MULTI-AGENT PLANNING FOR COORDINATED ROBOTIC WEED KILLING

BY

WYATT SPALDING MCALLISTER

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2020

Urbana, Illinois

Doctoral Committee:

Assistant Professor Girish Chowdhary, Chair
Professor Adam Davis
Professor Rayadurgam Srikant
Associate Professor Mohamed Ali Belabbas

Abstract

This work presents techniques for predictive modeling of weed growth, as well as an improved planning index to be used in conjunction with these techniques, for the purpose of improving the performance of coordinated weeding algorithms being developed for industrial agriculture. We demonstrate that the evolving Gaussian process method applied to measurements from the agents can predict the evolution of the field within the realistic simulation environment Weed World. In addition to prediction, this method provides physical insight into the seed bank distribution of the field. In this work we extend the evolving Gaussian process model in two important ways. First, we have developed a model that has a bias term, and we show how it is connected to the seed bank distribution. Secondly, we show that one may decouple the component of the model representing weed growth from the component which varies with the seed bank distribution, and adapt the latter online. We compare this predictive approach with one that relies on known properties of the weed growth model, and show that the evolving Gaussian process method gives better results, even without assuming this model information. Finally, we use an improved planning index, entropic value-at-risk (EVaR) in conjunction with the Whittle index, which allows a balanced trade-off between exploration and exploitation, and ensures model improvement when used with these various prediction schemes.

To my family and friends, without whose help and support this work would not have been possible.

Acknowledgments

Support for this work was provided by the USDA National Institute of Food and Agriculture (USDA NIFA 2018-67007-28379) and by the National Science Foundation Cyber Physical Systems program (NS 1820332).

I thank the TerraSentia robot team, whose past work in weed recognition, design of the mechanical weeding system, and field navigation, has made this work possible.

I thank all of my collaborators. Thanks to Joshua Whitman, with whom I collaborated on the evolving Gaussian process (E-GP) prediction method for Weed World. Regards to Allan Axelrod, whose insight on risk theory allowed us to use entropic value-at-risk (EVaR) for this domain. I greatly appreciate the diligent work done by Joshua Varghese on the development of the visualization for the new Python Weed World environment. Finally, I thank Dr. Girish Chowdhary, and Dr. Adam Davis for their insights and support in reviewing this work.

Table of Contents

Chapter 1

Introduction

Consider a team of mechanical-weeding robots managing herbicide-resistant weeds on any row-crop farm in the US. This team of robots needs to predict the weed growth across the whole farm in order to make intelligent decisions on robot coordination [1]. However, none of the robots can observe the whole field, and even together, they can only observe a very limited part of the field at any given time. Furthermore, the robots do not have high bandwidth access to a high-performance computing (HPC) facility, nor would such a facility be able to provide highly accurate predictions of weed growth without real-time access to the data being gathered through the weeding process, as weed distribution depends on the field and the day. The number of emerged seedlings at a given location in the field is governed by the latent seed bank density within the soil at that location, which varies over different fields. For this reason, it is necessary to leverage past data on a given field to form a predictive model, and estimate the field state online using limited data.

Our team of robots learns a predictive model of the specific field using dense weed demographic data from before canopy closure, either at planting or harvest time. We then use sparse measurements from the robots after canopy closure to estimate the state of the field online. This enables the team of robots to more intelligently coordinate their actions, leveraging our earlier work [1], rather than resorting to grid search over the whole field. Our work addresses the key challenge facing multi-agent field-robotic systems operating in spatiotemporally changing domains: managing the trade-off between learning a robust generalized model, and learning a model which is specialized in order to most accurately explain the current observations of the agents. We use the evolving Gaussian process (E-GP) model [2], trained on ecologically realistic simulations of weed growth [1]. This model can predict spatiotemporal dynamics throughout a continuous domain using a finite linear dynamical system embedded in a feature space. We advance this modeling technique by showing the model may be decoupled into two components: the first is based on the general weed growth dynamics, and the second captures the unique properties of the spatial distribution of weed seeds across the current field. Another unique benefit of the E-GP approach is that it offers physical insight into properties of the spatiotemporally evolving domain.

The E-GP model is completely data-driven, making no assumptions regarding how weeds grow. To benchmark the performance of the E-GP method, we developed another prediction scheme, Gaussian Seed Bank (G-SB), which uses a single time invariant Gaussian Process in conjunction with statistics about weed growth to make weed density predictions based directly on an online estimate of the underlying seed bank density. An E-GP model trained on weed growth data not only outperforms G-SB, but does so without the benefit of assumptions on the weed growth dynamics. We also show that a revised planning index, using Entropic value-at-risk (EVaR) in conjunction with the Whittle index, improves planning performance over the previous planning scheme in [1], and may be used with the above prediction methods ensure that actions taken by the planner lead to improvement in the prediction.

1.1 The Herbicide Resistant Weed Problem

Weed management has historically relied on a combination of crop rotation, mechanical weed control, and the use of herbicides [3]. The evolution of herbicide-resistant weeds, coupled with the fact that new herbicide discovery has ceased in the past 30 years, has resulted in a crisis for agricultural weed management [4, 5]. Crop losses due to herbicide resistant weeds are approximately half a billion per year, and may climb to $100 billion per year when chemical control is lost [6]. Evolution of resistance to multiple sites of herbicide action is accelerating, especially in the southern and north-central U.S. grain production regions [7]. Increasingly, farmers are only one site-of-action away from total loss of chemical control. For example, the five-way multiple resistant waterhemp (Amaranthus tuberculatus [Moq.] Sauer) in Illinois is now one gene away from total loss of chemical control [8]. Seeds with bred-in herbicide resistance are exacerbating the herbicide resistance problem in soybean production [9]. An alternative to chemical weeding is mechanical weeding.

Mechanical weed management most often targets young weeds, including germinating seeds and seedlings that are extremely vulnerable to physical damage. Before crop planting, superficial soil disturbance and subsequent soil cultivation can remove germinated weeds. However, hand weeding of young weeds at the two-leaf growth stage is difficult and impractical at scale. Mechanized inter-row cultivation has disadvantages, such as soil compaction due to use of heavy machinery, and an inability to work after the crop canopy closes.

Drones are ineffective for collecting data during much of the crop season, as canopy closure removes aerial visibility of the ground. A team of collaborative low-cost and lightweight mechanical weeding ground robots (termed here as *agbots*, illustrated in Figure 1.1) may be used to control herbicide-resistant weeds. Such a team of *agbots* can target weeds within and between crop rows, as opposed to tractors, combines, and planters, which cannot be used after the crop canopy closes. The *agbots* are ideal for working in dense fields, since they are small enough to drive over plants without damaging them, and do not compact the soil. This approach necessitates algorithms for managing large fields with the least number of robots.

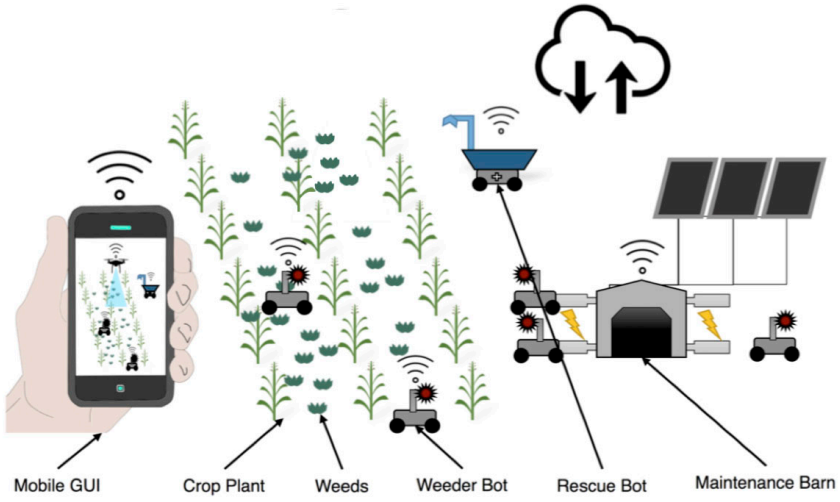


Figure 1.1: The *agbot* solution for robotic mechanical weed control is a dynamically configured team of weeder bots, and automated maintenance barns for persistent autonomous weed control, leveraging collaboration.

(a) TerraSentia

(b) Prototype Possible Weeding Apparatus

Figure 1.2: Prototypes of TerraSentia Robot and Weeding Apparatus in the field, which we use as the basis for our simulation parameters.

Termination of weed seedlings within the critical weed-free period [10], where crops are most vulnerable, is essential to preventing crop yield losses in corn and soybeans [11]. For many crops, weeding may be done under the canopy, and therefore under conditions of partial environmental information. To address this issue, several companies, such as TerraSentia [12], shown in Figures 1.2a and 1.2b, Ecorobotix [13], and Naio-Technologies [14], have developed small agricultural robots for autonomous weeding. For robots like these to be employed at scale, multi-agent planning strategies must solve the problem of coordination in field environments with limited observations.

This work builds upon past work presented at the International Conference on Intelligent Robots (IROS) 2018 [1]. In [1], a realistic simulation environment, Weed World, was developed and used to demonstrate the feasibility of a coordinated strategy for robotic weeding of fields with a range of characteristics. However, the coordination strategy used the average value for the weed density of previously weeded rows for the value of weeding unobserved rows. This current work will utilize evolving Gaussian processes (E-GP) [2] to proactively estimate the weed density of unobserved rows using past data, enabling more accurate planning strategies to be developed.

1.2 Mathematical Background

1.2.1 Gaussian Processes

A time continuous stochastic process is a Gaussian process if its trajectory over any finite time horizon, T, is a multivariate normal distribution. Here, X_t is the value of the process at location $(x, y) \in \Omega \times \Omega$.

$$\left(X_{t_1}, ..., X_{t_k}\right) \sim N\left(\mu, \Sigma\right) \tag{1.1}$$

The mean and variance of a Gaussian process can be written as follows.

$$\mu_{\tau,x} = w_\tau^T \phi(x) \tag{1.2}$$

$$\phi(x) = \begin{bmatrix} k(x, c_1) & \dots & k(x, c_{N_{\text{cent}}}) \end{bmatrix}^T \tag{1.3}$$

$$\sigma_{x,y} = \langle \phi(x), \phi(y) \rangle_{\mathcal{H}} \tag{1.4}$$

These kernels, k, and associated features, ϕ, satisfy the following law. Here, $\mathcal{H}$ is a reproducing kernel Hilbert space, N_{cent} is the number of centers, and Ω is the state space for cell x or y.

$$k \colon \Omega \times \Omega \to \mathbb{R}, \phi \colon \Omega \to \mathcal{H} : k(x, y) = \langle \phi(x), \phi(y) \rangle_{\mathcal{H}} \tag{1.5}$$

1.2.2 Evolving Gaussian Processes (E-GP)

Gaussian processes provide a model for the spatial distribution of a phenomenon across the entire time horizon of the process. However, there is no hard constraint on the temporal evolution of the process. In order to make precise predictions of the value of a process at a future time, a more general scheme is needed. One such scheme is evolving Gaussian processes (E-GP), which is a general prediction scheme [2] for spatio-temporally evolving phenomena, based on the kernel observer scheme [15]. Kernel observers is in turn built on an online Gaussian process (GP) framework [16], which creates an optimal sparse GP representation by updating a finite set of basis vectors in real time.

E-GP models spatio-temporally evolving processes by hierarchically separating the spatial and temporal dynamics. A set of radial basis function (RBF) kernels is trained over the space for each time step, and then the transition dynamics between them are learned. The E-GP model is given by the linear evolution and measurement equations, where the mean, $\mu_{\tau,x}$, of the process is measured as $y_{\tau,x}$, $\hat{A}$ is a linear transition operator, and w are weights in the kernel space.

$$w_{\tau+1} = \hat{A} w_\tau \tag{1.6}$$

$$y_{\tau,x} = \mu_{\tau,x} = (w_\tau)^T \phi(x) \tag{1.7}$$

The online GP framework utilized by E-GP leads to efficient computation. The hierarchical separation of the spatial and temporal dynamics further improves efficiency. This framework is able to easily use past observations to compute $\hat{A}$. Given the temporal model provided by $\hat{A}$, an online state estimate using current observations can provide w_τ, enabling prediction.

The procedure for E-GP is as follows:

1. After picking the kernel and estimating the bandwidth hyper-parameter σ, find an optimal basis vector set C using the algorithm in [16].
2. Use Gaussian process inference to find weight vectors for each time-step in the training set(s), generating the sequence $w_\tau, \tau = 1, ..., T$ for each system.
3. Using the weight trajectory, use matrix least-squares with the equation $\hat{A}[w_1, w_2, ..., w_{T-1}] = [w_2, w_3, ..., w_T]$ to solve for $\hat{A}$.

1.2.3 Bandit Problems

Consider a Markov decision process on a finite state space, $\mathbb{X}$. Let there be a set of discrete decision times, $t \in [T]$. The action set is binary, $a \in \{0, 1\}$. We assume that action $a = 0$ results in zero reward, and action $a = 1$ gives reward as follows.

$$R\left(x\left(t\right), a = 1\right) = \gamma^t r\left(x\left(t\right)\right) \tag{1.8}$$

If action $a = 0$ is taken, the state remains fixed. State transitions are instantaneous, and depend only on the current state.

In general we may have a system in which there are N bandit processes operating simultaneously, indexed by $i \in [N]$. In this case, the state space becomes $x \in \mathbb{X}_1 \times ... \times \mathbb{X}_N$. At each time instant, action $a_i = 1$ is applied to M bandits at a time, and action $a_i = 0$ is applied to all others. As before, only the bandits for which $a_i = 1$ is applied will change state.

The reward is rewritten a follows.

$$R\left(x_i\left(t\right), a_i = 1\right) = \gamma^t r_i\left(x_i\left(t\right)\right) \tag{1.9}$$

We choose between the M arms at every time step so as to maximize the infinite horizon expected return.

$$J\left(x_0\right) = \lim_{T \to \infty} \mathbb{E}\left[\sum_{t=0}^{T-1} \gamma^t r_i\left(x_i\left(t\right)\right) \mid x\left(0\right) = x_0\right] \tag{1.10}$$

We assume that we know the state x_i of each arm, the reward function r_i for that arm, as well as the discount factor γ.

1.2.4 The Gittins Index

The Gittins index, $G_i\left(x_i\right)$, is known to be an optimal index for the bandit problem [17].

$$G_i\left(x_i\right) = \sup_{\tau} \left\{ \frac{\mathbb{E}\left[\sum_{t=0}^{\tau} \gamma^t r_i\left(x_i\left(t\right)\right) \middle| x_i\left(0\right) = x_i\right]}{\mathbb{E}\left[\sum_{t=0}^{\tau} \gamma^t \middle| x_i\left(0\right) = x_i\right]} \right\} \tag{1.11}$$

Here, τ is the stopping time for the arm. The optimal policy is then that which pulls the M arms with the highest Gittins index at each time step.

$$a\left(t\right) = \arg \overset{(M)}{\max_{i \in [N]}} G_i\left(x_i\right) \tag{1.12}$$

This index can be interpreted as the reward per time, or rate of increase of reward, telling us that the policy accruing reward most efficiently is optimal.

1.2.5 Restless Bandits and The Whittle Index

In general, arms can change state at every time step, whether or not action $a_i = 1$ is applied. The reward for each arm now depends on both the state and action, as passive actions, $a_i = 0$, can give distinct rewards.

$$R\left(x_i\left(t\right), a_i\left(t\right)\right) = \gamma^t r_i\left(x_i\left(t\right), a_i\left(t\right)\right) \tag{1.13}$$

For each state $x_i\left(t\right)$ and action $a_i\left(t\right)$, we model the change in reward as:

$$\Delta r_i\left(x_i\left(t\right), a_i\left(t\right)\right) = r_i\left(x_i\left(t\right), a_i\left(t\right)\right) - r_i\left(x_i\left(t-1\right), a_i\left(t-1\right)\right) \tag{1.14}$$

Consider the Lagrange formulation of the problem, which takes argument λ, thought of as a charge to pull each arm. The Whittle index, $W_i\left(x_i\left(t\right)\right)$, is proposed [18].

$$J\left(\lambda\right) = \lim_{T\to\infty} \mathbb{E}\left[\sum_{t=0}^{T-1} \gamma^t \left(r_i\left(x_i\left(t\right)\right) - \lambda a_i\left(t\right) + \Delta r_i\left(x_i\left(t\right), a_i\left(t\right)\right)\right)\right] \tag{1.15}$$

$$W_i\left(x_i\left(t\right)\right) = \arg\min_{\lambda} : \left\{ \begin{array}{l} r_i\left(x_i\left(t\right), 1\right) - \lambda + P\left(x_i\left(t\right), 1\right) \Delta r_i\left(x_i\left(t\right), 1\right) \\ = r_i\left(x_i\left(t\right), 0\right) + P\left(x_i\left(t\right), 0\right) \Delta r_i\left(x_i\left(t\right), 0\right) \end{array} \right\} \tag{1.16}$$

Here, $P\left(x_i\left(t\right), a_i\left(t\right)\right)$ is the probability of a given state under an action. This is simply the infimum over the charge λ that makes pulling or not pulling each arm give equal return at a given time t. The Whittle index policy is that which chooses the M arms with the highest Whittle index at each time step. In certain conditions, this is asymptotically optimal [19].

$$a\left(t\right) = \arg\max_{i\in[N]}^{(M)} W_i\left(x_i\right) \tag{1.17}$$

1.3 Weed Growth Models

Our work includes a modeling solution for common waterhemp (*Amaranthus tuberculatus*). This model was chosen both for its computation efficiency in the use-case of a real-time dynamic weed simulation, and for its realism during the critical weed-free period [20]. Should offline data be available, there exist a wealth of innovative modeling techniques in weed science. We present a brief overview of various techniques here.

In [21], a formal taxonomy of weed modeling methods is presented. Several key model factors are outlined. These include the biology of weed species being modeled, the dynamics inherent to the weed population being modeled, extrinsic factors impacting the growth of that weed population, and economic factors which the weed growth will impact. Each model may differ in the way uncertainty is tracked, the modeling techniques utilized, and the means by which the model is validated.

Each weed species has a unique life cycle. Within this life cycle, viable seeds germinate before emergence, and mature plants produce new seeds to deposit in this seed bank. Key biological factors in a weed population include the portion of weeds in each stage of the life cycle, the portion of the seed bank which is viable, and the total biomass of the population.

The accuracy of growth models may depend on the density of weeds and other plants within the environment, and the competition for nutrients. This dependence necessitates a model of the weed population dynamics. These dynamics will vary based on the density of weeds and its effect on growth rate and seed production, competition between various weed species and between crops and weeds, the spatial distribution of the seed bank as influenced by the environment, and the herbicide resistance of weed populations.

Any weed population exists in a complex environment influenced by weather and resources, biological ecosystems, and human management of the weeds. Common environmental factors are light, moisture, and nutrients and their impact on both the growth of weeds and the effectiveness of management strategies, animals and diseases which impact the weed population, and management strategies including herbicides and mechanical tillage.

Many weed growth models explore the economic impact of weed populations. Some possible economic factors are the effect of weeds on crop yield, the optimal threshold of weed population density before treatment, and the best combination of herbicides and other treatment methods for a given population of weeds.

Various models have several key properties which assist in categorization. The first of these is uncertainty, or the way in which model confidence is computed, such as a sensitivity analysis on the effect of model parameters on model output, or a statistical variance incorporated into model predictions. The second of these is the technique used for modeling, such as a differential equation formulation, or a so called "matrix model" in which the temporal dynamics for each portion of the weed population is computed.

Weed growth models are validated via a comparison with long term data collected on a real weed population. Possible data for model comparison include the emerged seedling density, the seed bank density, the growth rate of seedling and seed bank densities, and the growth rate of seedlings in height and stem width. Validation is often cost prohibitive, and the set of long term data sets within the literature for a given species may be limited.

1.4 Coordinated Robotic Planning Algorithms

In [22], a formal taxonomy of cooperative robotic planning is presented. According to this taxonomy, the problem considered here is an asynchronous, homogeneous, centralized problem. This means that agents are assigned tasks at different times, agents have identical capabilities, and agents are assigned tasks from a centralized planner. This work also presents the problem domain of foraging, where robots move through an environment and collect objects or information. In this case, the foraging problem is framed in terms of recognizing and killing weeds while moving through the complex and uncertain field environment.

Past work has explored multi-robot task allocation (MRTA) in stochastic domains [23–25], leveraging both spatial constraints and predictive information to perform optimization. In [26], a formal taxonomy of MRTA is presented. Three major distinctions are made. The first is between single-robot problems, where each pool of tasks is managed by a separate robot, and multi-robot problems, in which each task pool is shared between multiple robotic agents. This problem is a multi-robot problem, as all agents cooperate to weed the field together in order to complete the weeding task more efficiently. This approach allows agents to adapt to changes in the environment, working together to effectively handle regions with more weeds.

The next distinction is between preemptive task allocation, in which optimization is performed continuously in an on-line manner, and agents may take over another agent's task or switch to another task before completion, and non-preemptive task allocation, in which tasks must be completed before a new task is assigned. A non-preemptive planning strategy is used, ensuring rows are completed before an agent is assigned a new row. This allows agents to plan once the task has been completed, allowing them to focus on navigation and plant recognition while in the row.

Another distinction is between single-agent tasks, where each task is performed by one agent, and multi-agent tasks, in which each task must be performed by multiple agents. Here, each robot is assigned to one row, so this problem is a single-agent task scenario, with multiple agents collaborating to complete a pool of single-agent tasks.

In [26], the problem of time-extended on-line assignment, in which multiple robots pick single-agent tasks from a pool larger than the number of agents, and complete them in a non-preemptive manner, is considered. This algorithm is an implementation of that proposed in [26] for the time-extended on-line assignment problem, which initially assigns each robot to the most suitable task, and then assigns robots to the most suitable task from the pool as they become available.

Chapter 2

Methods

2.1 Simulation Environment

The simulation environment, called Weed World (shown in Figure 2.1), was developed to allow large-scale simulations of coordinated weeding algorithms for multi-robot planning in uncertain environments [1]. This environment incorporates a realistic weed growth model (described in 2.2), as well as a framework for multi-agent collaboration, which enables a scalable amount of agents to easily share information. In this environment, the crops are assumed to be arranged in evenly spaced vertical rows.

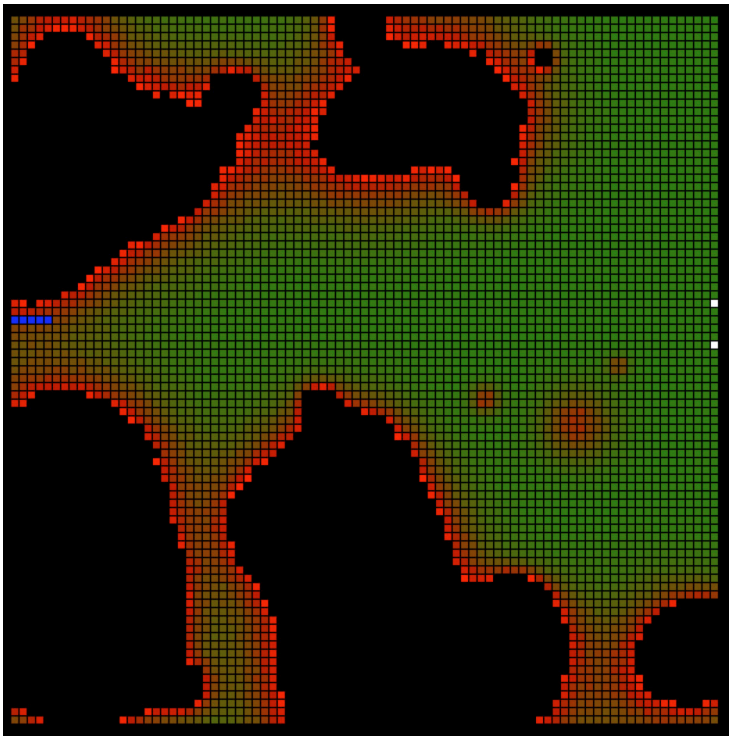


Figure 2.1: Simulation environment, Weed World, created in Python. Each cell represents a small 0.8 m square portion of the field. The colors of the squares represent weed seed bank density. Darker colors represent higher density. The agents are shown in solid blue.

2.2 Weed Generation

The weed growth model is composed of a matrix of cells (each representing 0.8 m^2 for a total of 0.4 hectares) evolving according to a random process, forming a cellular automata model [27]. This study will determine the number of robots needed per acre for effective weed management. We expect the algorithm to scale to larger fields in a straightforward manner. Seeds emerge from a finite, fixed seed bank, according to a time-inhomogeneous Poisson process. The parameters used, summarized in Table 2.1, are aligned with the growth model for the common waterhemp weed species specified in [28]. The density of the seed bank in each cell is $S(x, y, t)$, which is equal to S_0 (between 600 and 1560 seeds per cell) on average at time $t = 0$. The initial seed bank density in each cell $S_0(x, y)$ is chosen so that the Gini coefficient of concentration (GCC) between all the cells is between 0.31 and 0.35, which ensures that the relative density of weeds aligns with that seen in real experiments [28]. In order to achieve this distribution, we first give each cell a random density between zero and 20 percent of S_0, and then we create 50 patches of weeds with random centers and random radii up to 20 cells long, and fill those patches with an additional S_0 number of weeds distributed normally around each center.

The evolution of several fields is shown in Figure 2.2.

Table 2.1: Seed Bank Density Parameters: Consistent with Those Found in [28]

Parameter	GCC	S_0	Np. Patches	Patch Radius
Range	[0.31,0.35]	[600,1560]	50	[0, 20]

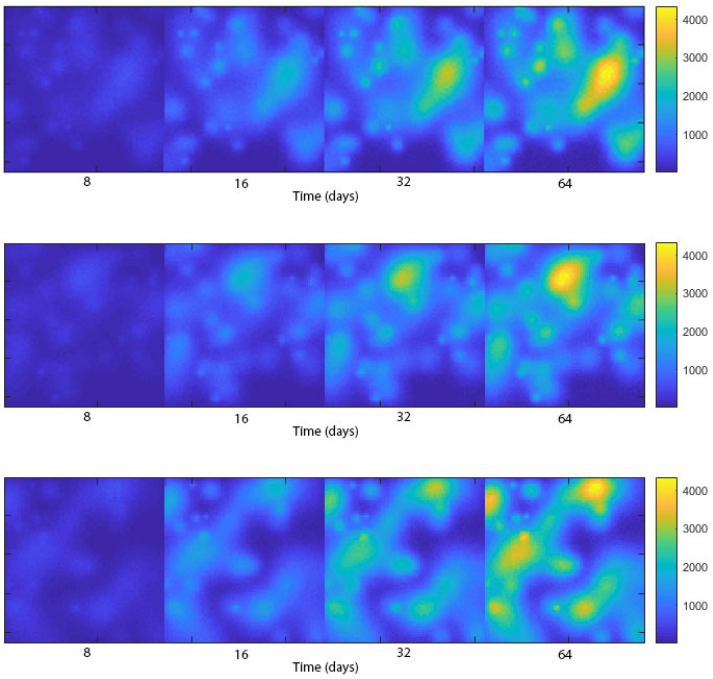


Figure 2.2: The evolution of different fields over time, as simulated in the Weed World environment. Differing seed bank distributions result in differing evolutions.

After initializing the simulation, a certain number of days, d_0, are allowed to elapse before the agbots begin weeding. The number of emerging weeds in each cell, N_{emerge}, is a randomly generated Poisson variable with mean, $\lambda\left(x,y,t\right)$, such that 90 percent of the seed bank, $S\left(x,y,t\right)$, emerges in T_{total}, which is two months. This emergence rate is aligned with past work [29–33], all of which present measurements of the seed bank densities for various species of weeds, and provide an analysis of weed growth models for these species.

$$\lambda_t\left(x,y,t\right) = \frac{0.9 \cdot \Delta t \cdot S\left(x,y,t\right)}{T_{\text{total}}}, \quad \lambda_0 = \frac{0.9 \cdot d_0 \cdot S_0}{T_{\text{total}}} \tag{2.1}$$

$$N_{\text{emerge}}\left(x,y,t\right) = \text{Poi}\left(\lambda_t\left(x,y,t\right)\right) \tag{2.2}$$

$$S\left(x,y,t\right) = \max\left\{0, S_0 - \sum_{t'=t_0}^{t} N_{\text{emerge}}\left(x,y,t'\right)\right\} \tag{2.3}$$

The weed density in each cell, $\zeta\left(x,y,t\right)$, grows as seeds emerge from the seed bank.

$$\zeta\left(x,y,t\right) = \sum_{t'=t_{\text{last weeded}}}^{t} N_{\text{emerge}}\left(x,y,t'\right) \tag{2.4}$$

The maximum weed height in each cell, $\delta(x,y,t)$, increases at a fixed, upper-bounded, rate of Γ inches per day. This assumption is valid up until the point at which weeds grow explosively, and mechanical weeding becomes impossible.

$$\delta\left(x,y,t\right) = \left(\frac{t - t_{\text{last weeded}}}{60 \cdot 60 \cdot 24}\right)\Gamma \tag{2.5}$$

2.3 Evolving Gaussian Processes (E-GP)

This work makes use of a prediction scheme, evolving Gaussian processes (E-GPs), introduced in [2], which in turn is built on kernel observers [34]. We utilize this predictive mechanism to perform inference with a single machine learning model over multiple systems. This model is a differentially constrained hierarchic modeling method that layers a linear transition model on the weights of a kernel-based model.

One benefit of E-GP is that, by separating the spatial and temporal dynamics hierarchically, and by using a linear transition model on the weights, the learning problem becomes tractable, while complex spatiotemporal behaviors can still be captured in a relatively low-complexity model. Secondly, the linear transition model provides physical insights into the system, which enables the design of observers and controllers [34,35].

We frame the E-GP problem as predictive inference of a time-varying stochastic process, which is assumed to lie in a reproducing kernel Hilbert space (RKHS). This is a Hilbert space where we can define a kernel, $k\colon \Omega \times \Omega \to \mathbb{R}$, as well as a smooth map, $\phi\colon \Omega \to \mathcal{H}$, such that $k(x,y) = \langle \phi(x), \phi(y) \rangle_{\mathcal{H}}$. We approximate the Hilbert space by $\hat{\mathcal{H}}$, generated by an approximate feature map, $\hat{\phi}$, which is detailed in [2].

The entire model is given by Equations 2.6 and 2.7, where w_T are the weights of the GP at time T, A is the linear transition operator in the weight space, and $\hat{\phi}$ is the feature map (kernel matrix) which transforms w_T to the input space.

$$w_T = Aw_{T-1} \tag{2.6}$$

$$y = (w_T)^T \hat{\phi} = (Aw_{T-1})^T \hat{\phi} \tag{2.7}$$

The primary novelty of the evolving Gaussian processes method is learning an $\hat{A}$ matrix for *multiple* systems. An outline of the E-GP method follows.

1. After picking the kernel and estimating the bandwidth hyperparameter σ (we use a maximum likelihood approach, although there are other approaches), we then find an optimal basis vector set C using the algorithm in [16].

2. We then use Gaussian process inference to find weight vectors for each time-step in the training set(s), generating the sequence $w_\tau, \tau = 1, \ldots, T$ for each system.

3. We use matrix least-squares regression with the pseudoinverse operation to solve the matrix equation $\hat{A}[w_1, w_2, ..., w_{T-1}] = [w_2, w_3, ..., w_T]$ for $\hat{A}$.

4. Finally, to generate a multi-system model, we concatenate the weight trajectories from each similar system in the least-squares computation of $\hat{A}$. That is, we let $W_\theta = [w_1^{(\theta)}, w_2^{(\theta)}, ..., w_{n-1}^{(\theta)}]$, and $W'_\theta = [w_2^{(\theta)}, w_3^{(\theta)}, ..., w_n^{(\theta)}]$, be the weight trajectory, and next weight trajectory for some parameter. We then solve the least-squares problem, again using lease squares regression with the pseudoinverse. $\hat{A}[W_{\theta_1}, \ldots, W_{\theta_n}] = [W'_\theta, \ldots, W'_{\theta_n}]$

Some modifications were necessary to implement E-GP with the weed growth data. The weed density in a fixed area has an upper bound on the seed bank density for that area, which it approaches asymptotically (in expectation) according to an exponentially decaying rate of growth. This means that the weed growth is better modeled by an $w_{\tau+1} = \hat{A}w_\tau + \hat{b}$, where the addition of a constant bias term means the system can approach asymptotes other than zero. In order to solve for this bias term, one can solve the following least squares problem.

$$\begin{bmatrix} \hat{A} & \hat{b} \end{bmatrix} \begin{bmatrix} w_1 & w_2 & \ldots & w_{T-1} \\ 1 & 1 & \ldots & 1 \end{bmatrix} = \begin{bmatrix} w_2 & w_3 & \ldots & w_T \end{bmatrix} \tag{2.8}$$

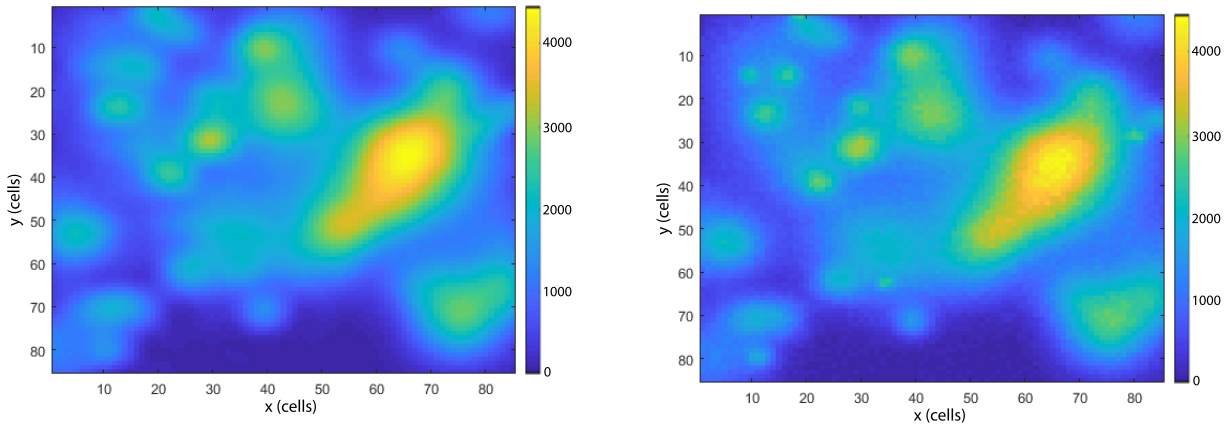
As stated previously, we conceived our model with the insight that the bias term would be chosen to ensure the fixed point of the weight trajectory was correlated to the seed bank of the field. The fixed point can be written as follows.

$$w = \hat{A}w + \hat{b} \Rightarrow w = \left(I - \hat{A}\right)^{-1} \hat{b} \tag{2.9}$$

In the original space this becomes the following.

$$y = (I - \hat{A})^{-T} \hat{b} \hat{\phi}(x) \tag{2.10}$$

We present a comparison of the above quantity for a sample field, using the individualized $\hat{A}w + \hat{b}$ model, with the terminal state of the weight trajectory for that field. We use the individualized model here because, while the generalized model gives high prediction performance, its $\hat{A}$ matrix only captures information about the average seed bank density. As shown in Figure 2.3a and Figure 2.3b, we observe that these are strongly correlated, as we would expect. This shows that the decoupled model has the benefit of giving us physical insight into the relative density of weeds within the field, which can be used to gain insight into the seed bank. Though this fact may not improve prediction performance, and though the seed bank and seedling densities are not perfectly correlated in a real scenario [36], it is relevant to the interests of weed scientists, who could use this analysis to gain insight into the seed bank for fields.



(a) $\left(I - \hat{A}\right)^{-1}\hat{b}$: Fixed point for a sample field.

(b) Terminal weed density for a sample field.

Figure 2.3: Comparison of Fixed Point and Terminal Weed Density

It may be the case that dense measurements throughout the field are required to accurately compute $\hat{A}$ and $\hat{b}$ for the E-GP. However, during the weeding process, we may only have access to sparse data from agent observations of their immediate surroundings. We want to leverage previous data on a given field location, possibly from planting or harvest time, when the canopy is not closed, to train $\hat{A}$ and $\hat{b}$, before computing an online state estimate using only the sparse data available during weeding. To do this, we leverage a Kalman filter state estimate approach which is described here.

Initially, we are given a set of agent observations for the total emerged weed density, $\bar{\zeta}(x, y, t)$. We concatenate these into a measurement matrix, M. We let the weights of the E-GP be the state of the Kalman filter. In [2], we showed the kernel matrix of the E-GP for a given set of measurements, represented by M, is just the observation matrix of the Kalman filter, H. In order to estimate the weights for the E-GP online, we simply use the E-GP model to predict the change in the weights at each time-step, and then refine the weight estimate by running the Kalman filter.

For the Kalman filter, we initialize the E-GP weights, as well as the covariance matrices, P_0, Q_0, R_0 for the error in the state estimate, and the process and observation noise respectively. Here, N_{cent} is the number of centers for the E-GP, which is 400. These parameters are chosen empirically to provide efficient training.

$$w_0 = 0_{N_{\text{cent}}} \in \mathbb{R}^{N_{\text{cent}}} \tag{2.11}$$

$$P_0 = 0.25^2 \mathbb{I}_{N_{\text{cent}}} \in \mathbb{R}^{N_{\text{cent}} \times N_{\text{cent}}} \tag{2.12}$$

$$Q_0 = 0.25^2 \mathbb{I}_{N_{\text{cent}}} \in \mathbb{R}^{N_{\text{cent}} \times N_{\text{cent}}} \tag{2.13}$$

$$R_0 = 0.25^2 \mathbb{I}_{N_{\text{cent}}} \in \mathbb{R}^{N_{\text{cent}} \times N_{\text{cent}}} \tag{2.14}$$

2.4 Predictive Baseline

The prediction process leveraging E-GP does not rely on prior knowledge of the weed growth model. However, we wish to compare the E-GP prediction to a more simplified model-based approach, which will take advantage of known properties of the weed growth model. This highlights the advantages of blending first-principles models with simple sparsified Bayesian nonparametric models. This is accomplished by proposing a candidate emergence model.

$$\zeta(x, y, t) = S_0(x, y) \left(1 - 0.1^{\frac{t}{2 \text{ mo.}}}\right) \tag{2.15}$$

This model satisfies the following boundary conditions, such that there are no emerged weeds at the initial time, and 90 percent of the weeds in the seed bank, represented by the constant term, $S_0(x, y)$, emerges within two months, [29–33].

$$\zeta(x, y, 0) = 0 \tag{2.16}$$

$$\zeta(x, y, 2 \text{ mo.}) = 0.9 S_0(x, y) \tag{2.17}$$

Now, in order to predict the total emerged weed density, $\zeta(x, y, t)$, we use a single time-invariant Gaussian process. This process takes in the observations, $\bar{\zeta}(x, y, t)$, and predicts $\hat{S}_0(x, y)$, which is then used to compute an estimate $\hat{\zeta}(x, y, t)$.

To accomplish this, we simply aggregate the observed measurements, $\bar{\zeta}(x, y, t)$, from the agents, and use them to predict the constant as $\hat{S}(x, y)$.

$$\hat{S}_0(x, y) = \frac{\bar{\zeta}(x, y, t)}{\left(1 - 0.1^{\frac{t}{2 \text{ mo.}}}\right)} \tag{2.18}$$

We train a Gaussian process for $\hat{S}_0(x, y)$ at the observed locations and use it to predict $\hat{S}_0(x, y)$ over the full space. This is then used to compute an estimate, $\hat{\zeta}(x, y, t)$, in the full space.

$$\hat{\zeta}(x, y, t) = \hat{S}_0(x, y) \left(1 - 0.1^{\frac{t}{2 \text{ mo.}}}\right) \tag{2.19}$$

2.5 State, Action, and Reward Model

Here, $N_{\text{dim}} = 85$ is the number of rows, N_{agents} is the number of agents, $Y_{\text{len.}} = 64$ m is the length of each row, and $R_W(x, y, t)$ is the reward per cell (x, y) at time t.

The environmental state, S, depends on the x and y positions of each agent in I. The action, $a_i(t)$, is defined to be the target row chosen by each agent.

$$S \equiv \{1, ..., N_{\text{dim}}\} \tag{2.20}$$

$$I \equiv \left\{1, ..., N_{\text{agents}}\right\} \tag{2.21}$$

$$x_i(t) \in S, \quad y_i(t) \in S \quad \forall i \in I \tag{2.22}$$

$$a_i(t) \in A \equiv S \tag{2.23}$$

Since we require agents to finish the rows they begin, only the x location is relevant for the state. This expedites computation, enabling more efficient planning within this complex domain.

$$x_i(t) \in S \quad \forall i \in I \tag{2.24}$$

We choose the reward associated with a cell to be the maximum height in that cell, prioritizing the regions with the tallest weeds, which prevents weeds from exceeding the maximum height our system can weed and thus causing major yield loss.

The reward for each row is the sum of the reward for each cell in the row, $R_W(x, y, t)$.

$$S \equiv \{1, ..., N_{\text{dim}}\} \tag{2.25}$$

$$I \equiv \left\{1, ..., N_{\text{agents}}\right\} \tag{2.26}$$

$$R_W(x, y, t) = \delta(x, y, t) \quad \forall x \in S, y \in S, \quad \forall i \in I \tag{2.27}$$

The agents keep track of the estimated density and maximum height for each observed cell, using this to estimate a total scalar reward for each row. This is the only required information for the reward.

$$R_i\left(a_i(t)\right) = \sum_{y=1}^{N_{\text{dim}}} R_W\left(a_i(t), y, t\right) \quad a_i(t) \in A \tag{2.28}$$

2.6 Previous Planning Algorithm

In [1], a time delayed reward was used, where each agent receives its reward after completing a row. The planned operation time for a given row is the sum of the time it takes to move to the proposed row, $T_{\text{to row}}$, the time it takes to move down it, $T_{\text{down row}}$, and the time it takes to weed all the cells in the row, $T_{\text{weed row}}$.

$$T_i\left(x_i\left(t\right), a_i\left(t\right)\right) = T_{\text{to row}} + T_{\text{down row}} + T_{\text{weed row}} \tag{2.29}$$

$$T_{\text{to row}} = \frac{\left(a_i\left(t\right) - x_i\left(t\right)\right)}{v} \tag{2.30}$$

$$T_{\text{down row}} = \frac{Y_{\text{len.}}}{v} \tag{2.31}$$

$$T_{\text{weed row}} = \max\left\{2\min, T_{\text{kill}} \sum_{y(t)=1}^{N_{\text{dim}}} \zeta\left(x, y, t\right)\right\} \tag{2.32}$$

It was observed in [1] that the Gittins index, $G\left(X_i\right)$, is known to be optimal metric for planning on tasks with an uncertain termination time and *known* statistics [17].

Here, x is the state of the bandit arm, τ is the stopping time, and r is the reward.

$$G\left(x\left(t\right)\right) = \sup_{\tau}\left\{\frac{\mathbb{E}\left[\sum_{t=0}^{\tau}\gamma^t r\left(x\left(t\right)\right)\middle| x\left(0\right) = x\right]}{\mathbb{E}\left[\sum_{t=0}^{\tau}\gamma^t\middle| x\left(0\right) = x\right]}\right\} \tag{2.33}$$

For our domain, the termination time, τ, is not a planning parameter, but is equal to $T_i\left(x_i\left(t\right), a_i\left(t\right)\right)$, and the reward is delayed until that time. This resulted in a heuristic index based on the Gittins index, $\bar{G}_i\left(a, x\right)$, as follows:

$$\bar{G}_i\left(a, x\right) = \frac{\gamma^{T_i\left(x_i(t), a_i(t)\right)} R_i\left(a_i\left(t\right)\right)}{\sum_{t=0}^{T_i\left(x_i(t), a_i(t)\right)}\gamma^t} \tag{2.34}$$

2.7 Revised Planning Algorithm

Our goal is to improve the planning algorithm from [1], which placed full confidence in the estimated reward for each row, to more effectively address the uncertainty of our model of this dynamic field environment. We want to find an optimization index which yields improved explore-exploit performance.

The difference between this work and [1] is that for unobserved rows, we now have a prediction for the reward, which varies across the field, and may grow more accurate as we explore the space further. We need to account for the value of information gained for a candidate row via predictive inference.

Entropic value-at-risk (EVaR) [37] is a principled way to optimize with regard to the reward and information gain. The parameter $\alpha \in (0, 1)$ is our confidence in our reward estimate, and X is the reward distribution.

$$\text{EVaR}\left[X; 1-\alpha\right] := \inf_{\eta>0}\left\{\frac{1}{\eta}\ln\left(\mathbb{E}_P\left[e^{\eta X}\right]/\alpha\right)\right\} \tag{2.35}$$

EVaR is an index based on the Chernoff bound.

$$\mathbb{P}\left(X \geq \text{EVaR}\left[X; 1-\alpha\right]\right) \leq \alpha \tag{2.36}$$

As in [38], we set our confidence α as follows:

$$\alpha = e^{-D_{KL}(Q||P)} \tag{2.37}$$

Here, Q and P are the distributions for the emergence time before and after a given measurement respectively, and D_{KL} is the Kullback-Leibler divergence, which uniquely quantifies the information gain of Q relative to P [39]. We obtain a modified version of the equation for EVaR:

$$\text{EVaR}\left[X; 1-\alpha\right] := \inf_{\eta>0} \frac{1}{\eta}\left(\ln\left(\mathbb{E}_P\left[e^{\eta X}\right]\right) + D_{KL}\right) \tag{2.38}$$

Equation 2.38 highlights that EVaR is a probabilistically meaningful optimization equation that includes an "exploration bonus" based on the information gain, represented by the D_{KL} term. In this way we have a probabilistically meaningful and inquisitive planning algorithm. In addition, EVaR is linear just like the expected value, meaning that linear transformations of the problem space will not result in different solutions. This is not the case for other exploration bonus methods [40–43]. Finally, EVaR exhibits the properties of strong monotonicity [44] and stop-loss ordering [45], which both speak to the increased ability to discern optimality amongst similarly valued random variables, as compared to the expected value, average value-at-risk, and value-at-risk.

While most work in financial mathematics assumes a constant confidence parameter, α, in our case, this parameter changes with each measurement made. However, the partial stochastic ordering [46] guarantees of strong monotonicity and stop-loss ordering hold at each time step, and since the value of α is constant throughout the field at each time step, these ordering guarantees hold at each time step. Finally, as sampling goes to infinity, the information gain converges to zero; therefore, our confidence function α converges to one, and our EVaR based index converges to the Gittins index.

In order to compute the KL divergence, we create a distribution for the emergence time for each cell, T_{emerge}.

We keep track of the time since we have weeded a cell using the current weeding time.

$$T_{\text{since weed}}\left(x, y, t\right) = t - T_{\text{last weed}}\left(x, y\right) \tag{2.39}$$

During an observation, we calculate the time since emergence and the emergence time as below, based on the observed maximum height.

$$T_{\text{since emerge}}\left(x, y, t\right) = \frac{\hat{\delta}\left(x, y, t\right)}{\Gamma} \tag{2.40}$$

$$\begin{aligned} T_{\text{emerge}}(x,y,n) = T_{\text{since weed}}(x,y,t) \\ - T_{\text{since emerge}}(x,y,t) \end{aligned} \tag{2.41}$$

We then keep an average of the emergence time for each visit to a cell.

$$\bar{T}_{\text{emerge}}(x,y) = \frac{1}{N_{\text{visits}}} \sum_{n=1}^{N_{\text{visits}}} T_{\text{emerge}}(x,y,n) \tag{2.42}$$

At each time step in the simulation, we make a histogram of $\bar{T}_{\text{emerge}}(x,y)$ at times t and $t-1$ and construct probability mass functions P, and Q. These distributions are used to compute the KL divergence, which then goes into the information gain term within EVaR, where $|P|$ is the size of P.

$$D_{KL}(P||Q) = \sum_{i=0}^{|P|} P(i) \left(\log\left(P(i)\right) - \log\left(Q(i)\right) \right) \tag{2.43}$$

Supposing that the empirical estimates are unbiased, both P and Q will converge to the same true pmf in the limit of observations. Therefore, $D_{KL}(P||Q)$ would converge to zero, causing EVaR to converge to the Gittins index [47] over time.

Let $\hat{\zeta}_0(x,y,t)$ be the density estimated by the predictor, based on the model for weed growth. This estimate may not be accurate, since squares are weeded periodically, setting their density to zero. To address this issue, the final estimates for the height and density are found based on the recorded number of weeds removed from each cell.

$$\hat{\zeta}(x,y,t) = \hat{\zeta}_0(x,y,t) - \sum_{t'=t_0}^{t_{\text{last weeded}}} N_{\text{weeds}}\left(x,y,t'\right) \tag{2.44}$$

$$\hat{\delta}(x,y,t) = 1 + \left(\frac{t - t_{\text{last weeded}}}{60 \cdot 60 \cdot 24} \right) \Gamma \tag{2.45}$$

This provides an estimate of the reward.

$$\hat{R}_i\left(a_i(t)\right) = \sum_{y=1}^{N_{\text{dim}}} \hat{\delta}(x,y,t) \quad a_i(t) \in A \tag{2.46}$$

Since our new prediction scheme is able to estimate the density in each cell, we instead optimize over the *reward density*. This ensures that our system accrues reward most efficiently both spatially and temporally, in order to better leverage E-GP for planning improvement. Using this information on the density enables our planner to prioritize cells with a denser seed bank, in which weeds will emerge more rapidly after being weeded, rather than simply exploiting cells with taller weeds as done in the old "greedy" algorithm.

$$\bar{R}_i\left(a_i(t)\right) = \sum_{y=1}^{N_{\text{dim}}} \hat{\delta}(x,y,t)\, \hat{\zeta}(x,y,t) \quad a_i(t) \in A \tag{2.47}$$

We then compute a new Gittins index using EVaR.

$$\hat{G}_i\left(a,x\right)=\frac{\gamma^{T_i\left(x_i(t),a_i(t)\right)}\mathrm{EVaR}\left[\bar{R}_i\left(a_i\left(t\right)\right);1-\alpha\right]}{\sum_{t=0}^{T_i\left(x_i(t),a_i(t)\right)}\gamma^t} \tag{2.48}$$

Finally, we note that for restless bandit problems, or those in which the reward of each bandit arm varies at each time step, the Whittle index, $W_i\left(x_i\left(t\right)\right)$, is introduced in [18] to solve this problem. This index provides a balanced trade-off between exploiting a known reward and allowing it to evolve forward in time.

$$W_i\left(x_i\left(t\right)\right)=\arg\min_{\lambda}:\left\{\begin{array}{l} r_i\left(x_i\left(t\right),1\right)-\lambda+P\left(x_i\left(t\right),1\right)\Delta r_i\left(x_i\left(t\right),1\right) \\ =r_i\left(x_i\left(t\right),0\right)+P\left(x_i\left(t\right),0\right)\Delta r_i\left(x_i\left(t\right),0\right)\end{array}\right\} \tag{2.49}$$

Here, λ can be thought of as a tax on allowing the reward of each arm to evolve forward without exploitation. Additionally, we model the change in reward, $\Delta r_i\left(x_i\left(t\right),a_i\left(t\right)\right)$, as follows.

$$\Delta r_i\left(x_i\left(t\right),a_i\left(t\right)\right)=r_i\left(x_i\left(t\right),a_i\left(t\right)\right)-r_i\left(x_i\left(t-1\right),a_i\left(t-1\right)\right) \tag{2.50}$$

We finally have a heuristic version of the Whittle index for our problem

$$\hat{W}_i\left(a,x\right)=\arg\min_{\lambda}:\left\{\begin{array}{l} \frac{\gamma^{T_i\left(x_i(t),a_i(t)\right)}\mathrm{EVaR}\left[\bar{R}_i\left(a_i(t)\right);1-\alpha\right]}{\sum_{t=0}^{T_i\left(x_i(t),a_i(t)\right)}\gamma^t}- \\ \left(\lambda+\frac{\gamma^{T_i\left(x_i(t),a_i(t)\right)}\mathrm{EVaR}\left[\Delta\bar{R}_i\left(a_i(t)\right);1-\alpha\right]}{\sum_{t=0}^{T_i\left(x_i(t),a_i(t)\right)}\gamma^t}\right)=0\end{array}\right\} \tag{2.51}$$

We are then able to choose the action which maximizes the index for every agent.

$$a_i\left(t\right)=\arg\max_{a_i(t)}\left(\hat{G}_i\left(a,x\right)\text{ or }\hat{W}_i\left(a,x\right)\right)\quad a_i\left(t\right)\in A \tag{2.52}$$

2.8 Overall Architecture

We assume an architecture in which agents are assigned rows from a centralized planner which has access to a global environmental model. This model aggregates the observations from all the agents at every step in the simulation. The overall process is as follows.

1. When agent i completes a row, it becomes idle and waits for a new row to be assigned according to Equation 2.52.
2. It proceeds to the new row and weeds it.
3. At every time step t during weeding, agent i sends its observations on visited rows to the centralized planner.
4. This planner uses each new observation to update the E-GP predictive model, using the new predictive information to plan the next agent assignment.

Before any row assignment, the E-GP prediction is made, and then predictive inference is used to compute the planning index, which then goes into the bandit planner described in Equation 2.52, as summarized in Figure 2.4. We assume a discretized and fixed time step, in which observations are aggregated, processed, and sent to the centralized planner. Communication delays may change this time step slightly in a real-world scenario. However, as long as the time step is much smaller than the time taken to weed a row, the time step will not affect decision making.

In practice, communication delays can be minimized by assigning a team of agents to each region of a larger field, and using our method to coordinate their actions. This work presents a study on one-acre fields, in which the assumption of reliable communication is reasonable.

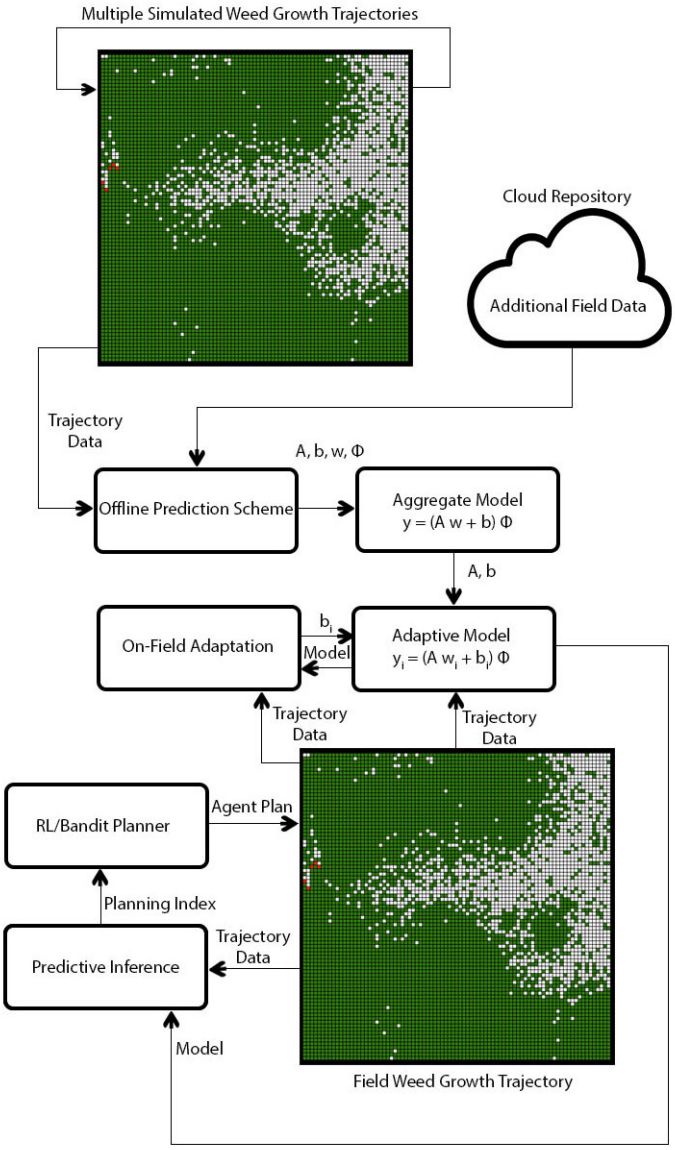


Figure 2.4: **System Block Diagram:** The red squares within the portion of the figure showing the field represent the robotic agents.

Chapter 3

Original Planner

3.1 Experiment Plan - Original Planner

Nine experiments are conducted, each with 100 trials with varying initial parameters shown in Table 3.1, in a simulated field of 0.4 hectares, gridded in 0.8 m^2 cells. Each trial is run for 4 days of simulated time. The algorithm is run in the case of full environmental information, $r_{\text{obs}} = \infty$, where the planner has complete knowledge of the reward for each cell within the environment, in order to establish an ideal benchmark. The algorithm is then run on the worst case scenario in terms of the information the planner has, where the observation radius, $r_{\text{obs}} = 0$. In this case, the planner effectively runs without the benefit of coordinated information gathering. Finally, the algorithm is run with an observation radius, $r_{\text{obs}} = 1$, to see how performance improves in the case of partial environmental information when information about neighboring rows is used. For these experiments, Monte Carlo runs are conducted to determine the feasibility of the method with respect to the change in the number of agents, N_{agent}, the days allowance, d_0, and the initial seed bank density for each field, S_0.

Here, r_{obs} is the observation radius, N_{agent} is the number of agents, d_0 is the days of allowed weed growth before weeding, and S_0 is the initial seed bank density. An X denotes a parameter for a Monte Carlo run over the ranges shown in the last column. Here, for S_0, 1080 is chosen as the median between 600 and 1560, the real-world range determined in [28], for the weed species of interest.

For Experiments 1 - 3, we used the values 15 for N_{agent}, and 1 for d_0. These values are chosen to ensure there are enough agents to weed the field, and that the robots have enough time to do so before the weeds grow too tall. The mean and standard deviation of the average reward for the environment (total reward over all the cells) are plotted at each time step. This provides a clear comparison between the different cases ($r_{\text{obs}} = \infty, 1, 0$), and showcases the effect of environmental information on weeding performance.

In Experiments 4 - 6, Monte Carlo runs are performed over a range of values for the parameters N_{agent}, and S_0, for nominal values of the parameter $d_0 = 1$. A heat map of the terminal reward (total reward over all the cells at the end of the simulation) is plotted for each set of parameters. This reveals how the number of agents affects performance for fields with various seed bank densities.

In Experiments 7 - 9, Monte Carlo runs are performed over a range of values for the parameters N_{agent}, and d_0, for nominal values of the parameter $S_0 = 1080$, and the results are displayed in the same manner. This shows the effect of the number of agents on the performance for fields with various days allowance. In all experiments, the robots are assumed to plan utilizing a global environmental model, which contains the current location of all the agents, as well as all the information available to the agents at a given time.

Table 3.1: Initial Experimental Parameters: Here, r_{obs} is the observation radius, N_{agent} is the number of agents, d_0 is the days of allowed weed growth before weeding, and S_0 is the initial seed bank density. An X denotes a parameter for a Monte Carlo run over the ranges shown in the last column. Here, 1080 is chosen as the median between 600 and 1560, the real-world range determined in [28]. The values of 15 for N_{agent}, and 1 for d_0, are chosen to ensure there are enough agents to weed the field, and that the robots have enough time to do so before the weeds grow too tall.

Exp.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Range**
r_{obs}	∞	1	0	∞	1	0	∞	1	0	[0,∞]
N_{agent}	15	15	15	X	X	X	X	X	X	[5,20]
d_0	1	1	1	1	1	1	X	X	X	[1,6]
S_0	1080	1080	1080	X	X	X	1080	1080	1080	[600,1560]

3.2 Comparison of Partial Environmental Information and Zero Information Cases With the Case of Full Environmental Information

As seen in Figure 3.1, partial environmental information, $r_{obs} = 1$, improves performance over the case of zero information, $r_{obs} = 0$. *The case of* $r_{obs} = 1$ *is able to weed the entire field, converging to a nominal value of* -300 *for the terminal reward.* In the case of zero information, $r_{obs} = 0$, when weeds become sparse, new weeds grow faster than the planner with $r_{obs} = 0$ can find and kill them. This is evidenced by the significant drop in performance from the case of $r_{obs} = 1$. For these experiments, nominal values for the seed bank density (1080 seeds per cell), days allowance (1 day), number of agents (15), and agent speed (1 m per second) were used.

3.3 Performance for Varying Number of Agents and Seed Bank Density: $d_0 = 1$ day, with $r_{obs} = \infty, 1, 0$

In Experiments 4-6, contour maps of the terminal reward for 100 trials with number of agents and seed bank density are shown for the cases of $r_{obs} = \infty, 1, 0$. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. The blue circles represent the data points and the black circles represent the maximum seed bank density the system can weed for each number of agents. The connecting line was added to help distinguish the black and blue circles.

In each experiment, as the seed bank density, S_0, increases, more agents are needed to completely weed the field. As seen in Figure 3.2, for the case of $r_{obs} = \infty$, observe that with more than 8 agents, seed bank densities of up to 1560 seeds per cell can be handled by the system. As evidenced by Figures 3.3 and 3.4, 10 and 12 agents are required to weed fields with the same maximum seed bank density, for the cases of $r_{obs} = 1$ and $r_{obs} = 0$ respectively. For the case of $r_{obs} = 0$, Figure 3.4 shows that the performance for a given number of agents is affected more strongly by the seed bank density. However, there is still a strong positive correlation between the initial seed bank density and the required number of agents for this density.

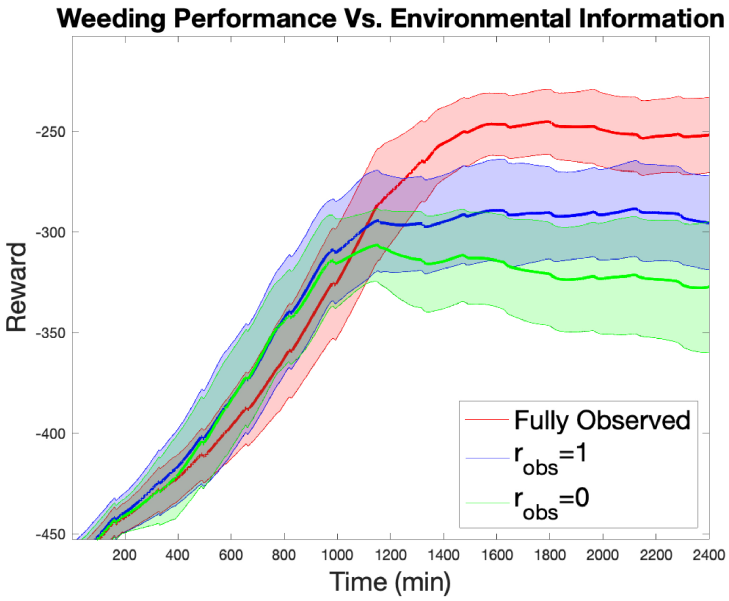


Figure 3.1: **Weeding Performance vs. Environmental Information:** The weeding performance over time is plotted for the case of full environmental information, partial environmental information, $r_{obs} = 1$, and zero environmental information, $r_{obs} = 0$. *The case of* $r_{obs} = 1$ *is able to weed the entire field, converging to a nominal value of* -300 *for the terminal reward. The case of* $r_{obs} = 0$ *shows a significant drop in performance.*

These results show that given enough agents, though information sharing improves performance, the system will succeed in weeding the field for the entire range of seed bank densities and amounts of environmental information tested.

3.4 Performance for Varying Number of Agents and Days Allowance: $S_0 = 1080$ seeds/cell with $r_{obs} = \infty, 1, 0$

In Experiments 7-9, contour maps of the terminal reward for 100 trials with number of agents and days allowance are shown show the cases of $r_{obs} = \infty, 1, 0$. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. The blue circles represent the data points and the black circles represent the maximum days allowance the system can weed for each number of agents. The connecting line was added to help distinguish the black and blue circles.

As days allowance, d_0, increases, more agents are needed to completely weed the field. As shown in Figures 3.5, 3.6, and 3.7, for each case of $r_{obs} = \infty, 1, 0$ with more than thirteen agents, the system can handle a days allowance up to 2, when weeds start to grow higher than the maximum height which the system is capable of weeding, and the system starts to fail for any number of agents. *These results show that the system cannot handle more than 2 days of missed work under any amount of environmental information, given the fact that it can only kill weeds which are smaller than* 7.62 *cm.*

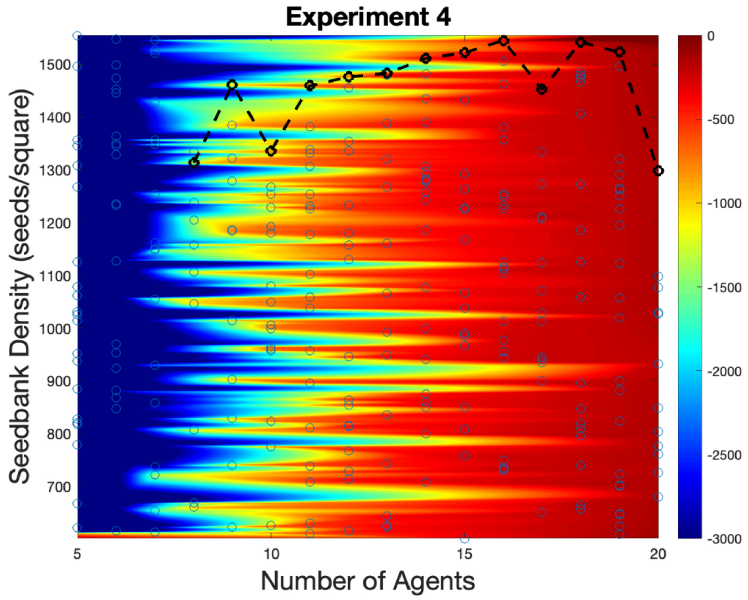


Figure 3.2: **Number of Agents vs. Seed Bank Density,** $r_{obs} = \infty$**:** The contour map of the terminal reward for 100 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. The blue circles represent the data points and the black circles represent the maximum seed bank density the system can weed for each number of agents. The connecting line was added to help distinguish the black and blue circles. *With more than 8 agents, the case of* $r_{obs} = \infty$ *is able to weed the field for the entire range of seed bank densities, performing equally well for any seed bank density.*

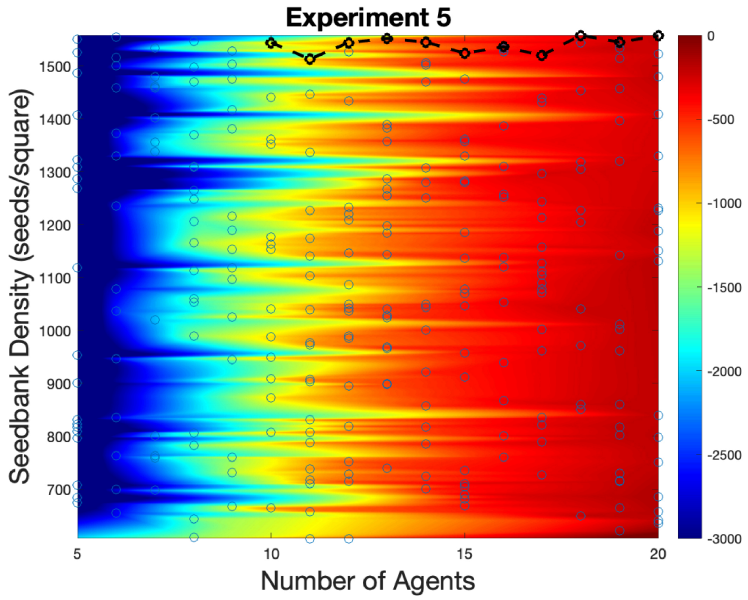


Figure 3.3: **Number of Agents vs. Seed Bank Density,** $r_{obs} = 1$**:** The contour map of the terminal reward for 100 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. The blue circles represent the data points and the black circles represent the maximum seed bank density the system can weed for each number of agents. The connecting line was added to help distinguish the black and blue circles. *With more than 10 agents, the case of* $r_{obs} = 1$ *is able to weed the field for the entire range of seed bank densities, performing equally well for any seed bank density.*

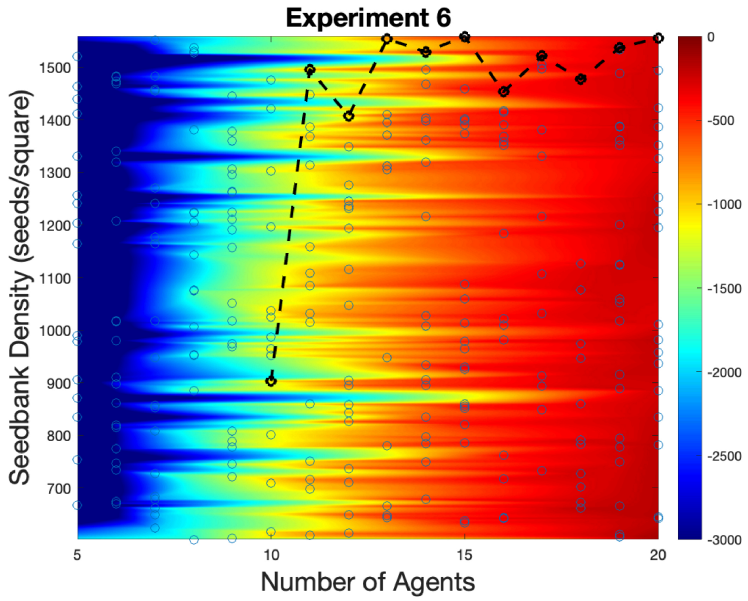


Figure 3.4: **Number of Agents vs. Seed Bank Density,** $r_{obs} = 0$**:** The contour map of the terminal reward for 100 trials with varying numbers of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. The blue circles represent the data points and the black circles represent the maximum seed bank density the system can weed for each number of agents. The connecting line was added to help distinguish the black and blue circles. *With more than 12 agents, the case of* $r_{obs} = 0$ *is able to weed the field for the entire range of seed bank densities. Though the system performance is more strongly affected by the seed bank density than in Experiment 4, the system still performs well with enough agents.*

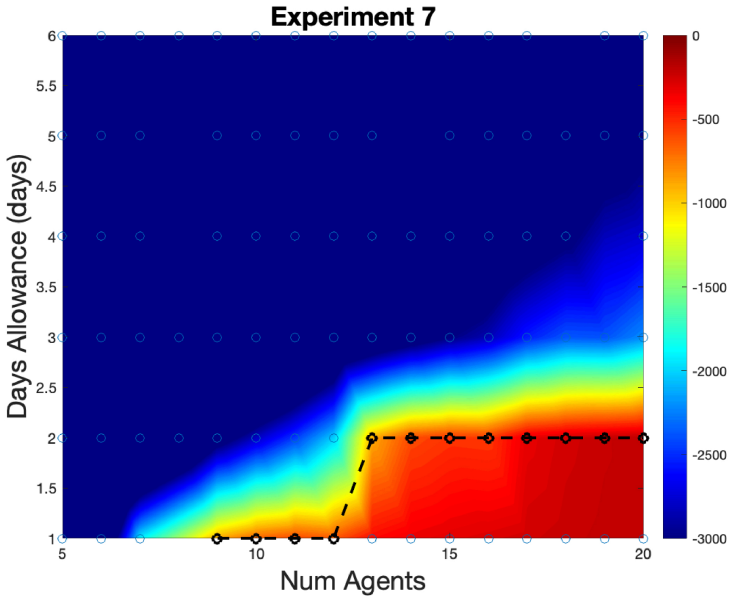


Figure 3.5: **Number of Agents vs. Days Allowance,** $r_{obs} = \infty$**:** The contour map of the terminal reward for 100 trials with varying numbers of agents and days allowance is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. The blue circles represent the data points and the black circles represent the maximum days allowance the system can weed for each number of agents. The connecting line was added to help distinguish the black and blue circles.. *With more than thirteen agents, the case of* $r_{obs} = \infty$ *is able to weed any field, if the days allowance is less than 2, when weeds initially grow larger than the system can eliminate.*

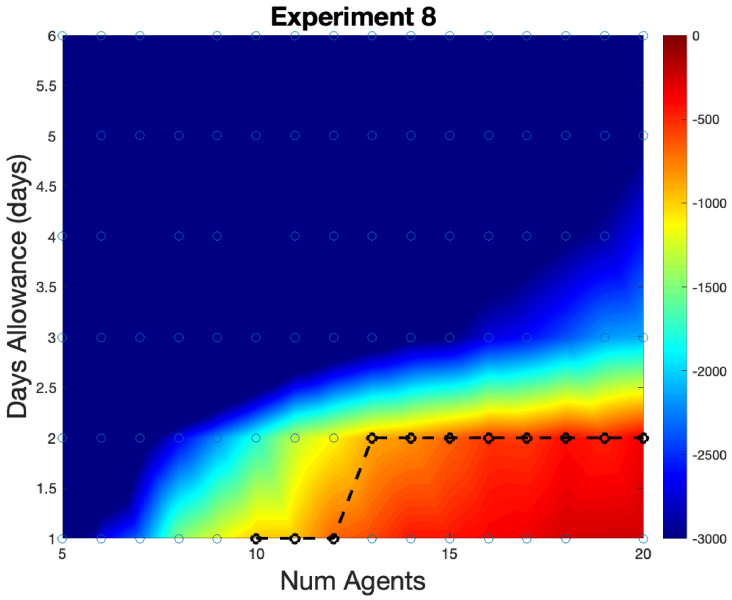


Figure 3.6: **Numbers of Agents vs. Days Allowance,** $r_{obs} = 1$**:** The contour map of the terminal reward for 100 trials with varying number of agents and days allowance is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. The blue circles represent the data points and the black circles represent the maximum days allowance the system can weed for each number of agents. The connecting line was added to help distinguish the black and blue circles. *With more than thirteen agents, the case of* $r_{obs} = 1$ *is able to weed any field, if the days allowance is less than 2, when weeds initially grow larger than the system can eliminate.*

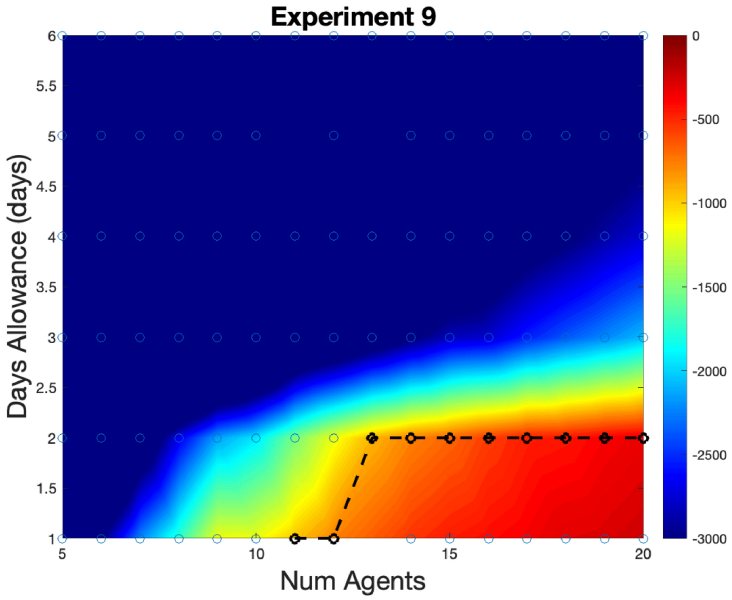


Figure 3.7: **Numbers of Agents vs. Days Allowance,** $r_{obs} = 0$**:** The contour map of the terminal reward for 100 trials with varying number of agents and days allowance is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. The blue circles represent the data points and the black circles represent the maximum days allowance the system can weed for each number of agents. The connecting line was added to help distinguish the black and blue circles. *With more than thirteen agents, the case of* $r_{obs} = 0$ *is able to weed any field, if the days allowance is less than 2, when weeds initially grow larger than the system can eliminate.*

3.5 Summary of Results

In Experiments 1 - 3, it is found that the use of information sharing between the agents significantly improves performance. In Experiment 4, it is found that in the fully observed case, $r_{obs} = \infty$, the system can weed fields with any of the seed bank densities tested without a drop in performance as long as more than 8 agents are used. This fully observed case corresponds to the real-world scenario when there are enough external sensors in the field, confirming that coordination provides a feasible solution to the weeding problem when agents have enough information available. In Experiment 5, it is found that for the case of $r_{obs} = 1$, with more than 10 agents, the system exhibits comparable performance to the fully observed scenario, while it takes 12 agents to achieve this for the case of $r_{obs} = 0$. This suggests that even without full environmental observation, information sharing strategies can be utilized to allow the system to adapt to a rage of field environments, as long as additional agents are used. However, for $r_{obs} = 0$, the performance is more strongly affected by the seed bank density, verifying the intuition that information sharing between agents significantly improves system performance.

In Experiments 7 - 9, it is seen that with more than 13 agents, fields with a days allowance of up to 2 days may be weeded in every case. Beyond this point, weeds grow too large for the system to eliminate before the weeding process starts. However, as more agents are added, performance continues to improve even when the system fails. This suggests that with enough agents, all weeds that are not too tall to be eliminated at the start of the simulation may be weeded by the system.

As summarized in Tables 3.2 and 3.3, a description of the hardware and system characteristics found to be critical to the design of the coordinated weeding system is presented. First, for the weeding apparatus, the parameters of two minutes, and 7.62 cm, were chosen for the maximum time to weed each cell, and the maximum allowed height for weeds killed, respectively. These conservative estimates were made based on the prototype for the weeding apparatus in development. Decreasing the maximum weeding time, and increasing the maximum weed height, would improve system performance. Also, in previous work [1], the speed of the robot was found to not affect system performance, as it becomes unimportant once the field is fully infested with weeds. It was determined in this work that the system cannot handle more than two days of weed growth before the weeding process starts. This is due to the hardware constraints above, as well as the model developed for weed growth using agricultural studies. This constraint would change if the maximum allowed height for weeds was increased.

Finally, the contribution of this work was to show that the system using coordination can successfully weed fields with any average seed bank density, as long as weeds are not initially too high for the system to kill. This is true for the range of seed bank densities considered, which is realistic across a range of weed species studied by crop scientists.

Table 3.2: Hardware Characteristics

Property	**Rational**
Maximum Time to Weed One Cell (T_{KILL})	Based on current prototypes for the robotic weeding apparatus, a conservative estimate of two minutes for this parameter was made. Decreasing it would certainly improve performance.
Maximum Allowed Height for Weeds Killed	Based on current prototypes for the robotic weeding apparatus, a conservative estimate of 7.62 cm for this parameter was made. Increasing it would certainly improve performance.
Maximum Robot Speed	As shown in previous work [1], this becomes unimportant once the field is fully infested with weeds.

Table 3.3: System Characteristics

Property	**Rational**
Maximum Number of Days the System Can Go Without Weeding Field	Given hardware constraints, as well as the model developed for weed growth using agricultural studies, the hard limit is two days, when weeds initially become too tall for the system to handle. If the maximum allowed height for weeds killed was increased, this would change.
Maximum Average Seed Bank Density the System Can Weed	The contribution of this work was to show that the system using coordination can successfully weed fields with any average seed bank density, as long as weeds are not initially too high for the system to kill. This is true for the range of seed bank densities considered, which is realistic across a range of weed species studied by crop scientists.

3.6 Elasticity Analysis and Design Criteria

For each of the Monte Carlo runs, we show the graph of the elasticity of the maximum seed bank density and days allowance to the number of agents, determined by the black data points. This measure was chosen because the fact that it is normalized to be at most positive one gives an easy way to compare the relative effect of parameters across different experiments. This was calculated according to Equation 3.1. We use Y_i, Y_{i-1} as the parameters of interest (either the seed bank density or days allowance), and X_i, X_{i-1} are the number of agents, for data points $i, i-1$ respectively.

$$S_{Y_i X_i} = \left| \frac{(Y_i - Y_{i-1})/Y_{i-1}}{(X_i - X_{i-1})/X_{i-1}} \right| \tag{3.1}$$

The elasticity analysis of Experiments 4-9 is presented below. As shown in Figure 3.8, for Experiments 4-6, the maximum seed bank density the system can weeds for the case of $r_{obs} = \infty$, and $r_{obs} = 1$ is much more elastic to the number of agents than the case of $r_{obs} = 0$. This is as expected, since information sharing between agents causes performance to increase more drastically when new agents are added.

Also, as seen in Figure 3.9, for Experiments 7-9, in every case, $r_{obs} = 0, 1, \infty$, after 2 days, weeds grow higher than the system can successfully eliminate, and therefore this parameter is only elastic to the number of agents until this point.

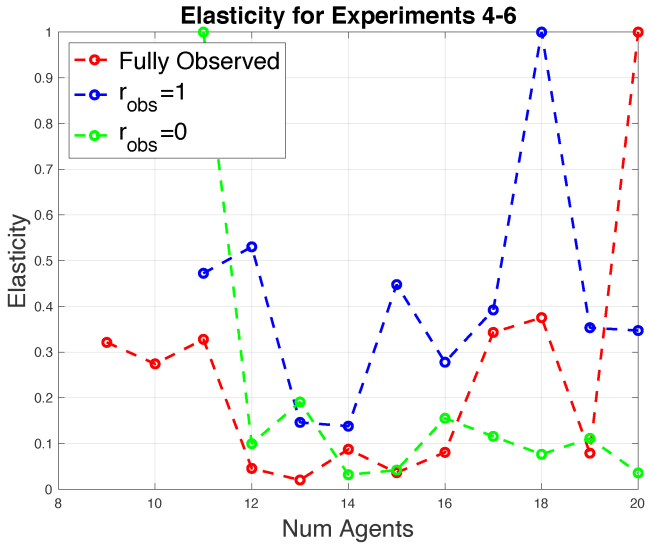


Figure 3.8: **Number of Agents vs. Seed Bank Density:** The plot of the elasticity of the maximum seed bank density the system can weed with respect to the number of agents. *In the case of $r_{obs} = \infty$ and $r_{obs} = 1$, the maximum seed bank density is elastic to the number of agents, since information sharing strongly improves performance. In the case of $r_{obs} = 0$, the maximum seed bank density is far less elastic to the number of agents, since information sharing between agents is not used, and thus does not cause as large a change in system performance.*

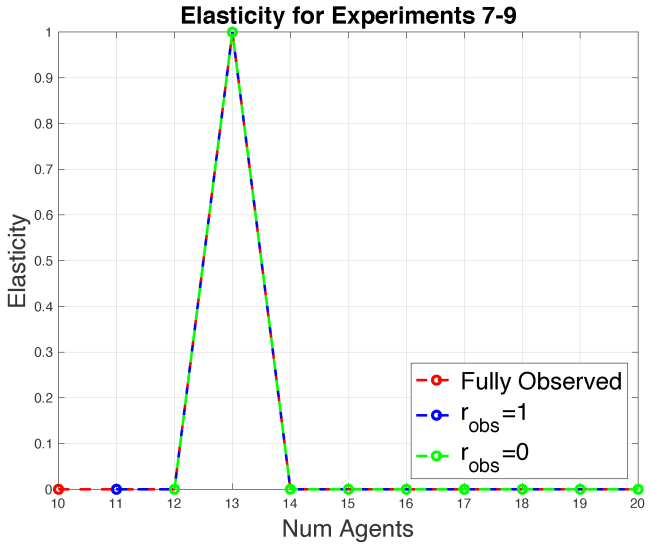


Figure 3.9: **Numbers of Agents vs. Days Allowance:** The plot of the elasticity of the maximum days allowance the system can weed with respect to the number of agents utilized. *In all cases, $r_{obs} = \infty, 1, 0$, the maximum days allowance is inelastic to the number of agents. After enough agents are used, the system can weed fields with days allowance up to 2 days.*

Chapter 4

Improved Planner

4.1 Experiment Plan - Improved Planner and Predictor

This section presents outline for the experiments to test for improvement over the planning algorithm in [48] (Section 2.6), detailed in Section 2.7. In [48], it was determined that wider ranges of observation for each agent improved performance in every case, by comparing the case where observation radius $r_{\text{obs}} = 1$ to the case where $r_{\text{obs}} = 0$, which had worse performance. It was also observed that beyond a critical point of two days for the number of days of weed growth allowed before the simulation starts (denoted d_0), the system could not succeed regardless of the number of agents used. Therefore, for these experiments, we set $d_0 = 1$ and set the observation radius $r_{\text{obs}} = 1$. These experiments compare the performance of the prediction scheme from [48] with the Gittins index used for the planning index, with the new prediction scheme using EVaR in the planning index. We do not use the fully-observed scenario, $r_{\text{obs}} = \infty$ with EVaR, as this would make exploration irrelevant since there is no uncertainty in the environmental model. We instead use the old fully observed scenario, which uses the Gittins index computed based on observations of the entire field. We also disregard the case $r_{\text{obs}} = 0$, which our algorithm outperforms, as this makes information gain impossible/irrelevant, since each cell is weeded and reset to zero as it is observed, and so exploration is of no benefit, as no information is shared between the agents.

We conduct eight experiments, each with 1000 trials with varying initial parameters shown in Table 4.1, in a simulated field of 0.4 hectares, gridded in 0.8 m^2 cells. Each trial is run for 2 days of simulated time, as this was found to be the critical weeding period for the environment. The algorithm is run with the original prediction scheme, and the planner using the Gittins index, where the planner has knowledge of each cell adjacent to one of the agents, in order to establish a benchmark of previous performance using the new Python Weed World simulation. The algorithm is then run with the EVaR planning index, with the old prediction scheme using the average reward in previously weeded rows, in order to observe how the revision to the planner affects performance. Next, the algorithm is run with full observability of the environment provided to the planner, to see how performance improves. Last, we compare to the baseline case of a lawn mower pattern.

In Table 4.1, an X denotes a parameter for a Monte Carlo run over the ranges shown in Table 4.2. For S_0, the real-world range is between 600 and 1560 for the weed species of interest [28]. We chose to double this range for our simulations in order to find an upper bound on the seed bank density that can be weeded for each algorithm with the given number of agents. This provides a more interesting comparison between the algorithms, and allows us to more accurately determine the sensitivity of each algorithm to the weed density.

Table 4.1: Initial Experimental Parameters: Here, N_{agent} is the number of agents, S_0 is the initial seed bank density of each field. An X denotes a parameter for a Monte Carlo run over the ranges in Table 4.2. We use the nominal value of 15 for N_{agent} to ensure there are enough agents to weed the field. Here, 1080 is the median of the real-world range for S_0, which is determined to be between 600 and 1560, for the weed species of interest [28]. The original planning and prediction methods are from [48].

Exp.	1	2	3	4	5	6	7	8
Plan	Orig.	EVaR	Orig.	Mower	Orig.	EVaR	Orig.	Mower
Predict	Orig.	Orig.	Full Obs.	None	Orig.	Orig.	Full Obs.	None
N_{agent}	15	15	15	15	X	X	X	X
S_0	1080	1080	1080	1080	X	X	X	X

Table 4.2: Ranges for Experimental Parameters in Monte Carlo Runs: N_{agent} is the number of agents, d_0 is then number of days of weed growth before the start of weeding, and r_{obs} is the observation radius. The real-world range for the initial seed bank density of the field, S_0, is between 600 and 1560, for the weed species of interest [28], but we chose to double this range in order to test the upper limits of our algorithms.

Parameter	**Range**
r_{obs}	1
d_0	1
N_{agent}	[5,20]
S_0	[600,3120]

For Experiments 1 - 4, we used the values 15 for N_{agent}, and 1080 for S_0. These values are chosen to ensure that there are enough agents to weed the field, and that the robots have enough time to do so before the weeds grow too tall. The mean and standard deviation of the average reward for the environment (total reward over all the cells divided by the number of agents) are plotted at each time step. This provides a clear comparison between the different cases (the original prediction scheme versus fully observed case, the Gittins index versus the EVaR index, and the lawn mower patter), and showcases the effect of each method used on weeding performance. The general trend in the performance shows the stages of weeding for each algorithm tested, and the maximum possible reward each can achieve.

In Experiment 5-8, Monte Carlo runs are performed over a range of values for the parameters N_{agent}, and S_0, for the lawnmower pattern. A heat map of the terminal reward (total reward over all the cells at the end of the simulation divided by the number of agents) is plotted with respect to initial seed bank density S_0 and number of agents N_{agents}. This shows the feasibility of the method with respect to changes in these parameters.

Within every Monte Carlo run, we consider a success to be a case in which an algorithm is able to keep the total maximum height within each cell of the field per agent under the value 1000. This is an adjustable threshold, which is chosen empirically, and represents the point at which we consider the agbot system to have reached an acceptable level of weeding efficiency per robot.

4.2 Comparison of Algorithm with Original Planner With the Algorithms Using Updated Planner, Planner With Full Environmental Information, Lawn Mower Pattern, $N_{\text{agent}} = 15$, $S_0 = 1080$, $d_0 = 1$ day, $r_{\text{obs}} = 1$

In Figure 4.1, the mean and standard deviation of the average reward for the environment (total reward over all the cells divided by the number of agents) are plotted at each time step for each algorithm in Experiments 1-4. Though we see slight statistical differences in numerical performance, the general trend in the reward is the same for all of these algorithms, which all succeed in fields with these nominal parameters. The lawn mower pattern does worse than any of the others, as we expect.

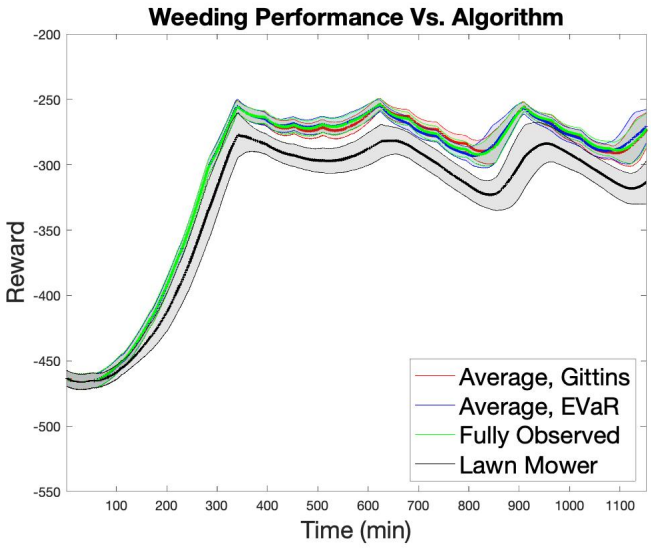


Figure 4.1: **Weeding Performance vs. Algorithm:** The mean and standard deviation of the average reward for the environment (total reward over all the cells divided by the number of agents) are plotted at each time step. The general trend in the reward is the same for all of these algorithms, which all succeed in fields with these nominal parameters. The lawn mower pattern does worse than any of the others, as we expect.

4.3 Performance for Varying Number of Agents and Seed Bank Density: $d_0 = 1$ day, with $r_{\text{obs}} = 1$, Original Planner with The Gittins Index

In Figure 4.2, we examine the case when the Gittins index is used for the planner. The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents.

The maximum density that can be handled with eight agents is significantly less than the maximum of the real-world range of the seed bank. When more agents are used, the system is able to succeed over the entire range of seed bank densities tested. This shows that improvements to the performance of the planning algorithm could allow farmers to use eight or fewer agents instead of nine or more.

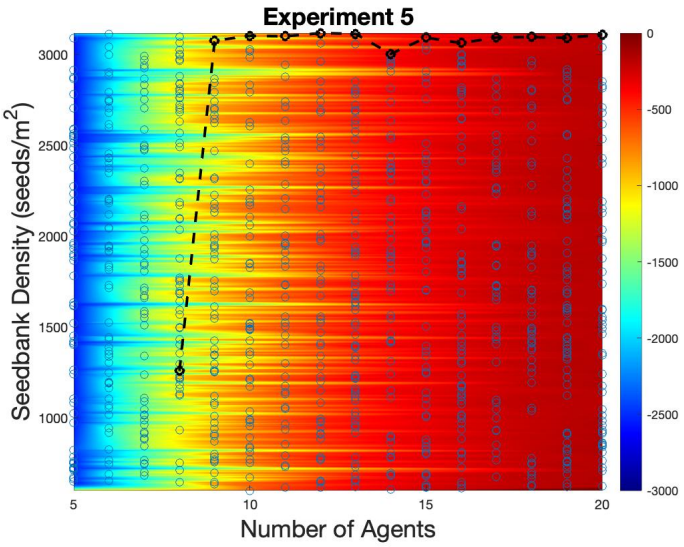


Figure 4.2: **Number of Agents vs. Seed Bank Density, Gittins Index, Partially Observed:** The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents. This figure shows that *the maximum density that can be handled with eight agents is significantly less than the maximum of the real-world range of the seed bank.* When more agents are used, the system is able to succeed over the entire range of seed bank densities tested. This shows that improvements to the performance of the planning algorithm planner could allow farmers to use eight or fewer agents instead of nine or more.

4.4 Performance for Varying Number of Agents and Seed Bank Density: $d_0 = 1$ day, $r_{\mathbf{obs}} = 1$, New Planner

In Figure 4.3, we examine the case when EVaR is used for the planner. The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents.

In this case, the system is able to succeed with only eight agents for the entire real-world range of seed bank densities, though not for the entire worst-case range. When more than eight agents are used, the system is able to succeed over the entire range of seed bank densities tested. *Our experiment shows that in most real-world scenarios, EVaR will yield stronger performance at a lower cost.*

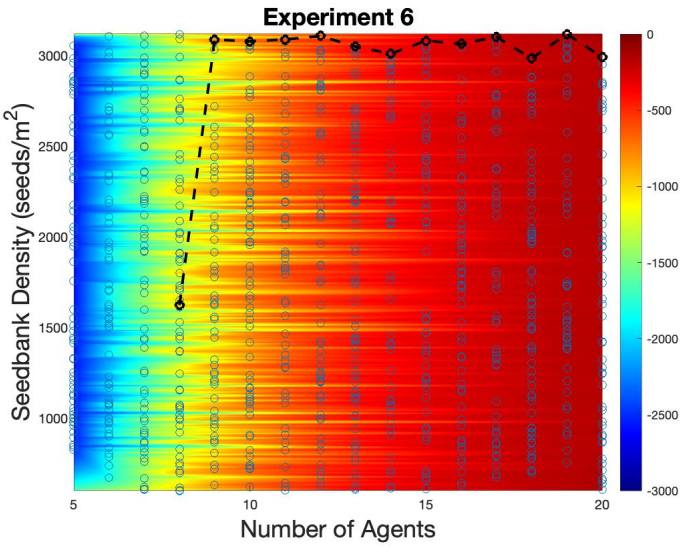


Figure 4.3: **Number of Agents vs. Seed Bank Density, EVaR Index, Partially Observed:** The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents. In this case, the system is able to succeed with only eight agents for the entire real-world range of seed bank densities, though not for the entire tested range. When more than eight agents are used, the system is able to succeed over the entire range of seed bank densities tested. *Our experiment shows that in most real-world scenarios, EVaR will yield stronger performance at a lower cost*

4.5 Performance for Varying Number of Agents and Seed Bank Density: $d_0 = 1$ day, $r_{\mathbf{obs}} = 1$, Planner with Full Environmental Information

In Figure 4.4, we examine the case with full environmental information. The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents.

Full environmental information is unrealistic, but serves as a useful comparison point. We find that the system can handle extremely high seed bank densities with only eight agents, though quite the entire tested range. Seven agents are still insufficient to succeed at any realistic seed bank density even with full observability. This suggests that *eight agents is a critical number for one-acre fields.*

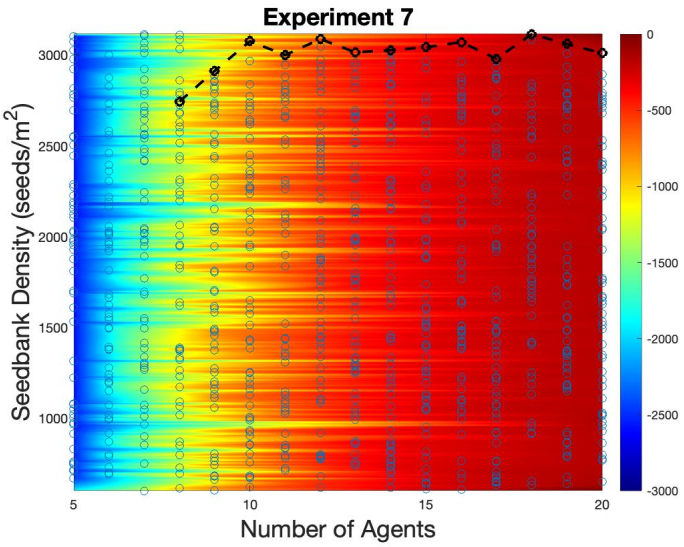


Figure 4.4: **Number of Agents vs. Seed Bank Density, Full Environmental Information:** The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents. We find that the system can handle extremely high seed bank densities with only eight agents, though quite the entire tested range. Seven agents are still insufficient to succeed at any realistic seed bank density even with full observability. This suggests that *eight agents is a critical number for one-acre fields*

4.6 Performance for Varying Number of Agents and Seed Bank Density: Lawn Mower Pattern, $d_0 = 1$ day, $r_{\mathbf{obs}} = 1$

In Figure 4.5, we examine the case of the baseline planner using a lawn mower pattern. The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents.

From this figure, we see that with eight agents, the system fails for every seed bank density tested. Nine agents are needed for success. This demonstrates that the lawn mower pattern does much worse than any algorithm tested in the main body of the text.

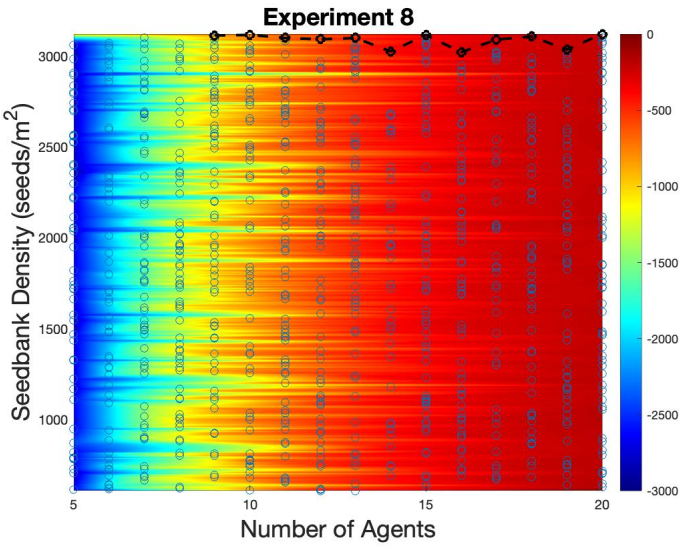


Figure 4.5: **Number of Agents vs. Seed Bank Density, Gittins Index, Partially Observed:** The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents. *With eight agents, the system fails for every seed bank density tested.* Nine agents are needed for success. This demonstrates that the lawn mower pattern does much worse than any algorithm tested in the main body of the text.

4.7 Discussion

These results show that compared to previous algorithms, the EVaR-based index enables teams of robots to weed fields with higher average background seed bank densities with the same number of agents. We see that EVaR even exhibits comparable performance to the fully observed scenario, over the real-world range of seed bank densities from 600 to 1560 for this weed species. Since our reward function is only concerned with the height of weeds, and since weeds emerge quickly after being destroyed and begin growing at a consistent rate, models of the seed bank density and growth patterns are not necessary for high performance in most realistic fields. Only when the field becomes unrealistically dense does the fully observed case start to outperform the case of EVaR with partial information.

The performance increase due to EVaR is caused by the fact that the information gain term directs the system to rapidly explore the environment before exploiting. By building a model of the emergence time throughout the field, the system obtains more accurate estimate of the average reward density. This leads to better decision making, and higher performance.

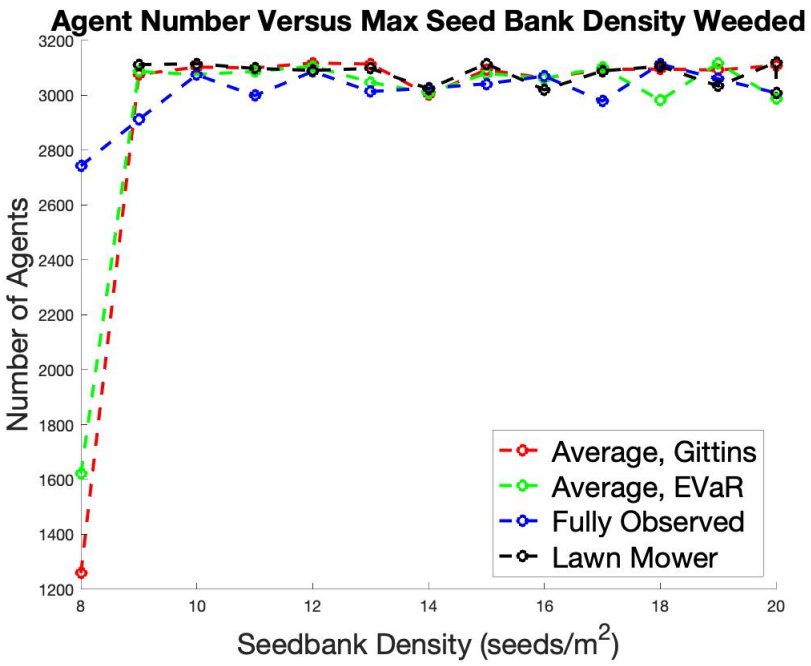


Figure 4.6: **Number of Agents vs. Seed Bank Density, Summary:** We compare the maximum seed bank density weeded for a given number of agents for each algorithm.

It is important to note that the Weed World simulation environment establishes a worst-case scenario for weeding performance. We assume weeds grow aggressively, without growth being curbed by competition from crops or bad weather. Robotic agents use conservative estimates for speed and weeding time, which make effective use of each agent critical. The goal of the system is to ensure weeds never grow taller than the mechanical system can eliminate, and enter into a regime of explosive growth. For weeds like common waterhemp, which grow to have a large and strong stem, this can mean full yield loss, due to inability to harvest the crop with combine harvesters. Though, as shown in Figure 4.6 the use of EVaR only saves one robot per acre, and only does so for fields with high density, it guarantees that any field in this real-world range of seed bank densities can be managed without the risk of explosive weed growth.

Chapter 5

Improved Predictor

5.1 Experiment Plan

This section presents an outline for the experiments to test for how the new E-GP prediction scheme improves planning performance over [48] (Section 2.6), detailed in Section 2.3. In [48], it was determined that wider ranges of observation for each agent improved performance in every case, by comparing the case where observation radius $r_{\text{obs}} = 1$ to the case where $r_{\text{obs}} = 0$, which had worse performance. It was also observed that beyond a critical point of two days for the number of days of weed growth allowed before the simulation starts (denoted d_0), the system could not succeed regardless of the number of agents used. Therefore, for these experiments, we set $d_0 = 1$ and set the observation radius $r_{\text{obs}} = 1$. These experiments compare the performance of the new E-GP prediction scheme 2.3 with the planning algorithm described in 2.7, with the G-SB prediction scheme described in 2.4 and using this same planner, as well as the fully observed scenario. We do not use the fully observed scenario, $r_{\text{obs}} = \infty$ with EVaR, as this would make exploration irrelevant since there is no uncertainty in the environmental model. We also disregard the case $r_{\text{obs}} = 0$, which our algorithm outperforms, as this makes information gain impossible/irrelevant, since each cell is weeded and reset to zero as it is observed, and so exploration is of no benefit.

We conduct seven experiments, each with 1000 trials with varying initial parameters shown in Table 4.1, in a simulated field of 0.4 hectares, gridded in 0.8 m^2 cells. Each trial is run for 2 days of simulated time, as this was found to be the critical weeding period for the environment. The algorithm is first run with full observability of the environment provided to the planner, using the old reward, to establish a baseline for comparison. Next, the algorithm is run with the E-GP predictor and the planner using EVaR, in order to showcase improvement. We then run the algorithm with the G-SB predictor and this same planner, to show how much the E-GP predictor itself improves performance. Finally, the previous two experiments are repeated for the cases in which the planner uses the Whittle index in conjunction with EVaR.

In Table 5.1, an X denotes a parameter for a Monte Carlo run over the ranges shown in Table 5.2. For S_0, the real-world range is between 600 and 1560, for the weed species of interest [28]. We chose to double the this range in order to find an upper bound on the seed bank density that can be weeded for each algorithm with the given number of agents. This provides a more interesting comparison between the algorithms, and allows us to more accurately determine the sensitivity of each algorithm to the weed density.

In Experiments 1-2, we use the values of 15 for N_{agent}, and 1080 for S_0. These values are chosen to ensure that there are enough agents to weed the field, and that the robots have enough time to do so before the weeds grow too tall. The mean and standard deviation of the percent planning error of the E-GP and G-SB prediction schemes when used with a planner leveraging the Whittle index are plotted at each time step.

Table 5.1: Initial Experimental Parameters: Here, N_{agent} is the number of agents, S_0 is the initial seed bank density of each field. An X denotes a parameter for a Monte Carlo run over the ranges in Table 4.2. We use the nominal value of 15 for N_{agent} to ensure there are enough agents to weed the field. Here, 1080 is the median of the real world range for S_0, which is determined to be between 600 and 1560, for the weed species of interest [28]. The original planning and prediction methods are from [48].

Exp.	1	2	3	4	5	6	7
Plan	Whittle	Whittle	Orig.	EVaR	EVaR	Whittle	Whittle
Predict	E-GB	G-SB	Full Obs.	E-GP	G-SB	E-GP	G-SB
N_{agent}	15	15	X	X	X	X	X
S_0	1080	1080	X	X	X	X	X

Table 5.2: Ranges for Experimental Parameters in Monte Carlo Runs: N_{agent} is the number of agents, d_0 is then number of days of weed growth before the start of weeding, and r_{obs} is the observation radius. The real-world range for S_0 is between 600 and 1560, for the weed species of interest [28], but we chose to double this range in order to test the upper limits of our algorithms.

Parameter	**Range**
r_{obs}	1
d_0	1
N_{agent}	[5,20]
S_0	[600,3120]

In Experiments 3-7, Monte Carlo runs are performed over a range of values for the parameters N_{agent}, and S_0. For each Monte Carlo experiment, a heat map of the terminal reward (total reward over all the cells at the end of the simulation) is plotted with respect to initial seed bank density S_0 and number of agents N_{agents}. This shows the feasibility of the method with respect to changes in these parameters.

Within every Monte Carlo run, we consider a success to be a case in which an algorithm is able to keep the total maximum heights from each cell in the field, per agent, under the value 1000. This is an adjustable threshold, which is chosen empirically, and represents the point at which we consider the agbot system to have reached an acceptable level of weeding efficiency per robot.

5.2 Comparison of Algorithms with E-GP Predictor Versus G-SB Predictor, Using Planner with the Whittle Index, $N_{\text{agent}} = 15$, $S_0 = 1080$, $d_0 = 1$ day, $r_{\text{obs}} = 1$

In Figure 5.1, the mean and standard deviation of the percent planning error are plotted at each time step for each algorithm in Experiments 1-2. The solid lines represent the mean and the transparent regions represent the standard deviation around this mean. The time axis is truncated, as the errors for both algorithms converge to a steady state value after they gain information from the agents as exploration of the field is conducted.

We observe that both prediction schemes converge to a low error. As the E-GP predictor is pretrained, the starting error in this case is much lower. Though the G-SB predictor converges to a lower error, as model information is used here, the initial error in this case is much higher and behaves erratically as the prediction scheme is trained.

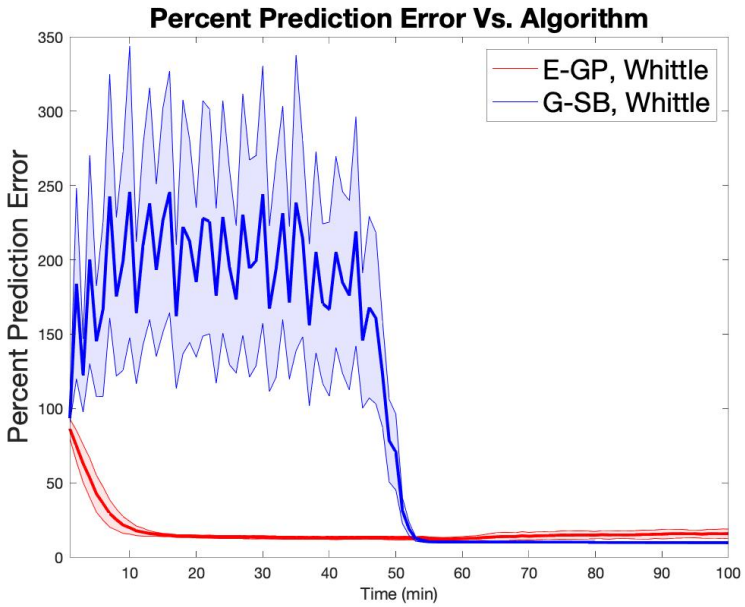


Figure 5.1: **Percent Planning Error vs. Algorithm:** The mean and standard deviation of the percent planning error are plotted at each time step for each algorithm in Experiments 1-2. *We observe that both prediction schemes converge to a low error. As the E-GP predictor is pretrained, the starting error in this case is much lower. Though the G-SB predictor converges to a lower error, as model information is used here, the initial error in this case is much higher and behaves erratically as the prediction scheme is trained.*

5.3 Performance for Varying Number of Agents and Seed Bank Density: $d_0 = 1$ day, with $r_{\mathbf{obs}} = 1$, Greedy with Full Environmental Information

In Figure 5.2, we examine the case with full environmental information, and the greedy reward which does not use the density. The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents.

With eight agents, the system succeeds for most densities, but not the entire tested range.

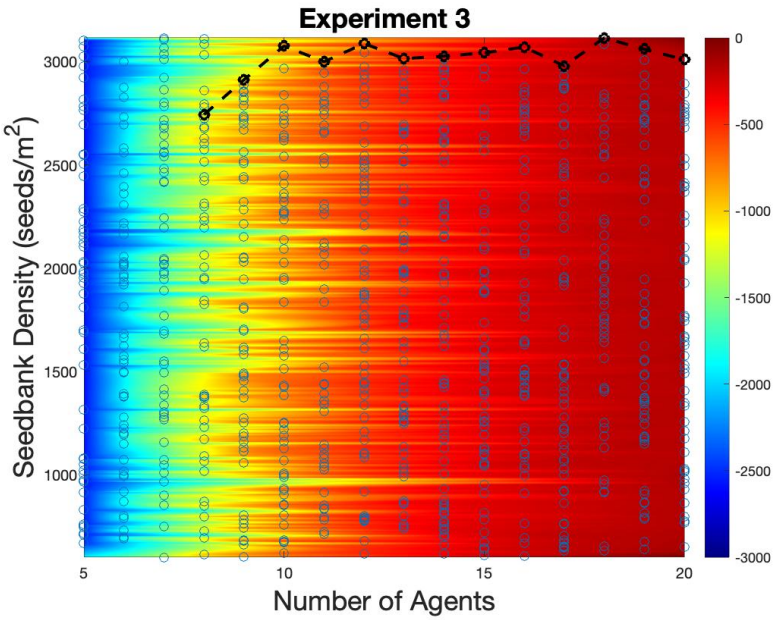


Figure 5.2: **Number of Agents vs. Seed Bank Density, Full Environmental Information:** The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents. *With eight agents, the system succeeds for most densities, but not the entire tested range.*

5.4 Performance for Varying Number of Agents and Seed Bank Density: $d_0 = 1$ day, with $r_{\text{obs}} = 1$, Updated Planner and E-GP Predictor

In Figure 5.3, we examine the case when EVaR is used for the planner, with the E-GP as the new predictor. The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents.

In this case, the system succeeds in every case when eight agents are used. This shows that E-GP offers an improvement in system performance, even over the fully observed case.

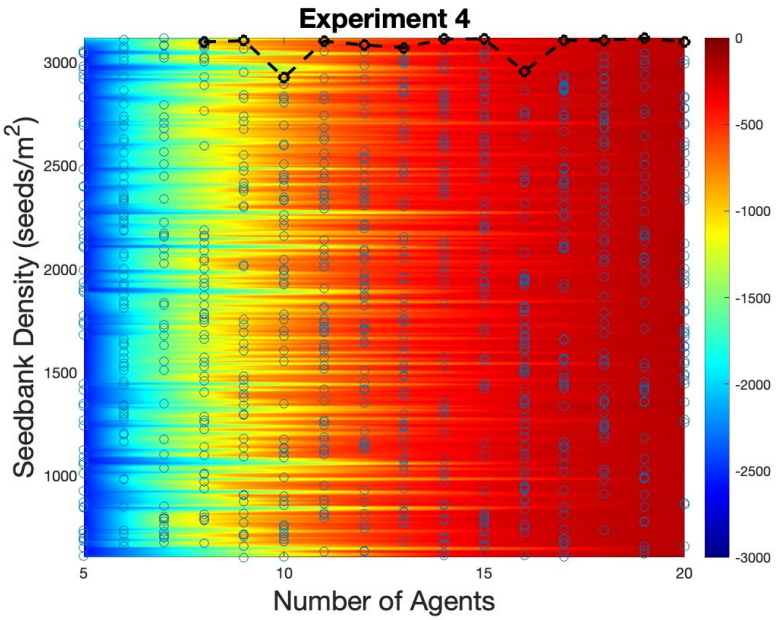


Figure 5.3: **Number of Agents vs. Seed Bank Density, EVaR Index with E-GP Predictor:** The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents. *With eight agents or more, the system succeeds in every case, showing that E-GP offers an improvement in system performance, even over the fully observed case.*

5.5 Performance for Varying Number of Agents and Seed Bank Density: $d_0 = 1$ day, with $r_{\text{obs}} = 1$, Updated Planner and G-SB Predictor

In Figure 5.4, we examine the case when EVaR is used for the planner, and G-SB for the predictor. The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents.

With eight agents, the system succeeds for densities significantly higher than the real-world range, but not as high as the case of E-GP. This shows that the E-GP predictor offers a significant improvement, even when the same planner is used.

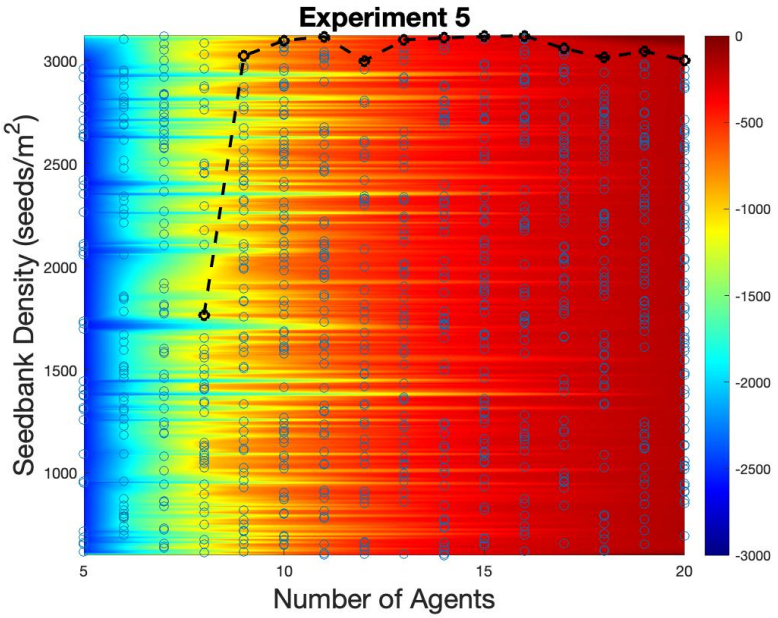


Figure 5.4: **Number of Agents vs. Seed Bank Density, EVaR Index with G-SB Predictor:** The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents. *With eight agents, the system succeeds for densities significantly higher than the real-world range, but not as high as the case of E-GP.* This shows that the E-GP predictor offers a significant improvement, even when the same planner is used.

5.6 Performance for Varying Number of Agents and Seed Bank Density: $d_0 = 1$ day, with $r_{\mathbf{obs}} = 1$, Whittle Index with E-GP

In Figure 5.5, we examine the case with full environmental information, and the greedy reward which does not use the density. The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents.

With eight agents, the system still succeeds for the entire tested range, suggesting that the use of E-GP saturates system performance, even without the use of the Whittle index.

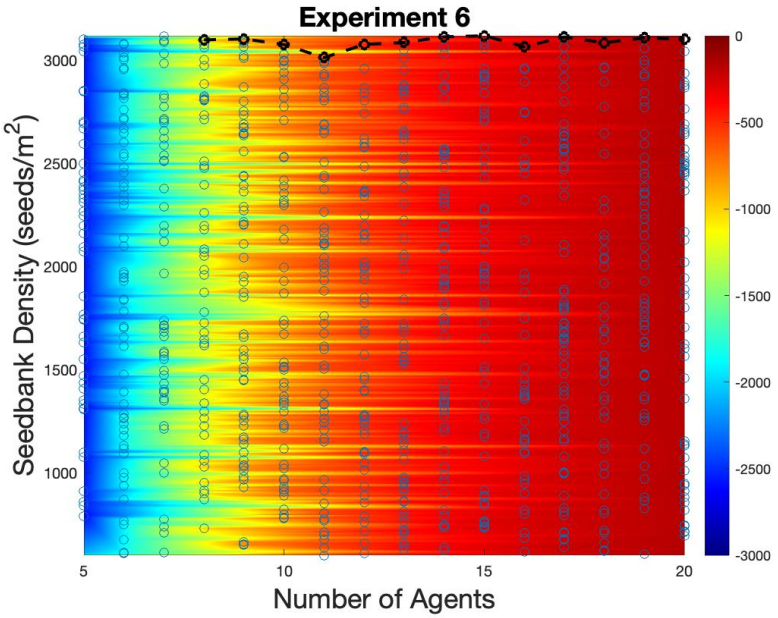


Figure 5.5: **Number of Agents vs. Seed Bank Density, Whittle Index with E-GP:** The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents. *With eight agents, the system still succeeds for the entire tested range, suggesting that the use of E-GP saturates system performance, even without the use of the Whittle index.*

5.7 Performance for Varying Number of Agents and Seed Bank Density: $d_0 = 1$ day, with $r_{\mathbf{obs}} = 1$, Whittle Index with G-SB

In Figure 5.6, we examine the case with full environmental information, and the greedy reward which does not use the density. The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents.

With eight agents, the system succeeds for a significantly higher range of densities than the previous case utilizing EVaR without the Whittle index.

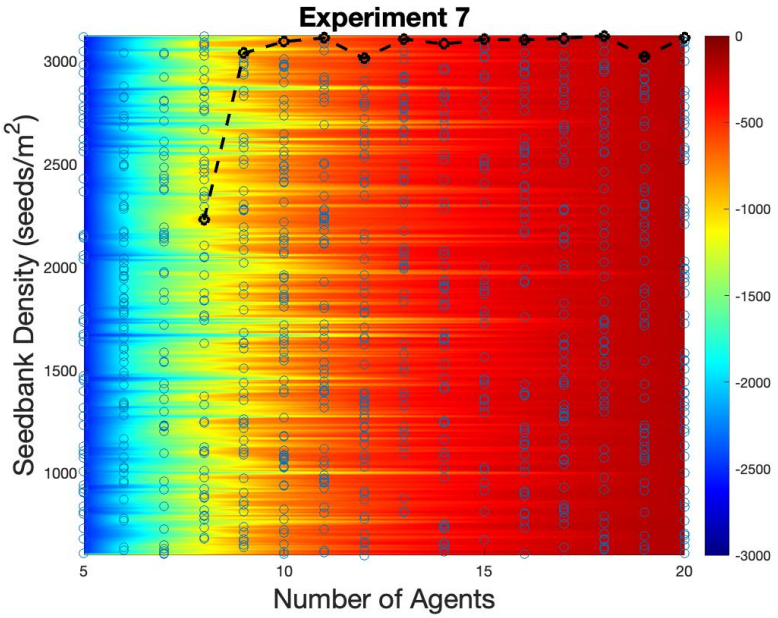
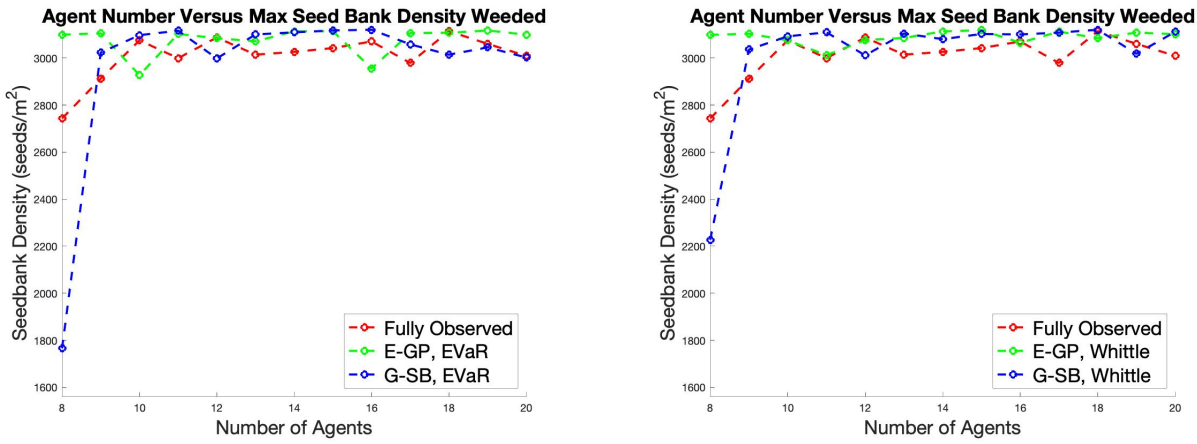


Figure 5.6: **Number of Agents vs. Seed Bank Density, Whittle Index with G-SB:** The contour map of the terminal reward for 1000 trials with number of agents and seed bank density is shown. The red end of the spectrum represents a terminal reward of zero, meaning the field has been weeded completely, and the blue end represents a high nonzero terminal reward, a strong failure case. Each blue circle represents a simulated trial, and the black dashed line represents the maximum seed bank density the system can weed for each number of agents. *With eight agents, the system succeeds for a significantly higher range of densities than the previous case utilizing EVaR without the Whittle index.*

5.8 Discussion

These results show that the use of E-GP significantly improves planning performance over the G-SB scenario. In fact, E-GP even does slightly better than the fully observed case. This may be due to the fact that cells with high density have faster emergence time, and thus grow in height more quickly after being weeded. By leveraging predictive information, E-GP can proactively prevent growth by prioritizing dense cells before the weeds there grow tall, rather than simply exploiting cells that currently have taller weeds, as in the fully observe case with the old "greedy" planning index.

Both the E-GP and G-SB predictor have unique advantages and disadvantages in a given environment. The E-GP predictor relies on past data on a given field, taken when the canopy is not occluded, either at planting or harvest time. If dense measurements over this time period are not possible, using this method may not be appropriate. Similarly, the G-SB predictor relies on the underlying assumption that the average emergence time for a given seed bank density is known for a particular field. This information may also not be available if the particular weed species within the field are unknown. In many situations, it may still be necessary to use EVaR with a course prediction scheme relying on the average observed reward for previously weeded rows, until sufficient information has been collected. By using the weeding system over multiple years on the same field, weeding performance can be improved over time, as a model for the weed population within that field is learned by the system.



(a) EVaR Index

(b) Whittle Index

Figure 5.7: Comparison of E-GP and G-SB Predictors to Greedy Fully Observed Case

Table 5.3: Maximum Seed Bank Density Weeded With Eight Agents

Full Obs.	**E-GP, EVaR**	**G-SB, EVaR**	**E-GP, Whittle**	**G-SB, Whittle**
2743	3098	1766	3099	2227

As summarized in Figure 5.7a and Figure 5.7b, and Table 5.3, the E-GP prediction scheme is able to handle seed bank densities in the entire range tested with only eight agents, whether or not the Whittle index is utilized with EVaR. The use of the Whittle index in conjunction with the G-SB prediction scheme improves performance significantly over the case where only EVaR is used. The greedy fully observed scenario has performance in between these two methods, for the reason described above, and is shown in both figures for ease of comparison.

Chapter 6

Conclusions

In this work, we develop a prediction strategy which is effective across many simulated fields. This strategy leverages both a predictive model based on past data on the field, and an efficient online state estimate which allows accurate density predictions using only the sparse data available during the crop season. By extending the original E-GP model in [2] we were able to develop a prediction strategy which performs well empirically for the weed growth application. By adding a bias term to the model presented in [2], we better account for the variation in the temporal evolution of weed growth across the field domain, and decouple this variation from the underlying spatiotemporal dynamics of weed growth. This is what enables the model to account for the variation in plant growth resulting from the unique seed bank distribution within each field. This model also provides physical insight into the seed bank density of the field. In fact, the fixed point of the model for a given field has a nearly identical shape to the seed bank density for that field. This would allow weed scientists to gain insight on how the seed bank is distributed in fields. Our more general E-GP method outperforms a baseline predictor relying on the weed growth model. We integrate these prediction strategies with a planning index using EVaR in conjunction with the Whittle index, in order to improve upon the decision making performance of our previous work in [1].

Chapter 7

Future Work

Our long term goal is that this work will showcase an autonomy framework which will both contribute to solving the agricultural problems facing humanity in the future, and provide insight to future researchers as they create new generations of intelligent infrastructure to address other world issues. As a domain to test our algorithms in a real agricultural environment, we will utilize the TerraSentia robot [12], developed by EarthSense. First, existing computer vision algorithms will be used to detect the weed density in videos taken with the TerraSentia robot in fields at the University of Illinois at Urbana-Champaign, and compared with human counts performed by crop scientists on those same fields. The GPS data from the TerraSentia robot will be used to match measurement locations. Next, our prediction scheme will be used on this observed data in order to validate our model for weed growth used in the Weed World simulation environment.

Finally, after the weeding apparatus for the robot, the communications framework which enables multi-robot coordination, and the online version of the computer vision algorithm for weed detection, are all completed, we will implement our path planning algorithm on board a team of TerraSentia robots. Experiments validating our algorithm's performance will be completed. Our final work will also detail relevant aspects of the design of the TerraSentia robot.

References

[1] W. McAllister, D. Osipychev, G. Chowdhary, and A. Davis, "Multi-agent planning for coordinated robotic weed killing," in *Intelligent Robots and Systems (IROS), 2018 IEEE/RSJ International Conference on*. IEEE, 2018.

[2] J. Whitman and G. Chowdhary, "Learning dynamics across similar spatiotemporally-evolving physical systems," in *Conference on Robot Learning*, 2017, pp. 472–481.

[3] D. L. Shaner, "Lessons learned from the history of herbicide resistance," *Weed Science*, vol. 62, no. 2, pp. 427–431, 2014.

[4] I. Heap, "The international survey of herbicide resistant weeds," *Heap, I*, 2017. [Online]. Available: http://www.weedscience.org

[5] B. D. Maxwell, M. L. Roush, and S. R. Radosevich, "Predicting the evolution and dynamics of herbicide resistance in weed populations," *Weed Technology*, vol. 4, no. 1, pp. 2–13, 1990.

[6] M. Livingston, J. Fernandez-Cornejo, and G. B. Frisvold, "Economic returns to herbicide resistance management in the short and long run: The role of neighbor effects," *Weed Science*, vol. 64, no. sp1, pp. 595–608, 2016.

[7] M. V. Bagavathiannan and J. K. Norsworthy, "Multiple - herbicide resistance is widespread in roadside palmer amaranth populations," *PloS One*, vol. 11, no. 4, p. e0148748, 2016.

[8] C. M. Evans, "Characterization of a novel five-way-resistant population of waterhemp (amaranthus tuberculatus)," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2016.

[9] J. Gressel, A. J. Gassmann, and M. D. Owen, "How well will stacked transgenic pest/herbicide resistances delay pests from evolving resistance?" *Pest Management Science*, vol. 73, no. 1, pp. 22–34, 2017.

[10] B. D. Booth, S. D. Murphy, and C. J. Swanton, *Weed ecology in natural and agricultural systems*. CABI Pub., 2003.

[11] E. R. Page, D. Cerrudo, P. Westra, M. Loux, K. Smith, C. Foresman, H. Wright, and C. J. Swanton, "Why early season weed control is important in maize," *Weed Science*, vol. 60, no. 3, pp. 423–430, 2012.

[12] TerraSentia, "Terrasentia robot - earthsense, inc." *TerraSentia*, 2017. [Online]. Available: https://www.earthsense.co/home/

[13] ecorobotix, "Technology for the environment," *ecorobotix*, 12 2018. [Online]. Available: https://www.ecorobotix.com/en/

[14] N. Technologies, "Autonomous weeding for agricultural robots - naio technologies," *Naio Technologies*, 2017. [Online]. Available: https://www.naio-technologies.com/en/

[15] H. A. Kingravi, H. R. Maske, and G. Chowdhary, "Kernel observers: Systems-theoretic modeling and inference of spatiotemporally evolving processes," in *Advances in Neural Information Processing Systems*, 2016, pp. 3990–3998.

[16] L. Csato and M. Opper, "Sparse on - line gaussian processes," *Neural computation*, vol. 14, no. 3, pp. 641–668, 2002.

[17] R. Weber, "On the gittins index for multiarmed bandits," *The Annals of Applied Probability*, pp. 1024–1033, 1992.

[18] P. Whittle, "Restless bandits: Activity allocation in a changing world," *Journal of applied probability*, vol. 25, no. A, pp. 287–298, 1988.

[19] R. R. Weber and G. Weiss, "On an index policy for restless bandits," *Journal of Applied Probability*, vol. 27, no. 3, pp. 637–648, 1990.

[20] M. R. Hall, C. J. Swanton, and G. W. Anderson, "The critical period of weed control in grain corn (zea mays)," *Weed Science*, vol. 40, no. 3, pp. 441–447, 1992.

[21] N. Holst, I. Rasmussen, and L. Bastiaans, "Field weed population dynamics: a review of model approaches and applications," *Weed Research*, vol. 47, no. 1, pp. 1–14, 2007.

[22] Y. U. Cao, A. S. Fukunaga, and A. Kahng, "Cooperative mobile robotics: Antecedents and directions," *Autonomous Robots*, vol. 4, no. 1, pp. 7–27, 1997.

[23] J. Niko-Mora, "Stochastic scheduling," Updated version of article in *Encyclopedia of Optimization*, C. A. Floudas and P. M. Pardalos, eds. Kluwer, 2001., 2005. [Online]. Available: http://halweb.uc3m.es/jnino/eng/pubs/ssche.pdf

[24] E. Frazzoli and F. Bullo, "Decentralized algorithms for vehicle routing in a stochastic time-varying environment," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, vol. 4. IEEE, 2004, pp. 3357–3363.

[25] A. Munoz-Meléndez, P. Dasgupta, and W. Lenagh, "A stochastic queueing model for multi-robot task allocation," in *International Conference on Informatics in Control, Automation and Robotics*, 2012, pp. 256–261.

[26] B. P. Gerkey and M. J. Matarić, "A formal analysis and taxonomy of task allocation in multi-robot systems," *International Journal of Robotics Research*, vol. 23, no. 9, pp. 939–954, 2004.

[27] B. Chopard and M. Droz, *Cellular automata*. Springer, 1998.

[28] D. Mulugeta and D. E. Stoltenberg, "Seed bank characterization and emergence of a weed community in a moldboard plow system," *Weed Science*, pp. 54–60, 1997.

[29] D. Nordby, R. Hartzler, and K. Bradley, "Biology and management of waterhemp," *Glyphosate, Weeds, and Crop Sciences, Purdue University Extension, publication GWC-13.12*, 2007.

[30] B. J. Schutte and A. S. Davis, "Do common waterhemp (amaranthus rudis) seedling emergence patterns meet criteria for herbicide resistance simulation modeling?" *Weed Technology*, vol. 28, no. 2, pp. 408–417, 2014.

[31] R. Werle, L. D. Sandell, D. D. Buhler, R. G. Hartzler, and J. L. Lindquist, "Predicting emergence of 23 summer annual weed species," *Weed Science*, 2014.

[32] B. A. Sellers, R. J. Smeda, W. G. Johnson, J. A. Kendig, and M. R. Ellersieck, "Comparative growth of six amaranthus species in Missouri," *Weed Science*, vol. 51, no. 3, pp. 329–333, 2003.

[33] M. J. Horak and T. M. Loughin, "Growth analysis of four amaranthus species," *Weed Science*, vol. 48, no. 3, pp. 347–355, 2000.

[34] H. Kingravi, H. Maske, and G. Chowdhary, "Kernel observers: Systems theoretic modeling and inference of spatiotemporally varying processes," in *Advances in Neural Information Processing Systems (NIPS)*, Barcelona, Spain, 2016, accepted.

[35] H. A. Kingravi, H. Maske, and G. Chowdhary, "A systems - theoretic approach for data-driven modeling and control of spatiotemporally evolving processes," in *2015 54th IEEE Conference on Decision and Control (CDC)*, 2015, pp. 7365–7370.

[36] A. S. Davis, K. A. Renner, and K. L. Gross, "Weed seedbank and community shifts in a long-term cropping systems experiment," *Weed Science*, vol. 53, no. 3, pp. 296–306, 2005.

[37] A. Ahmadi-Javid, "Entropic value-at-risk: A new coherent risk measure," *Journal of Optimization Theory and Applications*, vol. 155, no. 3, pp. 1105–1123, 2012.

[38] A. Axelrod, L. Carlone, G. Chowdhary, and S. Karaman, "Data-driven prediction of evar with confidence in time-varying datasets," in *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 5833–5838.

[39] A. Hobson, "A new theorem of information theory," *Journal of Statistical Physics*, vol. 1, no. 3, pp. 383–391, 1969.

[40] R. Houthooft, X. Chen, Y. Duan, J. Schulman, F. De Turck, and P. Abbeel, "Vime: Variational information maximizing exploration," in *Advances in Neural Information Processing Systems*, 2016, pp. 1109–1117.

[41] R. Allamaraju, H. Kingravi, A. Axelrod, G. Chowdhary, R. Grande, J. P. How, C. Crick, and W. Sheng, "Human aware uas path planning in urban environments using nonstationary mdps," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 1161–1167.

[42] D. Pathak, P. Agrawal, A. A. Efros, and T. Darrell, "Curiosity-driven exploration by self-supervised prediction," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 16–17.

[43] J. Achiam and S. Sastry, "Surprise-based intrinsic motivation for deep reinforcement learning," *arXiv preprint arXiv:1703.01732*, 2017.

[44] A. Ahmadi-Javid and M. Fallah-Tafti, "Portfolio optimization with entropic value-at-risk," *European Journal of Operational Research*, 2019.

[45] M. Denuit, J. Dhaene, M. Goovaerts, and R. Kaas, *Actuarial theory for dependent risks: measures, orders and models*. John Wiley & Sons, 2006, p. 149, citing stop-loss ordering.

[46] M. Denuit, J. Dhaene, M. Goovaerts, and R. Kaas, *Actuarial theory for dependent risks: measures, orders and models*. John Wiley & Sons, 2006, pp. 105–106, citing stochastic ordering.

[47] A. Axelrod and G. Chowdhary, "A dynamic risk form of entropic value at risk," in *AIAA Scitech 2019 Forum*, 2019, p. 0392.

[48] W. McAllister, D. Osipychev, G. Chowdhary, and A. Davis, "Agbots: Weeding a field with a team of autonomous robots," *Computers and Electronics in Agriculture*, vol. 163, p. 104827, 2019.