Sample size calculations for phytosanitary testing of small lots of seed

An investigation of various options

Stephen E. Lane[1], Rose Souza Richards[2], Claire McDonald[2], and Andrew P. Robinson[1]

[1]CEBRA, The University of Melbourne
[2]Ministry for Primary Industries, New Zealand

March 19, 2018

cebra
Centre of Excellence for
Biosecurity Risk Analysis

Executive Summary

New Zealand has a highly valued and internationally respected seed growing and seed export industry due to its disease-free status and the ability to provide additional growing seasons for Northern Hemisphere producers. The success of the industry is dependent on the ability to import seed of a wide variety of species and from different production areas. The Ministry for Primary Industries (MPI) is responsible for managing the biosecurity risks associated with seed imports which includes the importation of high value seeds as small seed lots.

At present the only option available for seed importers of small seed lots (under 2500 seeds) wishing to import into New Zealand, is to source the seed from countries that are declared free from the regulated diseases listed in the Import Health Standard (IHS) 155.02.05: Seeds for Sowing [8]. This requirement limits the number of countries/suppliers that are eligible to supply the various commodities.

In order to maximise the sustainability and growth of the New Zealand seed export industry, an alternative seed testing protocol designed specifically for importing small seed lots is required. The sampling protocol must be flexible enough to help facilitate frequent imports of different volumes of seeds, species, and various geographic locations.

In this report, we provide a detailed discussion on an approach which may be used to manage the *pathway-level* risk of contaminated seed being imported into New Zealand. In any system reliant on sampling to detect contamination, contaminated product may be missed and subsequently imported; this is known as *leakage*. The proposed approach that we detail provides, at a minimum, an architecture for interrogating the effect of such leakage. In particular, we show the impacts of lot size on possible leakage rates. Further, the approach that we detail may be used alongside detailed pest risk analyses to determine an acceptable leakage rate for a pathway, and subsequently provide a mechanism for calculating a sample size required for testing on a lot-by-lot basis. The system that we detail thus manages the whole-of-pathway risk, rather than the risk on an individual lot basis, as per the current system.

Additionally, this report provides a summary of the relevant methods that are described in the *International standards for phytosanitary measures: ISPM 31, methodologies for sampling of consignments* [5] for calculating sample sizes in relatively small lots. Whilst we acknowledge that the New Zealand's Ministry for Primary Industries already has such a methodology at its disposal, we use this report to stress its importance, and further provide an example software application that can be tailored to MPI's specific import requirements.

The full list of recommendations from this report are provided in the List of Recommen-

dations, however at a minimum we suggest that MPI continue to utilise and expand upon the ISPM 31 methodology (which uses the Hypergeometric distribution) of sample size calculation for imports of small seed lots. We also suggest that consideration be made into the effects of leaked contaminated seed, and that the novel approach of controlling such leakage be entertained as a way to manage pathway-level risk of seeds for sowing.

List of Recommendations

Contents

1

Introduction

New Zealand's Ministry for Primary Industries (MPI) is responsible for the import requirements to all viable seeds and products containing viable seeds for sowing from species that are eligible to be imported into New Zealand. Current procedures are designed to mitigate the risk of introduction of seed-borne (and seed-transmitted) diseases and pests, and also to manage the risk of contamination by regulated weed seeds and new organisms that are associated with this pathway. Most of the specified sampling and testing requirements stipulate a minimum of 2000 seeds for viruses in order to achieve a nominal 95% confidence of selecting and detecting diseased seeds at a contamination rate of at least 0.15%. Sampling is destructive, and this high sample size makes importation of small quantities of seed into New Zealand not feasible. For the purposes of this work a small seed lot is defined as a lot which contains 2500 seeds or less.

At present the only option available for seed importers of small seed lots is to source the product from countries[1] that are declared free from the regulated diseases listed in the Import Health Standard (IHS) 155.02.05: Seeds for sowing [7]. This requirement limits the number of suppliers that are eligible to supply the different commodities. To date, there is no option available for importers to modify sampling and testing protocols for seed lots smaller than 2500 seeds. To maximise the sustainability and growth of the New Zealand seed export industry, an alternative testing protocol designed specifically for importing small seed lots is required. The sampling protocol must be flexible enough to help facilitate frequent imports of different volumes of seeds, species, and various geographic locations.

This report presents different options that may be applied for a sampling protocol targeting only small seed lots. In Chapter 2, we describe sampling without replacement in further detail, and suggest that future sample size calculations be based on a less conservative statistical assumption than at present. Chapter 3 proposes a method of calculating sample sizes, in which we minimise the rate at which contaminated seed escapes detection; this is extended by Chapter 4 which details a case study based on the method. Chapter 5 raises the possibility of setting inspection protocols based on a good record of 'clean' imports, and Chapter 6 provides concluding remarks. The rest of this chapter presents an overview of the current system.

[1] Or areas, and other places of production.

1.1 The Current New Zealand Seed Import System

The current system is described in Figure 1.1. Seed lots are sorted, inspected, and a sample is taken for testing by an official sampler. The sample is tested by a National Plant Protection Organisation (NPPO) approved laboratory, which issues a phytosanitary certificate if the test is clear. This sampling occurs entirely offshore. If the test is clear and a phytosanitary certificate is issued, then the lot can be exported to New Zealand.

The bulk of the seed testing is performed offshore, and some testing is performed in New Zealand. Onshore, samples are processed by a single laboratory: the Plant Health and Environment Laboratory, in Auckland. If the test is negative, then the lot can be released. If the test is positive, then the lot is either destroyed or re-exported[2].

Offshore test results are not provided for individual lots — if the testing has been performed by a NPPO approved laboratory, and the lot is accompanied by an appropriate phytosanitary certificate, no further testing is required [7]. In other words, with testing for seed-borne diseases being conducted offshore, positive test results (i.e., test results for which a pathogen is found) are not provided to MPI, and only lots that pass the Import Health Standard (IHS) are presented. We will revisit this issue in Chapter 5.

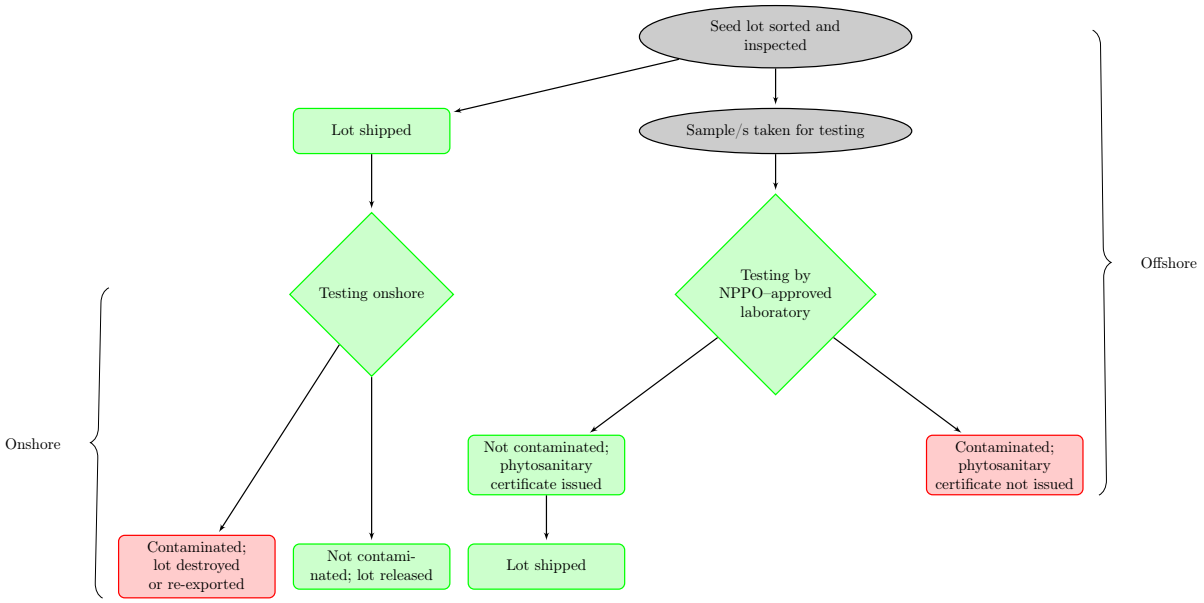


Figure 1.1: Description of the current seed-testing/import system.

1.2 Assumptions for the Current System

The current system sets a minimum sample size to be tested when importing seed for sowing for commodities where a testing regime is required[3]. A sample is selected, tested for various diseases, and the lot is either certified as 'clean' or free from the regulated pest(s).

[2]That is, the lot is sent back offshore, at the importers expense.

[3]Current sampling regimes adopted internationally can use methods detailed in [5] for sampling consignments, but in general are dependent on the pathogen type, infestation level, plant species imported, country of origin and volume of the seed lot imported.

The decision for MPI, assuming validated diagnostic tests for the specific pathogen and host plant are available for use, is how many seeds need to be tested to make the certification that the seed lot is free from the regulated pest with satisfactory confidence, without destroying the complete lot.

The current procedure to calculate how many seeds are required for that sample is statistically equivalent to the calculation from a sampling process in which individual seeds are selected, tested and returned to the lot, and may be sampled again[4]. This is not the exact process — seeds are in fact selected and not replaced after testing — however it is a useful approximation. This approximation enables a fixed number of seeds for testing to be specified in the IHS, which simplifies the IHS specifications.

When the lot size (N) to be imported is much larger than the number of seeds (n) specified by the IHS to be tested, the aforementioned approximation has little effect on the statistical properties of the sampling, specifically on the sample size required to be tested. When the lot size is small however, improvements to the process can be made; these improvements will be explored in more detail in Chapter 2. For now, we provide an example on how the sample size for testing is determined under the current system.

1.2.1 Description of the Process

Consider a lot that contains contaminated seed at a non-zero rate p, from which a sample of size n (of our choosing) is drawn and tested. There are two outcomes: the sample contains one or more contaminated seeds, or it contains no contaminated seeds. Assuming the sampling was performed correctly, and the laboratory tests are sensitive enough to detect even very low levels of the regulated pest in the representative sample, the lot is released if the selected sample contains no contaminated seeds; otherwise, if found to be positive, it is restricted from importation.

The decision for MPI is the choice of sample size n, such that the selected sample contains (with high probability) at least one of the contaminated seeds in the lot. The number of seeds in the selected sample has a particular mathematical form that allows the calculation of the sample size — it can be described by the *Binomial distribution* — all that is required is to specify p, the rate of contamination; ϵ, the efficacy of the detection; and C, the probability that the sample contains at least one contaminated seed. Details of these calculations can be found in Appendix A.

Note: The efficacy of the detection method, ϵ, is considered to be 100% throughout this report; if a lower efficacy is required, Appendix A provides further details.

1.2.2 Example: *Zea mays*

The IHS for *Zea mays* specifies all the requirements for importation, including testing, of *Zea mays* seeds intended for sowing in New Zealand. For example, a representative sample of a minimum of 3000 seeds for use in ELISA or PCR testing is required to test for Maize chlorotic mottle virus (MCMV). We calculate the figure of 3000 seeds required to test for MCMV by specifying: (i) the probability of a seed being contaminated as $p = 0.1\%$; and

[4]This procedure gives rise to the Binomial distribution.

(ii) the probability that the sample selected contains at least one contaminated seed as $C = 0.95$.

2

Improving the Calculation of Sample Size for Testing Under the Current System

In this chapter we detail an approach to calculating sample sizes for selecting (at random) seeds for testing. This approach was used for *Cucurbitaceae* seeds for sowing imports of between 1000 and 5000 seeds [7] at the time this project was initiated, however the current IHS stipulates a minimum of 2000 seeds[1] to be tested. We provide a description of the calculations in this chapter, from which we develop and provide details on a web-based calculator developed as a case-study, specifically tailored towards MPI's requirements.

Laboratory testing for seed-borne diseases requires destruction of the selected seed as discussed in Section 1.2. This type of selection of seeds for testing is known as sampling *without* replacement. When sampling without replacement, the selection of each seed is no longer independent of the other seeds selected. We can take advantage of this knowledge when calculating sample sizes. This knowledge will generally lead to smaller sample sizes required for testing. The impact on sample size is most clearly seen when the number of seeds in the lot is small relative to the number of seeds required for testing. With large lot sizes, the difference in calculated sample sizes between the method in this chapter, and that in Section 1.2.1 is negligible. When the lot size is small relative to the amount required for testing however, some savings (on sample size) can be made using the approach presented in this chapter.

It is important to note that if any reduction in sample sizes required for testing is made, the minimum quantity required for successful laboratory testing is maintained.

2.1 Description of the Process

We begin with a similar situation to that presented in Section 1.2.1. A lot contains contaminated seed at a non-zero rate p, from which a sample of size n (of our choosing) is drawn and tested. The difference here is that we are required to specify the cutoff *number* of contaminated seed in the lot instead of the cutoff contamination rate. We do this by

[1] For lots under 10000 seeds, a composite sample may be taken across all lots in the consignment.

multiplying the lot size N by the specified contamination rate p. If the outcome is not an integer, then we round down to make it an integer. Rounding down makes the test more conservative, see below for an example. As before (see Section 1.2.1), we must specify C, the probability the selected sample contains at least one contaminated seed.

The number of contaminated seeds selected for sampling can be described by a *Hypergeometric distribution*. We leave the technical details to Appendix B, but provide the following example to explain the process.

2.2 Example

An incoming lot contains $N = 2500$ *Zea mays* seeds. To demonstrate that the lot is free from Sugarcane mosaic virus, the IHS stipulates a negative result from an NPPO approved method assuming that the proportion of contaminated seeds in the lot is $p = 0.15\%$, and the probability that the selected sample contains at least one contaminated seed is $C = 0.95$. These specifications lead to the IHS requirement of a minimum sample of $n = 2000$ seeds to be selected.

Following the procedure described in Section 2.1, and Appendix B, we need to calculate the total number of contaminated seeds in the lot. In a lot of size $N = 2500$, if $p = 0.15\%$ we expect $0.0015 \times 2500 = 3.75$ contaminated seeds in the lot. Clearly we cannot have a fraction of an contaminated seed: it is either contaminated or not. There are two ways to resolve this: we could round up to 4 contaminated seeds, or round down to 3. Rounding up implies we are targeting a *higher* contamination rate ($4/2500 = 0.16\%$). This results in a smaller sample size requirement, which is a more liberal procedure — this is entirely justified, as a higher contamination rate is easier to detect. Instead, we err towards conservatism and round down the number of contaminated seeds to $m = 3$[2].

From Equation (B.3), we find that a sample size of $n = 1579$ is required. Compare this to the 2000 seeds required currently by the IHS; 500 seeds would remain for sowing by the importer (under the current IHS), as opposed to 921 using the method in this chapter. While this is only a modest saving, it is a saving nonetheless; moreover, the saving comes without loss in the confidence of the outcome. The method is easy to implement, as is shown in the next section, and we recommend its consideration. Further numerical examples are given in Table C.1 (Appendix C).

We have shown here that considerable savings (in terms of number of seeds tested) for the importer may be made by adopting the method described in this chapter. However, we note that any reduction in sample size needs to be considered with reference to minimum quantities required for testing in the laboratory. For example, 350 seeds may be required to be tested using the method of this chapter; however, the laboratory may require a minimum quantity of 500 seeds to guarantee an appropriate sensitivity of the test. In this case, the number of seeds to be tested would need to be 500 seeds or higher, irrespective of the sample size calculation.

Recommendation 1. *MPI should continue to consider sample size calculations based on the Hypergeometric distribution (corresponding to sampling without replacement) instead of the Binomial distribution (corresponding to sampling with replacement) for small seed*

[2]Note that the *realised* targeted contamination rate is now 0.12%, due to rounding down.

lots. Any changes to sample sizes should remain consistent with laboratory requirements for minimum quantities to maintain high sensitivity.

2.3 Web-based Sample Size Calculator (Hypergeometric)

We have developed a web-based sample size calculator using the method described in this chapter[3]. The calculator uses seed imports of interest to MPI as a case study to demonstrate how easily the calculations can be implemented. The application is built using the `shiny` package in R; example code is distributed with this project.

The application is run from a standard R session. Make sure to start the session from the folder in which both the `ui.R` and `server.R` code are located. Use the command `shiny::runApp()` (Code chunk 2.1) from the R prompt; the application will open in your system's default web browser.

R Chunk 2.1.

```
#>shiny::runApp()
```

The application will open in the default browser, similar to that shown in Figure 2.1. After opening, we can select the species of the seed in the lot using the 'Seed species' dropdown box (Figure 2.2). Figure 2.3 shows how to change the number of seeds in the lot — as a result, the number of seeds required for each individual test is shown. In this example, we have selected *Zea mays* as the seed species, and there are 3500 seeds in the lot. The output shows that there are seven seed-borne diseases that require testing, and that High plains virus requires a minimum of 2210 seeds to be sampled for testing.

Figure 2.3 shows the targeted contamination rate, p (design prevalence[4]), and probability of selecting at least one contaminated seed, C. These are the current values as specified in the IHS. Figure 2.4 shows an example of the spreadsheet used to enter these requirements as used by the application. For each test required for a seed species, a row is entered into the spreadsheet containing: the species name; the pest/disease name; and the design contamination rate from the IHS.

2.4 Consequences of Per Lot Sampling

In the next chapter, we will discuss the concept of leakage, which we define as contaminated seeds that are not detected by sampling, and which are consequently imported.

The amount of leakage in a system such as that discussed in this chapter is generally hidden. That is, consideration of the effects of leakage when it occurs is often only considered

[3]There are numerous web-based sample size calculators, however the one we have developed is specifically tailored towards MPI's interests.

[4]The design prevalence is a lower bound on the contamination rate we wish to detect (recalling smaller contamination rates are more difficult to detect).

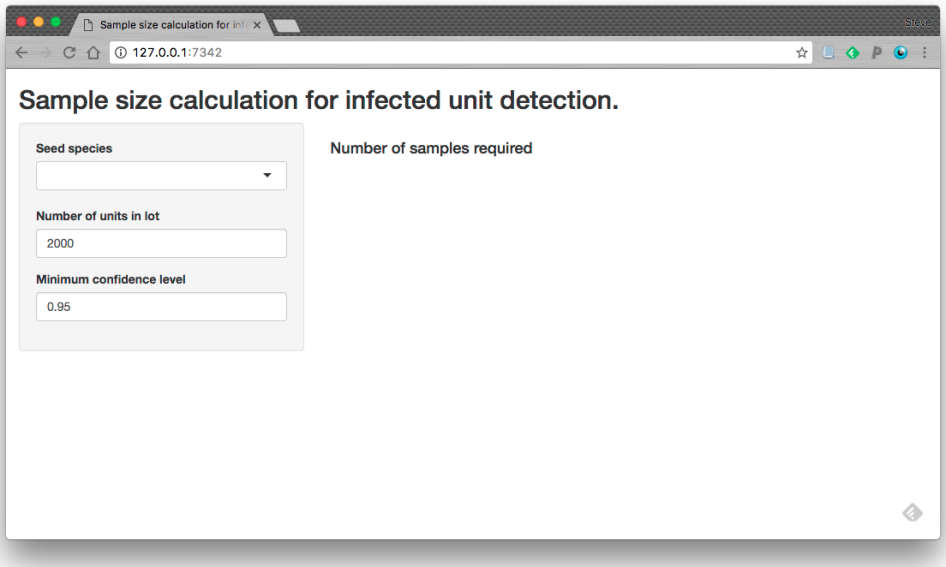


Figure 2.1: Initial screen of the interactive sample size calculator.

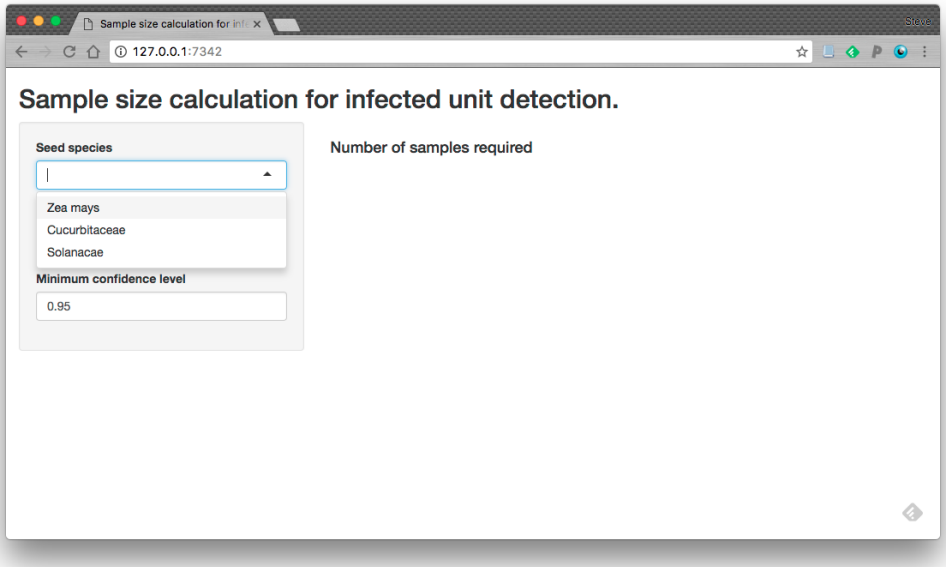


Figure 2.2: Selecting the species of seed in the lot.

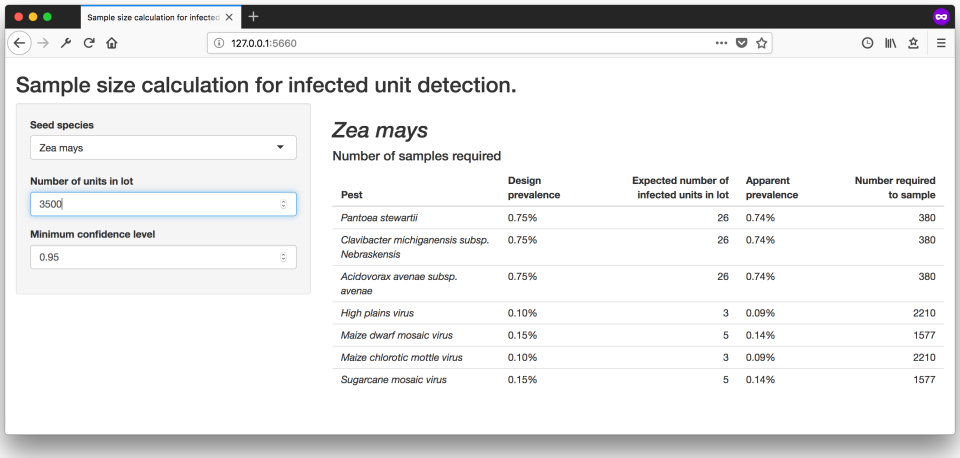


Figure 2.3: Changing the number of seeds in the lot. Note that the number of seeds required for each test is displayed as a result.

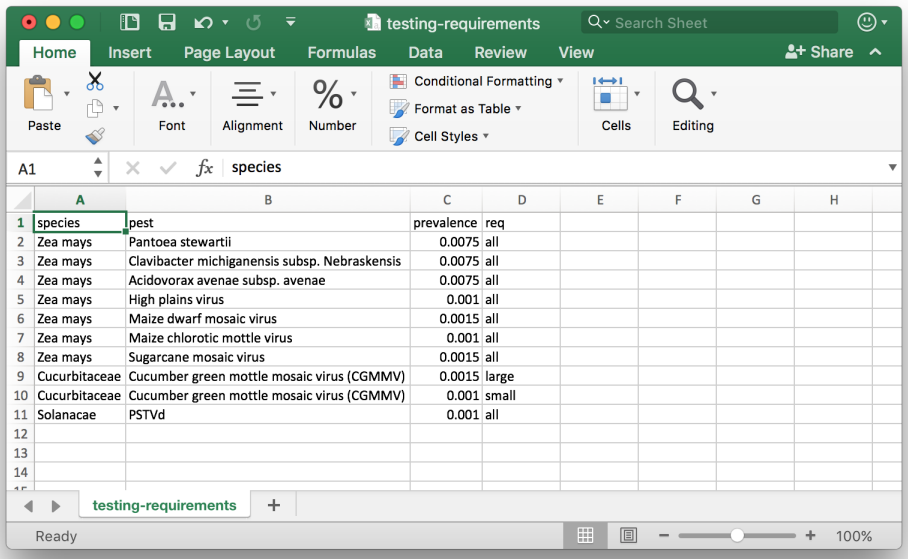


Figure 2.4: Example spreadsheet for specifying required tests and design contamination rates for input into the sample size calculator.

indirectly. Both the method of this chapter, and the current approach (Section 1.2) provide a *per lot* confidence. What this means is that on average, only a proportion C of contaminated lots will actually have the contamination detected. In other words, a proportion $1 - C$ of contaminated lots will leak the contaminated seeds.

The consequences of leakage are clearly linked with the number of seeds in the incoming lot (whether large or small). If a contaminated lot is undetected, then the number of contaminated seeds leaked is larger for a large seed lot, and smaller for a small seed lot. The amount of leakage is discussed in the next chapter.

3

Consideration of Sample Design by Amount of Leakage

In this chapter, we discuss some implications of the sampling schemes considered in the previous chapters (and any sampling scheme in general); in particular, we focus on the rate of leakage of contaminated seeds. Leakage — which we define as contaminated seeds that are not detected by sampling, and which are consequently imported — is possible in any system that relies on sampling to detect contamination; leakage occurs (on average) from $100 \times (1 - C)$ percent of contaminated lots imported (see Section 2.4) *that are sampled for inspection*[1]. Regulators generally recognise that "Zero risk is unattainable and undesirable" [2]. Even if zero risk were attainable, no regulatory jurisdiction would have the resources to search all cargo perfectly. Furthermore, when sampling is destructive (as in the motivating case for this study), 100% inspection is impossible.

The analysis presented in this chapter may be used to provide a different rationale for the sample size chosen for testing. Instead of setting the probability of finding contamination on a per lot basis as done in the preceding chapter, the long-run average leakage rate (exposure) may be managed. Chapter 4 provides a case study using this approach for *Capsicum annuum* and tests for Pepper chat fruit viroid (PCFVd). Importantly, this is a different way to think about how to manage the risk of the pathway, but does not necessarily lead to a more efficient approach. Any reduction in sample size from that calculated under the Hypergeometric assumption (Chapter 2), will necessarily result in a smaller *per lot* confidence by necessity; we explore this consequence further in Section 3.3.

3.1 Leakage Rate

The leakage rate is the proportion of contaminated seeds in the lot that are not detected. Leakage is affected by three factors, namely: (i) the contamination rate of seeds, (ii) the lot size, and (iii) the sample size for testing. With low contamination rates, the leakage rate is small because contamination is rare, and with high contamination rates, contaminated lots will be detected with high probability. Consequently, leakage is again small. At some

[1]Note that not all lots are inspected. Our discussion focuses on pathways with 100% inspection, and as long as the results are interpreted conditional on inspected lots, they hold for pathways with risk-based sampling.

intermediate rate of contamination, conditional on the lot size and sample size, there is a single, unique maximum leakage rate.

We now consider the average (long-run) leakage rate; that is, the leakage rate that would be expected to occur over many imported lots. We assume that the contamination rate p is fixed, but the number of contaminated seeds *within* each lot is variable[2].

Note: The *average leakage rate* needs to be interpreted relative to the pathway volume: whilst contaminated seed may get through in low rates, in large lots there will still be a large number of contaminated seeds that go undetected. For example, consider a lot of 100,000 seeds that is contaminated with $p = 0.1\%$. Suppose that a sample of 3000 is taken for testing. If the test shows that the sample is not contaminated, the remaining seeds in the lot are released, along with 100 contaminated seeds.

Note: The *absolute* number of leaked contaminated seeds for the pathway should also be considered as part of an assessment of the impact of leakage (see Chapter 4 for further details).

Note: For this chapter, we define a pathway as seed for sowing arriving as small lots of specific species. However, we note that the average leakage rate could be utilised independent of lot size.

Figure 3.1 displays the flow of contaminated seed through the sampling/testing system. Beginning with a lot of N seeds, if the lot is not contaminated then there is, obviously, 0 leakage. If the lot is contaminated, then a sample of size n may or may not contain the contaminated seed. Assuming the sensitivity of the laboratory tests to detect pathogens is near perfect, if the sample contains contaminated seed then they will be detected — again, there will be 0 leakage. If the sample does not contain contaminated seed, then conditional on the contamination rate p, leakage will occur with $p \times (N - n)$ (on average) contaminated seeds being leaked.

The average leakage rate can be calculated from the flow diagram in Figure 3.1 (by following the path that leads to the red coloured terminal node). Let a be the leakage rate, i.e. the proportion of released seed from a contaminated lot that is missed due to sampling, T^- denote a sample that doesn't contain contaminated seed and L^+ denote a contaminated lot. Then the average leakage rate is:

$$E(a) = p\frac{N-n}{N}\Pr(T^-|L^+). \tag{3.1}$$

Here, $\Pr(T^-|L^+)$ can be interpreted as the proportion of times we allow contaminated seed to leak — the blue box 'Not detected' on the middle pathway in Figure 3.1 has this probability. A full explanation of this result is provided in Appendix D.

[2]At the present stage we do not know how many lots are actually contaminated; here we merely assume that some of them will be at some specified rate.

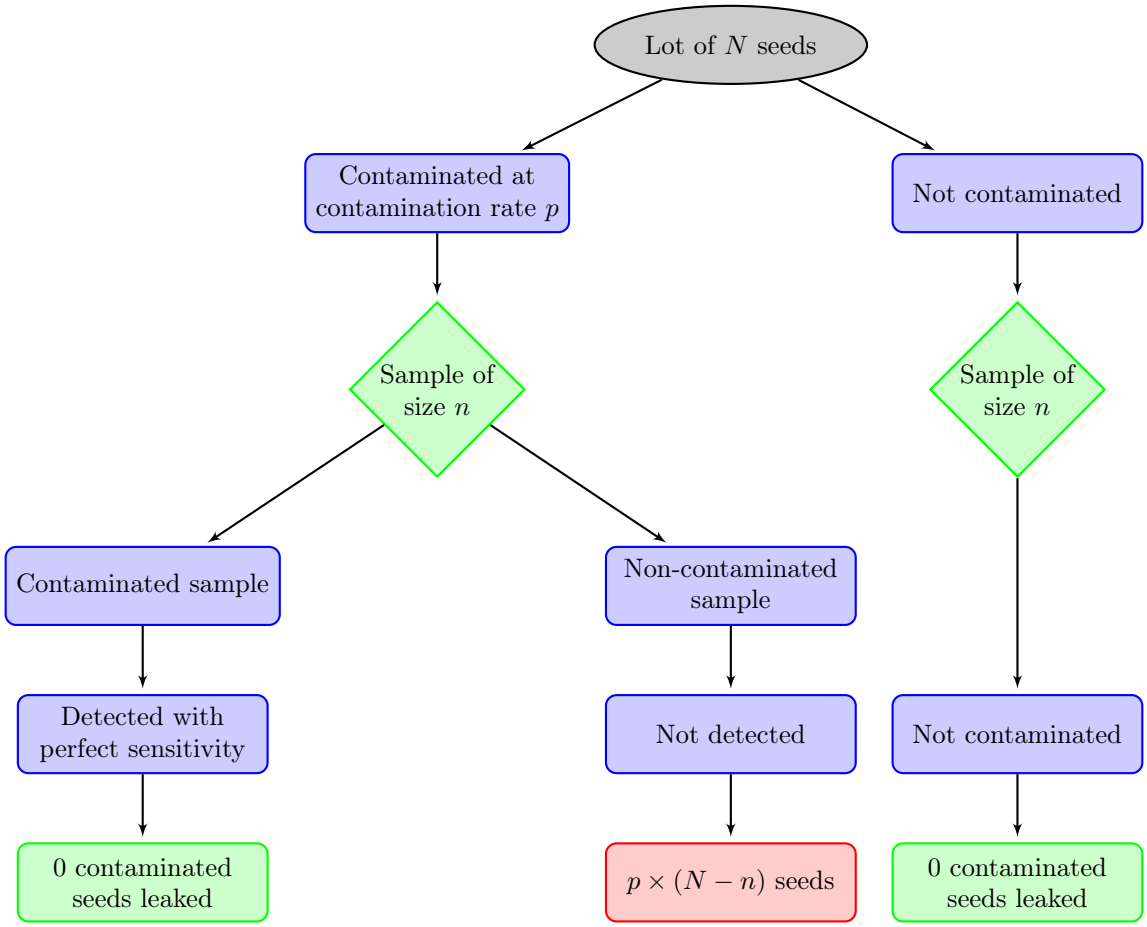
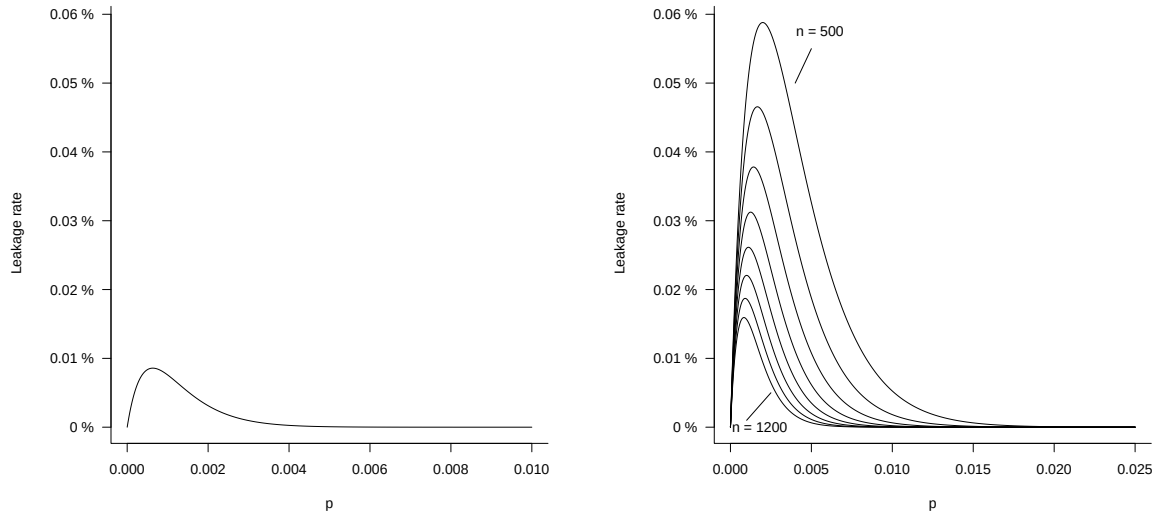


Figure 3.1: Flow of contaminated seeds through the sampling/testing system.

Example of Leakage Rates

Recall the example given in Section 2.2. For a lot of size 2500, and design contamination rate $p = 0.15\%$, we found the required sample size to be taken was $n = 1579$.



(a) Sample size of $n = 1579$

(b) Sample sizes $n \in \{500, 600, \ldots, 1200\}$

Figure 3.2: Leakage rate for lots of size $N = 2500$, for various infection contamination rate p and sample size n. Each curve in 3.2b relates to a different sample size for testing, and shows the amount of leakage is larger with smaller sample sizes..

Figure 3.2a shows the average leakage rate as a function of the lot contamination rate, with a sample size of $n = 1579$. For any lot contamination rate, the average leakage rate $a < 0.009\%$. Figure 3.2b shows the average leakage rate as a function of the lot contamination rate, with sample sizes $n \in \{500, 600, \ldots, 1200\}$. These curves were produced using Equation (3.1).

3.2 Sample Size Calculations for Maximum Average Leakage Rate of the Pathway

As discussed earlier in the chapter, we could consider using the leakage rate to set the sample size. As shown in Figure 3.2, a maximum average leakage rate amounts to finding (approximately) an incoming contamination rate p and confidence setting C, tailored to the lot size. That is, we use the maximum average leakage rate (i.e. the top of a leakage curve in Figure 3.2) to find our desired settings of the contamination rate p, and confidence, C.

Note: The development in this section provides a *pathway-level*[3] leakage cap, as opposed to a *per lot* leakage cap.

This sampling protocol requires us to consider the *impact* that a contaminated seed may have. One such specification may be:

Consideration of the specific virus/bacteria that may be present. Specific transmission rates and the economic costs of allowing leakage of contaminated seeds may be considered. It is outside the scope of this project to set rates in this fashion for all possible seed species, however we do develop a case study for *Capsicum annuum* and Pepper chat fruit viroid (PCFVd) (Chapter 4).

Note: In this situation, consideration of the sample size on *per lot* confidence should also be taken into account; see Section 3.3.

A method for setting the sample size can now be derived, by considering an *allowable* leakage rate. Knowing that there is a maximum average leakage rate for a fixed lot size N and sample size n (Figure 3.2), the idea is to choose n so that the maximum average leakage rate *for the pathway* is below this allowable leakage rate.

Consider Figure 3.2b, and assume that we set the maximum allowable leakage rate for the pathway to be 0.03%[4]. Choosing the sample size such that the average maximum leakage rate is below 0.03% equates to choosing n, such that the resulting curve is below 0.03%. In this example, some curves in Figure 3.2b lie below 0.03% and others lie above; thus, we can set the sample size from the interval $[500, 1200]$. For details on the algorithm to achieve this, see Appendix E.

To continue the example started above, if we set the maximum allowable leakage rate for the pathway to be 0.03%, we find that for a lot of size N=2500, a sample size of 823 would result in a maximum average leakage rate of 0.03%.

Note: A sample size chosen in this fashion *is dependent* upon the incoming lot size. For every lot size and allowable leakage rate, a new sample size *must be* calculated.

[3] In this report, we consider a pathway to be defined at the species/pest level.
[4] Note that the allowable leakage rate is required to be set by the regulator.

3.3 Effect of Controlling the Leakage Rate on Per Lot Confidence

Setting the sample size by controlling leakage equates to finding a corresponding design contamination rate p and confidence C to set the *per lot* sample size, as mentioned earlier in the chapter. Furthermore, controlling the maximum average leakage rate of the pathway provides *pathway-level* assurance, not *per lot* assurance.

As Figure 3.2 shows, given an incoming lot size N, and fixed sample size n, the average leakage rate has a unique maximum. Recall the previous example, where we required the maximum average leakage rate to be 0.03%; we found that for a lot of size N=2500, a sample size of n=823 was required to control the maximum average leakage rate. This leakage rate occurs when the incoming contamination rate is equal to p=0.122%.

We now have the required design parameters to calculate confidence (C) in our sampling; using Equation (B.2) from Appendix B, we find that our confidence is C=70.3. In words, we would say:

> In order to control the maximum average leakage rate of the pathway in consignments of N=2500 at 0.03%, a sample size of n=823 is required. Assuming that the design contamination rate is p=0.122%, a sample size of n=823 from a lot of size N=2500 will give us C=70.3% confidence in detecting contamination on a *per lot basis*, if it is present.

Note that the confidence in a single lot is C=70.3, which is lower than the standard 95% confidence. This is not necessarily an issue, as the lower confidence is for a single lot detection, and based off the assumption of a single design contamination rate of p=0.122%. It is important to note that leakage over multiple lots will be controlled *irrespective* of the true contamination rate.

3.4 Advantages and Disadvantages

In this section, we highlight some of the advantages and disadvantages to using the maximum average leakage rate for the pathway to set sample sizes for testing.

Advantages

By considering the maximum average leakage rate of the pathway, the impact of sampling for testing on a *per lot* basis is made clear. This is useful information, especially in an environment for which managing the risk of a single lot is the dominant paradigm.

Controlling the maximum average leakage rate for the pathway may reduce the sample size requirements for destructive testing; this was demonstrated in the example provided in Section 3.2. However as a result, Section 3.3 shows that this reduced sample size comes at a cost of loss of confidence on a per lot basis.

Controlling the maximum average leakage rate for the pathway does not require specific design parameters to be fixed for each lot — that is, we do not need to assume an

underlying contamination rate, nor specify a level of confidence. Whilst we do acknowledge that setting the maximum average leakage rate for the pathway may possibly be difficult in itself (see Chapter 4), the advantage of doing so comes from careful analysis of the consequences of leakage, and the impacts that the biosecurity system as whole may have on the possible establishment of diseases. As was shown in Section 3.3, we can use this information to find an equivalent set of design parameters.

Disadvantages

Controlling the maximum average leakage rate for the pathway may result in lower confidence for some lots when assuming low contamination rates (as at present). However, the trade-off here, is that for larger lots, the confidence will generally be higher. Changes to (per lot) confidence will require close consultation between industry and the regulator, however this is outside the scope of this project.

Setting the maximum average leakage rate for the pathway may be difficult itself. Consideration will need to be made of the impact of leaked seed, and possible costs to the environment, as well as possible economic losses. A transparent and evidence-based way of setting the leakage rate may be to consider the seed-to-seedling transmission rate of the virus/bacteria in question, as we will explore further in Chapter 4. Clearly, if the transmission risk is lower, it may be possible to allow higher leakage rates, but at what level is uncertain and out of scope for this report, beyond the example given in the next chapter.

Controlling the maximum average leakage rate for the pathway would represent a paradigm shift. An advantage of the current system is that it is well known and accepted — even if it may be seen to penalise small seed lots. The sample size requirements are specified in the International Standards for Phytosanitary Measures 31 [5], Table 1, Appendix 2.

Recommendation 2. *MPI should note the impact of testing on a per lot basis on the long-run leakage of contaminated seeds. Leakage analysis could be considered as a tool to understand risk more generally.*

3.5 Web-based Sample Size Calculator (Leakage)

We have developed a web-based sample size calculator using the method described in this chapter, similar to that in Section 2.3. The application is built using the `shiny` package in R; example code is distributed with this project.

The application is run from a standard R session. Make sure to start the session from the folder in which both the `ui.R` and `server.R` code are located (these are provided in the `leakage-mpi` folder). Use the command `shiny::runApp()` (Code chunk 3.1) from the R prompt; the application will open in your system's default web browser.

R Chunk 3.1.

```
#>shiny::runApp()
```

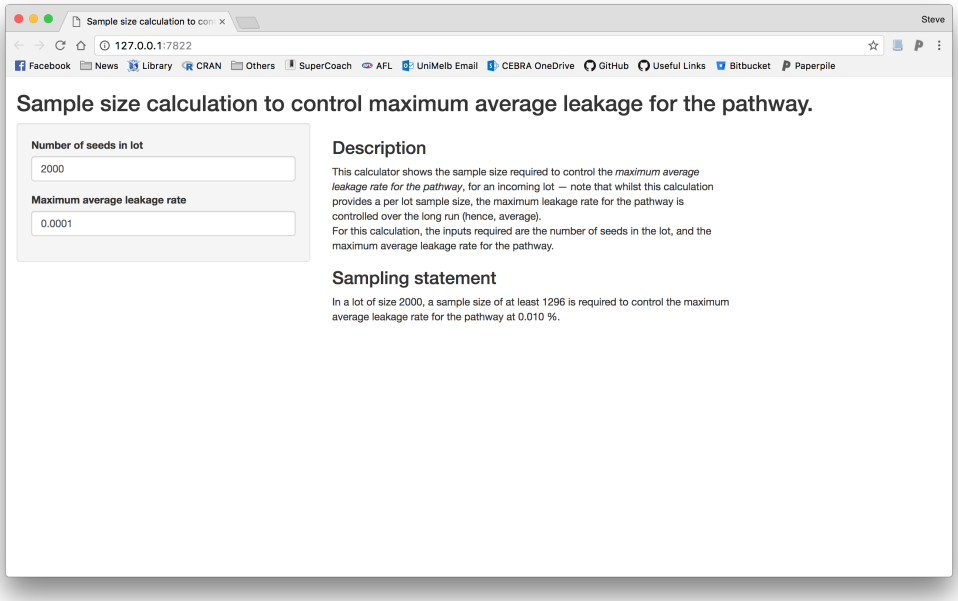


Figure 3.3: Initial screen of the interactive sample size calculator to control the maximum average leakage rate.

The application will open in the default browser, similar to that shown in Figure 3.3. To calculate the sample size to control maximum average leakage rate for the pathway, enter the appropriate parameters into the **Number of seeds in lot** and **Maximum average leakage rate** boxes.

4

Case Study: Setting the Maximum Average Leakage Rate for the Pathway Based on Seed-to-seedling Disease Transmission Rates

In this chapter we describe a method for setting the maximum average leakage rate for the pathway, based in part on a consideration of the seed-to-seedling disease transmission rate.

Consider the situation in which we know that for a particular pathway, there are k lots per year, of equal size N[1]. If the seed-to-seedling transmission rate of a particular disease is t_r, then in combination with the *propagule pressure* of the pathway, $k \cdot N$, we would expect $t_r \cdot k \cdot N$ seeds *capable* of transmitting the disease.

Suppose now that in a one year period, we are willing to allow at most c contaminated seeds to possibly transmit the disease per year[2]. The maximum number of contaminated seeds, c, may be chosen by considering the economic costs of eradication, for example; setting c is outside the scope of this project.

Equating the number of seeds capable of transmitting the disease to the maximum number of seeds we are willing to allow the possibility of transmission provides a mechanism for setting the maximum average leakage rate of the pathway per year, say a_m:

$$a_m = \frac{c}{t_r \cdot k \cdot N} \tag{4.1}$$

Note: In Equation (4.1), setting the maximum average leakage rate of the pathway per year depends upon the *propagule pressure* of the pathway; that is, the volume of seed being imported per year. If this changes, then the maximum average leakage rate of the pathway will change.

Note: For this discussion, we have used one year for the duration over which we consider leakage. This decision was arbitrary, and could change dependent upon the partic-

[1] It is unlikely that these will be exactly equal to N, but we just require them to be close enough.
[2] Remember: zero-risk is undesirable, and unattainable.

ular pathway being considered. Similarly, the risk of establishment could be used in place of the seed-to-seedling transmission rate.

4.1 Example: *Capsicum annuum*

As an example, we will use *Capsicum annuum* and Pepper chat fruit viroid (PCFVd). Let's assume that on this particular pathway, historical records show that there is on average 20 lots of 2500 seeds per year. The seed-to-seedling transmission rate of PCFVd in *Capsicum annuum* is recorded in the Import Risk Analysis: Tomato and Capsicum Seed for Sowing from all Countries (Draft) [6] as 19%. From the preceding information, approximately $0.19 \times 20 \times 2500 = 9500$ seeds would be capable of transmitting PCFVd, if contaminated, per year.

Assume that we are willing to allow the possibility of at most $c = 1$ seed contaminated with PCFVd to transmit the disease per year. We would thus set the maximum average leakage rate for the pathway at $1/9500 = 0.011\%$.

Using the method described Chapter 3, for a lot of size 2500 and average maximum leakage rate of the pathway equal to 0.011%, we should sample 1458 seeds for testing. To clarify: this sample size is calculated for a lot of size 2500; for different size lots, the calculations should be redone.

Note that the maximum average leakage rate of the pathway is dependent on the propagule pressure as described earlier. For example, suppose that instead, the *Capsicum annuum* pathway had on average 20 lots of 10000 seeds per year. Assuming that we are still willing to allow only $c = 1$ possible transmission per year, the maximum average leakage rate of the pathway would drop to 0.003%. The sample size required for testing in this case would now become $n = 5830$.

4.2 Effect of Variation in Lot Size

The consideration of maximum average leakage rate of the pathway for setting sample size has assumed an incoming lot size N that does not vary, or at the least, does not vary greatly. The choice of maximum average leakage rate for the pathway was chosen to address a specific idea of *propagule pressure* over a fixed duration — 20 lots over a year, each lot consisting of 2500 seeds. In this example, a total of 29160 seeds would be sampled over the 20 lots. If instead, we were to consider 10 lots of 5000 seeds, with the same maximum average leakage rate for the pathway, each lot would require a sample size of $n = 2057$ to be tested, for a total of 20570 seeds sampled over the 10 lots.

The discount in total number of seeds is due to the combination of larger lot sizes (and hence fewer number of lots) and larger sample sizes required for testing. For a fixed maximum average leakage rate for the pathway this leads to a higher probability of finding contaminated seed, if present, in the larger lots. Table 4.1 shows this relationship for varying lot sizes and maximum average leakage rate for the pathway.

The derivation so far assumes that the regulator wishes to control the *absolute* number of contaminated seed that may be leaked over some period of time (chosen to be a year in

Table 4.1: Comparison of sample sizes for testing small seed lots by controlling the average maximum leakage rate for the pathway with varying lot sizes. The total amount before testing is held constant at 50000, with the number of lots and lot size varying. The **Sample size** column shows the number of seed required to be tested per lot, and the final column (**Total**) shows the total number of seed to be tested out of 50000.

Number of lots	Lot size	Sample size	Total
20	2500	1458	29160
19	2632	1502	28538
18	2778	1548	27864
17	2941	1597	27149
16	3125	1650	26400
15	3333	1706	25590
14	3571	1767	24738
13	3846	1831	23803
12	4167	1901	22812
11	4545	1976	21736
10	5000	2057	20570
9	5556	2146	19314
8	6250	2242	17936
7	7143	2347	16429
6	8333	2462	14772
5	10000	2590	12950
4	12500	2731	10924
3	16667	2889	8667
2	25000	3066	6132

our example). At first, it may seem that this poses an issue for wildly varying lot sizes; the details in Table 4.1 show that larger lot sizes result in less total seeds tested over the duration. However, it is important to note that the average maximum leakage rate for the pathway will still be maintained, and hence the absolute number of contaminated seed leaked. To see this, note that the derivation actually relies on the *total* number of seed imported over the duration of interest; this is the $k \cdot N$ factor in Equation (4.1). So long as this number is maintained, the absolute number of contaminated seed that may possibly be leaked is maintained.

To demonstrate, let us consider two very different scenarios: 20 lots of equal lot size $N = 2500$, and three lots containing $N = 2500, 10000, 37500$ seeds respectively. Assume that each of these scenarios spans a year, and that we wish to control the maximum average leakage rate of the pathway at 0.011% as per the previous example. Note that in these examples, the total volume is the same under each scenario, i.e. 50000.

We use simulation to test each of these scenarios multiple times, see Appendix F for details of the simulation. In particular, we looked at cases where the underlying contamination rate was one of $\{0.05\%, 0.075\%, 0.1\%, 0.125\%, 0.15\%\}$, and each lot was tested at the required sample size as per Section 3.2[3]. Table 4.2 shows the mean (standard deviation) of the number of leaked contaminated seed per year of the simulation. Recall that the example for the case study (Section 4.1) was to allow at most one contaminated seed transmit the disease per year, where the seed-to-seedling transmission rate of the disease was 19%; this translates to a maximum of $1/0.19 = 5.263$ contaminated seeds allowed to be leaked on average per year.

As is immediately seen in Table 4.2, the mean number of leaked contaminated seed is controlled by the sampling under both scenarios, and for all underlying contamination rates. It is important to note from Table 4.2 that smaller lot sizes generally result in larger leakage — this should not be surprising, as smaller lot sizes will have smaller samples drawn from them. This is important for the regulator however, as it shows that the risk is greater in lots of smaller volume, than in larger lots.

Table 4.2: Comparison of mean (sd) number of leaked contaminated seed over one year, where the total lot size is 50000. Calculations are performed via simulation, with the scenarios representing a consistent flow of lots (Scenario 1, 20 lots of 2500 seeds), and a varying flow (Scenario 2, lots of 2500, 5000, 37500 seeds). The actual incoming contamination rates are shown in the headings above each column.

	Mean (sd) number of leaked contaminated seed				
	0.050%	0.075%	0.100%	0.125%	0.150%
Scenario 1	5.0 (2.5)	5.3 (2.8)	4.8 (2.9)	4.2 (3.0)	3.4 (2.9)
Scenario 2	4.5 (7.2)	3.4 (8.0)	2.1 (7.0)	1.4 (6.4)	0.9 (5.2)

[3]See Appendix E for details

5

Continuous Sampling Plans for Small Seed Lots

A system that relies on per lot sampling ignores pathway information. As an example, an importer may have an incredibly clean pathway, yet is not rewarded by reduced testing effort; the pathway history is ignored, and each new lot is treated independently of the previous lots.

Controlling the maximum average leakage rate — as discussed in Chapter 3 — that can be translated into a per lot basis.

In this chapter we discuss the potential benefits if extra pathway information were made available to MPI. In particular, we discuss a protocol for sampling that uses historical testing information. Similar to a continuous sampling plan (CSP) [10, 11], a possible way to reduce the sample size requirements for small lots would be to set the probability that the sample contains at least one contaminated seed, C[1], with consideration of a supplier's good standing.

5.1 A Modified CSP for Seeds

A continuous sampling plan is a monitoring technique that prescribes: i) an inspection state, where all lots are inspected/sampled, and ii) a monitoring state, where lots are randomly inspected/sampled at a pre-specified rate. A particular version of CSP, CSP-3 [3], has been successfully implemented within the Department of Agriculture and Water Resources (DAWR) Compliance-Based Inspection Scheme CBIS [10, 1, 4, 9]. For low-risk pathways such as dried apricots, green coffee beans and cashews, DAWR uses the CSP-3 to monitor compliance. A description of CSP-3 is found in Appendix G.

A version of a CSP is now presented for monitoring small seed imports. We define the 'enhanced state' as requiring the probability of selecting a contaminated seed to be high, for example the current 95%. The 'monitoring state' would have a reduced probability of selecting a contaminated seed, for example, 80%. All suppliers would start off in the enhanced state. After a fixed number of lots that are not contaminated, suppliers would

[1] Or the contamination rate, p.

be moved to the monitoring state. All lots would require sample sizes for detection at the lower probability setting until a lot tests positive. As soon as a positive test is recorded, the supplier is returned to initial high setting of the enhanced state.

5.2 Requirements

There are some considerable data and information sharing requirements for this protocol to be implemented.

This protocol would require the collection of *all* attempted imports from each supplier. Required fields would include the supplier details; import (and attempted import) details; and the test results. Clearly this requirement would require strong collaboration between MPI, importers and suppliers. A large amount of information is required to be shared, which may be commercially sensitive — trust in the system would need to be demonstrated. There are at least two possible ways this data could be gathered:

- An intention to import notice could be lodged by importers in New Zealand. The lot could then be identified and tested; when the results are known, they are reported back to MPI. Importation is allowed if the tests are 'clean'.
- All testing is performed in New Zealand. All attempted imports are recorded, with 'clean' lots released.

Recommendation 3. *MPI should consider collecting attempted imports data. This would allow imports data to be used in setting sampling protocols.*

Recommendation 4. *MPI should consider using collected import data to simulate a variety of theoretical scenarios to determine the effect of different sampling approaches.*

6

Concluding Remarks

This report has investigated several options that could be available to MPI, that may reduce the destructive sampling burden that is currently of some issue for small seed imports. The key issue at stake is that the current system may be seen to penalise small seed importers, by requiring a (relatively) large sample of seed be withheld from a lot to determine that it is free of disease.

Chapter 2 discussed the use of the Hypergeometric assumption for small seed lots. We acknowledge that this is not novel, however we note that in the current IHS [8] there are no references to either ISPM 31 [5] or internal policies regarding the use of the Hypergeometric distribution for calculating sample sizes. As an extension to this chapter, we also provided a sample application for calculating sample sizes using the Hypergeometric assumption, tailored to MPI-specific requirements.

In Chapter 3 we discussed the idea of leakage. Leakage is of course inherent in any system that does not inspect each and every individual item. This chapter helps demonstrate the impact of leakage, and also provided a technique that could be used to set the design parameters for a *per lot* sampling system such as the one used when assuming Hypergeometric sampling. A case study in Chapter 4 demonstrates a possible method in setting the desired maximum average leakage rate.

Finally, Chapter 5 raised the possibility of using more granular pathway information to set sample size requirements for destructive testing. At the start of this project, it was thought that importers would not be forthcoming in sharing such detailed information. Recent discussions (Thomas, pers. comm.) have suggested that some suppliers/importers may be amenable to such a proposal. Such a proposal would be welcomed as a new proposal for ongoing work.

Acknowledgements

This report is a product of the Centre of Excellence for Biosecurity Risk Analysis (CEBRA). In preparing this report, the authors acknowledge the financial and other support provided by the Australian Department of Agriculture and Water Resources, the New Zealand Ministry for Primary Industries and the University of Melbourne.

Bibliography

[1] Tony Arthur et al. *Statistical Modelling and Risk Return Improvements for the Plant Quarantine Pathway*. Tech. rep. Australian Centre of Excellence for Risk Analysis, Nov. 2013.

[2] R Beale et al. *One Biosecurity: a Working Partnership*. Commonwealth of Australia, 2008.

[3] Harold F Dodge and Mary N Torrey. "Additional continuous sampling inspection plans". In: *Industrial Quality Control* 7.5 (1951), pp. 7–12.

[4] Jessica Holliday and Andrew Robinson. *Ornamental Finfish Import Reform Program: A sampling framework and trial for the surveillance program*. Tech. rep. Australian Centre of Excellence for Risk Analysis, May 2013.

[5] Internation Plant Protection Convention. *International standards for phytosanitary measures: ISPM 31, methodologies for sampling of consignments*. {Food and Agriculture Organization of the United Nations, Rome, Italy}. 2008.

[6] Ministry of Agriculture and Forestry. *Draft Import Risk Analysis: Tomato and Capsicum seed for sowing from all countries*. Apr. 2012.

[7] New Zealand Ministry for Primary Industries. *Import Health Standard: Seeds for Sowing*. Tech. rep. 155.02.05. Wellington, Oct. 2016.

[8] New Zealand Ministry for Primary Industries. *Import Health Standard: Seeds for Sowing*. Tech. rep. 155.02.05. Wellington, June 2017.

[9] Andrew Robinson and Yuko Hood. *Risk-Based Management for Imported Ornamental Fin-fish.* Tech. rep. Centre of Excellence for Biosecurity Risk Analysis, Apr. 2015.

[10] Andrew Robinson et al. *DAFF Biosecurity: Plant-Product Pathways*. Tech. rep. 1001J 1. Australian Centre of Excellence for Risk Analysis, May 2012.

[11] Andrew Robinson et al. *Plant Quarantine Inspection and Auditing across the Biosecurity Continuum*. Tech. rep. 1101C 1. Australian Centre of Excellence for Risk Analysis, Sept. 2013.

Appendix A

Derivation of Sample Size under Binomial Sampling

In this appendix we detail the derivation of the sample sizes as specified in the IHS for testing a seed lot quarantine pests. As discussed in Section 1.2, the current system assumes sampling with replacement, for which it follows that the number of contaminated seeds in a sample follows a Binomial distribution: $X \sim \text{Binomial}(n, p)$, where n is the size of the sample, and p is the probability that a seed is contaminated.

To find the required sample size, we specify the minimum probability (C) of detecting at least one contaminated seed in the sample. Mathematically, we want $\Pr(X \geq 1; n, p) = C$; equivalently, $1 - \Pr(X = 0; n, p) = C$. From the Binomial distribution, we find:

$$\begin{aligned} C &= 1 - \Pr(X = 0; n, p) \\ &= 1 - (1-p)^n \\ n \log(1-p) &= \log(1-C) \\ \Rightarrow n &= \log(1-C)/\log(1-p) \end{aligned} \tag{A.1}$$

In the example given in Section 1.2, we specify $p = 0.1\%$ and $C = 0.95$. Plugging these in to Equation (A.1), we find $n = 2994.234$, which is commonly rounded up to $n = 3000$.

A.1 Imperfect Detection

As noted in Section 1.2.1, the efficacy of the detection method may be less than 100%. In this case, it is a simple matter of adjusting the contamination rate p, for the level of efficacy. For example, letting the efficacy of the detection method by ϵ, the adjustment to Equation (A.1) becomes $n = \log(1-C)/\log(1-p \cdot \epsilon)$.

As an example, suppose that we knew that efficacy was ϵ =90%. This increases the required sample to $n = 3327.093$.

Appendix B

Derivation of Sample Size under Hypergeometric Sampling

The distribution of the number of contaminated seeds selected for sampling is correctly described by a Hypergeometric distribution, as opposed to the Binomial distribution assumed at present. This is particularly important when considering small seed lots.

In order to use the Hypergeometric distribution to calculate the needed sample size, we need to specify the *number*[1] of contaminated seeds in the lot, whereas using the Binomial distribution (Appendix A) requires the contamination *rate*. Thus, when adjusting current testing regulations, we will need to take the number of contaminated seeds in the lot as $m = \lfloor pN \rfloor$, where N is the number of seeds in the lot, and p is the design contamination rate as set by MPI. The rounding is required as the Hypergeometric distribution is a discrete distribution that amounts to sampling without replacement from a lot — clearly we can't have a fractional number of contaminated seeds present!

Let X be the number of contaminated seeds present in a sample of size n, that is drawn from a lot of size N containing m contaminated seeds. Then $X \sim \text{Hypergeometric}(x, m, N, n)$, with

$$\Pr(X = x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}, \text{ for } x \in 0, 1, \ldots, \min\{m, n\} \tag{B.1}$$

To determine the number of seeds that we need to sample from a lot for a given confidence C, we want to choose n such that the probability of including at least one contaminated seed in the sample is greater than or equal to C, assuming perfect detectability. Thus, we require a solution for n to the following:

$$\begin{aligned} C &\leq \Pr(X \geq 1) \\ &= 1 - \Pr(X = 0) \\ &= 1 - \frac{\binom{N-m}{n}}{\binom{N}{n}} \end{aligned} \tag{B.2}$$

[1] We note though, that this is in fact determined from the rate of contamination and the lot size.

A solution for n to Equation (B.2) can be found by a simple search over possible values for n. Alternatively, a reasonable approximation is

$$n \approx \left[1 - (1 - C)^{1/m}\right] [N - (m - 1)/2] \,. \tag{B.3}$$

The Hypergeometric approach is essentially a 'business as usual' option, in that the only alteration to the current approach is to require a smaller sample size, as permitted by statistical theory. Again we stress that whilst the sample size is reduced, quality (in the form of confidence) is maintained.

Appendix C

Comparison of Sample Sizes Between Binomial and Hypergeometric Methods

In this appendix we compare sample sizes calculated using: the Binomial approximation, as per the current system (Section 1.2); and the Hypergeometric as in Chapter 2. We do this for the standard design contamination rates of 0.1% and 0.15% for viruses, and 0.75% for bacteria.

For large lots and 95% probability of the sample containing at least one contaminated seed, sample sizes of 3000 for $p = 0.1\%$, 2000 for $p = 0.15\%$, and 400 for $p = 0.75\%$ are required to be tested.

Table C.1: Sample sizes for testing small seed lots using the Hypergeometric method of Chapter 2. The sample size using the Binomial approximation of Section 1.2 is shown at as the last row in the table.

	Required sample size under Hypergeometric assumption		
Lot size, N	$p = 0.1$	$p = 0.15$	$p = 0.75$
500	475	475	316
1000	950	950	348
1500	1425	1165	357
2000	1553	1263	361
2500	1941	1579	382
3000	1895	1581	381
3500	2210	1577	380
4000	2108	1572	379
4500	2372	1768	390
5000	2253	1740	388
7500	2611	1787	390
10000	2588	1810	391
15000	2715	1909	395
20000	2781	1900	394
Binomial	3000	2000	400

Note in Table C.1 that sample sizes do not grow linearly with increasing N, and may in fact get smaller[1]. This is because we need to use integer values for the number of contaminated seed in a lot using the method in Chapter 2. Section C.1 below explores this behaviour in detail.

C.1 Non-monotone Behaviour of Sample Sizes Using the Hypergeometric Distribution

An undesirable feature of using the Hypergeometric distribution method of Chapter 2, is the non-monotone behaviour of calculated sample sizes. As discussed earlier in this appendix, the contamination rate is used to calculate the *number* of contaminated seeds, which is required to be an integer. If the number of calculated contaminated seeds is not an integer, we round down so that the sample size is more conservative. This causes calculated sample sizes in some cases to be smaller even when the lot size is larger. Increased lot sizes would be expected to lead to increased sample sizes, so this non-monotone behaviour is undesirable.

Figure C.1 displays this behaviour for lot sizes ranging from 1999 to 6000, and a targeted contamination rate of 0.1%. For a lot of size 1999, the number of contaminated seeds is 1.99, which gets rounded down to 1; this leads to a conservative (inflated) sample size of $n = 1900$. If the lot contains one more seed, for a lot size of 2000, then the sample size required is $n = 1553$ — a difference of 347!

There are two possibilities to combat the non-monotone behaviour, each of which is shown in Figure C.1. The first change is to use a step function: when the sample size drops, hold the calculated sample size at the last maximum. This change is shown as a green line in Figure C.1. The alternative change is to use linear interpolation: where sample sizes begin to drop, interpolate between two consecutive local maximums. This change is shown as a blue line in Figure C.1.

[1]For example, the sample size required to detect at least 1 contaminated seed at 0.1% contamination rate with 95% confidence in a lot of size 2500 is 1941 and for a lot of size 3000 is 1895.

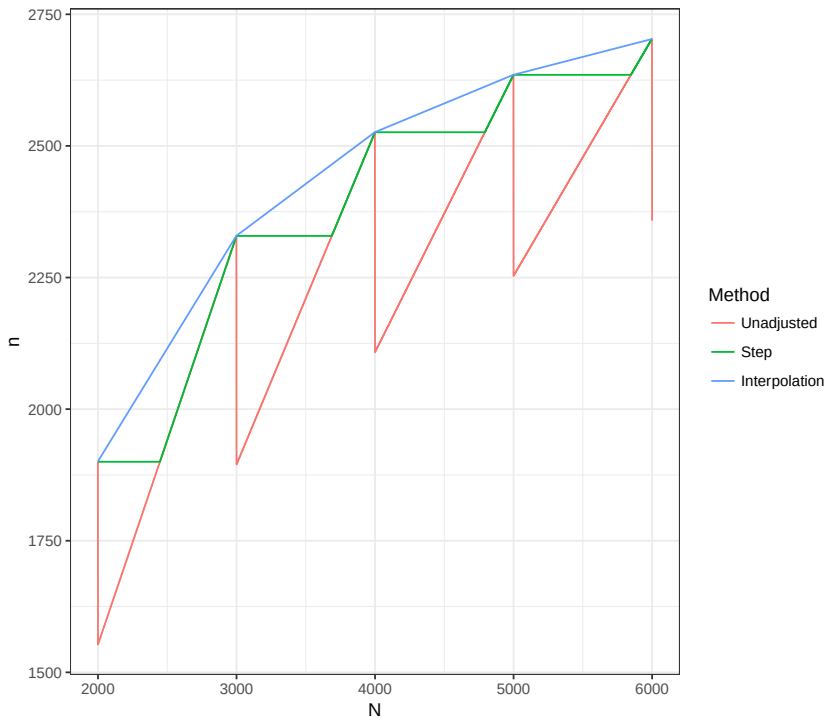


Figure C.1: Non-monotone behaviours of sample sizes calculated using the Hypergeometric dsitribution.

Appendix D

Derivation of Average Leakage Rate

In this appendix we derive the expected number of leaked seeds ($p \times (N-n)$, Section 3.1), and consequently the average leakage rate.

We first start with some definitions. Let the number of contaminated seeds in a lot be equal to M, where M is a random variable; we assume that the rate of contamination, p is fixed, and do not specify the distribution of M, other than it depends on p, through the probability density function $\Pr(M = m)$.

Next, let the number of contaminated seeds selected in the sample be X. We assume that the lot size N, is small relative to the sample size n, and assume that X has a Hypergeometric distribution (Chapter 2), $X \sim \text{Hypergeometric}(N, n, M)$, where $\Pr(X = x)$ is as in Equation (B.1). We further assume that $M < n$[1]. We further define a random variable Z, which is the event that a lot is released for importation:

$$Z = \begin{cases} 1 & \text{with probability } \Pr(X = 0) \\ 0 & \text{with probability } \Pr(X > 0) \end{cases} \tag{D.1}$$

The probabilities in Equation (D.1) are by definition. There are two possible outcomes resulting from testing the sample of n: if the test is positive, we have detected that the lot is contaminated, and thus no seed is allowed to be imported. On the other hand, if the test is negative, the lot is determined to be not contaminated, and is subsequently released for importation.

Now, let l be the *number* of leaked seeds;

$$l = \begin{cases} M, & Z = 1 \\ 0, & Z = 0 \end{cases} \tag{D.2}$$

and we are required to find $E(l|Z)$. Now, from Equation (D.2), we have that $E(l|Z) = \Pr(Z = 1)E(M|Z = 1) + \Pr(Z = 0) \times 0$. From Equation (D.1), $\Pr(Z = 1) = \Pr(X = 0)$, which we put aside for the moment.

[1] This is a mathematical convenience, but in applications is likely to be true. Furthermore, if $M > n$, detection is almost guaranteed

Considering $E(M|Z=1)$, we have:

$$\begin{aligned}
E(M|Z=1) &= \sum_{m=0}^{N} m \cdot \Pr(M=m|Z=1) \\
&= \sum_{m=0}^{N} m \cdot \sum_{x=0}^{M} \Pr(M=m|Z=1, X=x) \cdot \Pr(X=x|Z=1) \\
&= \sum_{x=0}^{M} \sum_{m=0}^{N} m \cdot \Pr(M=m|Z=1, X=x) \cdot \Pr(X=x|Z=1) \\
&= \sum_{x=0}^{M} \Pr(X=x|Z=1) \sum_{m=0}^{N} m \cdot \Pr(M=m|Z=1, X=x) \\
&= \sum_{x=0}^{M} \Pr(X=x|Z=1) E(K|Z=1, X=x) \\
&= \sum_{x=0}^{M} \Pr(X=x|Z=1)[p \cdot (N-n) + x] \\
&= p \cdot (N-n) \sum_{x=0}^{M} \Pr(X=x|Z=1) + \sum_{x=0}^{M} x \cdot \Pr(X=x|Z=1) \\
&= p \cdot (N-n) + \sum_{x=0}^{M} x \cdot \Pr(X=x|Z=1) \\
&= p \cdot (N-n)
\end{aligned} \tag{D.3}$$

Combing Equation (D.3) with Equation (D.2), we have the average leakage rate as:

$$\begin{aligned}
E(l|Z) &= \Pr(X=0) \cdot p \cdot \frac{N-n}{N} \\
&= \frac{\binom{N-M}{n}}{\binom{N}{n}} \cdot p \cdot \frac{N-n}{N}
\end{aligned} \tag{D.4}$$

Unfortunately, we are still left with the dependence on M in Equation (D.4). To complete the analysis, we need to make an assumption about the distribution of M, and we can then find the expectation of Equation (D.4) with respect to M. By Jensen's inequality, $g(E_M(l|Z) \leq E_M(g(M))$, so an upper bound for the leakage rate is

$$E(l|Z) = \frac{\binom{N-E(M)}{n}}{\binom{N}{n}} \cdot p \cdot \frac{N-n}{N}. \tag{D.5}$$

Assuming M has a Binomial distribution, $E(M) = p \cdot N$; if this is not an integer, we have two options: approximate $\Pr(X=0) \approx (1-p)^n$ (i.e. the Binomial approximation to the Hypergeometric); or linearly interpolate two Hypergeometric probabilities. Because we are calculating probabilities on the continuous scale, the Binomial approximation should be sufficient.

Appendix E

Algorithm for Sample Size to Control Average Leakage Rate

In this appendix we present a derivation of sample size to control the maximum average leakage rate. From Equation (D.5), we can calculate the p which gives the largest average leakage rate, given N and n. This is a standard optimisation problem and amounts to taking the derivative of Equation (D.5) with respect to p, setting to 0, and solving for 0. This is simply $p_{\max} = 1/(n+1)$; that is, for a fixed sample size n, the largest average leakage rate occurs if $p = p_{\max}$. Substituting this value back in to Equation (D.5) gives:

$$\begin{aligned} \max E(a) &= p_{\max}\frac{N}{N-n}(1-p_{\max})^n \\ &= \frac{1}{n+1}\frac{N}{N-n}\left(1-\frac{1}{n+1}\right)^n \end{aligned} \tag{E.1}$$

To find the optimal sample size n, subject to $\max E[a]$ being less than the allowable leakage, we do the following:

1. Set $n_0 = 1$;
2. Calculate $\max E(a)|n_0$ (by plugging n_0 into Equation (E.1).
3. Repeat the following until $\max E(a)|n_k$ is less than the allowable leakage rate:
   (a) Set $n_k = n_{k-1} + 1$;
   (b) Calculate $\max E(a)|n_k$.
4. Use n_k as calculated in Step 3.(b) as the optimal sample size.

Appendix F

Simulation Study Details for Control of Maximum Average Leakage Rate for the Pathway

In this chapter we provide details of the simulation study used to investigate the control of maximum average leakage rate for the pathway. In particular, we investigate how varying lot sizes may affect the overall leakage rate. In Section 4.2 we investigated two very different scenarios: 20 lots of equal lot size $N = 2500$, and three lots containing $N = 2500, 10000, 37500$ seeds. We assumed that each of these scenarios plays out over a year, and that we wish to control the maximum average leakage rate of the pathway at 0.011% as per the example in Section 4.1.

For each of the underlying contamination rates considered ($\gamma \in \{0.05\%, 0.075\%, 0.1\%, 0.125\%, 0.15\%\}$), we do the following:

1. For each lot in each scenario, we generate contaminated seed at rate γ, according to a Binomial distribution, $X \sim \text{Binomial}(N, \gamma)$, where X is the number of contaminated seed in the lot, and N is the size of the lot.

   For example, for the lot with 10000 seeds from Scenario 2, with $\gamma = 0.1\%$, we draw x contaminated seed from a Binomial$(10000, 0.1\%)$ distribution; thus the lot has x contaminated seed, and $(10000 - x)$ seed free of contamination.

2. We then simulate drawing a sample for testing from the lot, at the required sample size. For 10000 seeds, and maximum average leakage rate for the pathway at 0.011%, we require 2590 seeds for sampling.

3. We then count whether any contaminated seed are within this sample; if yes, the lot is rejected, and 0 contaminated seeds are released. If no, then the lot is accepted, and x contaminated seeds are released.

4. The number of contaminated seed in each lot are then summed, to give the total number of leaked contaminated seed in each scenario per year.

We repeated the above steps a large (4999) number of times, and calculated the mean and standard deviation of the number of leaked contaminated seed per year.

Appendix G

Continuous Sampling Plans

A continuous sampling plan is a monitoring technique that prescribes: i) an inspection state, where all lots are inspected/sampled, and ii) a monitoring state, where lots are randomly inspected/sampled at a pre-specified rate. In this Appendix, we describe a particular version of CSP, CSP-3 [3] in more detail. This description is an abridged version of [10]; see that document for more details.

The decision to be made is how a regulator would utilise the inspection history of a particular product/pathway. Assuming that a small amount of leakage is allowable, the CSP-3 allows a pathway to be in a monitoring state. In the monitoring state, the pathway is inspected at a reduced rate, for example each lot is inspected at random with probability f. A pathway would generally start off in the enhanced state, in which every lot is inspected/sampled. The CSP-3 is then:

1. Enhanced mode: inspect/sample all lots, until a certain number of compliant lots have been observed (for example, i consecutive lots pass inspection). Now, switch to:
2. Monitoring mode: inspect each lot at random, with a probability of f. If the lot fails inspection, the next four lots are all inspected. If all four lots pass inspection, stay in monitoring mode. If another lot fails inspection within k inspections, return to enhanced mode.

There are three parameters which control the behaviour of CSP-3: i, the number of consecutive lots required to pass inspection; k, the lag between failed inspections which signals a switch back to enhanced mode. Usually, $k = i$; and f, the rate at which lots are inspected during monitoring mode.

The values of i and f determine the *evidence* required to switch to monitoring mode, and the chances of detecting underlying changes in the rate of failures. A higher value of i can be interpreted as requiring more evidence that a pathway is 'clean', before switching to monitoring mode. A higher value of f results in a higher chance of detecting changes in the underlying failure rate.

There are clear trade-offs in setting the two parameters. Higher values of both i and f will mean more sampling, hence a higher cost. Lower values will mean less sampling, and may increase the amount of contaminated lots passing through the system. Setting these parameters can be achieved via simulation; see [10] for details.