Valuing Monitoring Networks for New Pathogens: The Case of Soybean Rust

Tiesta Thakur
Department of Applied Economics
University of Minnesota
St. Paul, MN 55108-6040
E-Mail: thaku035@umn.edu

Frances R. Homans
Department of Applied Economics
University of Minnesota

Robert G. Haight
U.S. Forest Service Northern Research Station

Terrance M. Hurley
Department of Applied Economics
University of Minnesota

Selected Paper
AAEA Annual Meeting
Washington D.C., August 5-7, 2018

Abstract

We estimate producers' value of information provided by alternative configurations of a monitoring network for an invasive wind-borne crop disease, soybean rust (*Phakopsora pachyrhizi*). Soybean producers use monitoring information to inform themselves about their risk of being infected and to make fungicide application decisions. The value of information is the expected gain in producers' profit associated with the use of the monitoring network, and we estimate this with a dynamic model of producer decision-making. We find that the value of the sentinel plot network increases with the number of sentinel plots, reaching a maximum at 400 sentinel plots. In addition, the optimal spatial arrangement of sentinel plots indicated by our model is substantially different from actual placements. Current sentinel plots are disproportionately placed in the Southern US where the risk of infection is high, but the amount of soybean is relatively small. Our estimates suggest more plots should be placed in the Corn Belt where the risk of an infection is lower, but where much more soybean is produced.

1 Introduction

In 2004, soybean rust (*Phakopsora pachyrhizi*) arrived in the United States (US), likely carried from South America by the winds of Hurricane Ivan (Isard et al. (2005)). Soybean rust is a fungal plant pathogen that cannot overwinter in temperate climates found in most of the US, but thrives throughout the year in the heat and humidity of the Gulf Coast, where sporulation in kudzu plants occurs throughout the year. Rust spores are very likely blown into the continental interior of North America throughout the year, although many more spores are transported in the late summer and early fall because of the high availability of inoculum in the South. When weather conditions are suitable, rust spores deposited on soybean fields can take hold and cause substantial yield loss. Kuchler et al. (1984) estimated that losses could be as high as \$7.1 billion per year if soybean rust becomes fully established, while Livingston et al. (2004) estimated that losses in the first year of infestation could run between \$640 million and \$1.3 billion.

Soybean rust is a recent example of a potentially catastrophic pathogen entering the US agricultural system and there are other plant diseases looming on the horizon. For example, the new races of wheat stem rust that recently emerged in Africa (Ug99) are of current concern, as are a number of pathogens that could affect corn and other crops (Rossman et al. (2006)). In the US, the National Plant 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 et al. (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, 2016), which includes a sentinel plot monitoring network. The sentinel plots in the monitoring network (more than 700 in 2007) are areas of early maturing soybeans grown specifically to detect rust. The ipmPIPE also included a web-based information technology system that provides farmers with direct access to information on confirmed cases of rust, options for management, and forecasts and expert commentary on disease progression throughout the growing season. A successor project, the integrated Pest Information Platform for Extension and Education (iPiPE) has integrated parts of the soybean rust detection program in a larger effort designed to detect a wider array of pathogens (http://ipmpipe.org/).

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 et al. (2009)) because farmers have limited incentives and ability to coordinate such efforts on their own. Still, the adverse effect of soybean rust on US production has been less than anticipated (Livingston (2010)), prompting questions about the value of the monitoring network as compared to its costs. Particularly relevant is the work of Roberts et al. (2009) which considers the case of the soybean rust network. They identify the value of the monitoring network as the benefit to farmers from making a better choice among three possible strategies, within a single planting season, due to a more accurate belief about the probability of infection. They find that the value of the network depends critically on the accuracy of the information and farmers' prior estimates of the probability of infection. For most reasonable parameter values, and from an *ex ante* perspective, they find that the value of the program exceeded its costs. In recent years, investment in the soybean rust platform has been scaled back to 87 sentinel plots in 75 counties in 2014 (USDA ipmPIPE restricted website).

In this paper, we focus on the value of a monitoring network such as the ipmPIPE program through modeling the optimal design of such a program. We build on the work of Roberts et al. (2006) and Roberts et al. (2009) and examine how the value of the network changes with additional sentinel plots included in the network and with re-allocations of the sentinel plots across space. Our model also allows us to ask how the value of the network may be sensitive to how long the system is in operation. We first characterize farmer behavior in the absence of a sentinel plot system. We then add a sentinel plot to the model so that farmers can use the sentinel plot to provide a within-season signal about the presence of the pathogen. Next, we model multiple years of the operation the network. Repeated observations of infection at a nearby sentinel plot allow farmers to learn about their vulnerability to infection even as they guard against it using a preventative fungicide. Then, we apply our model to US soybean acreage, starting with a set of county-specific prior probabilities of soybean rust infection. We then introduce the possibility of a sentinel plot located in each county, including the county with the farm field, and compute the resulting value of information of each farm field-sentinel plot pair. We impose limits on the number of sentinel plots and use a spatial optimization model to locate them across US counties to maximize the expected gain in producers' profit associated with the use of the monitoring network. These results yield the value of information for different scales of the program. Our complete model allows us to determine values of a sentinel plot network, both contemporaneous and over time, and to find out how to get the most out of a budget-constrained program. Our model can thus be used to assess the value of the current program as well as suggest changes to the current program that might improve efficiency.

2 Farmer Decision Model

In this section, we develop a farm-level decision model first in the absence of a sentinel plot network. We then introduce how a neighboring sentinel plot affects within-season decisions. Finally, we consider learning in a multi-season model.

2.1 Management without a Sentinel Plot Network

Following Roberts et al. (2009), we assume a farmer faces two possible states of the world, represented by i: there is (i=1) or there is not (i=0) a soybean rust infection in her farm field. Also like Roberts et al., we assume that a farmer can respond to the threat of a soybean rust infection with one of three management options: do nothing about the infection, 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 when a farmer does not experience a soybean rust infection and C be production cost exclusive of any effort to control soybean rust. Let λ_n be the proportion of yield lost in the event of an uncontrolled soybean rust infection. The returns to doing nothing about soybean rust (strategy N) can then be written as $\pi_{N0} = PY–C$ without an infection and $\pi_{N1} = PY(1 - \lambda_n)–C$ with one.

One strategy is to scout and apply a curative fungicide as needed. We denote this strategy as the curative strategy, labeled R. With this strategy, a farmer only applies fungicide when scouting reveals a soybean rust infection. It is costly to scout, and these costs must be incurred regardless of whether or not soybean rust is detected. If soybean rust is detected, then it is also costly to apply a curative fungicide. Furthermore, by the time soybean rust is detected, some damage to the crop will have already occurred. With $C_{sc} > 0$ equal to the cost of scouting, $C_c > 0$ equal to the cost of a curative fungicide, and λ_r equal to the proportion of yield loss with a curative fungicide in the event of an infection, the returns to this strategy can be written as $\pi_{R0} = PY - C - C_{sc}$ without an infection and $\pi_{R1} = PY(1-\lambda_r) - C - C_{sc} - C_c$ with one. Note that, while a curative treatment will not give perfect control, losses from the curative strategy will be less than losses from doing nothing in the event of an infection: $\lambda_r < \lambda_n$.

Choosing a prophylactic preventative strategy, labeled PP, instead of scouting or doing nothing about soybean rust reduces yield loss further because rust has little or no chance of damaging the crop. However, to be successful, a preventative fungicide must be applied before an infection occurs. Therefore, there is the possibility of applying a fungicide even when no infection would have occurred otherwise. Furthermore, the cost of a preventative fungicide treatment, C_p, is generally greater than the cost of a curative treatment: $C_p > C_c$. In the event of an infection, yield losses with a preventative treatment tend to be negligible (Johansson et al. (2006)). Thus, the returns to the prophylactic preventative strategy are written as $\pi_{PP0} = \pi_{PP1} = PY–C–C_p$ both with and without an infection.

The farmer's decision depends on the expected returns of the three strategies. If there is no chance of an infection, a farmer should clearly choose the "do nothing" strategy, N. If an infection is sure to happen, 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, the farmer's belief about the probability of experiencing an infection, ϕ^f, is crucial for choosing the best strategy. An important source of information for farmers is personal experience, which can be influenced by management choices. Doing nothing or scouting allows them to observe, first-hand, any infection in their fields. These strategies provides them with an opportunity to better understand the underlying risk of infection. However, with a preventative fungicide, the farmers forgo this learning opportunity because treatment prohibits the emergence of the disease if the inoculum reaches the crop.

2.2 Within-Season Decision-making with a Sentinel Plot Network

In Roberts et al. (2006) and Roberts et al. (2009), the monitoring network sends one of two signals to a farmer: a high or low risk of infection. Farmers use the signal to update their beliefs about the probability of infection during the season. Further, 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.

In our model, we assume that the signal a grower receives is whether or not an infection has been confirmed at the relevant sentinel plot. This signal is more precisely defined than the high or low risk signal used by Roberts et al. (2009). Signal quality is expressed as the correlation between infection at the sentinel plot and infection in the farmer's field, and this correlation diminishes with distance. Second, we change the timing of decision making so that the farmer chooses her strategy based on her prior probability of infection before the signal arrives. Third, we add a fourth distinct management option: apply a conditional preventative fungicide mid-season if an infection is confirmed at the relevant sentinel plot during the season. We add this second type of preventative treatment to better characterize observed farmer behavior (Hershman). The Conditional Preventative strategy (CP) is defined as a strategy in which the farmer bases the decision of whether or not to apply preventative fungicide on whether or not the relevant sentinel plot becomes infected. If the ipmPIPE signals infection in the relevant sentinel plot, a fungicide is applied. Otherwise, no fungicide is applied. Note that by definition, this strategy is available only when the ipmPIPE and its monitoring network exists.

Given the ipmPIPE system, a farmer faces one of the following four situations: (1) there is an infection both in the farmer's field and in the sentinel plot, (2) there is no infection in both locations, (3) there is an infection in the field and not in the sentinel plot and (4) there is no infection in field and there is an infection in the sentinel plot. Let j represent the occurrence of soybean rust 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. Thus, the the expected returns to the Conditional Preventative strategy are $E(\pi_{CP}) = \phi_{11}\pi_{PP} + \phi_{10}\pi_{N1} + \phi_{01}\pi_{PP} + \phi_{00}\pi_{N0}$. Note that there is some probability (ϕ_{10}) that an infection will occur while the field is unprotected because no infection is detected at the sentinel plot. Table 1 summarizes these scenarios and their corresponding probabilities.

The ipmPIPE provides farmers with information about the infection status of sentinel plots in the network. The extent to which this information helps farmers depends on how correlated the likelihood of infection in the farmer's field is to that of the sentinel plot. This correlation is expected to be decreasing in the distance between field and sentinel plot: a sentinel plot close to the farmer's field is likely to provide more useful information about the risk of infection in the field than one which is far away. However, nearness alone does not guarantee high correlation. Correlation also depends on the similarity in probability of infection between the field and sentinel plot. If a sentinel plot has conditions favoring the growth of soybean rust while a nearby field has hostile conditions for rust growth, then an infection in the sentinel plot does not necessarily imply a high risk of infection in the field. Thus, information from a sentinel plot is more useful to a farmer in revising beliefs if the sentinel plot faces a similar risk of infection as the farmer's field.

Let the subjective probability of infection in a sentinel plot be denoted by ϕ^s. We assume that beliefs about the probability of infection in any county are common knowledge. Let ρ be the correlation between the occurrence of infection in the farmer's field and sentinel plot. Then, each ϕ_{ij} can be expressed as a function of ρ, ϕ^f and ϕ^s in the following manner:[1]

$$\begin{aligned}
\phi_{11} &= \phi^f \phi^s + \rho\sqrt{\phi^f \phi^s (1-\phi^s)(1-\phi^f)} && (1)\\
\phi_{01} &= \phi^s (1-\phi^f) - \rho\sqrt{\phi^f \phi^s (1-\phi^s)(1-\phi^f)} && (2)\\
\phi_{10} &= \phi^f (1-\phi^s) - \rho\sqrt{\phi^f \phi^s (1-\phi^s)(1-\phi^f)} && (3)\\
\phi_{00} &= (1-\phi^s)(1-\phi^f) + \rho\sqrt{\phi^f \phi^s (1-\phi^s)(1-\phi^f)}. && (4)
\end{aligned}$$

We know that $\phi_{11}, \phi_{10}, \phi_{01}, \phi_{00} \geq 0$. Therefore,

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

We expect the correlation to be a decreasing function of distance *'d'* between the field and the sentinel plot. Although this sounds intuitive, there are no formal estimates of the relationship between correlation and distance. Therefore, we proceed by assuming the correlation to be inversely related to distance using the form $\frac{1}{1+e^{ad+b}}$, where a and b are parameters which will be defined shortly. Also, ρ must be less than the R.H.S of equation 5 (UB). Hence, we scale the correlation function to ensure that ρ cannot be greater than UB:

$$\rho = \frac{UB}{1+e^{ad+b}}. \tag{6}$$

We assume $\rho \in [0, 1]$ because a negative correlation would imply that a signal from the sentinel plot is systematically wrong. Instead, infections in far-distant sentinel plots would be uncorrelated with

[1] Derivation for the following equations are shown in Appendix 1

infections in the farm field.

2.3 Learning in a multi-season model

Until now, we discussed the farmer's decision making during a single growing season. However, their decision making capacity is expected to change with their personal experience because as their experience increases so does their understanding of their risks of rust infection. Their experience, in turn, is influenced by management choices. Doing nothing or scouting allows them to observe, first-hand, any infection in their fields. Hence, these two strategies provide farmers with an opportunity to better understand the underlying risk of infection. However, with a preventative fungicide, the farmers forgo this learning opportunity because treatment prohibits the emergence of the disease if the inoculum reaches the crop.

The ipmPIPE and its monitoring network provide an additional opportunity for a farmer to learn about the risk of infection in their field, irrespective of their chosen strategy. Because of the ipmPIPE system, the farmer knows if 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 the risk of infection in her field. The usefulness of this new information will depend on the correlation between the field and sentinel plot as outlined below. In essence, the ipmPIPE system provides many of the benefits of scouting without cost and allows a farmer to delay or completely forgo using a preventative treatment unless monitoring information suggests an infection has become likely (namely, if infections are identified at sentinel plots close to the farmer's field).

We now explain how our model captures a farmer's learning process about her probability of infection. We solve a multi-season decision problem wherein a farmer maximizes the expected utility of lifetime profits by choosing the best strategy for soybean rust control given beliefs about the risk of infection and the availability of the ipmPIPE system. Let growing seasons be denoted by t. Suppose the probability that Y infections are observed in Z years in a region h follows a binomial distribution with parameter y_t^h for year t, where h can either denote the farmer's field or the sentinel plot. Let the farmer's prior belief about the probability of a rust infection in region h be characterized by the beta distribution with shape parameters α_t and β_t: $b(y_t^h; \alpha_t^h, \beta_t^h)$ with α_t^h as the number of times an infection is observed and β_t^h as the number of times a year passes without an infection in location h.[2] Then the farmer's state of the world in time period t can be defined by:

- $\{\alpha_t^f, \beta_t^f\}$ without ipmPIPE.

[2]The beta distribution is an ideal way to describe a farmer's belief because it is constrained to the unit interval and is

- $\{\alpha_t^f, \beta_t^f, \alpha_t^s, \beta_t^s\}$ with ipmPIPE.

In the presence of the ipmPIPE and its monitoring network, the farmer has knowledge about the risk of infection in the sentinel plot in addition to knowledge about the risk in the farm field. Using the beta distribution b to characterize prior beliefs also implies that the posterior distribution $b(y_{t+1}^h; \alpha_{t+1}^h, \beta_{t+1}^h)$ equals $b(y_{t+1}^h; \alpha_t^h + 1, \beta_t^h)$ when the farmer observes an infection and $b(y_{t+1}^h; \alpha_t^h, \beta_t^h + 1)$ when no infection is observed.

The farmer's knowledge about her state of the world in $t+1$, or in other words her learning, depend on two factors: (i) the availability of the monitoring network and (ii) her optimal strategy for tackling the risk of infection in her field, given her beliefs.

- When the farmer chooses to scout (R) or do nothing about soybean rust (N), she gets first-hand information on whether or not her farm became infected. This is true even in absence of monitoring network. Thus, α and β are updated in $t+1$ when the strategy is either R or N, 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 field. But, with the monitoring network in place, she can refine her beliefs about the risk of infection in her field based on the information about infections in the relevant sentinel plot.

- Updating beliefs while choosing Conditional Preventative CP as the optimal strategy works in the same way as in case of PP. The only difference is that this strategy is available only when the monitoring network is present.

The knowledge about the state of the world in $t+1$ translates into an updated belief about the probability of infection in the farm field. This is because the expectation of the beta distribution $b(y_t^h; \alpha_t^h, \beta_t^h)$ is

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

and ϕ_t^f equals $E(y_t^f)$ and ϕ_t^s equals $E(y_t^s)$ by definition.[3] Hence, information on $\{\alpha_{t+1}^f, \beta_{t+1}^f\}$ and $\{\alpha_{t+1}^s, \beta_{t+1}^s\}$ in time period t characterizes the complete set of farmer's beliefs $\{\phi_{t+1}^f, \phi_{t+1}^s, \phi_{ijt+1}, \rho_{t+1}\}$ for the next time period.

a conjugate prior of the binomial distribution.

[3]Since experiences with soybean rust are not the same across seasons, ϕ^f, ϕ^s, the ϕ_{ij}s and ρ are time dependent and are therefore denoted as ϕ_t^f, ϕ_t^s, ϕ_{ijt} and ρ_t respectively.

The farmer's expected return for the four possible strategies can thus be written as follows:

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

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

$$\begin{aligned} E(\pi_{PP_t}) &= \frac{\alpha_t^f}{\alpha_t^f+\beta_t^f}\pi_{PP1} + \frac{\beta_t^f}{\alpha_t^f+\beta_t^f}\pi_{PP0} \\ &= \pi_{PP0} \end{aligned} \tag{10}$$

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

3 Dynamic Optimization

We now formalize our soybean rust problem using dynamic programming, where the state variables are the parameters of a beta distribution characterizing a farmer's belief about the probability of an infection. We use our model to learn how the optimal expected profit for a farmer improves in the presence of the ipmPIPE system. The value of monitoring is defined as this increase in farmer's expected profit due to the availability of the monitoring network. To compute the value of monitoring, we solve the dynamic programming problem both with and without the monitoring network.

3.1 Optimization without sentinel plot network

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 "do nothing" strategy N initially and the optimal strategy thereafter. Since choosing not to apply fungicide (doing nothing) 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 then will update their priors to $\{\alpha_t^f, \beta_t^f + 1\}$. The second argument differs from the first only because the farmer chooses the curative strategy 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 the prophylactic preventative strategy (PP). This strategy does not provide any new information. Therefore, α^f and β^f remain the same in the next time period.

3.2 Optimization with sentinel plot network

Next, we find the farmer's optimal value function when the monitoring network is available. We assume that the network is available only for L years. The sentinel plots allow farmers to gather evidence about their risk of infection regardless of the management strategy they choose. In addition, the ipmPIPE and its sentinel plots provide a within-season signal upon which farmers can condition preventative treatments. Therefore, the optimal value function with the monitoring network (subscripted M) can be written as

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

where

$$\begin{aligned}
\mathrm{E}(V_{M_{t+1}}(\alpha^f_{t+1},\beta^f_{t+1},\alpha^s_{t+1},\beta^s_{t+1})) &= \phi_{11t}V_{M_{t+1}}(\alpha^f_t+1,\beta^f_t,\alpha^s_t+1,\beta^s_t)+\phi_{00t}V_{M_{t+1}}(\alpha^f_t,\beta^f_t+1,\alpha^s_t,\beta^s_t+1)\\
&+ \phi_{10t}V_{M_{t+1}}(\alpha^f_t+1,\beta^f_t,\alpha^s_t,\beta^s_t+1)+\phi_{01t}V_{M_{t+1}}(\alpha^f_t,\beta^f_t+1,\alpha^s_t+1,\beta^s_t)
\end{aligned} \tag{15}$$

$$\begin{aligned}
\mathrm{E}(V'_{M_{t+1}}(\alpha^f_{t+1},\beta^f_{t+1},\alpha^s_{t+1},\beta^s_{t+1})) &= \phi^s_t\Big[\frac{\phi_{11t}}{\phi^s_t}V_{M_{t+1}}(\alpha^f_t+1,\beta^f_t,\alpha^s_t+1,\beta^s_t)\\
&+ \frac{\phi_{01t}}{\phi^s_t}V_{M_{t+1}}(\alpha^f_t,\beta^f_t+1,\alpha^s_t+1,\beta^s_t)\Big]\\
&+ (1-\phi^s_t)\Big[\frac{\phi_{00t}}{1-\phi^s_t}V_{M_{t+1}}(\alpha^f_t,\beta^f_t+1,\alpha^s_t,\beta^s_t+1)\\
&+ \frac{\phi_{10t}}{1-\phi^s_t}V_{M_{t+1}}(\alpha^f_t+1,\beta^f_t,\alpha^s_t,\beta^s_t+1)\Big]
\end{aligned} \tag{16}$$

are the expected value functions when the chosen strategy is either N or R (equation (15)), and when it is either CP or PP (equation (16)). If we look at equations (15) and (16) more closely, we find that they work out to be the same. However, they have different interpretations. At time period t when the farmer is choosing her optimal strategy, she knows that she will see the exact state she enters in $t+1$ if she chose R or N. But if she chose CP or PP, she will not directly observe her state of the world in $t+1$ as these measures would prevent the emergence of infection in time period t altogether. Instead now she predicts the state which she enters based on the information from sentinel plots. As a result, while equation (15) captures the actual expected value of the farmer, equation 16 is only an expectation of the expected value. The sentinel plots also provide an within-season signal to condition preventative treatments yielding an additional opportunity in the maximum of equation (14) that did not appear in equation (12). This strategy has the contemporaneous expected return of $E(\pi_{CP_t})$. Once the sentinel plots are no longer available (when $t \geq L$), the farmer's problem becomes identical to the problem in equations (12) and (13).

4 Farm Model Implications

Having these two models of decision making allows us to calculate the value to farmers of the monitoring provided by sentinel plots. Recall that 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 equations (12-13) or (14-16) 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_c, 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). The values used for λ_r, λ_n and δ were adopted from Roberts et al. (2006) and Roberts et al. (2009) and are summarized in Table 2. Their values for C_{sc}, C_c and C_p were for 2005 and are adjusted to account for inflation (Table 2). We assume ρ to be very high (≈ 1) at distances less than 80-100 mi. We also assume that the correlation drops after this point and that its value approaches 0 when the distance reaches 200 miles. Given these assumptions, and if we suppose correlation to be 0.01 at 150 mi, then parameters a and b from equation (6) take on the values 0.12 and -13.4, respectively.

Before turning to the primary questions of interest regarding the overall value of the sentinel plot monitoring network, we explore some of the basic implications of the farmer decision model. We solve the dynamic programming model for an infinite planning horizon using the parameter values in Table 1 and an assumed yield of 37 bushels/acre. Figure 3 shows the first four years of optimal strategies at each time period given the farmer's beliefs about the probability of infection. 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 the preventative strategy 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 the no-action strategy N. For the majority of states in this example, it is optimal for the farmer to choose the curative strategy, R.

Figure 4 shows how optimal strategies change in presence of the monitoring network. Suppose beliefs about the probability of rust infection are 0.58 for both the farmer's field and the sentinel plot ($\alpha_0^f = \alpha_0^s = 5.8$, $\beta_0^f = \beta_0^s = 4.2$). According to our model, when the sentinel plots are unavailable, the farmer chooses prophylactic preventative PP as the optimal strategy. When the sentinel plot is present in the same county as the farmer's field, i.e., when $d = 0$ and $\rho = 1$, she will use the information from the monitoring network since it will tell her whether her field will get infected or not. Therefore, whenever the sentinel plot is in the same county as the farmer's field, conditional preventative CP is the farmer's optimal strategy. However, the monitoring network becomes less informative with increase in distance d. This is because an increase in d decreases ρ and consequently the value of information from the monitoring network. Up to $d < 100$ mi, the model predicts that the farmer will continue with the conditional preventative strategy. Once the sentinel plot is placed at a distance $d >\approx 100$ mi., our model predicts that the farmer will no longer rely on the information from monitoring network and will revert back to the optimal strategy without a monitoring network, *PP*.

It should be noted that the threshold where the optimal strategy shifts from conditional preventative to prophylactic preventative depends on the shape of the underlying correlation function. Suppose there is no monitoring available and the belief about the probability of infection in the field (ϕ^f) is equal to 0.1. With this low probability, the farmer's optimal strategy is to do nothing about soybean rust. 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 therefore have a very low correlation ρ irrespective of the distance d between them (Figure 1). As a result, even if the sentinel plot is placed close to the field, say at d equal 5 miles, the sentinel plot does not provide valuable information about the likelihood of infection in the farmer's field. Hence, the farmer chooses to continue doing nothing about soybean rust even in the presence of the monitoring network.

5 Spatial Optimization Model

We implement our farm-level model assuming a representative farmer in each county, using the 2014 soybean acreage and yield reported for each county by USDA. The state-level data on soybean prices are also from USDA-NASS for 2014. We use county level estimates of the probability of a soybean rust infection from Bekkerman et al. (2008) to represent prior beliefs about the likelihood of a rust infection for the representative farmer in each county. To transform these estimates into valid parameters for a beta distribution, we calculate α_0^h and β_0^h such that $\frac{\alpha_0^h}{\alpha_0^h + \beta_0^h}$ equals the probability estimate and $\alpha_0^h + \beta_0^h = 10$, which ensures that all priors are based on the same number of years of potential rust infection observations prior to our evaluation.

Given yields, prices, costs and farmers' initial beliefs, we now allocate a fixed number of sentinel plots to the counties to maximize the value of information generated from them. We formulate this spatial optimization problem using a linear integer programming model. Let m $(1, 2...M)$ represent the counties with farm fields and n $(1, 2...N)$ represent counties with sentinel plots. Then, the value of monitoring for county m when sentinel plot is placed in county n is: $V_M(\alpha_0^m, \beta_0^m, \alpha_0^n, \beta_0^n, L) - V(\alpha_0^m, \beta_0^m) = v_{mn}$. We define z_n as a binary decision variable equal to 1 if county n is selected for a sentinel plot. Let x_{mn} be a binary decision variable equal to 1 if sentinel plot in county n is assigned to farmers in county m. Then the spatial optimization problem can be stated as:

$$\max_{x_{mn}, z_n \in \{0,1\}} \Sigma_{m=1}^{M} \Sigma_{n=1}^{N} v_{mn} x_{mn}$$
$$x_{mn} \leq z_n \quad \forall m, n$$
$$\Sigma_{n=1}^{N} x_{mn} = 1 \quad \forall m$$
$$\Sigma_{n=1}^{N} z_n \leq B.$$

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 one sentinel plot is assigned to each county. 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.

We identify 1360 counties for which we have all information required: initial beliefs, yields, acreage and prices for 2014. Hence, in our spatial optimization problem, $M = N = 1360$. We compute the value of monitoring for each pair of these 1360 counties, leading to 1360×1360 values of monitoring (v_{mn}). An exact solution to this spatial optimization problem is then solved using the branch and bound algorithm in GAMS with the Cplex solver (GAMS (2015)).

6 Results

In our model, factors such as the probability of infection and soybean production (yield and acreage) play important roles in the farmer's decision making process. Based on these factors as shown in Figures 5 and 6, we can predict the optimal strategy of farmers both with and without the monitoring network.[4] We start by studying figures 5, 6(i) and 6(ii) in order to understand the intuition behind our solution before moving on to discussing the solution itself.

When there is no monitoring network, farmers can choose either N, R, or PP to control rust infection in

[4]The areas in white are not included in this study.

their fields. Doing nothing about soybean rust (N) leads to maximum profit in the case of no infection and minimum profit when infection occurs. Therefore, it will be the most suitable for farmers who have a very low chance of rust infection. On the other hand, PP will be optimal for those who expect severe losses from soybean rust. This will generally be true for farmers with a very high risk of infection. But, when a farmer faces a high probability of infection along with comparatively low soybean yields, she may find that the costs of a preventative strategy outweigh its benefits. In this case, she will choose scouting (R) over the preventative strategy.[5]

We see that, while soybean in mostly grown in the Midwest (Figure 6), the probability of rust infection in the Midwest is typically less than 0.02. Hence, in Figure 9(i), we see that most of the northern states choose the "do nothing" strategy, N. The Southeast is generally characterized by a high probability of rust infection owing to its proximity to Gulf of Mexico. That is why we observe the choice of prophylactic treatment PP in Arkansas, Mississippi, Louisiana and coastal counties. The counties in which farmers choose preventative measures have an average probability of infection of 0.68 and average yield of 47.21 bu/acre in comparison to the corresponding values being 0.01 and 45.72 bu/acre for those who optimally do nothing about rust infection. The regions where scouting R is the optimal strategy are more moderate in nature. They face a risk of infection (0.39) which is high enough to abandon the N strategy but lower than the risk in areas using prophylactic treatment. These regions also have a comparatively low soybean yield (44.04 bu/acre). This low yield, when combined with moderate/high risk of infection, does not make the expected loss from a rust infection high enough to justify the extra cost of prophylactic treatment over scouting. For example, farmers in counties in northern Texas and Alabama and in northwestern South Carolina, which have comparatively higher probability of infection but low soybean yield, choose to scout.

Figures 5 and 6 also help us understand how the value of monitoring varies spatially across counties. The value of monitoring should increase with the total expected loss from rust infection and therefore should be increasing in the probability of rust infection, soybean acreage and soybean yield. Based on this logic, we can make a few easy predictions: first, the value of monitoring will not be large for states like Minnesota, Wisconsin, North and South Dakota owing to their very low likelihoods of infection. Second, the value of monitoring is not expected to be very big for most of the counties in southeastern states such as Alabama, Georgia and South Carolina which are characterized by low soybean acreage and yield. Third, the counties on the border of Arkansas and Mississippi should be among the major beneficiaries of

[5] The optimal strategy of a farmer is also a function of the soybean price she faces. Since we do not find much variance in price across counties, we focus on the latter factors when discussing the optimal strategy of farmers.

the monitoring network because they have high soybean acreages and yields and they experience a high likelihood of rust infection.

Figure 7 shows that our results match our intuition. This figure displays the dollar value of monitoring when all 1360 counties have sentinel plots. While the northern and southeastern states do not gain much from monitoring in general, Arkansas and Mississippi are among those who profit the most from sentinel plots. In this region, most of the counties have soybean grown on at least 60,000 acres while facing a risk of rust infection of at least 50%. Figure 7 also shows a high value of monitoring for many counties in the states of Nebraska, Illinois, Indiana and Ohio that do not have high probabilities of rust infection but are characterized by high enough yields and acreage to result in large expected losses from soybean rust.

Figure 7 can be interpreted as the solution to the spatial optimization problem when B takes its maximum value, i.e., when $B = 1360$. We now study the optimal placement of sentinel plots when the number of sentinel plots, B, decreases. Figure 8 shows how sentinel plots are allocated among counties to maximize the total value of monitoring in US when B equals 185 and 80 respectively. When B equals 185 (Figure 8(i)), sentinel plots are allocated to states with the highest values of monitoring such as Arkansas, Mississippi, Illinois and Ohio, as well as to southern Nebraska. Reducing the number of plots to 80 causes Nebraska, Ohio and Illinois to lose many of its sentinel plots. These regions, although characterized by high soybean production both in terms of yield and acreage, have comparatively low risks of rust infection (at most 25%). As a result, the expected loss from soybean rust is not big enough in these places to preserve the initial allotment of sentinel plots when B shrinks from 185 to 80. A big cluster of sentinel plots remains in Arkansas and Mississippi (Figure 8(ii)).

Figure 9 shows optimal management strategies with and without the monitoring network. There are two facts which are clear when we compare 9(i) and 9(ii). First, in presence of a spatially optimized monitoring network, most farmers switch to using Conditional Preventative as their optimal strategy.[6] Second, with optimally allocated sentinel plots farmers no longer use Curative treatments. Both these results clearly point out how the farmers prefer to use signals from monitoring network in order to make better pest management decisions. However, while these points do depict the benefits of a monitoring network to farmers, it should be remembered that these benefits also incorporate the gains from optimal spatial allocation of sentinel plots.

Figure 9 also demonstrates that counties derive spillover gains from nearby sentinel plots. In Figure 8(ii), 3 sentinel plots have been allocated close to the state of Nebraska. In Figure 9(ii), we see that this

[6] In fact, when N=1360 farmers only use Conditional Preventative.

area has 10 more counties that use information from nearby sentinel plots, i.e. choose CP, to tackle rust infection. Therefore, not only are the sentinel plots beneficial to the farmers in their own county, but they also benefit those in other counties. Such gains are feasible in our model when the correlation between the farm field and sentinel plot is high. This requires meeting two conditions (Figure 4): i) the counties should have similar probabilities of infection, and ii) the counties should not be separated by a very large distance. Here, both these conditions are met. Not only are these counties neighbours, but they also have similar beliefs about the probability of infection as seen in Figure 5. This makes the correlation sufficiently high for these counties to value the information from a nearby sentinel plot.

We now compare how the spatial optimization of sentinel plots could improve the total value of monitoring relative to actual placement of plots in 2014. Figure 10(i) maps counties with sentinel plots in 2014. There were 75 counties in total that contained sentinel plots. Out of these 75, only 39 counties are included in our data. Hence, for the comparison to be appropriate, we solve the spatial optimization problem with $B = 39$. Figure 10(ii) shows the optimal placement of sentinel plots. Examining the two figures (10(i)and 10(ii)) simultaneously shows that the actual sentinel plot locations in 2014 do not match the counties identified as optimal in our model. As predicted earlier, the optimal solution would place sentinel plots mostly in the north-west of Southeast and in the Midwest region. However, sentinel plots in 2014 were mostly clustered in the southern part of the Southeast region. Although this region has a moderate to high risk of infection, it has very low soybean production and is also far away from the main production areas. As a result, the quality of signal of infection coming from these far off sentinel plots are relatively poor for the main soybean producing areas. This results in a total value of monitoring of \$0.07 bn at maximum. On the other hand, the optimal value of monitoring post-spatial optimization is \$0.58 bn, which is 828.5% of the actual value captured. It should be noted that, while our findings show the inefficient allocation of sentinel plots, we do not wish to emphasize the magnitude of the estimated loss from this inefficiency. This is because our estimates of the total value of monitoring are dependant on the assumptions about the correlation function. A higher correlation across longer distances would lead to a higher value of monitoring based on the sentinel plot locations in 2014.

We now investigate the number of sentinel plots that will maximize the value to farmers in the US. Figure 10 shows how the optimized value of monitoring varies with budget B. When B is 80, the total value of monitoring is \$0.58 bn. This value keeps increasing at a decreasing rate until B is 400 where it attains a maximum of \$0.59 bn. No additional value is gained from sentinel plots beyond 400, i.e., the marginal benefit of sentinel plots becomes zero. Because we do not explore the costs associated with expansion and maintenance of the sentinel plots network, we cannot determine the optimal number of plots. However, Figure 10 still tells us that there is no additional benefit in having B greater than 400, assuming the sentinel plots have also been efficiently allocated to the counties.

7 Conclusion

There are many potentially catastrophic pathogens which enter the US agricultural system every year. In order to prevent damage from these pathogens, the government invests 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, we develop a dynamic model of farmer decisions regarding pest management and use it to estimate the value of the sentinel plot monitoring network and to optimize the spatial arrangement of sentinel plots. The paper uses the case of soybean rust and hence 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 pests better and also refines their beliefs about their risk of infection. 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 far 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 they 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 is relatively low. Our estimates suggest that more plots should be placed in the Corn Belt where the risk of an infection is lower, but where much more soybean is produced. Such a modification in sentinel plots arrangement could have increased the value of monitoring in 2014 by 829%. However, it is important to note that there are other potentially valuable uses of monitoring plots that we did not include in our model of farmers' decisions. For example, 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.

Our results also show that, given our assumptions, no more than 400 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 400 sentinel plots. In recent years, there have been decreases in the number of sentinel plots. However, most of the sentinel plots still remain in the South. This fact points us to a useful direction of research– estimating the optimal number of sentinel plots in ipmPIPE. 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.

Before we answer this question, it is important to better understand the spatial auto-correlation of infections since the marginal benefit curve is derived based on assumptions about the correlation function. We assume that the correlation of risk of infection is sigmoidal in the distance between sentinel plot and farmer's field and that it becomes zero at 150 mi from the farmer's field. These assumptions are not yet corroborated. Any change in these assumptions can cause significant changes in the value of monitoring and therefore in the suggested arrangement of sentinel plots. This leaves scope for improvement.

References

Quick stats national agricultural statistics service. http://www.nass.usda. gov/. Accessed: Jan, 2016.

United states department of agriculture ipmpipe restricted site. https://ipmpipe.zedxinc.com/login.html. Accessed: May, 2018.

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*, 311–331.

GAMS (2015). General algebraic modeling system. *GAMS Development Corporation, Release 24.6. Washington, DC, USA.*

Hershman, D. Personal communication. University of Kentucky Research and Education Center. 2013-02-03.

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. *AIBS Bulletin 55*(10), 851–861.

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.

Kuchler, F., M. Duffy, R. Shrum, W. 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.

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. *USDA Outlook Report No.(OCS04D02), April, www. ers. usda. gov/publications/OCS/APR04/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.

Miller, S. A., F. D. Beed, and C. L. Harmon (2009). Plant disease diagnostic capabilities and networks. *Annual review of phytopathology 47*, 15–38.

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.

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. Technical report, United States Department of Agriculture, Economic Research Service.

Rossman, A., K. Britton, D. Luster, M. Palm, M. Royer, and J. Sherald (2006). Main content area evaluating the threat posed by fungi on the aphis list of regulated plant pests. *Plant health progress*, 603–610.

Stack, J., K. Cardwell, R. Hammerschmidt, J. Byrne, R. Loria, K. Snover-Clift, W. Baldwin, G. Wisler, H. Beck, R. Bostock, et al. (2006). The national plant diagnostic network. *Plant Disease 90*(2), 128–136.

Tables

		Field	
		Infection	No infection
Sentinel Plot	Infection	ϕ_{11}	ϕ_{01}
	No infection	ϕ_{10}	ϕ_{00}

Table 1: States of the world and their probability of occurrence when ipmPIPE is available

Parameter	Value
Soybean Prices P	\$10.10/Bu
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

Table 2: Baseline parameters for the farmer decision model.

Figures

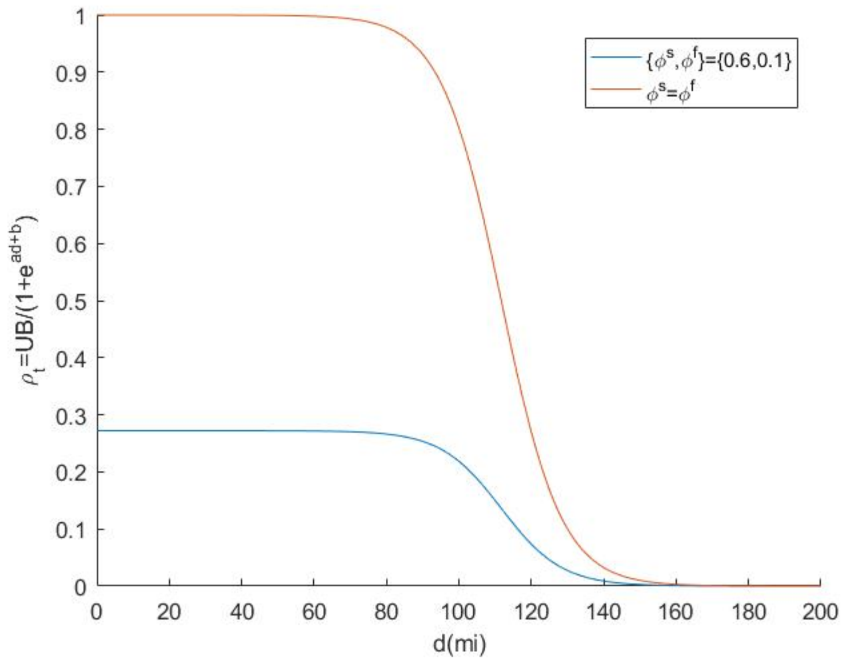


Figure 1: Correlation function ρ

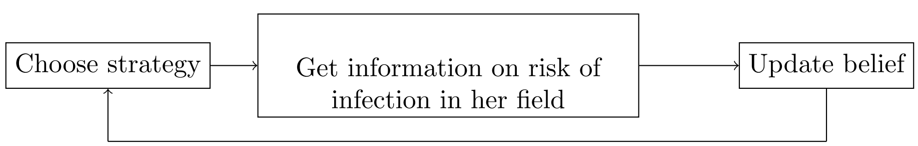


Figure 2: Timeline

Figure 3: Optimal management decisions for a farmer with no monitoring network.

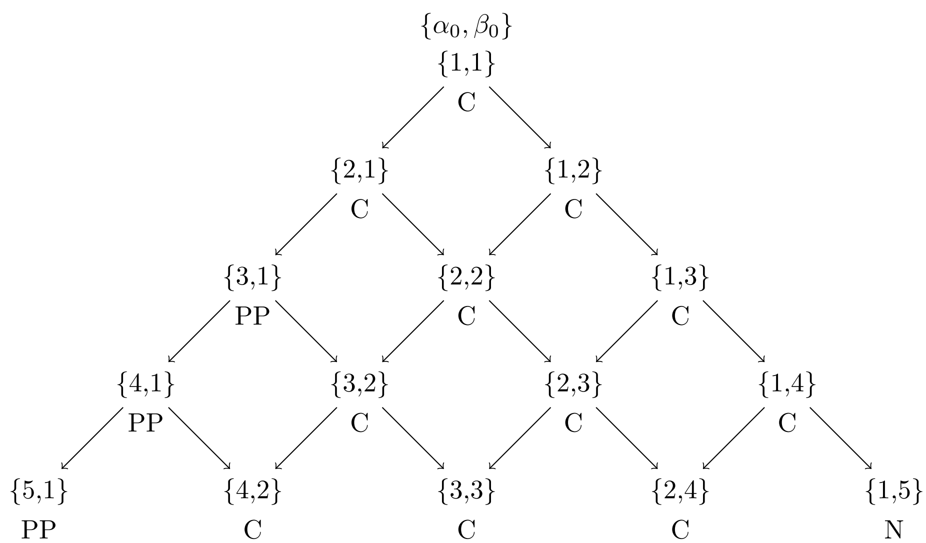
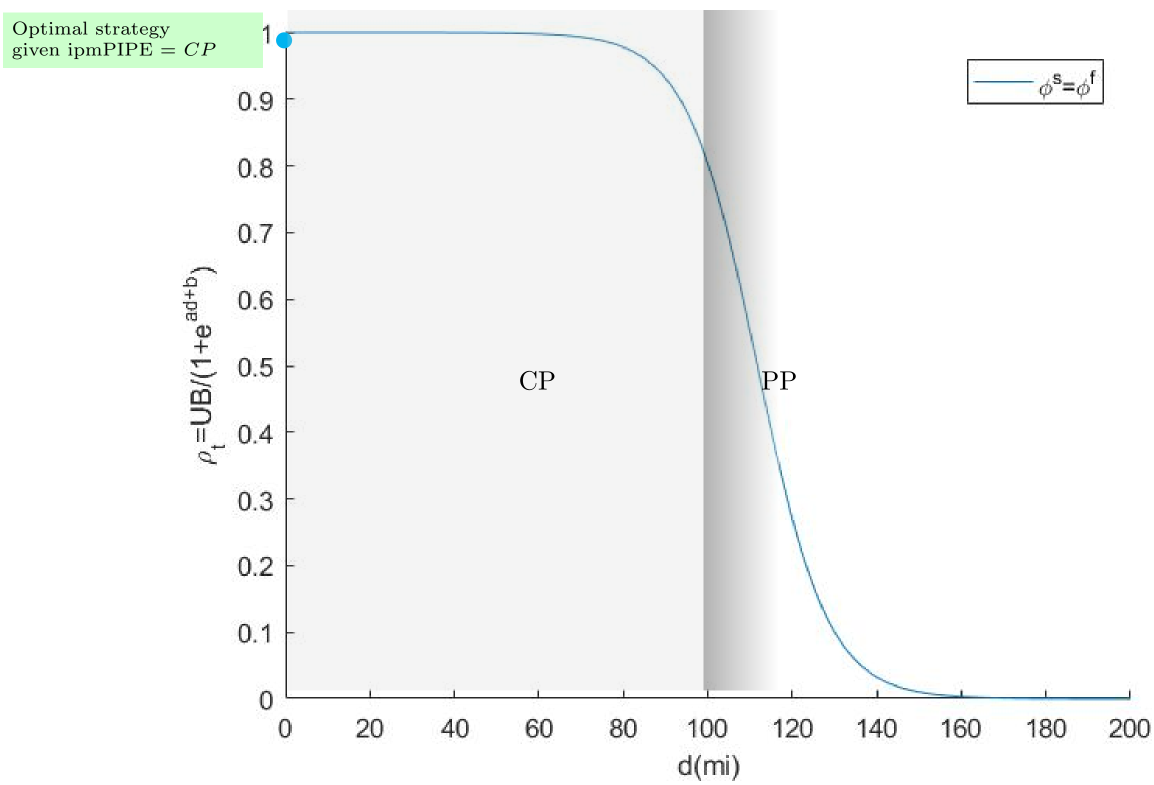


Figure 4: Optimal management decisions for a farmer with monitoring network.



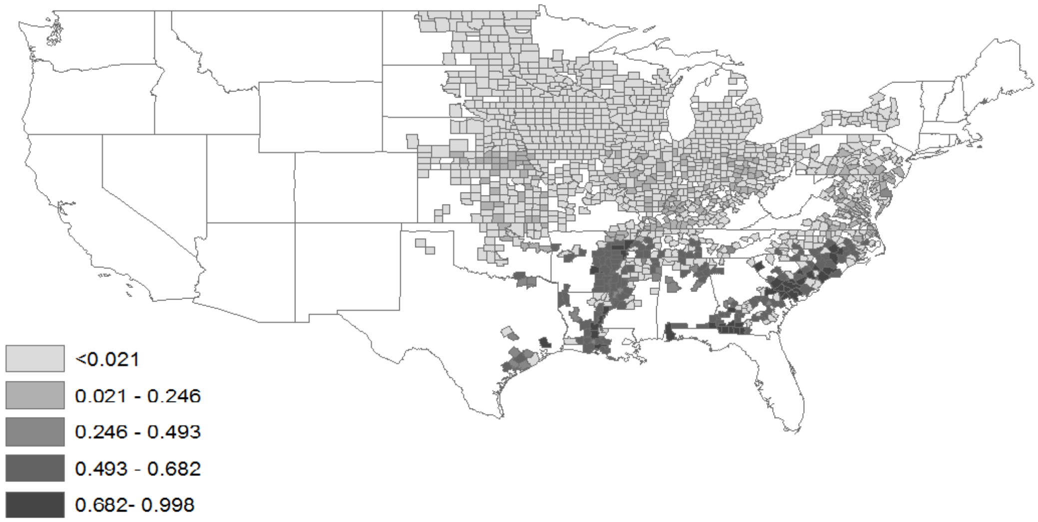


Figure 5: Probability of infection

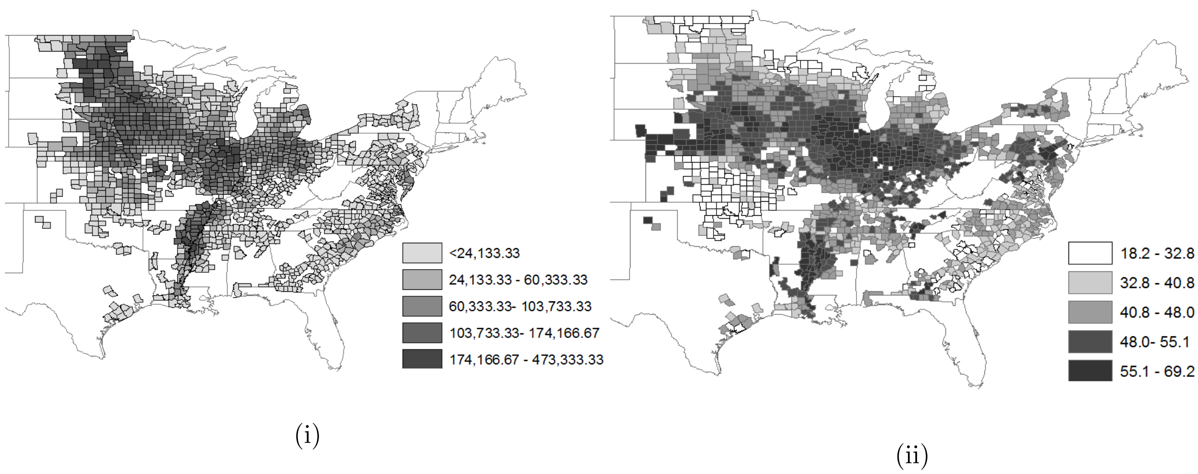


Figure 6: (i) Soybean acreage (in acres) and (ii) yield (bu/acre) for 2014

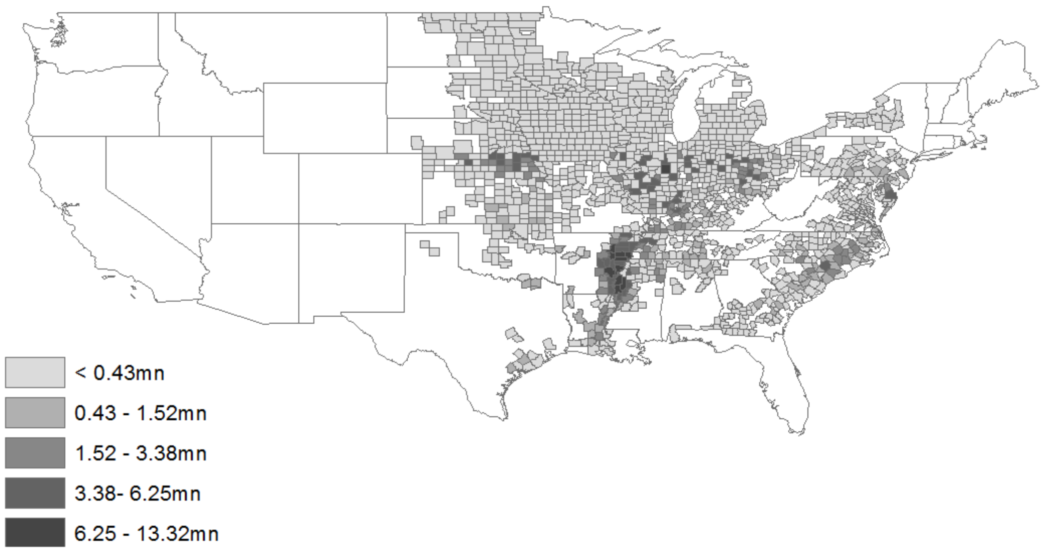
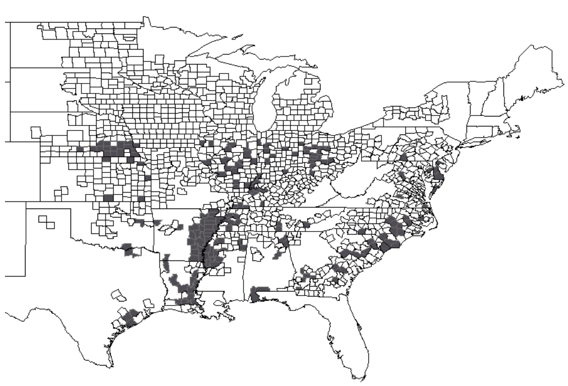
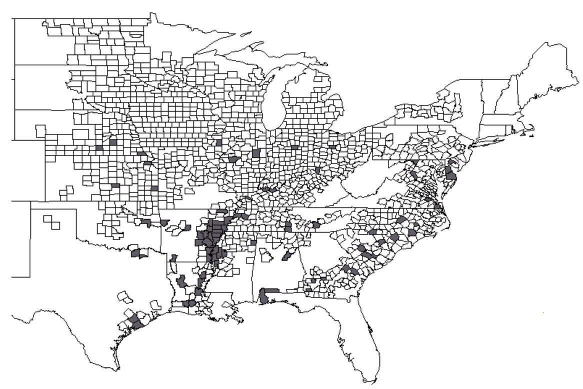


Figure 7: Value of monitoring when each county has a sentinel plot

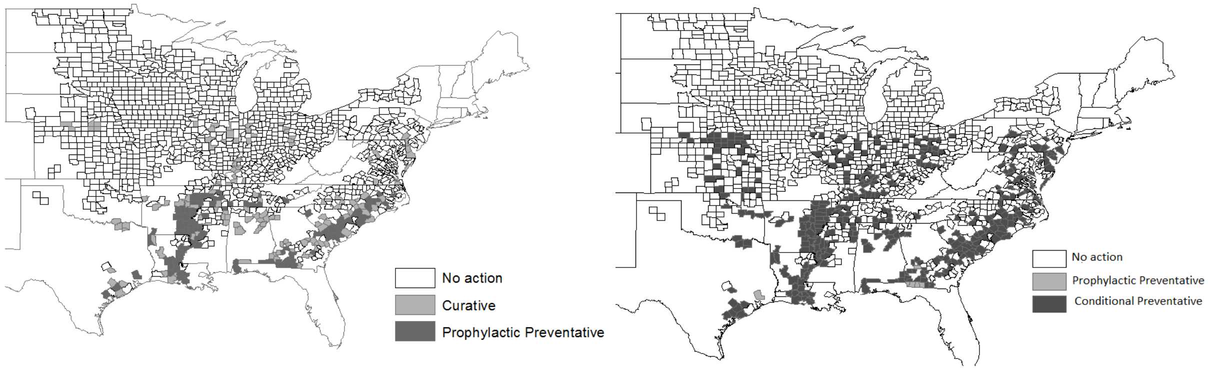


(i)



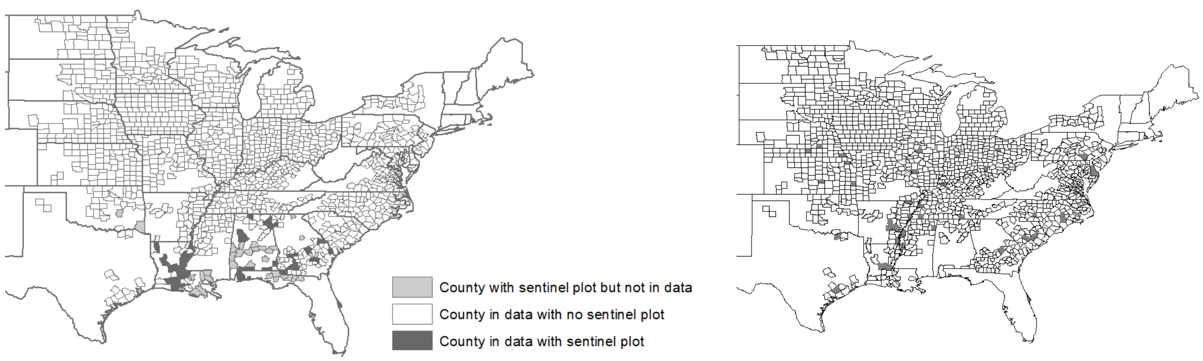
(ii)

Figure 8: Optimal sentinel plot locations when (i) B=185 and (ii) B=80.



(i) (ii)

Figure 9: Optimal strategies when (i) monitoring network is unavailable and (ii) post-spatial optimization when B=80.



(i) (ii)

Figure 10: Counties with (i) sentinel plots in 2014 (ii) sentinel plots post-spatial optimization when B=39.

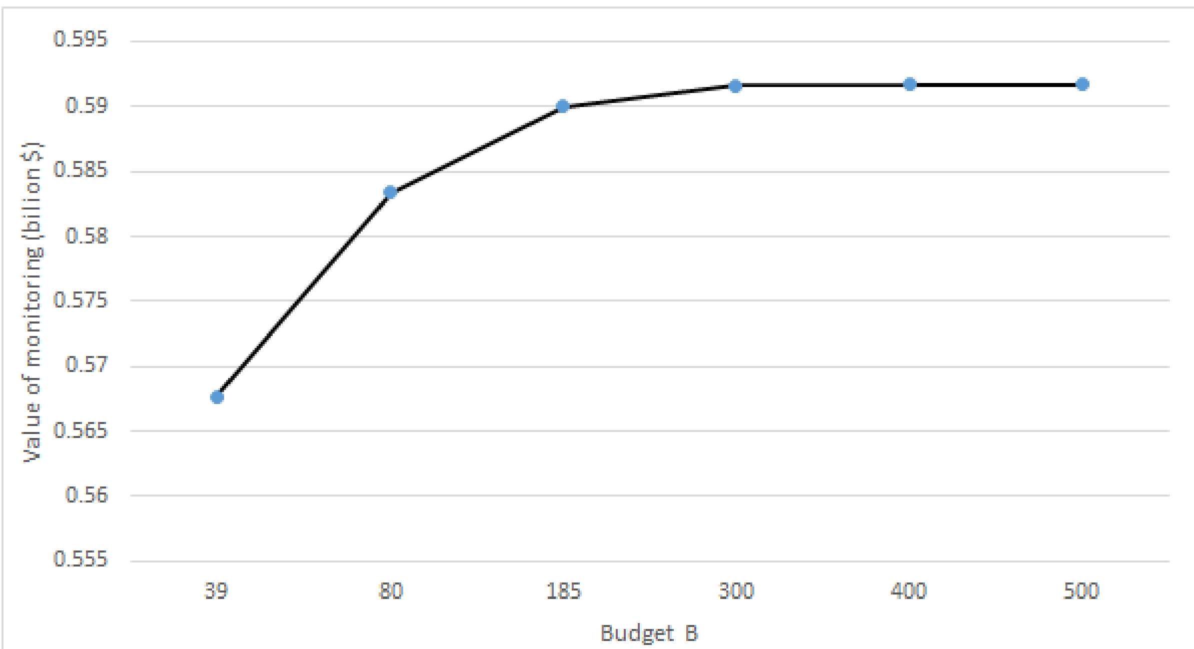


Figure 11: Value of monitoring by number of sentinel plots (budget)

Appendix 1

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 the probability that j=1. Hence

$$
\begin{aligned}
E(i) &= \phi^f && (17)\\
E(j) &= \phi^s && (18)\\
Var(i) &= \phi^f(1-\phi^f) && (19)\\
Var(j) &= \phi^s(1-\phi^s) && (20)\\
E(ij) &= \phi_{11} && (21)
\end{aligned}
$$

Therefore, correlation ρ can be expressed as

$$
\begin{aligned}
\rho &= \frac{Cov(i,j)}{\sqrt{Var(i)Var(j)}}\\
&= \frac{E(ij)-E(i)E(j)}{\sqrt{Var(i)Var(j)}}\\
&= \frac{\phi_{11}-\phi^s\phi^f}{\sqrt{\phi^s(1-\phi^s)\phi^f(1-\phi^f)}}\\
\implies \phi_{11} &= \phi^s\phi^f + \rho\sqrt{\phi^s(1-\phi^s)\phi^f(1-\phi^f)}. && (22)
\end{aligned}
$$

Also,

$$
\begin{aligned}
\phi_{11}+\phi_{10} &= \phi^f && (23)\\
\phi_{11}+\phi_{01} &= \phi^s && (24)\\
\phi_{00}+\phi_{11}+\phi_{10}+\phi_{01} &= 1. && (25)
\end{aligned}
$$

Using the above equations, we get:

$$
\begin{aligned}
\phi_{01} &= \phi^s(1-\phi^f) - \rho\sqrt{\phi^s\phi^f(1-\phi^s)(1-\phi^f)} && (26)\\
\phi_{10} &= \phi^f(1-\phi^s) - \rho\sqrt{\phi^s\phi^f(1-\phi^s)(1-\phi^f)} && (27)\\
\phi_{00} &= (1-\phi^s)(1-\phi^f) + \rho\sqrt{\phi^s\phi^f(1-\phi^s)(1-\phi^f)}. && (28)
\end{aligned}
$$