

## Dispersal in a hurry: Bayesian learning from surveillance to establish area freedom from plant pests with early dispersal

S. Low-Choy<sup>ab</sup>, Nichole Hammond<sup>c</sup>, Lindsay Penrose<sup>d</sup>, Chris Anderson<sup>e</sup>, Sharyn Taylor<sup>f</sup>

<sup>a</sup>Cooperative Research Centre for National Plant Biosecurity, Canberra, Australia; <sup>b</sup>Mathematical Sciences Discipline, Queensland University of Technology, Brisbane, Australia; <sup>c</sup>Department of Agriculture and Food, Perth, WA; <sup>d</sup>Department of Agriculture, Forestry and Fisheries, Canberra, Australia; <sup>e</sup>NSW Department of Primary Industries; <sup>f</sup>Plant Health Australia  
Email: [s.lowchoy@qut.edu.au](mailto:s.lowchoy@qut.edu.au)

**Abstract:** Declaration of area freedom from plant pests is crucial for the agricultural sector, since it promotes continuing domestic and international trade of crops at risk from exotic pests. Freedom from plant pests may also enhance environmental health, with indirect effects on agricultural productivity. Every year, several new exotic plant pest species are reported for the first time. In the face of this continual pressure and growing globalization, resources to undertake surveillance are limited. Design of surveillance is critical for determining how to allocate these limited resources.

Designs for surveillance that help assess area freedom have focused on the colonization process, captured by a prevalence model that does not accommodate dispersal. In this paper, we extend these designs to accommodate early dispersal from a few colonization points. This provides a basis for evaluating the effectiveness of surveillance over multiple sampling occasions.

To achieve this we harness a Bayesian statistical framework. Although there are some computational overheads, this provides several benefits: (i) an intuitive hierarchical structure that helps separate then link modelling components; (ii) the facility to incorporate expert knowledge; and (iii) inference that directly addresses the questions of farmers and biosecurity managers, in a way that the range of plausible outcomes is provided together with point estimates. Finally the Bayesian framework facilitates a natural cycle of learning, that readily incorporates new information – from surveillance snapshots – as it becomes available.

Firstly, we harness the natural hierarchical structure of the Bayesian statistical framework to separate the model—for the spatio-temporal dynamics of dispersal underlying prevalence—from the model for detection, which depends on prevalence.

Secondly, expert knowledge on both point estimates and variability can be explicitly incorporated as Bayesian prior distributions, and in each phase, these priors are updated into new posteriors as more surveillance data becomes available. This is important since much of the data informing design of surveillance for exotic plant pests relies heavily on expert judgment, especially during the early phases of plant biosecurity—when establishing area freedom.

Thirdly, the Bayesian posterior approach used here automatically answers the question of *If we detect nothing, how many infested plants could we have missed?* This approach provides a ready mechanism for including information about dispersal on the infested plants (both missed and detected).

Finally, the Bayesian framework facilitates an adaptive cycle of learning. We can apply Bayesian inference to analyze the first surveillance snapshot and learn about prevalence and detectability parameters. Then Bayesian predictions can be used to progress the pest status before analysis of the next snapshot. This flexibly provides a basis for incorporating new knowledge as it is obtained.

We utilized freely available software, that enjoy high utilization among non-statisticians and statisticians, for: exploratory data analysis, statistical modelling and visualization.

**Keywords:** *Surveillance design; plant pests; disease; area freedom; dispersal; Bayesian statistics*

## 1. INTRODUCTION

In plant biosecurity, the detection of new outbreaks of exotic plant pests will crucially determine whether the outbreak may feasibly be managed or eradicated, and hence the impacts and longer-term viability and resilience of the agricultural industry. Here we use “pest” in its official sense, as an organism that negatively affects plant health and/or productivity at a range of biological scales, e.g. pathogens, diseases, insects and feral animals. Early detection of such pests critically underpins many phases of the biosecurity continuum, by supporting: a *preventative* and holistic approach to plant health; continuing trade via evidence of *area freedom*; *emergency responses* to identify the point of origin and likely pathways; and monitoring to *delineate* the extent of outbreaks.

Each year, several outbreaks of new pests are reported in many countries such as Australia. Such diseases could potentially affect a large portion of the agricultural industry, both domestically and internationally. Here we consider how surveillance could be designed to support claims of area freedom by individual enterprises. The purpose of surveillance is therefore: to assess the probability that the plant pest is absent from an enterprise, given the resources allocated to search for, and detect, the pest. In particular we desire a measure of confidence or credibility about how accurately we have assessed this probability.

Several different quantitative techniques (Hester, et al. 2010) have been used to design surveillance in biosecurity. These can be broadly grouped as (i) sample size determination (SSD) for simple distributions, such as the binomial; (ii) simulation models, such as stochastic scenario trees (SSTs) and Bayesian Networks (BNs); (iii) and Bayesian inference. Simple probabilistic models are very popular, since they provide a closed form solution for SSD. Typically these assume that the number of (independently occurring) detections follows a binomial or hyper-geometric distribution, for a specified number of infested individuals, distributed homogeneously throughout a population of known size (Cannon and Roe 1982). However these techniques often do not account for the information available on accuracy of surveillance, or detectability, in terms of its sensitivity ( $Se$ ) to detect the pest when present, and specificity ( $Sp$ ) to avoid detecting the pest when absent.

With SSTs and BNs, detectability can be modeled using a sequence of probabilities for detection and reporting in the field, then in the laboratory (Martin, Cameron and Greiner 2007, Hood, Barry and Martin 2009). These models extend the simple SSD approaches, and define more complex probabilistic (likelihood) models. However, these approaches are all logically constrained to evaluate the chance of any surveillance dataset under fixed values of prevalence, and for known detectability. In contrast, Bayesian inference assesses the plausible levels of prevalence and detectability, from a specific set of surveillance data. The Bayesian approach has been previously used to analyze detections of disease in humans (Winkler and Smith 2004) or animals (Suess, Gardner and Johnson 2002) when prevalence is low, and errors in detection cannot be discounted. We chose the Bayesian inferential framework since it suits the biosecurity regulatory context.

Methodologies described so far correspond to a single “snapshot” of surveillance. However we would obtain greater confidence in area freedom if surveillance occurred over several sampling occasions. This allows some time for the pest outbreak to progress, so that it may become larger and therefore more easily detectable. Most often it is assumed that sampling occasions are independent, so that probabilities (e.g. of power to detect when present) simply multiply over sampling occasions (e.g. Martin, Cameron and Greiner 2007). Where pests spread quickly, there may be a strong relationship between the numbers of plants infested at consecutive sampling occasions. This can be captured using temporal correlation, e.g. via discounting approaches embedded within SSTs (Martin 2008) A spatio-temporally explicit approach considers all possible evolutions of pest levels between sampling occasions (Stanaway, Mengersen and Reeves in press). Here we consider a less computationally demanding approach, which analyzes the spatial snapshots of surveillance data as above, but also allow temporal evolution of pest levels between snapshots.

In Section 2, we present a Bayesian model for snapshot surveillance, which provides the basis for extension to multiple sampling occasions. In Section 3 we present an efficient method of incorporating these dynamics in a way that neatly harnesses the Bayesian learning cycle. In Section 4, we show how it applies to a simplified version of a case study on Myrtle Rust. We finish with a discussion in Section 5.

## 2. SNAPSHOT SURVEILLANCE

The classical approaches (such as SSD in the Frequentist paradigm, SSTs, BNs) address the question: *When the pest is present (at a specific level), and for fixed levels of false positive and false negative rates, then how likely are we to detect nothing?* Specifically, they evaluate the chance of the observed number of detections  $Y$  for particular fixed values of the model parameters for prevalence, true and false positive rates ( $\pi, \delta, \varphi$ ):

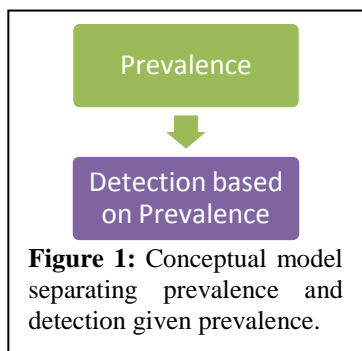
$$\Pr(\text{Data}|\text{Model parameters}) = \Pr(Y = 0|\pi, \delta, \varphi). \quad (1)$$

Bayesian inference reverses this logic by updating the prior state of knowledge about the model parameters – here prevalence and detectability – by incorporating surveillance data  $Y$ , to provide an updated or “posterior” probability for model parameters:

$$\Pr(\text{Model parameters}|\text{Data}) = p(\pi, \delta, \phi|Y = 0). \tag{2}$$

Thus Bayesian inference allows regulators to directly interpret the results of surveillance: *If we detect nothing, then what could the prevalence level plausibly be; how likely is it that the pest is present?* Indeed many people confuse these two questions as framed in Eqns (1) and (2). This logical error is so common that it has earned the moniker “inversion fallacy” (Gigerenzer and Hoffrage 1995).

We simplify the context and the corresponding model so that this paper can focus on the core idea of multiple sampling occasions. Future work will extend to more complex spatial structure of surveyed enterprises, and the sampling strategies to reflect this structure, e.g. risk-targeted sampling (Jarrad, et al. 2011).



Conceptually,  $X$ , the actual level of prevalence (green box, Figure 1) can be modeled separately from  $Y|X$ , detectability (purple box, Figure 1), which relates to the level of prevalence at the time. Then the Law of Total Probability (LoTP) integrates out the unobserved latent variables  $X$ :

$$p(y|\pi, \delta, \phi) = \int_x p(y|x, \delta, \phi) p(x|\pi) dx$$

Prevalence is governed by parameters  $\pi$ . Detectability is governed by parameters for specificity or one minus the false positive rate,  $\phi = 1 - \text{FPR}$ , and sensitivity or true positive rate,  $\text{TPR} = \delta$ .

A hierarchical Bayesian model is therefore formed, analogous to models for surveillance of populations in ecology (Royle and Dorazio 2008). The prior distribution is  $p(\theta) = p(\pi) \cdot p(\delta) \cdot p(\phi)$  where the parameter set is

by  $\theta = (\pi, \delta, \phi)$ . The posterior distribution is thus  $p(\theta|y) \propto p(y|\theta) \cdot p(\theta)$ .

### 2.1. Prevalence model

Suppose that an enterprise (e.g. a farm, plant nursery or glasshouse) is divided into  $j = 1, \dots, J$  areas (e.g. paddocks or sections). Each area comprises  $n_j$  plants. The total number of susceptible host plants across the enterprise is  $N = \sum_j n_j$ . Use  $\mu_j$  to denote the probability of colonization for each plant in area  $j$ , assuming that the pest is in the vicinity of the area. Since we are focusing on early detection we assume that  $\mu_j$  is small. Thus the number of plants colonized  $X_j$  can be modeled using a Binomial distribution. Otherwise, a hyper-geometric sampling model may be more appropriate to reflect sampling *without* replacement.

$$X_j \sim \text{Bin}(n_j, \mu_j) \text{ or } X_j \sim \text{HG}(n_j \mu_j, n_j(1 - \mu_j), n_j) \tag{5}$$

If all areas are equally susceptible to colonization then  $\mu_j$  will be constant across areas. We explicitly model the latent variables  $X_j$ , instead of using a zero-inflated formulation for detections  $Y_j$  (Branscum, Gardner and Johnson 2004) as this provides a convenient basis for propagating pest dispersal between sampling occasions (Section 3). If zero prevalence occurs more often than expected with a Binomial or hyper-geometric distribution, then a negative Binomial or zero-inflated Binomial distribution could also be considered.

### 2.2. Detection model

In each area, detectability is governed by the true and false positive rates of the search strategy, which in turn depend on the skill of the inspector. If an area is searched, then we record the number of detections  $Y_j = y$ . If the area is not searched, then this value is missing (or Not Available, denoted NA), so  $Y_j = \text{NA}$ . By definition the probability of no detections depends on whether the pest is present:

$$v_{ij} = \Pr(Y_{ij} = 0 | X_{ij} = x, \delta, \phi) = \begin{cases} 1 - \delta, & x = 1 \\ 1 - \phi, & x = 0 \end{cases} = (1 - \delta)^x + (1 - \phi)^{(1-x)} \tag{6}$$

Aggregated over  $n_j$  plants in a block gives  $v_j = \prod_i v_{ij} = (1 - \delta)^{x_j} (1 - \phi)^{n_j - x_j}$  where  $x_j = \sum_{ij} X_{ij}$ .

### 2.3. Prior models: target, selection and encoding

Information can be “elicited” from experts to define prior distributions in a Bayesian statistical model. We use two-parameter Beta distributions to describe the prior distributions of TPR and FPR (Suess, Gardner and Johnson 2002). Previous work elicited the mode, a value considered most likely by the expert, and an upper 95<sup>th</sup> percentile considered unlikely to be exceeded (5% chance) (Branscum, Gardner and Johnson 2004). We

also elicited a lower bound, then encoded the Beta distribution that “best” fit all elicited information (Low Choy, Mengersen and Rousseau 2008), allowing experts to choose asymmetric intervals (e.g. 0<sup>th</sup> of 5<sup>th</sup> percentiles). Questions were ordered to reduce cognitive bias (Low Choy, et al. 2010).

During elicitation of prevalence and detectability, we note it was imperative that the spatial and temporal units are clear. Hence to encode Eqn (6) we elicited the probability that a single infested plant would be detected. A low technology approach was used out of necessity, since elicitation was carried out during a telephone conference. Feedback was delivered via email. Alternatively software could be used to provide instantaneous feedback and help refine the expert assessments (Low Choy, et al. 2010).

### 3. SURVEILLANCE WITH EARLY DISPERSAL

#### 3.1. Strategy for modeling repeated surveillance

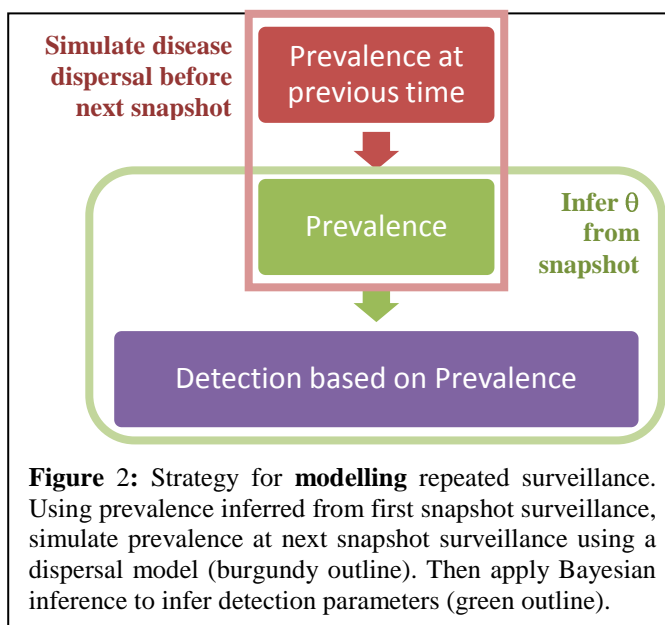
We can harness the Bayesian learning cycle to better match the way in which we learn from the data as it becomes available:

1. *Learn from the first surveillance snapshot* using Bayesian inference. Use simple assumptions on prevalence.
2. *Allow time for the pest to spread enough to improve chance of detection.* Simulate this via Bayesian prediction.
3. *Learn from the second surveillance snapshot* using Bayesian inference. Use outputs from steps 1 and 2 as priors.

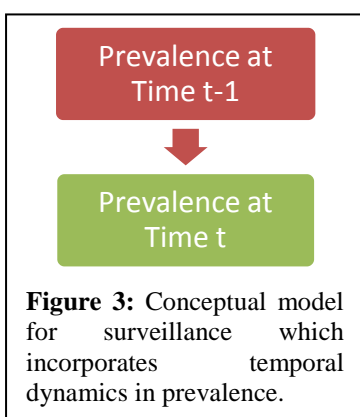
In Phase 1 we obtain baseline information from the first snapshot, and infer parameters  $\theta$  for prevalence and detectability. The need to establish area freedom allows a simple prevalence model initially (colonization without spread, parameterized by  $\pi$ ).

A considerable challenge to surveillance is the trade-off between: searching early enough to detect the pest before it escalates and is not eradicable, compared to waiting long enough so that expensive search effort is not “wasted” on searching “for a needle in a haystack”. In Phase 2 we have reframed this problem, by using the first snapshot as a baseline, and only then assessing how long should we wait before the next snapshot.

We can use the posteriors on model parameters from Phase 1 as priors in Phase 3. This learning cycle also allows us to refine models: for simulating dispersal between snapshots (Phase 2), and for inferring detectability and colonization rates from analyzing snapshots (Phase 3.)



**Figure 2:** Strategy for **modelling** repeated surveillance. Using prevalence inferred from first snapshot surveillance, simulate prevalence at next snapshot surveillance using a dispersal model (burgundy outline). Then apply Bayesian inference to infer detection parameters (green outline).



**Figure 3:** Conceptual model for surveillance which incorporates temporal dynamics in prevalence.

#### 3.2. Learning Phase 2: Temporal dynamics

Here we consider a “real-time” approach that keeps updating our knowledge as more information is accumulated through surveillance. The Bayesian cycle of learning provides an ideal framework, since its hierarchical nature has already separated the model for prevalence from detection (Figure 1). This separation is important since detectability relates to the level of prevalence at the time. This model structure provides a platform for incorporating as much detail as necessary on how prevalence evolves over time, so that prevalence depends on prevalence at the previous time (Figure 2) by specifying  $p(X_t|X_{t-1})$  where  $X_t = \{X_{tj}\}_{j=1,\dots,J}$ .

In the snapshot model for a single sampling occasion (Section 2), we obtained a posterior distribution for the latent prevalence process on the first sampling occasion,  $p(X_1|Y, \cdot)$  where  $X_1$  denotes the number of missed infested plants. (Here the dot signifies all model parameters.) In between sampling occasions, any pests that are present may disperse,

depending on the conditions during the interim. A simple model for dispersal between sampling occasions is an exponential model:

$$X_{t_j} \sim \text{Exp}(\lambda_j | X_{t-1,j}) \text{ so that } p(X_{t_j} | \lambda_j, X_{t-1,j}) = \lambda_j \Delta_t \exp\{-\lambda_j \Delta_t X_{t-1,j}\} \quad (7)$$

where  $\Delta_t$  is the time elapsed between surveillance occasions, and  $\lambda_j$  is the average rate of increase in number of plants infested per unit time interval. Thus for each location, this defines a time-dynamic model  $p(X_{t_j} | \lambda_j, X_{t-1,j})$  for pest prevalence in each section-area. If necessary, this model can be extended to also allow for colonization between snapshots. It replaces the simpler Phase 1 model based on colonization rate  $\pi$ .

If we elicit  $\Delta X$  as the percentage increase in  $X$  over time period  $t$ , then  $\Delta X = 100X_t/X_{t-1}$ . If we assume that the expected dispersal rate  $\lambda_j$  is constant over time (although the prevalence level may change), then using a simple method-of-moments approach (Low Choy, Mengersen and Rousseau 2008), we estimate  $\hat{\lambda}_j = \Delta X/\Delta_t$ .

This assumes that the colonization points that would have been missed on the first sampling occasion were few and far between. This assumption is not as constraining as it first appears: if the sampling occasions are too far apart, then the pest outbreak could escalate, making it unlikely that surveillance will report no detections, and the complex inference required to interpret zero detections would no longer be necessary.

### 3.3. Learning Phase 3: Incorporating temporal dynamics into the Bayesian cycle of learning

At the end of an initial sampling period with few colonizations, the posterior predictive distribution for the number of missed infested plants is obtained as  $p(X|Y, \cdot) = \int p(x|\theta)p(\theta|y) d\theta$ . Thus a series of posterior predictions can be generated via the posterior distribution of parameters, and the likelihood for prevalence  $\theta^{(m)} \sim p(\theta|y)$ ,  $x_1^{(m)} \sim p(x|\theta^{(m)})$  as defined in Section 2 and Eqn (5). For each infested area, we can further generate posterior predictions of the likely number of infested plants, accounting for dispersal (Eqn 7), via  $x_2^{(m)} \sim \text{Exp}(\lambda_j, x_1^{(m)})$ . These predictions form a prior for the level of infestation at the next sampling occasion.

### 3.4. Algorithm: Bayesian learning for early dispersal

In summary the algorithm involves:

1. Prior to the first round of surveillance, specify hyper-parameters in prior distributions:
  - a. How likely is each area (containing susceptible hosts) to be colonized? Specify  $p(\pi_j)$ .
  - b. What proportion of plants might be affected in a colonized area? Specify  $p(\mu_j)$ .
  - c. Specify hyper-parameters in prior distributions on detectability:  $p(\delta), p(\varphi)$ .
2. Undertake surveillance, and determine whether any pests were detected.
  - a. If so, then exit the area freedom phase, and enter the back-tracing or delineation phase.
  - b. Else set  $\{Y_j = 0\}$  for areas searched and fit model (4)-(6) to obtain posteriors for  $\delta, \varphi, \{\mu_j\}$ .
  - c. Retain posterior predictions of the number of missed infected plants,  $\{X_j\}$ , for each area.
3. For each subsequent round of surveillance:
  - a. Apply the dispersal model (Eqn 7) to output from Step 2c to define the prior for  $\{X_j\}$ .
  - b. Set the prior for parameters  $\delta, \varphi, \{\mu_j\}$ , using the previous round's posteriors (Step 2b).
  - c. Undertake the next round of surveillance and update the results (as per steps 2a-c above).

Bayesian posterior distributions were simulated using WinBUGS version 1.4.3 (Spiegelhalter, et al. 2003), with 1000 chains for burn-in, then 10,000 posterior simulations were retained after thinning by 10. Posterior predictions and graphics were produced in R (Ihaka and Gentleman 1996). Code is available from SLC.

## 4. APPLICATION

### 4.1. Case study

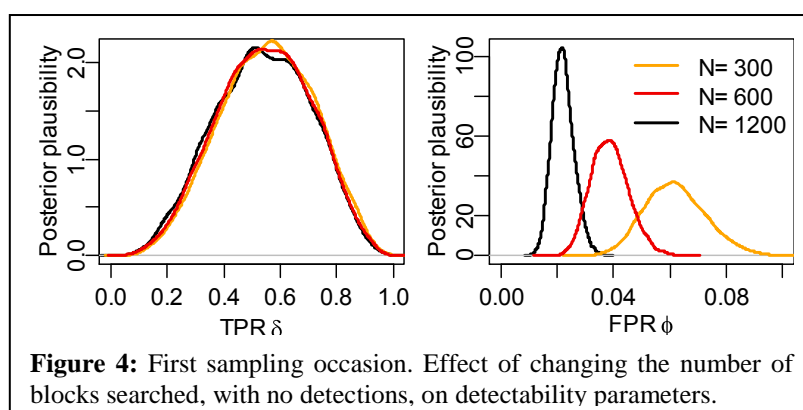
For Myrtle Rust we consider surveillance designs for areas comprising 3000 plants each. For simplicity here, we ignore further structure within and among areas. When searched, we assume that a block of adjacent plants was inspected in an area, e.g. 30, 600 or 1500 neighbouring plants (susceptible hosts). These dimensions were chosen since (a) we wanted to compare with the common advice to sample 600 plants (Cannon and Roe 1982), and (b) they were considered typical of nursery wholesalers or farms.

#### 4.2. Elicited detectability and prevalence

We consider inspectors of moderate skill, who are motivated, and have had training in detecting this disease on this host and/or experience with similar diseases or hosts, but require mentoring to confirm suspicious findings. An expert (LP) with experience in undertaking surveillance for similar diseases was consulted and provided the following assessments. TPR would most likely be 2 in 3, with a realistic range (95% chance) from 1 in 3 to 9 in 10. This was encoded using a Beta (6, 3.5) distribution, where elicited quantiles have cumulative probability 0.03 and 0.98. FPR would most likely range from 10-20% (with 95% chance) with a best estimate of 15%. This was encoded as a Beta (31,171); elicited quantiles had probability 0.01 and 0.96.

Probability of a plant being infested was set slightly higher than the usual threshold of 1 in 100, by setting the mode to 2 in a 100, with a plausible range (95% chance) from 5 in 1000 to 1 in 10. This was encoded using a Beta (2,50) distribution, where elicited quantiles have cumulative probability 0.03 and 0.97.

#### 4.3. Inference from the first sampling occasion



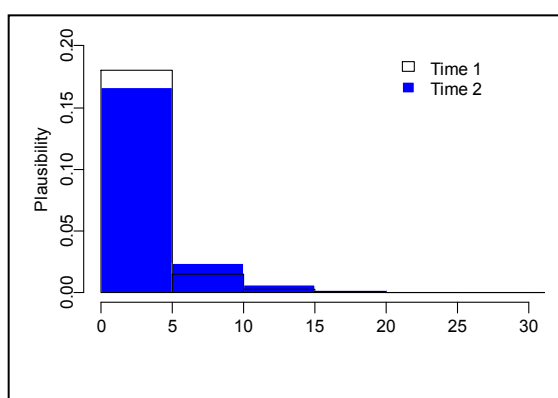
We considered block sizes of 30 plants, with 10, 20 or 40 blocks searched (one plant at a time), totalling 300, 600 or 1200 plants searched from 3000. Figure 3 shows: (a) no information gain on TPR (no detections recorded); and (b) as more blocks are searched, with zero detections, FPR reduces, and TNR rises. The pattern for prevalence was similar to FPR (not shown here), with the upper 95%

credible limit (Cr95) decreasing as more blocks are searched: 0.027, 0.016, 0.009. Similarly the Cr95 of the number of missed infested plants (out of 3000), also decreased from 74, 39, to 17.

#### 4.4. Elicited dispersal rate

After an estimated 4-6 weeks, Myrtle Rust symptoms may double the initial level of infestation, if appropriate environmental conditions (host, adequate heat, humidity, and wind or other dispersal mechanism) have occurred between sampling occasions. As per Section 3.2, we provide a conservative (high) estimate of the rate of dispersal to be  $2/4 = 0.5$  new infested plants per colonized plant per week.

#### 4.5. Inference from a second sampling occasion



Assuming the second sampling occasion occurs after 4 weeks, then the mean infested number of plants (12.08) is double the initial value (5.97). However the upper 95% credible limit for the infested number of plants at 46 is more than double that for the first sampling occasion (17), for 40 blocks searched. The disease was propagated between sampling occasions using the dispersal model (Eqn 7). At the next sampling occasion (4 weeks later), we also suppose zero detections in the searched blocks. Then the posterior distribution of the number of infested plants in the searched zone becomes only slightly higher. At Time 2, the upper 95<sup>th</sup> credible interval is 11 missed infested plants in the searched zone, compared to 8 at Time 1.

### 5. DISCUSSION AND CONCLUSIONS

The model presented here provides an intermediate option between models for area freedom that solely consider colonization and models for dispersal that explicitly model the spatial and temporal progression of the pest. This model focuses solely on the temporal evolution of pest numbers, ignoring spatial location.

We build on previous Bayesian approaches to modelling surveillance (Branscum, Gardner and Johnson 2004, Stanaway, Mengersen and Reeves in press), but have made several changes. We permit non-homogeneous risk of colonization across areas on a farm. The beta prior distributions for prevalence and detectability are encoded using an approach (Low Choy, et al. 2010) to reduce cognitive biases. We retain the latent variables for prevalence to provide a platform for extending into surveillance over multiple occasions.

We chose a simple exponential model for temporal evolution of the pest, but more complex models could easily be incorporated within this hierarchical framework to include: colonizations between snapshots and risk factors for colonization (e.g. habitat suitability) or dispersal. Current work on the Myrtle Rust case study is extending the model to another spatial scale, of rows within areas. In this context of surveillance in plant biosecurity, the Bayesian cycle of learning helps demonstrate and promote a gradual uptake of new knowledge (models and data). This facilitates “real-time” learning, to potentially improve delivery of information as surveillance results accumulate, for example during emergency responses.

## ACKNOWLEDGMENTS

We would like to thank the Surveillance Reference Group and for their insights on practical implementation and implications of surveillance designs, and the Consultative Committee for Emergency Plant Pests (CCEPP) for their questions which motivated this work. We also thank reviewers for helpful feedback.

## REFERENCES

- Branscum, A.J., I.A. Gardner, and W.O. Johnson. “Bayesian modeling of animal- and herd-level prevalences.” *Preventive Veterinary Medicine* 66 (2004): 101-112.
- Cannon, R. M., and R. T. Roe. *Livestock disease surveys: a field manual for veterinarians*. Australian Government Publishing Service, Canberra, 1982.
- Gigerenzer, G., and U. Hoffrage. “How to improve Bayesian reasoning without instruction: frequency formats.” *Psychological Review* 102 (1995): 684–704.
- Hester, Susie, Cindy Hauser, John Kean, Terry Walshe, and Andrew Robinson. “Post-border surveillance techniques: review, synthesis and deployment.” Milestone Report, Australian Centre of Excellence for Risk Analysis (ACERA), 2010.
- Hood, G. M., S. C. Barry, and P. A. J. Martin. “Alternative Methods for Computing the Sensitivity of Complex Surveillance Systems.” *Risk Analysis* 29, no. 12 (2009).
- Jarrad, Frith, et al. “Improved design method for biosecurity surveillance and early detection of non-indigenous rats.” *New Zealand Journal of Ecology* 35, no. 2 (2011): in press.
- Low Choy, S., J. Murray, A. James, and K. Mengersen. “Indirect elicitation from ecological experts: from methods and software to habitat modelling and rock-wallabies.” In *Handbook of Applied Bayesian Analysis*, edited by A. O'Hagan and M. West. Oxford University Press, Oxford, UK, 2010.
- Low Choy, S.J., K. Mengersen, and J. Rousseau. “Encoding expert opinion on skewed nonnegative distributions.” *Journal of Applied Probability and Statistics* 3 (2008): 1-21.
- Martin, P. A. J. “Current value of historical and ongoing surveillance for disease freedom: Surveillance for bovine Johne's disease in Western Australia.” *Preventive Veterinary Medicine* 84 (2008): 291-309.
- Martin, P. A. J., A. R. Cameron, and M. Greiner. “Demonstrating freedom from disease using multiple complex data sources 1: A new methodology based on scenario trees.” *Preventive Veterinary Medicine* 79 (2007): 71-97.
- Royle, J. A., and R. M. Dorazio. *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities*. Academic Press, Elsevier, London, 2008.
- Spiegelhalter, D. J., A. Thomas, N. G. Best, and D. Lunn. “WinBUGS version 1.4 user manual.” MRC Biostatistics Unit, Cambridge., 2003.
- Stanaway, M.A., K.L. Mengersen, and R. Reeves. “Hierarchical Bayesian modelling of early detection surveillance for plant pest invasions.” *Environmental and Ecological Statistics* 17 (in press): 1-23.
- Suess, E. A., I. A. Gardner, and W. O. Johnson. “Hierarchical Bayesian model for prevalence inferences and determination of a country's status for an animal pathogen.” *Preventive Veterinary Medicine* 55 (2002): 155-171.
- Winkler, R.L., and J. E. Smith. “On uncertainty in medical testing.” *Medical Decision Making* 24 (2004): 654-658.