may not translate directly to field trials.

Acknowledgments

The authors thank Dr. Markus Keller and his lab members, Lynn Mills and Zilia Khaliullina, at the Irrigated Agriculture Research and Extension Center (IAREC), Washington State University, for their invaluable support in collecting and sharing grape phenology data. This research was supported by USDA NIFA award No. 2021-67021-35344 (AgAID AI Institute).

References

Francisco Albornoz. Crop responses to nitrogen overfertilization: A review. *Scientia Horticulturae*, 205:79–83, June 2016.

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.

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.

Gregory Astill. Developing Automation and Mechanization for Specialty Crops: A Review of U.S. Department of Agriculture Programs: A Report to Congress, 2020.

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.

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), January 2019.

Michael Bain and Claude Sammut. A framework for behavioural cloning. In *Machine Intelligence 15*, pp. 103–129. Oxford University PressOxford, January 2000.

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.

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.

Jonathan Binas, Leonie Luginbuehl, and Yoshua Bengio. Reinforcement Learning for Sustainable Agriculture. In *Climate Change AI*. Climate Change AI, June 2019.

Kenneth J. Boote, James W. Jones, and Nigel B. Pickering. Potential Uses and Limitations of Crop Models. *Agronomy Journal*, 88(5):704–716, 1996.

Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. OpenAI Gym, June 2016.

Jacob T. Bushong, Jeremiah L. Mullock, Eric C. Miller, William R. Raun, and D. Brian Arnall. Evaluation of mid-season sensor based nitrogen fertilizer recommendations for winter wheat using different estimates of yield potential. *Precision Agriculture*, 17(4):470–487, August 2016.

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.

Mengting Chen, Yuanlai Cui, Xiaonan Wang, Hengwang Xie, Fangping Liu, Tongyuan Luo, Shizong Zheng, and Yufeng Luo. A reinforcement learning approach to irrigation decision-making for rice using weather forecasts. *Agricultural Water Management*, 250:106838, May 2021.

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.

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

Allard de Wit. ajwdewit/WOFOST_crop_parameters, January 2025. URL https://github.com/ajwdewit/WOFOST_crop_parameters.

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.

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.

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.

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.

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.

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.

Frank Forcella. Real-time assessment of seed dormancy and seedling growth for weed management. *Seed Science Research*, 8(2):201–210, June 1998.

Justin Fu, Katie Luo, and Sergey Levine. Learning Robust Rewards with Adversarial Inverse Reinforcement Learning, August 2018.

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.

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.

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.

Adam Gleave, Mohammad Taufeeque, Juan Rocamonde, Erik Jenner, Steven H. Wang, Sam Toyer, Maximilian Ernestus, Nora Belrose, Scott Emmons, and Stuart Russell. imitation: Clean Imitation Learning Implementations, November 2022.

Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor, August 2018.

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.

Matteo Hessel, Hubert Soyer, Lasse Espeholt, Wojciech Czarnecki, Simon Schmitt, and Hado van Hasselt. Multi-Task Deep Reinforcement Learning with PopArt. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01):3796–3803, July 2019.

Jonathan Ho and Stefano Ermon. Generative Adversarial Imitation Learning. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016.

Shengyi Huang, Rousslan Fernand Julien Dossa, Chang Ye, and Jeff Braga. CleanRL: High-quality Single-file Implementations of Deep Reinforcement Learning Algorithms, November 2021.

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.

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.

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.

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.

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.

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.

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.

Jens Kober, J. Andrew Bagnell, and Jan Peters. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, 32(11), August 2013.

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.

Leena Lindén, Hannu Rita, and Terhi Suojala. Logit Models for Estimating Lethal Temperatures in Apple. *HortScience HortSci*, February 1996.

D.h. Lorenz, K.w. Eichhorn, H. Bleiholder, R. Klose, U. Meier, and E. Weber. Growth Stages of the Grapevine: Phenological growth stages of the grapevine (Vitis vinifera L. ssp. vinifera)—Codes and descriptions according to the extended BBCH scale. *Australian Journal of Grape and Wine Research*, 1(2):100–103, 1995.

Malvern Madondo, Muneeza Azmat, Kelsey Dipietro, Raya Horesh, Michael Jacobs, Arun Bawa, Raghavan Srinivasan, and Fearghal O'Donncha. A swat-based reinforcement learning framework for crop management. *arXiv preprint arXiv:2302.04988*, 2023.

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.

Bhairav Mehta, Manfred Diaz, Florian Golemo, Christopher J. Pal, and Liam Paull. Active Domain Randomization. In *Proceedings of the Conference on Robot Learning*, pp. 1162–1176. PMLR, May 2020.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing Atari with Deep Reinforcement Learning, December 2013.

Sergi Munné-Bosch. Aging in Perennials. *Critical Reviews in Plant Sciences*, 26(3):123–138, June 2007.

Fernando Noguiera. Bayesian Optimization: Open source constrainted global optimization tool for Python, 2014. URL https://github.com/bayesian-optimization/BayesianOptimization.

Hiske Overweg, Herman N. C. Berghuijs, and Ioannis N. Athanasiadis. CropGym: a Reinforcement Learning Environment for Crop Management, April 2021.

Amber Parker, Inaki Garcia de Cortázar-Atauri, Isabelle Chuine, Gérard Barbeau, Benjamin Bois, Jean-Michel Boursiquot, Jean-Yves Cahurel, Marion Claverie, Thierry Dufourcq, Laurence Gény, Guy Guimberteau, Rainer W. Hofmann, Olivier Jacquet, Thierry Lacombe, Christine Monamy, Hernan Ojeda, Laurent Panigai, Jean-Christophe Payan, Begoña Rodriquez Lovelle, Emmanuel Rouchaud, Christophe Schneider, Jean-Laurent Spring, Paolo Storchi, Diego Tomasi, William Trambouze, Michael Trought, and Cornelis van Leeuwen. Classification of varieties for their timing of flowering and veraison using a modelling approach: A case study for the grapevine species *Vitis vinifera* L. *Agricultural and Forest Meteorology*, 180:249–264, October 2013.

Xue Bin Peng, Marcin Andrychowicz, Wojciech Zaremba, and Pieter Abbeel. Sim-to-Real Transfer of Robotic Control with Dynamics Randomization, March 2018.

Antonin Raffin, Ashley Hill, Adam Gleave, Anssi Kanervisto, Maximilian Ernestus, and Noah Dormann. Stable-Baselines3: Reliable Reinforcement Learning Implementations. *Journal of Machine Learning Research*, 22(268):1–8, 2021.

Antje Rohde and Rishikesh P. Bhalerao. Plant dormancy in the perennial context. *Trends in Plant Science*, 12(5):217–223, May 2007.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal Policy Optimization Algorithms, August 2017.

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.

Thomas R Sinclair and No'am Seligman. Criteria for publishing papers on crop modeling. *Field Crops Research*, 68(3):165–172, November 2000.

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.

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.

Camille Tevenart and Marielle Brunette. Role of Farmers' Risk and Ambiguity Preferences on Fertilization Decisions: An Experiment. *Sustainability*, 13(17), January 2021.

Howard Thomas, Huw Martin Thomas, and Helen Ougham. Annuality, perenniality and cell death. *Journal of Experimental Botany*, 51(352):1781–1788, November 2000.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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), January 2020.

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.

Supplementary Materials

The following content was not necessarily subject to peer review.

7 WOFOSTGym Modified Perennial Crop Growth Model

Prior works has modified the WOFOST CGM to model the growth of pear and jujube crops across multiple growing seasons de Wit (2024); Bai et al. (2019); Wang et al. (2022), establishing it as a strong foundation for developing a perennial crop simulator. Below, we outline the key modifications we made to WOFOST to support perennial crop modeling.

Perennial Phenology We primarily focused on modifying the phenology submodule within the WOFOST CGM to account for the substantial differences between annual and perennial crop phenology. Unlike annual crops, 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). To capture this behavior, we introduced parameters for dormancy induction based on day length, release temperature threshold, and minimum dormancy duration. In our 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 persist year-round, while the leaves and storage organs regrow each season subject to intercepted light and nutrient uptake. Crop organ death rates are modeled as a function of the development stage of the crop (Lindén et al., 1996). Notably, perennials exhibit a reduced seasonal growth as they age (Munné-Bosch, 2007). While the underlying mechanisms for reduced growth remain difficult to quantify, we model this decline empirically through increased maintenance respiration and decreased carbon conversion efficiency as a function of age (Zhu et al., 2021). See Figure 6 for a visual overview of how key crop processes impact one another throughout the course of a perennial crop simulation.

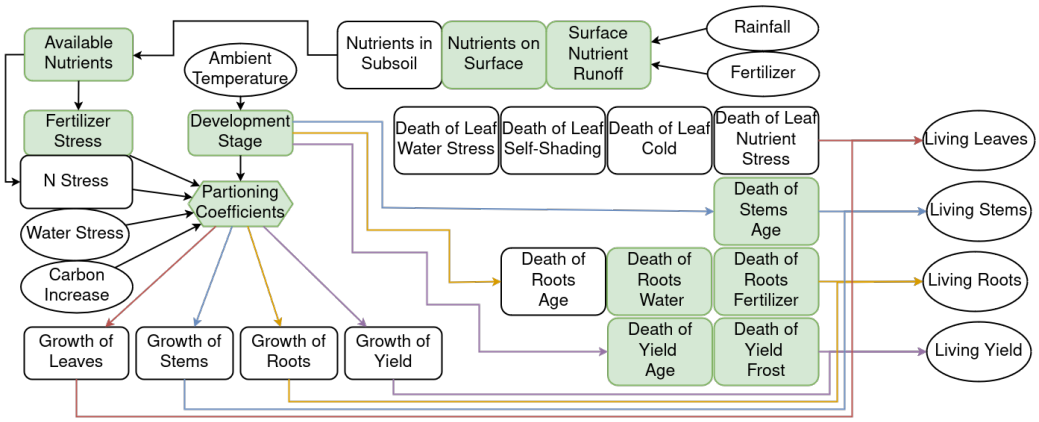

Figure 6: A simplified flowchart of the perennial crop growth in WOFOSTGym. Boxes highlighted in green denote additions or areas of substantial change to the underlying WOFOST CGM to support perennial crop growth. The development stage of the crop is driven by the daily ambient temperature. The development 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.

Modified Nutrient Module A multi-layer nutrient balance is important for modeling the effects of fertilization stressors on the roots, stems, and nutrient partitioning (Albornoz, 2016). We extend

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).

8 Available Crops

WOFOSTGym includes parameters for two perennial crops: pear and jujube, and 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. Each crop contains between *one and ten varieties*. WOFOST CGM parameters for each variety were calibrated empirically from field data (de Wit, 2025). By modeling each crop variety as a task, agromanagement decisions for multiple crop varieties can be optimized with multi-task RL (Hessel et al., 2019).

In addition to the high-fidelity models for 25 crops, WOFOSTGym also includes parameters for modeling the phenology of 31 grape cultivars. These cultivars are: Aligote, Alvarinho, Auxerrois, Barbera, Cabernet Franc, Cabernet Sauvignon, Chardonnay, Chenin Blanc, Concord, 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.

9 Crop Growth Models

CGMs are typically one of three types: mechanistic, empirical, or hybrid. Mechanistic models simulate canopy or nutrient level crop processes to validate scientific understanding of crop growth (Estes et al., 2013). Empirical models 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).

WOFOST is a single-year and multi-crop agroecosystem model (Jones et al., 2017). 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 shown that it can be used to accurately predict yield in perennial fruit trees with some small modifications to the base model Wang et al. (2022); Bai et al. (2019).

Given WOFOST's ability to simulate a wide variety of crop and soil dynamics, 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); de Wit et al. (2019). There are a wide variety of CGMs available for use. Table 5 outlines the desiderata used to select WOFOST as the CGM for WOFOSTGym.

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

Table 5: Different CGMs and their strengths and weaknesses for modeling high fidelity crop growth, interfacing with RL algorithms, and supporting perennial crop decision evaluation.