Journal of Agricultural and Resource Economics Preprint
ISSN: 1068-5502 (Print); 2327-8285 (Online)
doi: 10.22004/ag.econ.342180

Valuing Monitoring Networks for New Pathogens: The Case of Soybean Rust in the United States

Tiesta Thakur, Terrance M. Hurley, Frances R. Homans, and Robert G. Haight

The value of information to farmers from a soybean rust sentinel plot monitoring network is estimated. Monitoring provides value over time from farmers learning about their risk of infection and contemporaneously from farmers making more profitable fungicide applications. Optimizing the spatial configuration of the network reveals the value of information plateaus at around $920M with around 75 sentinel plots. The optimal spatial distribution of sentinel plots differs substantially from historical placements. Optimal sentinel plot placement would have included fewer plots in the Southern United States where the risk of infection is relatively high and more in the Midwest where soybean production is relatively high.

Key words: Bayesian Updating, Dynamic Optimization, Spatial Optimization

Introduction

Soybean rust (caused by the fungus *Phakopsora pachyrhizi*) arrived in the United States (US) in 2004 (Wrather and Koenning, 2009) likely carried from South America by Hurricane Ivan (Isard et al., 2005). The fungus cannot overwinter in temperate US climates but thrives year-round on kudzu in the Gulf Coast's heat and humidity (Sikora et al., 2014). During the soybean growing season, rust spores can blow into soybean fields in temperate zones creating a substantial yield loss risk. Researchers have estimated the aggregate loss to the US producers, consumers and society by assuming different extents of the severity and spread of the soybean rust infection. Kuchler et al. (1984) estimated that, 80% of the time, the annual net social loss could be as high as $1.4 billion even with a conservative estimate of the potential damage, while Livingston et al. (2004) estimated first-year expected loss could reach $1.3 billion. Yet, US soybean rust damage turned out to be less than feared. Wrather and Koenning (2009) report that production losses for 2005-2008, evaluated at average prices, imply financial losses of only $4.14 million annually.

Soybean rust is a recent example of a potentially catastrophic pathogen entering US agriculture with other plant diseases looming on the horizon (Rossman et al., 2006). Changing climatic conditions appear to be increasing the risk of plant pathogen infections, including soybean rust, at latitudes farther from the equator (Chaloner, Gurr, and Bebber, 2021). In the US, the National Plant

Tiesta Thakur is an assistant professor in the School of Public Policy, Indian Institute of Technology Delhi. Terrance M. Hurley (corresponding author, tmh@umn.edu) is a professor in the Department of Applied Economics, University of Minnesota. Frances R. Homans is a professor in the Department of Applied Economics, University of Minnesota. Robert G. Haight is a research forester emeritus at the USDA Forest Service, Northern Research Station.
The authors thank Scott Isard for valuable guidance throughout this project. We also thank Don Hershman, Loren Giesler, Ed Sikora and other participants at the workshop for "Developing a Cost-effective Rational Strategy to Detect Soybean Rust in the United States" for helpful insights into the practical aspects of managing soybean rust, Stephen Aultman for developing an early version of the dynamic programming model with Bayesian updating, and Mike Livingston, Scott Isard and Becky Epanchin-Niell for reading and commenting on earlier drafts of this paper. Support for this project was provided by USDA award #2007-39552-18643 and by the USDA Forest Service Northern Research Station. The authors also appreciate the financial support of the Minnesota State Agricultural Experiment Station (Projects: MIN-14-076, MIN-14-134 and MIN-14-120). All interpretations, conclusions, and any remaining errors are solely attributable to the authors.

Review coordinated by Dayton M. Lambert.

Diagnostic Network has been established to detect the arrival of new agricultural pathogens (Stack et al., 2006) and similar programs have been developed worldwide (Miller, Beed, and Harmon, 2009). The arrival of soybean rust in the continental US prompted the creation of the Integrated Pest Management Pest Information Platform for Extension and Education (ipmPIPE), which includes a sentinel plot monitoring network. The sentinel plots in the monitoring network are areas of early maturing soybean grown to detect rust. Initially, the ipmPIPE included a web-based information technology system that provided farmers with direct access to information on confirmed rust cases, management options, and forecasts with expert commentary on disease progression throughout the growing season.

A consortium of government agencies, agricultural trade organizations, and land-grant institutions provided the resources necessary to support the ipmPIPE. Such broad public and trade support may be justified by the public good aspects of the information provided (Miller, Beed, and Harmon, 2009) because farmers have limited incentives and ability to coordinate such efforts on their own. Still, since the adverse effects of soybean rust on US production were less than anticipated, there have been questions about the value of the network as compared to its costs (Livingston, 2010). Particularly relevant is the work of Roberts et al. (2009) in which the value of the network as the benefit to farmers from making better fungicide decisions is identified. They find that the value of the network depends on the accuracy of the information and farmers' prior estimates of the infection probability. For most reasonable parameter values, and from an ex-ante perspective, they find that the value of the program exceeded its costs. Investment in the soybean rust platform has been scaled back from more than 700 sentinel plots in 2007 to 75 in 2014 (ipmPIPE), though the program has also expanded monitoring and developed new tools for other crops and diseases.

This paper focuses on estimating a monitoring network's value when optimally designed. It examines how the presence of a sentinel plot at some distance from a farm affects farmer decisions and then proceeds to assess how the number of sentinel plots, optimally placed across space, affects network value. This paper brings into focus the importance of the correlation of the disease occurrence across space because the relevance of the appearance of the disease in the sentinel plot for a farm at any location depends, in part, on how far that farm is from the sentinel plot. For most US soybean growing regions, the most important vulnerability is long-distance rather than local dispersal (Isard et al., 2005). A key feature of the developed framework is its accommodation of the spatial correlations emerging from the shared vulnerability to a pathogen arriving from afar.

There is a rich optimization literature on surveillance and control systems for invasive species management (Büyüktahtakın and Haight, 2018). Modeling the population dynamics of invasive species is key part of system design (Epanchin-Niell and Hastings, 2010). While each invasive species has distinct population dynamics, it is common to divide invasion into a series of stages—introduction, establishment, and spread—and model the population dynamics as a function of barriers an invading species must overcome to move through these stages (Blackburn et al., 2011). Management strategies are associated with each stage and are intended to prevent a population from moving through them. Surveillance and control strategies are employed from the beginning of the establishment stage to the end of the spread stage with the goal of finding new populations of invaders and eradicating or containing them while minimizing the costs of surveillance, control of detected populations, and damage caused by undetected populations (Epanchin-Niell and Hastings, 2010). Stochastic models in the surveillance domain concentrate on the optimal surveillance location and intensity assuming the location of the invader is unknown and that control measures are automatically applied to eradicate or slow the spread of found populations (e.g., (Mehta et al., 2007; Epanchin-Niell et al., 2012)). Few models address the joint optimization of surveillance and control (Horie et al., 2013; Yemshanov et al., 2019), and they are limited to short (two-period) time horizons. Kıbış et al. (2021) jointly optimize surveillance and control with a multi-stage, spatial, stochastic programming model that includes a great deal of biological complexity, including the classification of the host population with respect to health conditions and transition of one disease class into another. In all these models, the parameters driving resource allocations include host density, infestation or establishment rates, population spread or growth rates, and the sensor detection rate as well as the costs of surveillance, control, and damage.

Two features of our decision environment for soybean rust distinguish it from the existing literature. First, the dynamics of soybean rust do not fit establishment and spread models typically used to design surveillance and control strategies. Soybean rust spores spread north from the Gulf region via wind events during the soybean growing season and infect soybeans if the temperature and moisture conditions are suitable. Because establishing soybean rust populations do not overwinter and spread in the primary soybean growing regions of the US, the parameter of interest is the establishment rate, which is uncertain and can be monitored over time to update estimates. The second feature of our decision environment is multiple stakeholders. Soybean farmers have choices about how they manage their soybean crops to maximize profits, and those choices depend on parameters such as expected soybean rust establishment rate, crop prices, and management costs, all of which vary over space. The model developed here addresses a central planner's problem of designing a system that maximizes the farmers' value of information provided by the monitoring network by providing information about the presence of soybean rust to soybean farmers who may use that information in their management decisions.

Farmer Decision Model

This section develops a farm-level decision model without a sentinel plot network. It then incorporates within-season management decisions based on sentinel plot information.[1] Finally, the model is extended to include multi-season learning about soybean rust risk.

Management without a Sentinel Plot Network

Following Roberts et al. (2009), farmers face two possible states of the world, represented by i=1 for rust infected or i=0 for uninfected farm fields. A farmer can respond to the threat of an infection in one of the three ways: do nothing, scout and apply a curative fungicide as needed, or apply a prophylactic preventative fungicide. Let P be the price of soybeans, Y be the soybean yield without a soybean rust infection and C be the production cost exclusive of any control effort. Let λ_n be the yield loss proportion with an uncontrolled infection. Do Nothing (N) returns are $\pi_{N0} = PY - C$ without an infection and $\pi_{N1} = PY(1 - \lambda_n) - C$ with one.

The second strategy, scout and apply a curative fungicide as needed, is labeled R. With this strategy, a farmer applies fungicide when scouting reveals a soybean rust infection. It is costly to scout and costs are incurred whether or not soybean rust is detected. If soybean rust is detected, a curative fungicide application is also costly. By the time soybean rust is detected, some crop damage will have occurred. With $C_{sc} > 0$ equal to the cost of scouting, $C_r > 0$ equal to the cost of a curative fungicide, and λ_r equal to the yield loss proportion with a curative fungicide and infection, this strategy's returns are $\pi_{R0} = PY - C - C_{sc}$ without an infection and $\pi_{R1} = PY(1 - \lambda_r) - C - C_{sc} - C_r$ with one. While a curative treatment does not give perfect control of an infection, losses will be less than doing nothing: $\lambda_r < \lambda_n$.

A Prophylactic Preventative strategy, labeled PP, reduces yield loss even further in the event of an infection because rust has little or no chance of damaging the crop. For success in implementing this strategy, a preventative fungicide must be applied before an infection so that a fungicide is applied even when no infection would have occurred. Furthermore, the cost of a preventative treatment, C_p, is generally greater than the cost of a curative treatment: $C_p > C_r$. In the event of an

[1] There may be value to information sharing among farmers in the absence of a sentinel plot network. Bekkerman et al. (2012) model a system in which neighbors treat their crops once soybean rust is detected in neighboring fields. We chose not to incorporate informal information sharing into our model and instead focused on systematic provision of information from detection of soybean rust on early-maturing varieties of soybean.

infection, yield losses with a preventative treatment tend to be negligible (Johansson et al., 2006),[2] so *PP* returns are $\pi_{PP} = \pi_{PP0} = \pi_{PP1} = P$Âů$Y$âĂŞ$C$âĂŞ$C_p$ both with and without an infection. The farmer's decision depends on the three strategies' expected returns. If there is no chance of infection, a farmer should choose *N*. If infection is guaranteed, returns depend on yield losses, fungicide costs, and scouting costs. Since a farmer chooses a strategy before knowing whether or not an infection will occur, her belief about the probability of an infection, ϕ^f, is crucial. An important source of information is personal experience which is influenced by management choices. Choosing *N* or *R* allows farmers to observe infections first-hand. These two strategies provide an opportunity to learn about the risk of infection. With a preventative fungicide, a farmer forgoes this opportunity because treatment precludes the observation of the disease when infection would have occurred.

Management with a Sentinel Plot Network

In Roberts et al. (2006, 2009), the monitoring network sends one of two signals: the risk of infection is either high or low. Farmers use the signal to update their beliefs about the probability of infection during the season. The signal may be of low, medium, or high quality: the quality of the signal determines the degree to which farmers update their beliefs. Farmers can respond with one of three strategies: do nothing about soybean rust, scout and apply a curative fungicide in the event of an infection, or apply a prophylactic preventative fungicide.

The model developed here differs in three ways. The first difference is how the signal is defined. Rather than modeling a high or low risk of infection, the model assumes the signal a farmer receives is whether an infection has been confirmed at the sentinel plot. Instead of characterizing the signal quality, the correlation between infection at the sentinel plot and infection in the farmer's field is assumed to decline with the distance between the two. These two features are more precisely defined than the signal and signal quality of Roberts et al. (2009) and make it possible to evaluate the benefits of proximity to the sentinel plot for use in the spatial optimization model. Second, the timing of decision-making is different so that the farmer chooses her strategy based on her prior probability of infection before the signal arrives. Third, a fourth management option is introduced: apply a conditional preventative fungicide mid-season if an infection is confirmed at the relevant sentinel plot during the season. This second type of preventative treatment is added to better characterize observed farmer behavior (D. Hershman, personal communication). The Conditional Preventative (*CP*) strategy is defined as a strategy in which the farmer bases the decision to apply preventative fungicide on whether the relevant sentinel plot becomes infected. This strategy is available only when the ipmPIPE exists. If the ipmPIPE signals infection in a relevant sentinel plot, a fungicide is applied. Otherwise, no fungicide is applied. Note that the Conditional Preventative strategy is composed of the actions reflected in *N* and *PP* strategies but is different from these two strategies in the timing of the action. With *N* and *PP*, the commitment to doing nothing or applying prophylactic preventative fungicide is made at the outset of the season. With *CP*, the choice between the two actions is made once the signal from the sentinel plot has been received.

Given the ipmPIPE system, a farmer faces one of the following four situations: (1) there is an infection in the farmer's field and in the sentinel plot, (2) neither has an infection, (3) there is an infection in the field and not in the sentinel plot, and (4) there is an infection in the sentinel plot, but not in the farmer's field. Let j represents the occurrence of an infection in the sentinel plot such that j equals 1 if an infection occurs, 0 otherwise. Let ϕ_{ij} be the subjective probability associated with each of the four scenarios described above. The expected return to the *CP* strategy is $E(\pi_{CP}) = \phi_{11}\pi_{PP} + \phi_{10}\pi_{N1} + \phi_{01}\pi_{PP} + \phi_{00}\pi_{N0}$. Note that there is some probability (ϕ_{10}) that the field becomes infected but is left unprotected because no infection is detected in the sentinel plot. The ipmPIPE provides information about a sentinel plot's infection status. How this information helps farmers depends on the correlation of infection between the farmer's field and the sentinel

[2] Johansson et al. (2006) find that the average yield impact of preventative treatment is -0.97%. To simplify our model, we assume a 100% efficacy of the preventative strategy.

plot. Sentinel plots closer to the farmer's field are likely to provide more accurate information about the infection risk. Thus, correlation is expected to be decreasing in the distance between a farmer's field and a sentinel plot.[3] However, nearness alone does not guarantee a high correlation. If a sentinel plot has climate conditions favorable for soybean rust growth while the farmer's field does not, then an infection in the sentinel plot does not necessarily imply a high risk of infection in the farm field.[4] Therefore, sentinel plot information is more useful for revising beliefs if the probability of infection in the sentinel plot is similar to the probability in the farmer's field.

The subjective infection probabilities in a sentinel plot and farmer's field are ϕ^s and ϕ^f respectively. Assume beliefs about the infection probability are common knowledge. Let ρ be the correlation between infection in a farmer's field and in a sentinel plot. Each ϕ_{ij} can then be expressed as:[5]

$$\phi_{11} = \phi^f\phi^s + \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} \tag{1}$$

$$\phi_{01} = (1-\phi^f)\phi^s - \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} \tag{2}$$

$$\phi_{10} = \phi^f(1-\phi^s) - \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} \tag{3}$$

$$\phi_{00} = (1-\phi^f)(1-\phi^s) + \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)}. \tag{4}$$

These subjective probabilities $\{\phi_{11},\phi_{01},\phi_{10},\phi_{00}\}$ are bounded by zero and one. Therefore,

$$\rho \le min\left\{\sqrt{\frac{\phi^s(1-\phi^f)}{\phi^f(1-\phi^s)}},\sqrt{\frac{\phi^f(1-\phi^s)}{\phi^s(1-\phi^f)}},1\right\}. \tag{5}$$

This correlation is expected to be a decreasing function of distance, d, between the farmer's field and a sentinel plot. It is also expected to be non-negative because an infection in a sentinel plot would either increase or have no effect on the likelihood of infection in a farmer's field. These features are captured in the following function for ρ:

$$\rho = \begin{cases} UB, & d \le d^* \\ \dfrac{UB}{1+e^{\theta(d-d^*)+\gamma}}, & d > d^*, \end{cases} \tag{6}$$

where UB is the upper bound on ρ given by equation 5 and θ, d^* and γ are parameters.

Learning in a Multi-Season Model

As shown empirically in Bekkerman, Goodwin, and Piggott (2008), a farmer's decision-making capacity is expected to change with personal experience because, as their experience increases, so does their understanding of rust infection risk. Their experience, in turn, is influenced by management choices. Choosing N or R allows them to observe any infection in their fields and learn about their risk of infection. However, with a preventative fungicide, a farmer forgoes learning because treatment inhibits disease emergence if rust spores reach the crop.

[3] Bekkerman, Goodwin, and Piggott (2008) find that the chance of rust infection is significantly affected by the infection status of the neighboring locations.

[4] Bekkerman, Goodwin, and Piggott (2008) find that the soybean rust infection risk is affected by factors such as infection in the previous year, the previous year's proportion of soybeans harvested planted, and the maturity group of the planted soybeans. Yang and Batchelor (1997) found that relative humidity and temperatures have large impacts on determining the spread of soybean rust.

[5] Derivations of equations 1-5 are available in the appendix.

The ipmPIPE and its monitoring network provide an additional opportunity for a farmer to learn about the infection risk in their field, irrespective of their chosen strategy. Due to the ipmPIPE system, a farmer knows when an infection happens in a sentinel plot. As a result, even if the farmer chooses to use a preventative treatment, information from the sentinel plot will help in refining the farmer's belief about their infection risk. The usefulness of this new information will depend on the correlation between the field and sentinel plot. In essence, the ipmPIPE system provides many scouting benefits without the cost and allows a farmer to delay or completely forgo a preventative treatment unless monitoring suggests an infection has become likely because an infection is identified at a sentinel plot close to the farmer's field.

To capture a farmer's learning about the probability of infection, a multi-season decision problem is solved where farmers maximize the expected present value of profit by choosing the best strategy for soybean rust control given their beliefs about the infection probability and the availability of the ipmPIPE system. Suppose the probability that X infections are observed in t years in location h follows a binomial distribution with parameter ω_t^h, where h can denote the farmer's field f or sentinel plot s. Suppose the farmer's prior rust infection probability belief in location h follows the beta distribution with shape parameter α_t^h equal to the number of times an infection is observed and β_t^h equal to the number of times an infection is not observed in location h: $b(\omega_t^h;\alpha_t^h,\beta_t^h)$. The farmer's belief about the state of the world in period t can be defined by $\{\alpha_t^f,\beta_t^f\}$ without ipmPIPE and $\{\alpha_t^f,\beta_t^f,\alpha_t^s,\beta_t^s\}$ with it because, with ipmPIPE, the farmer knows about the infection risk in the sentinel plot in addition to the risk in the farm field.[6]

Using the beta distribution to characterize prior beliefs also implies that the posterior distribution $b(\omega_{t+1}^h;\alpha_{t+1}^h,\beta_{t+1}^h)$ equals $b(\omega_{t+1}^h;\alpha_t^h+1,\beta_t^h)$ with an observed infection and $b(\omega_{t+1}^h;\alpha_t^h,\beta_t^h+1)$ without. The farmer's knowledge about the state of the world in t + 1 depends on two factors: (i) the availability of the monitoring network and (ii) the optimal strategy for tackling the risk of infection in the farm field, given her beliefs.

When the farmer chooses R or N, she gets first-hand information on whether her fields became infected. This is true even in the absence of the monitoring network. Thus, α_t^f and β_t^f are updated in t + 1, irrespective of the availability of monitoring network. When she chooses prophylactic measures, *PP*, she never gets to observe first-hand whether an infection would have happened in her fields. But, with the monitoring network in place, she can refine her beliefs, ϕ_{ij}s, based on the information about infections in the sentinel plot which may affect her future decisions. When the farmer chooses *CP*, she will get direct information from her own field only when the sentinel plot is not infected because, in that case, she will not apply a preventative fungicide. If an infection appears in the sentinel plot, she will apply preventative fungicide on her field and will not get direct information for adjusting her risk beliefs.

Beliefs about the farm's infection probability are updated with observations of the state of the world each time period. This is because the expectation of the beta distribution $b(\omega_t^h;\alpha_t^h,\beta_t^h)$ is

$$E(\omega_t^h) = \frac{\alpha_t^h}{\alpha_t^h + \beta_t^h} \tag{7}$$

where ϕ_t^f equals $E(\omega_t^f)$ and ϕ_t^s equals $E(\omega_t^s)$ by definition. Observations in time period t will allow these parameters to be updated to $\{\alpha_{t+1}^f,\beta_{t+1}^f\}$ and $\{\alpha_{t+1}^s,\beta_{t+1}^s\}$ which completely characterize the farmer's beliefs for the next time period based on equations 1- 5.[7] The farmer's contemporaneous

[6] That is, the marginal distributions for ϕ_t^f and ϕ_t^s are defined by beta distributions with parameters $\{\alpha_t^f,\beta_t^f\}$ and $\{\alpha_t^s,\beta_t^s\}$. These distribution are then linked via the correlation function to get the joint distribution of infection in the farmer's field and sentinel plot (see equation 1 - equation 5).

[7] Note that the marginal probabilities of infestation in the sentinel plot and the farmers' field are independent beta distributions, while the joint distribution links these two distributions through the correlation function, so the joint probabilities of infection need not be independent.

expected return for the four possible strategies can thus be written as follows:

$$E(\pi_{N_t}) = \frac{\alpha_t^f}{\alpha_t^f + \beta_t^f}\pi_{N1} + \frac{\beta_t^f}{\alpha_t^f + \beta_t^f}\pi_{N0} \tag{8}$$

$$E(\pi_{R_t}) = \frac{\alpha_t^f}{\alpha_t^f + \beta_t^f}\pi_{R1} + \frac{\beta_t^f}{\alpha_t^f + \beta_t^f}\pi_{R0} \tag{9}$$

$$E(\pi_{PP_t}) = \frac{\alpha_t^f}{\alpha_t^f + \beta_t^f}\pi_{PP1} + \frac{\beta_t^f}{\alpha_t^f + \beta_t^f}\pi_{PP0}$$

$$= \pi_{PP} \tag{10}$$

$$E(\pi_{CP_t}) = \phi_{11t}\pi_{PP} + \phi_{10t}\pi_{N1} + \phi_{01t}\pi_{PP} + \phi_{00t}\pi_{N0}. \tag{11}$$

Dynamic Optimization with Infection Probabilities

Calculating the value of ipmPIPE, including the opportunity it provides to learn about the risk of infection over time, is accomplished using dynamic programming where the state variables are the parameters of a Beta distribution characterizing a farmer's belief about infection probabilities.

Optimization Without Sentinel Plot Information

Without the monitoring network, the optimal value function for the farmer is

$$V_t(\alpha_t^f, \beta_t^f) = \text{Max}\begin{cases} \text{E}(\pi_{N_t}) + \delta\text{E}(V_{t+1}(\alpha_{t+1}^f, \beta_{t+1}^f)), \\ \text{E}(\pi_{R_t}) + \delta\text{E}(V_{t+1}(\alpha_{t+1}^f, \beta_{t+1}^f)), \\ \text{E}(\pi_{PP_t}) + \delta V_{t+1}(\alpha_t^f, \beta_t^f), \end{cases} \tag{12}$$

where

$$\text{E}(V_{t+1}(\alpha_{t+1}^f, \beta_{t+1}^f)) = \phi_t^f V_{t+1}(\alpha_t^f + 1, \beta_t^f) + (1 - \phi_t^f)V_{t+1}(\alpha_t^f, \beta_t^f + 1) \tag{13}$$

is the expected value function when there are opportunities to learn and $\delta \in (0,1)$ is the discount factor. The first argument in the maximum of equation 12 reflects a farmer's expected return from choosing the strategy N initially and the optimal strategy thereafter. Since choosing N lets the farmer observe whether or not an infection occurs, the farmer's optimal future strategies will depend on what is observed, and this yields the expected value function in equation 13. With probability ϕ_t^f, the farmer will observe an infection and update their priors to $\{\alpha_t^f + 1, \beta_t^f\}$. With probability $1 - \phi_t^f$, the farmer will not observe an infection and will update their priors to $\{\alpha_t^f, \beta_t^f + 1\}$. The second argument differs from the first only because the farmer chooses R initially. With this choice, the farmer still has the opportunity to learn, yielding the discounted expected value in equation 13 for the remainder of time. The third argument is different from the first two because the farmer chooses PP. This strategy does not provide any new information. Therefore, α^f and β^f remain the same in the next time period.

Optimization with Sentinel Plot Information

A sentinel plot network allows farmers to gather information about their infection risk regardless of management strategy and provides a new within-season signal for conditional preventative fungicide treatments. Assuming the network is available for L years, the optimal value function with the monitoring network (subscripted M) can be written as

$$V_{M_t}(\alpha_t^f,\beta_t^f,\alpha_t^s,\beta_t^s) = \text{Max}\begin{cases} \text{E}(\pi_{N_t}) + \delta\text{E}(V_{M_{t+1}}(\alpha_{t+1}^f,\beta_{t+1}^f,\alpha_{t+1}^s,\beta_{t+1}^s)), & \\ \text{E}(\pi_{R_t}) + \delta\text{E}(V_{M_{t+1}}(\alpha_{t+1}^f,\beta_{t+1}^f,\alpha_{t+1}^s,\beta_{t+1}^s)), & \\ \text{E}(\pi_{PP_t}) + \delta\text{E}(V'_{M_{t+1}}(\alpha_{t+1}^f,\beta_{t+1}^f,\alpha_{t+1}^s,\beta_{t+1}^s)), & \\ \text{E}(\pi_{CP_t}) + \delta\text{E}(V''_{M_{t+1}}(\alpha_{t+1}^f,\beta_{t+1}^f,\alpha_{t+1}^s,\beta_{t+1}^s)), & t < L \\ V(\alpha_t^f,\beta_t^f), & t \geq L \end{cases} \tag{14}$$

where

$$\begin{aligned} \text{E}(V_{M_{t+1}}(\alpha_{t+1}^f,\beta_{t+1}^f,\alpha_{t+1}^s,\beta_{t+1}^s)) &= \phi_{11t}V_{M_{t+1}}(\alpha_t^f+1,\beta_t^f,\alpha_t^s+1,\beta_t^s) \\ &+ \phi_{00t}V_{M_{t+1}}(\alpha_t^f,\beta_t^f+1,\alpha_t^s,\beta_t^s+1) \\ &+ \phi_{10t}V_{M_{t+1}}(\alpha_t^f+1,\beta_t^f,\alpha_t^s,\beta_t^s+1) \\ &+ \phi_{01t}V_{M_{t+1}}(\alpha_t^f,\beta_t^f+1,\alpha_t^s+1,\beta_t^s) \end{aligned} \tag{15}$$

$$\begin{aligned} \text{E}(V'_{M_{t+1}}(\alpha_{t+1}^f,\beta_{t+1}^f,\alpha_{t+1}^s,\beta_{t+1}^s)) &= \phi_t^s(V_{M_{t+1}}(\alpha_t^f,\beta_t^f,\alpha_t^s+1,\beta_t^s)) \\ &+ (1-\phi_t^s)(V_{M_{t+1}}(\alpha_t^f,\beta_t^f,\alpha_t^s,\beta_t^s+1)) \end{aligned} \tag{16}$$

$$\begin{aligned} \text{E}(V''_{M_{t+1}}(\alpha_{t+1}^f,\beta_{t+1}^f,\alpha_{t+1}^s,\beta_{t+1}^s)) &= \phi_t^s(V_{M_{t+1}}(\alpha_t^f,\beta_t^f,\alpha_t^s+1,\beta_t^s)) \\ &+ (1-\phi_t^s)\left[\frac{\phi_{00t}}{1-\phi_t^s}V_{M_{t+1}}(\alpha_t^f,\beta_t^f+1,\alpha_t^s,\beta_t^s+1)\right. \\ &+ \left.\frac{\phi_{10t}}{1-\phi_t^s}V_{M_{t+1}}(\alpha_t^f+1,\beta_t^f,\alpha_t^s,\beta_t^s+1)\right] \end{aligned} \tag{17}$$

are the expected value functions when the chosen strategy is either N or R (equation 15), when it is PP (equation 16), and when it is CP (equation 17). Recall that, while updating of α_t^f and β_t^f occurs only if preventative fungicide is not applied, the probabilities ϕ_{ijt} depend on ϕ_t^s and ρ_t, and thus are always updated.

The sentinel plot's provision of a within-season signal for conditioning preventative treatments results in an opportunity in equation 14, with a contemporaneous expected return of $E(\pi_{CP_t})$, not appearing in equation 12. Once the sentinel plots are no longer available ($t \geq L$), the farmer's problem becomes identical to equation 12.

Farm Model Implications

Having these two models permits the calculation of the value to farmers of the monitoring provided by sentinel plots. The value of monitoring at any time is the difference in the expected value of managing soybean rust with and without information. Therefore, at $t = 0$, the value of information is: $V_{M_0}(\alpha_0,\beta_0) - V_0(\alpha_0,\beta_0)$. Solving the optimization problem in equations 12-13 or 14-17 requires information on expected yield losses in the event of an infection both with and without curative treatments (λ_r,λ_n), expected soybean yields and prices (Y and P), scouting and fungicide costs (C_{sc}, C_r, and C_p), a farmer's initial beliefs about the probability of a soybean rust infection $\{\alpha_0,\beta_0\}$, the number of years that the monitoring network is available (L), the distance and correlation between infection at the farmer's field and sentinel plot (d,ρ), and the discount factor (δ). L is assumed

Table 1. Parameters for the Farmer Decision Model

Parameter	Value
Scouting Costs C_{sc}	\$8.68/Acre
Cost of Curative Fungicide C_c	\$17.86/Acre
Cost of Preventative Fungicide treatment C_p	\$33.15/Acre
Yield loss with no treatment λ_n	25%
Yield loss with Curative Fungicide treatment λ_r	7%
Discount Factor δ	0.94

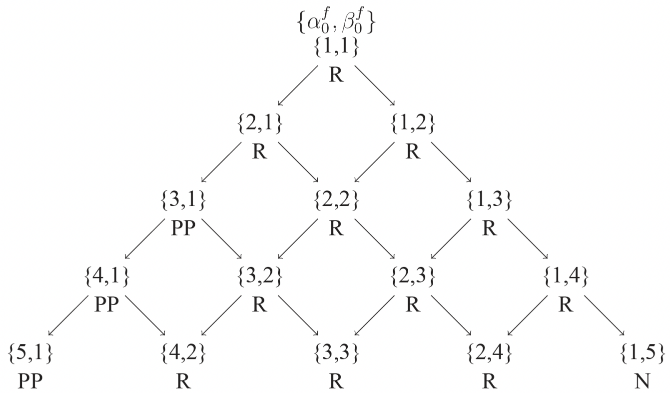

Figure 1. Optimal Management Decisions for a Farmer without a Monitoring Network

Notes: The numbers in the brackets correspond to α_t^f, the number of times an infection is observed, and β_t^f, the number of times an infection is not observed in the farmers field. Optimal actions R (Curative), PP (Prophylactic Preventative), and N (No Action) are noted below each α, β pair at time steps 0-4.

to be 4 years. The values used for λ_r, λ_n and δ were adopted from Johansson et al. (2006) and are summarized in Table 1. Their values for C_{sc},C_r and C_p were for 2005 and are adjusted to account for inflation. The distance between county centroids, d, are calculated based on the Haversine formula. As a result, a field and sentinel plot in the same county have $d = 0$. Correlation, ρ, is assumed to be close to one at distances less than 80 miles. It drops after this point and approaches 0 when the distance reaches 250 miles. Given these assumptions, the parameters θ, d^* and γ from equation 6 take on the values 0.12, 0 and -13.4, respectively.

As a prelude to estimating the value of the sentinel plot monitoring network, the implications of the farmer decision model are explored. The dynamic programming model is solved for an infinite planning horizon using the parameter values in Table 1 and an assumed yield of 37 bushels/acre. Figure 1 shows the first five years of optimal strategies given the farmer's beliefs about the probability of infection in the absence of a monitoring network. The vertices of the decision tree are labeled according to the state variables α_t^f and β_t^f. At the first vertex, α_0^f and β_0^f are both equal to 1 and the beta distribution is equivalent to the uniform distribution. Branches to the left represent observed infections and branches to the right are years without infections. This figure shows that the first node where it is optimal for farmers to choose *PP* is with a prior belief described by {3, 1}. This corresponds to the first 2 observations being infections. Alternatively, if a farmer starts with a uniform prior and sees no infections for the next four years {1, 5}, then it is optimal to choose N. For the majority of states in this example, it is optimal for the farmer to choose R.

Figure 2 shows an example of how the optimal strategy changes with the monitoring network. Suppose beliefs about the rust infection probability are 0.6 for both the farmer's field and the sentinel plot ($\alpha_0^f = \alpha_0^s = 6$, $\beta_0^f = \beta_0^s = 4$). With these parameters, the optimal strategy without a sentinel plot is *PP*. With a sentinel plot in the same county as the farmer's field (i.e., $d = 0$), CP is the farmer's optimal strategy because the farmer will know with certainty whether or not her field will

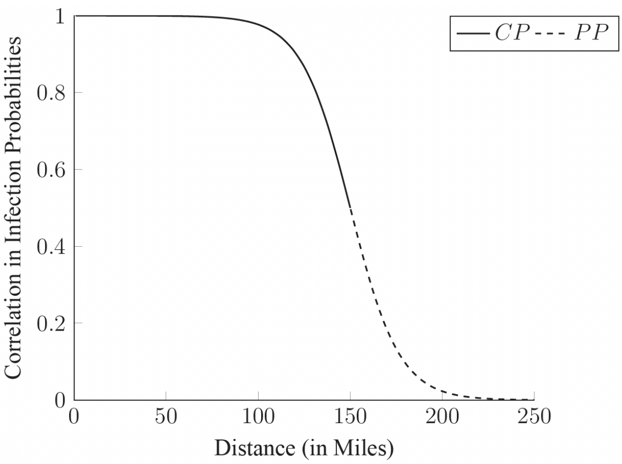

Figure 2. Optimal Management Decisions in the Presence of a Monitoring Network

Notes: Here, the prior belief about infection probability on the farm, ϕ_f, is equal to the prior belief about infection probability on the Sentinel Plot, ϕ_s. CP is the Conditional Preventative strategy and PP is the Prophylactic Preventative strategy.

be infected based on the infection status of the sentinel plot. As the monitoring network becomes less informative with an increase in distance d, ρ decreases as does the value of sentinel plot information. Up to $d < 150$ mi, the model predicts that the farmer will continue with CP. Once the sentinel plot is further than 150 mi, the optimal strategy becomes PP.

It should be noted that the threshold where the optimal strategy shifts away from CP depends on the shape of the underlying correlation function. As an example, suppose there is no monitoring and the belief about the infection probability in the field (ϕ^f) is 0.1. With this low probability, the farmer's optimal strategy is N. Now suppose there is a nearby sentinel plot with ϕ^s equal to 0.6. The farmer's field and the sentinel plot are characterized by very different beliefs and have a relatively low correlation irrespective of the distance between them. Even if the sentinel plot is placed close to the field, the sentinel plot does not provide valuable information about the likelihood of infection, making N optimal even with the monitoring network.

Spatial Optimization Model

Output from the dynamic programming model is used to identify the placement of a chosen number of sentinel plots that maximizes the present value of aggregate expected profit. The model assumes a representative farmer in each county with 2014 county-level soybean acreage and yield along with state-level soybean prices.[8] County-level estimates of the probability of a soybean rust infection from Bekkerman, Goodwin, and Piggott (2008) are used for farmer's prior beliefs. Calculating α_0^h and β_0^h such that $\frac{\alpha_0^h}{\alpha_0^h + \beta_0^h}$ equals the probability estimate and $\alpha_0^h + \beta_0^h = 10$ ensures priors are based on the same number of years of rust infection observations prior to this evaluation.[9][10] Summary statistics are in Table 2.

A fixed number of sentinel plots are allocated across counties to maximize the value of information generated from them. Let $m \in \{1,2...M\}$ represent the counties with farm fields and

[8] Available from US Department of Agriculture (2014) .

[9] We work with the data from 2014 because this was the last year we were able to acquire data from ipmPIPE on soybean rust infestation. Therefore, we do not have information on the location and number of sentinel plots after 2014 that we could insightfully compare to our model's results after 2014.

[10] Soybean rust was first diagnosed in 2004 while our study is for 2014. Therefore, there are 10 observations on rust infection available to each farmer.

Table 2. Summary Statistics in 1,360 Counties: Soybean Prices, Soybean Yields, Soybean Acreage, and Farmers' Prior Beliefs about the Probability of Soybean Rust Infection

Variable	Average	Standard Dev	Max	Min
Price ($/bu)	10.096	0.356	11	9.37
Yield (bu/acre)	45.664	9.365	69.2	18.2
Acreage ('000 acres)	49.275	50.454	473.33	0.3
Prior Probability	0.096	0.21	0.99	0

$n \in \{1,2...N\}$ represent counties with sentinel plots. The value of monitoring for county m with a sentinel plot placed in county n, $v_{mn} = V_M(\alpha_0^m, \beta_0^m, \alpha_0^n, \beta_0^n, L) - V(\alpha_0^m, \beta_0^m)$, as well as the optimal management strategy is computed for every possible combination of counties. The spatial optimization problem is then solved using linear integer programming. The binary decision variable z_n equals 1 if county n is selected for a sentinel plot. Let x_{mn} be a binary decision variable equal to 1 if the sentinel plot in county n is assigned to farmers in county m. The spatial optimization problem can then be stated as:

$$\max_{x_{mn}, z_n \in \{0,1\}} \Sigma_{m=1}^{M} \Sigma_{n=1}^{N} v_{mn} x_{mn} \tag{18}$$

$$x_{mn} \leq z_n \quad \forall m,n$$

$$\Sigma_{n=1}^{N} x_{mn} = 1 \quad \forall m$$

$$\Sigma_{n=1}^{N} z_n \leq B.$$

The objective is to maximize the value of information across all M counties. The first constraint stipulates that the assignment of county n to county m can take place (i.e., $x_{mn} = 1$) only if county n is selected for a sentinel plot (i.e., $z_n = 1$). The second constraint requires that each county m with a farm field is assigned exactly one county n with a sentinel plot. The third constraint stipulates that the number of counties selected for sentinel plots must be less than the budget, B, where the budget is defined as the number of sentinel plots in the system.

Complete information (i.e., initial prior beliefs, yields, acreage, and prices) for 1,360 counties was available for analysis. Hence, in the spatial optimization problem, $M = N = 1{,}360$. The value of monitoring for each pair of these 1,360 counties leads to 1,360 × 1,360 values of monitoring (v_{mn}). An exact solution to the spatial optimization problem is solved using the branch and bound algorithm in MATLAB using its mixed-integer linear programming (MILP) solver (The MathWorks Inc, 2022). The net present value of aggregate expected profit without sentinel plots is subtracted from the optimum with sentinel plots to obtain the value of a spatially optimized sentinel plot network.

Results

First, equation 12 is solved for optimal strategies in each county in the absence of a sentinel plot network. Figure 3a shows that farmers choose different strategies in different parts of the United States, with the N strategy favored in the Upper Midwest and PP favored in the South. Then, equation 18 is solved with sentinel plots placed in each county, i.e., $B = 1360$. The total value of monitoring is $918 million, and Figure 3b shows how this value is distributed across counties in the United States. Note that when each county contains a sentinel plot, CP is always the optimal strategy. County-level soybean production is shown in Figure 4a and infection probabilities in Figure 4b. These affect how severe expected losses might be from a soybean rust infection and thus influence the optimal pest management strategy.[11,12] Expected losses can be moderate to high where production

[11] Revenue generated from production is also key. However, with prices varying at the state level, model results are driven more by variation in county-level production than price.

[12] The areas in white with no county borders are not included in this study.

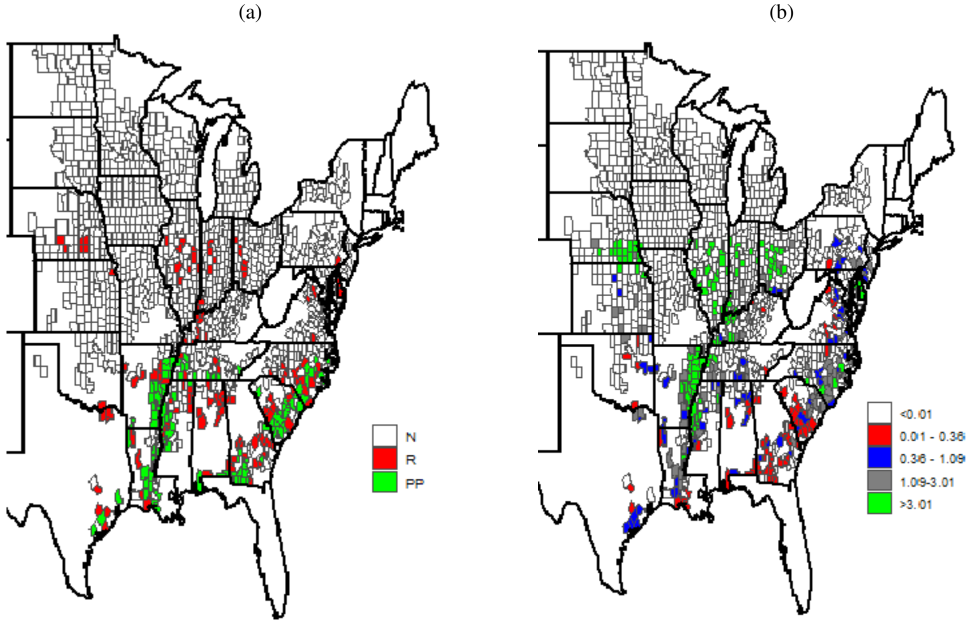

Figure 3. (a) Optimal Strategy in the Absence of a Monitoring Network. (b) Value of Monitoring ($ million) when Every County has a Sentinel Plot

Notes: N, *R* and *PP* stand for Do Nothing, Curative and Prophylactic Preventative strategies respectively.

is high even if the probability of soybean rust infection is low. Examples include counties in southern Nebraska, Illinois, central Indiana and western Ohio. Another situation with high expected losses include locations with high infestation probability and moderate production as occurs along the Mississippi-Arkansas border. In our data, expected losses are higher with higher probabilities of infection and moderate levels of production. Farmers in this region choose to prevent the severe expected loss in the absence of a monitoring system by using a preventative strategy. Farmers with a relatively low expected loss due to a low risk of infection will choose *R* over *PP*. Those who face either a negligible probability of infection or have very low soybean production will choose *N*. Correspondingly, farmers who choose *PP* in the absence of a monitoring network will benefit the most from its establishment while those who choose *N* will benefit the least. This is evident from Figure 3b.[13]

Next, equation 18 is re-solved for the optimal placement of sentinel plots when the number of sentinel plots, *B*, is constrained. Figure 5 shows the optimal placement of sentinel plots when *B* equals 185. Counties that benefited the most from having a sentinel plot (shown in Figure 3b) retain their sentinel plots after *B* is reduced to 185; a significant number of plots are allotted to the Mississippi-Arkansas border, Illinois, Ohio, Indiana, Southern Nebraska, Louisiana and North Carolina. Each of these regions are allocated more than 10 for a total of 115 out of the budgeted 185 sentinel plots. These counties use the costlier options, *PP* and *R*, in the absence of monitoring.

The next optimization is conducted to compare to the actual sentinel plot arrangement of 2014. In that year, almost all the 75 sentinel plots were clustered in the southern part of the southeastern region (e.g., Alabama, Florida, Georgia, and Louisiana) as shown in Figure 6a. Out of these 75, 39 are present among the 1,360 counties with data necessary for optimization. Therefore, the spatial optimization problem is solved for $B = 39$; its result is shown in the Figure 6b. Examining these

[13] The average production and probability of soybean rust infestation are 2.47M bu and 0.01 for farmers who choose *N* in the absence of a monitoring network; the corresponding values for farmers with *PP* is 1.98M bu and 0.39 and, for *R*, 2.23M bu and 0.68.

(a) (b)

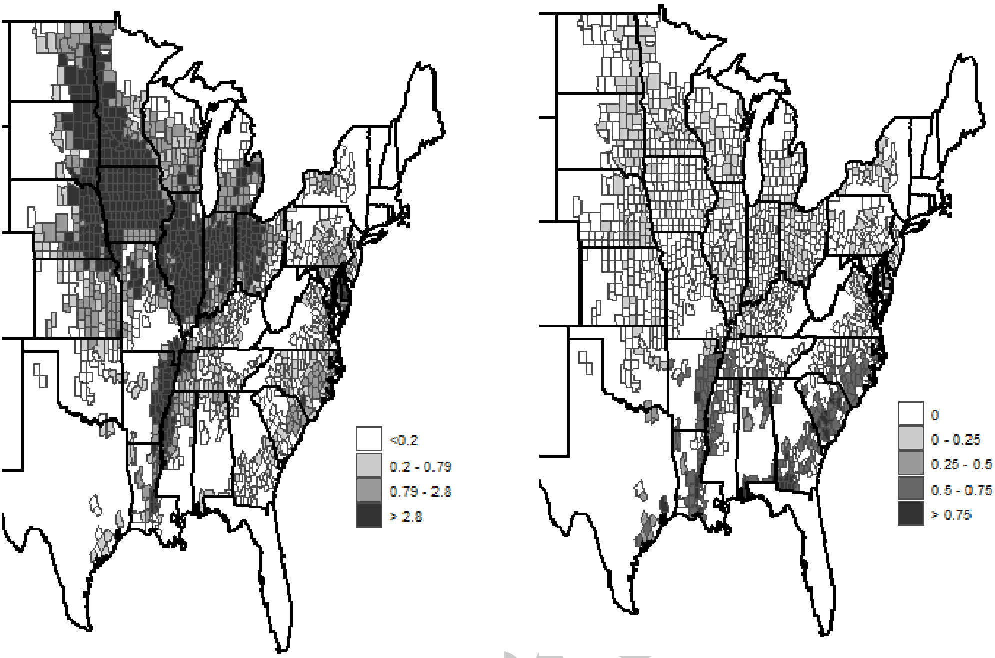

Figure 4. (a) Soybean Production (thousand bu), and (b) Prior Probability of Soybean Rust Infection

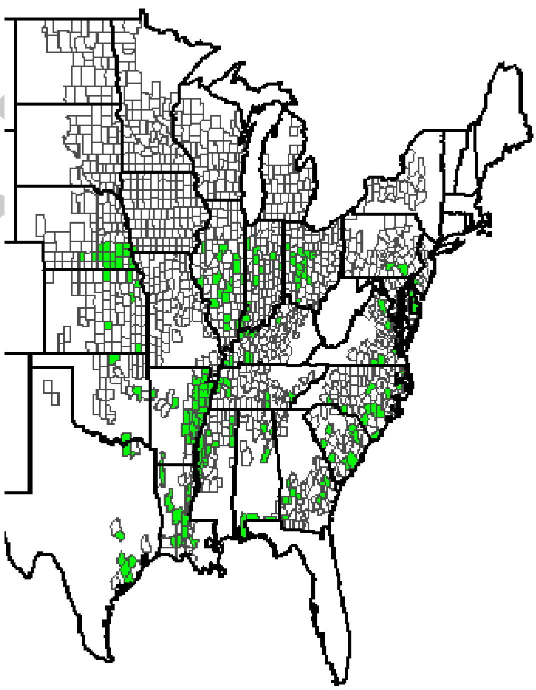

Figure 5. Optimal Sentinel Plot Arrangement when the Sentinel Plot Budget *B* equals 185

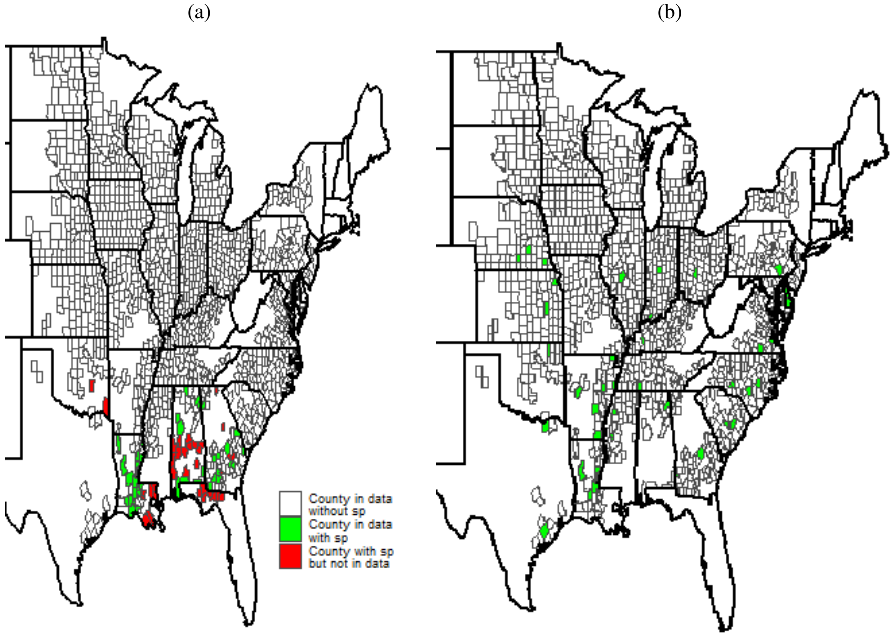

Figure 6. (a) Sentinel Plots (sp) in 2014, and (b) Optimal Sentinel Plot Placements when the Sentinel Plot Budget *B* is 39

figures confirms that the actual sentinel plot locations in 2014 do not match the counties identified by the spatial optimization model which places sentinel plots further north. The optimal allocation is more dispersed than the actual arrangement of 2014. Midwestern states are allotted 10 sentinel plots. Arkansas is allotted 6 while the Carolinas, Virginia, Delaware and Pennsylvania together are allotted 10 out of 39 sentinel plots. None of these states had any sentinel plots in the actual arrangement of 2014. Although the southern part of the southeast region has a comparatively high risk of infection, it has low soybean production and is also far away from the main production areas. As a result, the quality of the signal of infection coming from these far-off sentinel plots is poor for the main soybean-producing regions in the US. The actual placement in 2014 would have resulted in a total value of monitoring of, at most, \$114M. On the other hand, the value from an optimally designed monitoring network with 39 sentinel plots is \$908M, which is about eight times the value achieved with the actual set of plots. In fact, simply placing 39 sentinel plots in counties with the maximum soybean production would generate a minimum of \$117M assuming no spatial correlation in soybean rust infestation and \$456M with the baseline correlation model.

Figure 7 shows the optimal management strategies when *B* equals 185 and 39. For 1,012 counties, the expected losses from soybean rust are small enough, either due to low levels of production or low probabilities of infection or both, so that *N* is optimal whether *B* equals 185 or 39. *R* is no longer an optimal strategy for any county. When *B* equals 185, *CP* is optimal for 344 counties and *PP* is optimal for 4 counties. When *B* equals 39, *CP* is optimal for 336 counties and *PP* is optimal for 12 counties. These results show that the vast majority of farmers would switch from *R* and *PP* to *CP* as their optimal strategy in the presence of a spatially optimized monitoring network. They prefer to use signals from the monitoring network to make better pest management decisions. However, while these results confirm the benefits of a monitoring network to farmers, it should be remembered that these benefits incorporate the gains from the optimal spatial allocation of sentinel plots.

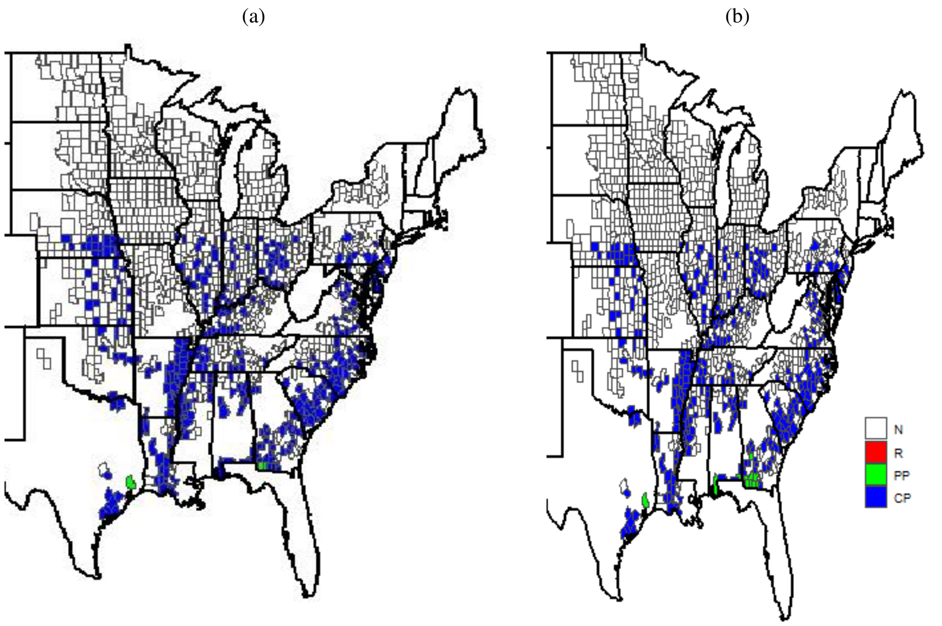

Figure 7. Optimal Actions when the Sentinel Plot Budget *B* equals (a) 185 and (b) 39.

Note N, *R*, *PP* and *CP* stand for Do Nothing, Curative, Prophylactic Preventative and Conditional Preventative strategies respectively.

Figures 6b and 7b also demonstrate how counties derive spillover benefits from nearby sentinel plots. Figure 7b shows that many counties around a sentinel plot county use *CP* as their optimal strategy. In fact, shrinking the sentinel plot budget hardly changes the optimal pest management strategy away from *CP* for majority counties. This means that sentinel plots are beneficial to the farmers in their own county and to farmers in other counties. Such gains are feasible in the model when the correlation between the farm field and the sentinel plot is high. This requires meeting two conditions: i) the counties should have similar infection probabilities, and ii) a large distance should not separate counties. Here, both these conditions are met. Not only are these counties neighbors, but they also have similar beliefs about the probability of rust infection. This makes the correlation sufficiently high for these counties to value the information from a neighboring sentinel plot.

The sensitivity of the value of the monitoring network to its size (i.e., number of counties with sentinel plots) is explored in Table 3. As *B* increases from 39 to 300, the total value of monitoring increases from \$0.78B to \$0.92B. Hardly any additional value is gained from adding sentinel plots beyond 75 —the marginal benefit of sentinel plots becomes negligible. Since the costs associated with the expansion and maintenance of the sentinel plot network are not available, the optimal number of plots is unknown. However, the results are still informative because they suggest where steep gains in benefits to additional sentinel plots can be found.

Sensitivity Analysis

Figure 8 shows the correlation functions used for sensitivity analyses and Table 4 gives parameter values for these functions. The ‘Benchmark’ function is the baseline used to generate the reported results. Based on the position relative to the Benchmark curve and whether they cross over, the functions are called ‘Right Shift’ and ‘Pivot.’ Table 5 compares the monitoring value (million dollars) derived from the spatial optimization of 39 sentinel plots (column I) to the value which would be realized from the actual 2014 arrangement of 39 sentinel plots (column II) in our data

Table 3. Total Value of Monitoring as a Function of the Number of Sentinel Plots

Number of Sentinel Plots	Value of Monitoring ($ M)
10	778.54
25	883.29
39	907.76
50	913.22
75	916.75
80	916.78
185	917.66
300	917.72
400	917.72

Table 4. Parameter Values Used in Sensitivity Analysis Scenarios

Correlation Scenario	θ	d^*	γ
Benchmark	0.075	0	−11.25
Pivot	0.035	0	−5.625
Right Shift	0.075	50	−11.25

Notes: The correlation, ρ, between infection in the farmer's field and infection in the sentinel plot is modeled as $UB/(1 + e^{(\theta(d-d^*)+\gamma)})$ for distances less than d^*. Relative to the benchmark, a decrease in θ and an increase in γ reduces the correlation within 150 miles and increases it farther away. This scenario is labeled "Pivot." Relative to the benchmark, increasing d^* extends the distance for which correlation is high. This scenario is labeled "Right Shift." These functions are illustrated in Figure 8.

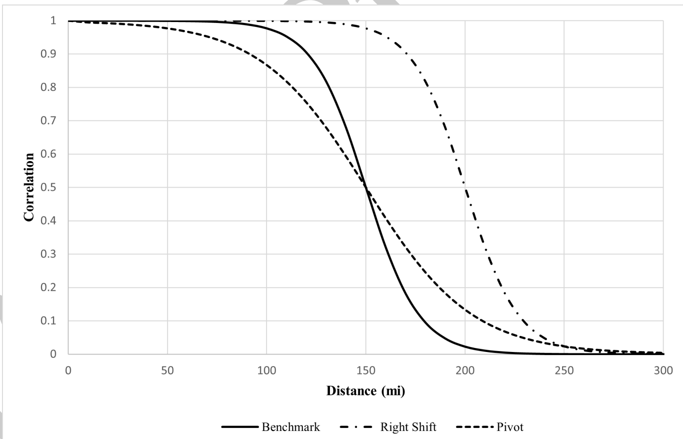

Figure 8. Shapes of the Correlation Functions used in the Sensitivity Analysis

for different correlation functions. Column II shows using the arrangement of the 39 actual sentinel plots in 2014 will capture the value of monitoring in the range of $92 to $188M which amounts to 10.5% to 20.7% (Column III) of the optimal. These values are much smaller than the total value that could be realized with the optimal spatial allocation of sentinel plots.[14]

[14] The data includes only 39 out of the 75 counties that had sentinel plots in 2014. Including the other 36 sentinel plot counties that are not in the data is not expected to change the spatial optimization results significantly. We can predict this based on our solution to the spatial optimization problem when B=80. The model still places sentinel plots relatively north and in a scattered fashion.

Table 5. Results of the Sensitivity Analysis for 39 Sentinel Plots

Correlation Scenario	Optimized Value of Monitoring ($ million)	Value Realized with Sentinel Plots Placed in 2014 ($ million)	Value Realized in 2014 as a Percentage of Optimal
Benchmark	907.76	114.33	12.6%
Pivot	877.19	92.17	10.5%
Right Shift	911.33	188.45	20.7%

Table 6. Moran's I for Optimal Spatial Arrangements of 39 Sentinel Plots with Different Correlation Functions and the Actual Arrangement in 2014

Scenario	Moran's I
Benchmark	1.00×10^{-6}
Pivot	9.80×10^{-7}
Right Shift	1.32×10^{-6}
Sentinel plot arrangement in 2014	1.04×10^{-5}

Table 6 reports the degree of spatial auto-correlation for a given spatial arrangement of sentinel plots using Moran's I.[15] This table makes two key points. First, the values of I across different correlation functions are very similar. Therefore, the spatial optimization model always disperses the sentinel plots irrespective of the underlying correlation function. Second, I is higher for the actual sentinel plot arrangement of 2014 by an order of magnitude. This means that sentinel plots were much more clustered in 2014 than suggested by the spatial optimization model because the weights are inverse of the distance between the counties (Figure 6).

Table 7 shows the percentage loss in the monitoring value when the underlying correlation function is misdiagnosed. Suppose the row depicts the actual correlation function and the correlation function in the column is wrongly assumed to be true and consequently used to determine the optimal spatial arrangement of sentinel plots. Then Table 7 gives the percentage loss in the value of monitoring from using a sub-optimal sentinel plot arrangement. For example, the second entry in the first row is 0.59; the interpretation is that there would be a 0.59% loss in the monitoring value due to a sub-optimal spatial arrangement of sentinel plots based on the wrongly identified correlation function 'Pivot' instead of the correct function 'Benchmark.' The diagonal elements correspond to the cases where the underlying correlation functions have been correctly used leading to no losses in the value of monitoring. The table clearly shows that the losses in monitoring value from misdiagnosing the underlying correlation function are modest. The greatest loss (of 2.37%) occurs when the underlying correlation function is assumed to be the 'Right Shift' when the actual correlation is 'Pivot.'

The sensitivity analyses show that, when maximizing farmers' benefits from a monitoring network, misdiagnosing the correlation function is not as big a mistake as designing the network sub-optimally. The policymakers should take into account the fact that most of the beneficiaries of the monitoring system, the soybean farmers, operate further north while rust infection is most likely in the regions close to the Gulf of Mexico. As a result, clustering sentinel plots in the south, far away from the primary soybean production region, reduces the accuracy of the information on soybean

[15] Global Moran's I, which is a measure of clustering of spatial data, is calculated using the following formula (Moran, 1948):

$$I = \frac{N}{W} \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} (z_i - \bar{z_i}) w_{ij} (z_j - \bar{z_j})}{\sum_{i=1}^{N} (z_j - \bar{z_j})^2}, \tag{19}$$

where z_i equals 1 if county i is assigned a sentinel plot and 0 otherwise. N is the total number of counties (N = 1360). $\bar{z}$ is the average of z_i and equals $\frac{B}{N}$. w_{ij} is the spatial weight assigned to pair of counties i and j. In our study, it equals the inverse of the distance between counties i and j when $i \neq j$, and 0 otherwise. W is the sum of weights ($W = \sum_i^N \sum_j^N w_{ij}$). Weights w_{ij} have been row normalized to sum to 1.

Table 7. Sensitivity Analysis Results for 39 Sentinel Plots: Percentage Loss in the Value of Monitoring with Incorrect Correlation Functions

	Benchmark	Pivot	Right Shift
Benchmark	0	0.59	0.40
Pivot	1.57	0	2.37
Right Shift	0.16	0.62	0

Notes: This table shows results when parameters for the scenarios in the columns are used to determine the optimal sentinel plot locations when the actual correlation functions are those in the scenarios in the rows.

rust infection and makes the monitoring system less beneficial to farmers. Consequently, the 2014 spatial allocation of sentinel plots was likely to have been sub-optimal.

Conclusion

Potentially catastrophic plant pathogens threaten to enter the US every year. To prevent damage from these pathogens, the government and other organizations can invest in sentinel plot monitoring networks to help predict and therefore control the risk of plant diseases. However, the net benefit of such monitoring networks is greatly debated because they are costly to maintain and, for many pathogens such as soybean rust, very few incidents of infection have been reported throughout the US mainland. In this paper, a dynamic model of farmer pest management decisions is developed and used to estimate the value of a sentinel plot monitoring network with an optimal spatial arrangement. This model could be used to optimally allocate sentinel plots, and to estimate the value of a monitoring network, for emerging pathogens or for a resurgence of soybean rust under new climatic conditions. The model could also be used to evaluate the relative efficiency of heuristics such as prioritizing counties with high levels of production or randomizing sentinel plot placement.

The paper uses the case of soybean rust and estimates the value of ipmPIPE, the sentinel plot monitoring network for soybean rust. Farmers benefit from the information provided by the monitoring network since it helps them manage soybean rust better within the growing season and refine their beliefs about the infection risk. They value information more when their expected loss (average soybean production times the risk of infection) from infection is high given the quality of information. The quality of information from a sentinel plot decreases when the sentinel plot is placed farther away from a farmer's field. As a result, when the sentinel plots are placed in the Deep South while soybean is mostly being grown in the North, the total welfare from monitoring will be lower than when some are placed further north. Current sentinel plots are disproportionately placed in the Southern US where the risk of infection is high, but the amount of soybean produced is low. Our estimates suggest that sentinel plots should be placed in the Corn Belt as well where the risk of infection is lower, but where more soybean is produced. The analysis suggests that such a modification in sentinel plot arrangement could have increased the value of monitoring in 2014 by over 800%.

It is important to note that there are other potentially valuable uses of monitoring plots that are not included in the farmer decision model. For example, local dispersal during the relevant stage of soybean growth occurs in the southern part of the United States and it may be possible to slow this spread meaningfully during the season through a choice to treat early in the season. If infections in sentinel plots prompt growers to respond quickly to infection, they may have an added benefit in the South where local dispersal is commonplace. In addition, monitoring plots in the South not only inform local soybean growers about management decisions but also inform aerobiology forecasts of soybean rust spread in the Corn Belt, which inform growers about when and where to scout for soybean rust infections. Further, monitoring plots for soybean rust in the South also provide information about the presence of other soybean diseases. Further work is needed to account for these additional uses and update the valuation and optimal location of monitoring plots.

The results also show that only about 75 sentinel plots are needed to maximize the value of ipmPIPE. There have been multiple years, especially in the initial years of its life, when ipmPIPE had many more than 75 sentinel plots. In recent years, there has been a substantial decrease in the number of sentinel plots. Most of the sentinel plots remain in the South. This fact points to a useful direction for further researchâĂŞ estimating the optimal number of sentinel plots in ipmPIPE and how it should change over time as farmers gain more knowledge of their infection risk. With data on the cost of maintenance and establishment of sentinel plots for each county, one can estimate the marginal cost of monitoring. The optimal number of sentinel plots can then be estimated by equating the marginal cost of monitoring to its marginal value.

[First submitted June 2022; accepted for publication February 2024.]

References

Bekkerman, A., B. K. Goodwin, and N. E. Piggott. 2008. "Spatio-Temporal Risk and Severity Analysis of Soybean Rust in the United States." *Journal of Agricultural and Resource Economics* 33(3):311–331. doi: 10.22004/ag.econ.46564.

Bekkerman, A., N. E. Piggott, B. K. Goodwin, and K. Jefferson-Moore. 2012. "A Market-Based Mitigation Program for Wind-Borne Diseases." *Agricultural and Resource Economics Review* 41(2):175âĂŞ188. doi: 10.1017/S1068280500003336.

Blackburn, T. M., P. Pyšek, S. Bacher, J. T. Carlton, R. P. Duncan, V. Jarošík, J. R. Wilson, and D. M. Richardson. 2011. "A Proposed Unified Framework for Biological Invasions." *Trends in Ecology & Evolution* 26(7):333–339. doi: 10.1016/j.tree.2011.03.023.

Büyüktahtakın, I. E., and R. G. Haight. 2018. "A Review of Operations Research Models in Invasive Species Management: State of the Art, Challenges, and Future Directions." *Annals of Operations Research* 271(2):357–403. doi: 10.1007/s10479-017-2670-5.

Chaloner, T. M., S. J. Gurr, and D. P. Bebber. 2021. "Plant Pathogen Infection Risk Tracks Global Crop Yields Under Climate Change." *Nature Climate Change* 11(8):710–715. doi: 10.1038/s41558-021-01104-8.

Epanchin-Niell, R. S., R. G. Haight, L. Berec, J. M. Kean, and A. M. Liebhold. 2012. "Optimal Surveillance and Eradication of Invasive Species in Heterogeneous Landscapes." *Ecology Letters* 15(8):803–812. doi: 10.1111/j.1461-0248.2012.01800.x.

Epanchin-Niell, R. S., and A. Hastings. 2010. "Controlling Established Invaders: Integrating Economics and Spread Dynamics to Determine Optimal Management." *Ecology letters* 13(4):528–541. doi: 10.1111/j.1461-0248.2010.01440.x.

Horie, T., R. G. Haight, F. R. Homans, and R. C. Venette. 2013. "Optimal Strategies for the Surveillance and Control of Forest Pathogens: A Case Study with Oak Wilt." *Ecological Economics* 86:78–85. doi: 10.1016/j.ecolecon.2012.09.017.

ipmPIPE. "ipm Pest Information Platform for Extension and Education." http://www.ipmpipe.org/. Accessed: Jan, 2016.

Isard, S. A., S. H. Gage, P. Comtois, and J. M. Russo. 2005. "Principles of the Atmospheric Pathway for Invasive Species Applied to Soybean Rust." *BioScience* 55(10):851–861. Publisher: American Institute of Biological Sciences. doi: 10.1641/0006-3568(2005)055[0851:POTAPF]2.0.CO;2.

Johansson, R. C., M. J. Livingston, J. Westra, and K. Guidry. 2006. "Simulating the US Impacts of Alternative Asian Soybean Rust Treatment Regimes." *Agricultural and Resource Economics Review* 35(1):116–127. doi: 10.1017/S1068280500010091.

Kıbış, E. Y., İ. E. Büyüktahtakın, R. G. Haight, N. Akhundov, K. Knight, and C. E. Flower. 2021. "A Multistage Stochastic Programming Approach to the Optimal Surveillance and Control of the Emerald Ash Borer in Cities." *INFORMS Journal on Computing* 33(2):808–834. doi: 10.1287/ijoc.2020.0963.

Kuchler, F., M. Duffy, R. D. Shrum, W. M. Dowler et al. 1984. “Potential Economic Consequences of the Entry of an Exotic Fungal Pest: The Case of Soybean Rust.” *Phytopathology* 74(8):916–920. doi: 10.1094/Phyto-74-916.

Livingston, M., R. Johansson, S. Daberkow, M. Roberts, M. Ash, and V. Breneman. 2004. *Economic and Policy Implications of Wind-Borne Entry of Asian Soybean Rust into the United States*. US Department of Agriculture Outlook Report No.(OCS04D02).

Livingston, M. J. 2010. “US Soybean Producer Perceptions and Management of Soybean Rust in the United States under the USDA Pest Information Platform for Extension and Education.” *Agricultural and Resource Economics Review* 39(3):547–560. doi: 10.1017/S1068280500007504.

Mehta, S. V., R. G. Haight, F. R. Homans, S. Polasky, and R. C. Venette. 2007. “Optimal Detection and Control Strategies for Invasive Species Management.” *Ecological Economics* 61(2-3):237–245. doi: 10.1016/j.ecolecon.2006.10.024.

Miller, S. A., F. D. Beed, and C. L. Harmon. 2009. “Plant Disease Diagnostic Capabilities and Networks.” *Annual Review of Phytopathology* 47:15–38. doi: 10.1146/annurev-phyto-080508-081743.

Moran, P. A. P. 1948. “The Interpretation of Statistical Maps.” *Journal of the Royal Statistical Society. Series B (Methodological)* 10(2):243–251. doi: 10.1111/j.2517-6161.1948.tb00012.x.

Roberts, M. J., D. Schimmelpfennig, M. J. Livingston, and E. Ashley. 2009. “Estimating the Value of an Early-Warning System.” *Review of Agricultural Economics* 31(2):303–329. doi: 10.1111/j.1467-9353.2009.01439.x.

Roberts, M. J., D. E. Schimmelpfennig, E. Ashley, M. J. Livingston, M. S. Ash, U. Vasavada et al. 2006. *The Value of Plant Disease Early-Warning Systems: A Case Study of USDA’s Soybean Rust Coordinated Framework*. United States Department of Agriculture, Economic Research Service.

Rossman, A. Y., K. Britton, D. Luster, M. Palm, M. H. Royer, and J. Sherald. 2006. “Evaluating the Threat Posed by Fungi on the APHIS List of Regulated Plant Pests.” *Plant Health Progress* 7(1):61. doi: 10.1094/PHP-2006-0505-01-PS.

Sikora, E. J., T. W. Allen, K. A. Wise, G. Bergstrom, C. A. Bradley, J. Bond, D. Brown-Rytlewski, M. Chilvers, J. Damicone, E. DeWolf, A. Dorrance, N. Dufault, P. Esker, T. R. Faske, L. Giesler, N. Goldberg, J. Golod, I. R. G. GÃşmez, C. Grau, A. Grybauskas, G. Franc, R. Hammerschmidt, G. L. Hartman, R. A. Henn, D. Hershman, C. Hollier, T. Isakeit, S. Isard, B. Jacobsen, D. Jardine, R. Kemerait, S. Koenning, M. Langham, D. Malvick, S. Markell, J. J. Marois, S. Monfort, D. Mueller, J. Mueller, R. Mulrooney, M. Newman, L. Osborne, G. B. Padgett, B. E. Ruden, J. Rupe, R. Schneider, H. Schwartz, G. Shaner, S. Singh, E. Stromberg, L. Sweets, A. Tenuta, S. Vaiciunas, X. B. Yang, H. Young-Kelly, and J. Zidek. 2014. “A Coordinated Effort to Manage Soybean Rust in North America: A Success Story in Soybean Disease Monitoring.” *Plant Disease* 98(7):864–875. doi: 10.1094/PDIS-02-14-0121-FE.

Stack, J., K. Cardwell, R. Hammerschmidt, J. Byrne, R. Loria, K. Snover-Clift, W. Baldwin, G. Wisler, H. Beck, R. Bostock, C. Thomas, and E. Luke. 2006. “The National Plant Diagnostic Network.” *Plant Disease* 90(2):128–136. doi: 10.1094/PD-90-0128.

The MathWorks Inc. 2022. “Mixed Integer Linear Programming.” https://www.mathworks.com/help/optim/ug/intlinprog.html.

US Department of Agriculture. 2014. “Quick Stats.” Washington, DC: US Department of Agriculture, National Agricultural Statistical Service. Available online on https:/www.nass.usda.gov/Quickstats/ [Accessed December 1, 2017].

Wrather, A., and S. Koenning. 2009. “Effects of Diseases on Soybean Yields in the United States 1996 to 2007.” *Plant Health Progress* 10(1). doi: 10.1094/PHP-2009-0401-01-RS.

Yang, X. B., and W. D. Batchelor. 1997. “Modeling Plant Disease Dynamics Using Neural Networks.” *AI Applications (USA)* 11(3):47–55.

Yemshanov, D., R. G. Haight, C. Chen, N. Liu, C. J. K. MacQuarrie, F. H. Koch, R. Venette, and K. Ryall. 2019. "Managing Biological Invasions in Urban Environments with the Acceptance Sampling Approach." *PLOS ONE* 14(8):e0220687. doi: 10.1371/journal.pone.0220687.

Appendix

Let i be defined as the occurrence of infection in the field and the probability of infection ($i = 1$) is ϕ^f. Similarly, j denotes infection in the sentinel plot, and ϕ^s is the probability that j=1. Hence

$$E(i) = \phi^f \tag{A1}$$

$$E(j) = \phi^s \tag{A2}$$

$$Var(i) = \phi^f(1-\phi^f) \tag{A3}$$

$$Var(j) = \phi^s(1-\phi^s) \tag{A4}$$

$$E(ij) = \phi_{11} \tag{A5}$$

Therefore, correlation ρ can be expressed as

$$\begin{aligned}\rho &= \frac{Cov(i,j)}{\sqrt{Var(i)Var(j)}} \\ &= \frac{E(ij)-E(i)E(j)}{\sqrt{Var(i)Var(j)}} \\ &= \frac{\phi_{11}-\phi^f\phi^s}{\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)}}\end{aligned}$$

$$\implies \phi_{11} = \phi^f\phi^s + \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)}. \tag{A6}$$

Also,

$$\phi_{11}+\phi_{10} = \phi^f \tag{A7}$$

$$\phi_{11}+\phi_{01} = \phi^s \tag{A8}$$

$$\phi_{00}+\phi_{11}+\phi_{10}+\phi_{01} = 1. \tag{A9}$$

Using the above equations, we get:

$$\phi_{01} = (1-\phi^f)\phi^s - \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} \tag{A10}$$

$$\phi_{10} = \phi^f(1-\phi^s) - \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} \tag{A11}$$

$$\phi_{00} = (1-\phi^f)(1-\phi^s) + \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)}. \tag{A12}$$

We know that $\phi_{00},\phi_{01},\phi_{10},\phi_{11} \in [0,1]$ as they are probabilities. If we ensure that all the probabilities are greater than 0, then equation A9 ensures that the probabilities will lie between 0 and 1.

$$\phi_{11} \geq 0$$

$$\implies \phi^f\phi^s + \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} \geq 0$$

$$\implies \frac{-\phi^f\phi^s}{\phi^f(1-\phi^f)\sqrt{\phi^s(1-\phi^s)}} \leq \rho \tag{A13}$$

$$\begin{aligned} &\phi_{00} \geq 0 \\ \implies &(1-\phi^f)(1-\phi^s) + \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} \geq 0 \\ \implies &\frac{-(1-\phi^f)(1-\phi^s)}{\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)}} \leq \rho \end{aligned} \tag{A14}$$

$$\begin{aligned} &\phi_{10} \geq 0 \\ \implies &\phi^f(1-\phi^s) + \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} \geq 0 \\ \implies &\sqrt{\frac{\phi^f(1-\phi^s)}{\phi^s(1-\phi^f)}} \geq \rho \end{aligned} \tag{A15}$$

$$\begin{aligned} &\phi_{01} \geq 0 \\ \implies &(1-\phi^f)\phi^s + \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} \geq 0 \\ \implies &\sqrt{\frac{\phi^s(1-\phi^f)}{\phi^f(1-\phi^s)}} \geq \rho \end{aligned} \tag{A16}$$

Since we assume that $\rho \geq 0$, equations A13 and A14 are automatically satisfied. Equations A15 and A16 along with the fact that $\rho \leq 1$ gives the upper bound (UB) on ρ,

$$UB = min\left\{\sqrt{\frac{\phi^s(1-\phi^f)}{\phi^f(1-\phi^s)}}, \sqrt{\frac{\phi^f(1-\phi^s)}{\phi^s(1-\phi^f)}}, 1\right\}.$$