



Prepared for the U.S. Department of Energy  
under Contract DE-AC05-76RL01830

PNNL-16636

# An Environmental Sampling Model for Combining Judgment and Randomly Placed Samples

LH Sego  
KK Anderson  
BD Matzke  
K Sieber  
S Shulman

J Bennett  
M Gillen  
JE Wilson  
BA Pulsipher

August 2007



**Pacific Northwest**  
NATIONAL LABORATORY

*Proudly Operated by **Battelle** Since 1965*

## DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY

*operated by*

BATTELLE

*for the*

UNITED STATES DEPARTMENT OF ENERGY

*under Contract DE-AC05-76RL01830*

Printed in the United States of America

Available to DOE and DOE contractors from the  
Office of Scientific and Technical Information,  
P.O. Box 62, Oak Ridge, TN 37831-0062;  
ph: (865) 576-8401  
fax: (865) 576-5728  
email: [reports@adonis.osti.gov](mailto:reports@adonis.osti.gov)

Available to the public from the National Technical Information Service,  
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161  
ph: (800) 553-6847  
fax: (703) 605-6900  
email: [orders@ntis.fedworld.gov](mailto:orders@ntis.fedworld.gov)  
online ordering: <http://www.ntis.gov/ordering.htm>



This document was printed on recycled paper.

(9/2003)

# An Environmental Sampling Model for Combining Judgmental and Randomly Placed Samples

LH Sego <sup>1</sup>	J Bennett <sup>2</sup>
KK Anderson <sup>1</sup>	M Gillen <sup>2</sup>
BD Matzke <sup>1</sup>	JE Wilson <sup>1</sup>
K Sieber <sup>2</sup>	BA Pulsipher <sup>1</sup>
S Shulman <sup>2</sup>	

1. Pacific Northwest National Laboratory
2. National Institute for Occupational Safety and Health

August 2007

Prepared for the U.S. Department of Energy  
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory  
Richland, Washington 99352

**Author's Note**

The Combined Judgmental and Random sampling methodology discussed herein is a work in progress. This report describes the progress that was completed as of August, 2007. Additional refinements and investigations of the model that are currently underway will be made available in subsequent reports and/or journal articles.

Landon Sego  
July 13, 2009

# An environmental sampling model for combining judgmental and randomly placed samples

Landon Sego<sup>1</sup>, Kevin Anderson<sup>1</sup>, Stanley Shulman<sup>2</sup>, Brett Matzke<sup>1</sup>, Brent Pulsipher<sup>1</sup>,  
Karl Sieber<sup>2</sup>, James Bennett<sup>2</sup>, Matt Gillen<sup>2</sup>, John Wilson<sup>1</sup>

1. Pacific Northwest National Laboratory
2. National Institute for Occupational Safety and Health

August, 2007

## Summary

In the event of the release of a lethal agent (i.e. *Bacillus anthracis*) inside a building, law enforcement and public health responders take samples to identify and characterize the contamination. Sample locations may be rapidly chosen based on available incident details and professional judgment. To achieve greater confidence of whether or not the decision area was contaminated, or to certify that detectable contamination is not present after decontamination, we developed a Bayesian model for combining the information gained from both judgmental and randomly placed samples. The approach permits investigators to determine the sample size required to achieve a defined level of statistical confidence that a decision area is free of detectable contamination, when none of the samples indicate the presence of contamination. We present herein the model, an example of how it may be used, a discussion of its sensitivity to parameter inputs, and recommendations for its practical use.

**Key words:** environmental sampling; judgment sampling; acceptance sampling; Bayesian modeling; anthrax; bioterrorism.

## 1 Introduction

In the event of the release of a lethal biological or chemical agent inside a building, emergency responders take samples to identify and characterize the contamination. The information obtained from sampling is used to assess the risk to public health, formulate and validate remediation plans, and eventually, to certify the safety of the building for reoccupancy. During the initial response to the release, sample locations are rapidly chosen based on professional judgment. For example, investigators at the National Institute of Occupational Safety and Health (NIOSH) successfully used professional judgment approaches to quickly identify the presence of *Bacillus anthracis* contamination during the 2001 anthrax incidents. Judgment samples are usually taken from areas where detectable contamination is deemed most likely to exist. However, decisions that are based solely upon the results of judgment samples rely heavily on the accuracy of the "judgment." To achieve greater confidence of whether a room or zone was

contaminated, or to certify the cleanliness of a room after decontamination, it is desirable to augment judgment samples with samples that are taken from randomly selected locations.

Sometimes judgment samples are referred to as "targeted" samples. We use the two terms interchangeably. We define a judgment (or targeted) sample to be any sample whose location is determined by prior belief or knowledge and not in a random fashion. Random samples are sometimes called "probabilistic" samples. This terminology implies the commonly held belief that probabilistic statements can *only* be made based upon the results of randomly placed samples. However, in this paper we present a Bayesian model that gives rise to a formal probabilistic statement regarding the cleanliness of the sampling area based upon the information obtained from both judgment and random samples. We presume that judgment samples are taken in areas that are more likely to be contaminated than areas available for random sampling. Consequently, if none of the judgment samples reveal detectable contamination, we can leverage that information to reduce the number of random samples required to achieve the desired level of confidence that the room is clean.

The methodology presented herein is especially suited for use in an area of the building where contamination is deemed unlikely. This may be because the area is removed from the source of contamination or because the area has been decontaminated. In either case, the objective is to demonstrate, with high probability, that a high percentage of the room does not contain detectable contamination, given that none of the samples reveal contamination. If at any time during the sampling process, one of the samples indicates the presence of contamination, the room is declared to be dirty and no further samples for this particular design will be taken. It may be desirable to then implement a different type of sampling plan to characterize the spread and location of the contamination. However, this is a separate problem and is not discussed here.

This paper is arranged as follows. In Section 2, we present the Bayesian model and methodology for combining judgment and randomly placed samples. In Section 3, we present an example of how the methodology could be used inside a small office. In Section 4, we examine the impact of model assumptions and compare the model to existing approaches. In Section 5, we provide recommendations for application of the methodology and give our conclusions.

## 2 Methodology

To implement the sampling design, a grid of square cells will be superimposed on the surfaces of the room. Each square cell in the grid corresponds to a possible sample point, and the size of the cells depends on the collection method. For example, the National Institute of Occupational Safety and Health (NIOSH 2002) describes three different types of surface sampling: 1) sterile swab sampling, which uses about 100 cm<sup>2</sup> per sample, 2) sterile surface wipes, which uses about 1 ft<sup>2</sup> (929 cm<sup>2</sup>) per sample, and 3) vacuum sampling, which may involve considerably larger surface areas. The measured outcome will be presence or absence of contamination, as determined by the limit of detection. At the present time we do not account for classification errors (false positives or false negatives) in this methodology. Judgment samples are considered more likely to contain contamination than random samples. Therefore, the locations of the judgment samples describes a rough, dichotomous dispersion model, where the cells chosen for

judgment samples are modeled as “higher risk” and the remaining squares available for random sampling are “lower risk.” The high-risk cells do not necessarily have to be contiguous. We assume that the judgment samples form a complete census of the high-risk cells.

The proposed method uses a Bayesian approach to determine the posterior predictive distribution of the number of unsampled grid cells that are contaminated, given that all judgment and random samples do not indicate contamination. A summary of the notation used to describe the model is presented in Table 1.

**Table 1:** Summary of model notation

Symbol	Description	Range
$\theta$	probability that a judgment sample is contaminated	$0 < \theta < 1$
$\alpha$	shape parameter of the Beta distribution	$> 0$
$\beta$	another shape parameter of the Beta distribution	$> 0$
$p(\theta)$	Beta density function with shape parameters $\alpha$ and $\beta$	$\geq 0$
$r$	a judgment sample is $r$ times more likely to be contaminated than a random sample	$\geq 1$
$P_J$	the <i>a priori</i> marginal probability that all judgment samples will come back negative: $P(X = 0)$	$0 \leq P_J \leq 1$
$n_1$	number of judgment samples	positive integer
$n_2$	number of random samples	positive integer
$N$	total number of grid cells in the room	positive integer
$X$	number of judgment samples that indicate contamination	$0, 1, \dots, n_1$
$Y$	number of random samples that indicate contamination	$0, 1, \dots, n_2$
$Z$	number of unsampled, lower risk grid cells that are contaminated	$0, 1, \dots, N - n_1 - n_2$
$\gamma$	desired proportion of room that does not contain detectable contamination	$0 < \gamma \leq 1$
$C$	desired probability (confidence) that $100 \times \gamma\%$ of the room does not contain detectable contamination	$0 < C \leq 1$

The outcome of each sample is modeled as a Bernoulli random variable, where the probability that contamination is detected in a judgment sample is  $\theta$ , and the probability that contamination is detected in a random sample is  $\theta/r$  (for  $r \geq 1$ ). Because we assume the Bernoulli outcomes are independent, the number of judgment cells that indicate contamination follows a binomial distribution:  $X|\theta \sim \text{Bin}(n_1, \theta)$ . The distribution of the number of random samples that indicate contamination is also binomial:  $Y|\theta \sim \text{Bin}(n_2, \theta/r)$ . Likewise, the distribution of the number of

unsampled, low-risk grid cells that are contaminated is binomial:  $Z|\theta \sim \text{Bin}\left(N - n_1 - n_2, \theta/r\right)$ .

The beta distribution with shape parameters  $\alpha$  and  $\beta$  is used to model the prior distribution of  $\theta$ . In the present model,  $r$  is fixed. However, the methodology could be extended to include a prior distribution for  $r$ .

The prior distribution for  $\theta$  is specified by the shape parameters  $\alpha$  and  $\beta$  of the Beta distribution. During clearance sampling we typically assume the sampling area has a low likelihood of being contaminated. Consequently, we desire to put more prior weight on values of  $\theta$  that are close to 0 and less weight on values of  $\theta$  that are close to 1. Fixing  $\alpha = 1$  and choosing a suitable  $\beta > 1$  results in a family of decreasing, convex Beta densities that puts more weight toward 0. (Note  $\alpha = 1$  and  $\beta = 1$  results in a uniform distribution). Letting  $\alpha < 1$  and  $\beta > 1$  also produces decreasing, convex Beta densities, but this can result in so much weight being placed very close to 0 that it becomes difficult to numerically integrate the marginal and posterior predictive densities. We recommend letting  $\alpha = 1$  and  $\beta \geq 1$  because 1) this parameter space gives rise to marginal and posterior predictive densities that can be integrated numerically, and 2) it is more conservative than letting  $\alpha < 1$ .

There are a number of ways that prior knowledge about the sampling area could be used to elicit the value of  $\beta$ . We recommend estimating the *a priori* probability that all judgment samples will not indicate contamination,  $P(X = 0) \equiv P_J$ , and then solve for the value of  $\beta$  that satisfies the equation. Another approach would be to estimate the probability the entire room does not contain detectable contamination,  $P(X = 0, Y = 0, Z = 0)$  and solve for the corresponding value of  $\beta$ .

The second option is less desirable because it can easily lead to priors that dominate the observed data by producing Beta densities that are nearly degenerate at zero. Using the first option will make it less likely to overstate the *a priori* confidence in the cleanliness of the sampling area. Having thus specified  $P_J$ , the value of  $\beta$  is given by the solution to the following equation:

$$\begin{aligned} P_J = P(X = 0) &= \int_0^1 P(X = 0|\theta) p(\theta) d\theta \\ &= \int_0^1 (1-\theta)^{n_1} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} d\theta \\ &= \frac{\Gamma(\alpha + \beta)\Gamma(\beta + n_1)}{\Gamma(\beta)\Gamma(\alpha + \beta + n_1)} \end{aligned} \quad (2.1)$$

When  $\alpha = 1$ , equation (2.1) is readily solved algebraically:

$$\beta = \frac{n_1 P_J}{1 - P_J}. \quad (2.2)$$



Using Bayes' theorem, we can derive the posterior predictive distribution of  $Z$ , the number of unsampled, lower risk grid cells that are contaminated, given that none of the judgment or random samples reveal contamination. This distribution can then be used to determine the number of random samples required to achieve a high probability,  $C$ , that at least a large fraction,  $\gamma$ , of the sampling area does not contain detectable contamination. In particular, we want to find the smallest number of random samples,  $n_2$ , that satisfies the following:

$$P\left(Z \leq (1-\gamma)N \mid X=0, Y=0\right) = \frac{\int_0^1 \theta^{\alpha-1} (1-\theta)^{\beta+n_1-1} \sum_{z=0}^{\lfloor (1-\gamma)N \rfloor} \left[ \binom{N-n_1-n_2}{z} \left(\frac{\theta}{r}\right)^z \left(1-\frac{\theta}{r}\right)^{N-n_1-z} \right] d\theta}{\int_0^1 \theta^{\alpha-1} (1-\theta)^{\beta+n_1-1} \left(1-\frac{\theta}{r}\right)^{n_2} d\theta} \geq C \quad (2.3)$$

where  $\lfloor (1-\gamma)N \rfloor$  denotes the floor of  $(1-\gamma)N$ , i.e. the largest integer that is less than or equal to  $(1-\gamma)N$ . The derivation of equation (2.3) is given in the Appendix.

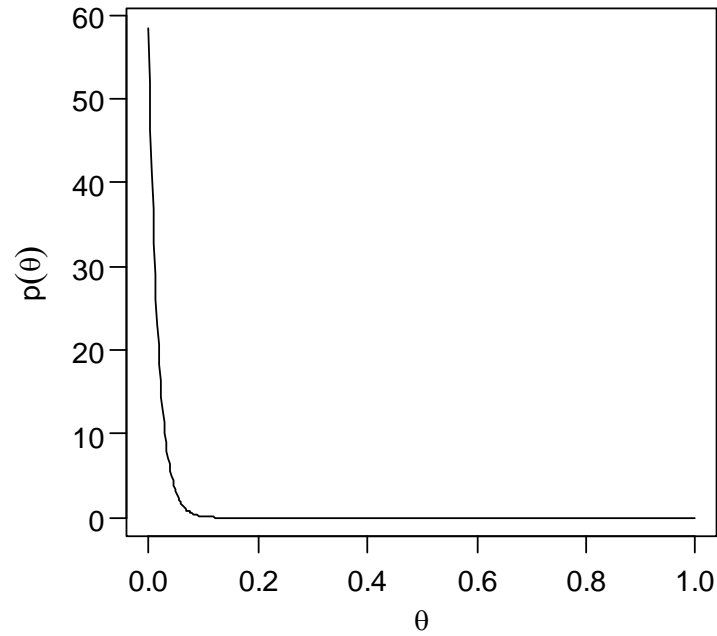
### 3 Example

Suppose an  $8 \times 8 \times 8$  ft room we wish to sample has been divided into 3568 square cells of size  $100 \text{ cm}^2$ . Assuming none of the samples indicate contamination, we wish to determine the number of random samples required to obtain a 95% probability that 99% of the room is clean. Suppose we identify  $n_1 = 25$  high-risk grid cells from which we will take judgment samples. We will assume judgment cells are  $r = 3$  times more likely to be contaminated than the remaining cells that will be used for random sampling. Since we have a fairly high level of *a priori* confidence that the room is clean, we will also suppose there is a 70% chance the judgment samples will contain no detectable contamination, i.e.  $P_J = 0.70$ . Using equation (2.2), we find that  $\beta = 58\frac{1}{3}$ . The resulting prior density with  $\alpha = 1$  and  $\beta = 58\frac{1}{3}$  is shown in Figure 1.

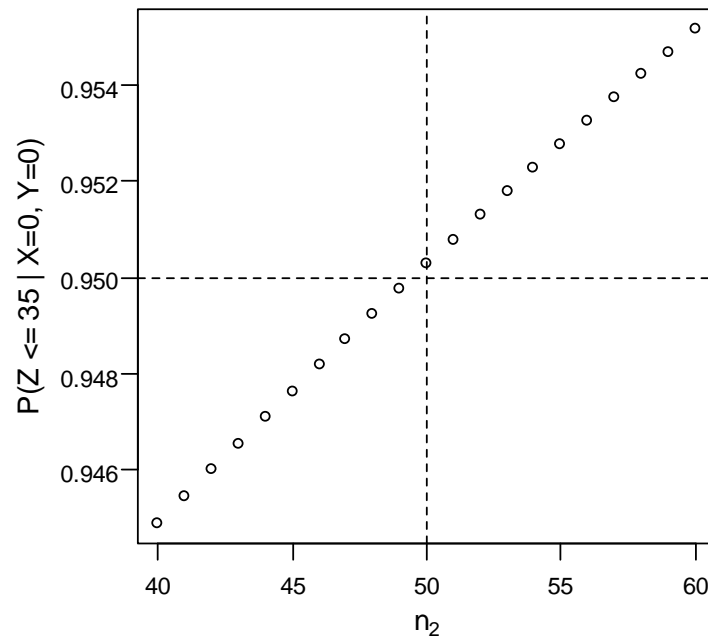
Note that the prior puts most of the probability close to 0, which is consistent with our original assumption that the room has a low risk of being contaminated. Now that the prior has been chosen, using equation (2.3), we can identify the number of random samples,  $n_2$ , that satisfies

$$\begin{aligned} P(Z \leq (1-\gamma)N \mid X=0, Y=0) &= P(Z \leq (1-0.99) \cdot 3568 \mid X=0, Y=0) \\ &= P(Z \leq 35 \mid X=0, Y=0) \geq 0.99. \end{aligned}$$

Figure 2 demonstrates the relationship between  $n_2$  and  $P(Z \leq 35 \mid X=0, Y=0)$ .



**Figure 1:** Prior density for  $\theta$  when  $\alpha = 1$  and  $\beta = 58\frac{1}{3}$ .



**Figure 2:** Relationship between  $n_2$  and  $P(Z \leq 35 | X = 0, Y = 0)$  when  $n_1 = 25$ ,  $\alpha = 1$ ,  $\beta = 58\frac{1}{3}$ ,  $r = 3$ , and  $N = 3568$ . The plot demonstrates  $n_2 = 50$  is the smallest number of random samples that satisfies  $P(Z \leq 35 | X = 0, Y = 0) \geq 0.950$ .

To summarize this example, if we make the *a priori* assumption that 1) the probability that all 25 judgment samples will be negative is 0.70 and 2) judgment samples are 3 times more likely to be contaminated than random samples, and, if, after sampling, none of the 25 judgment samples nor the 50 random samples indicate contamination, then we conclude there is a 95% probability that at least 99% of the room does not contain detectable contamination.

## 4 Discussion

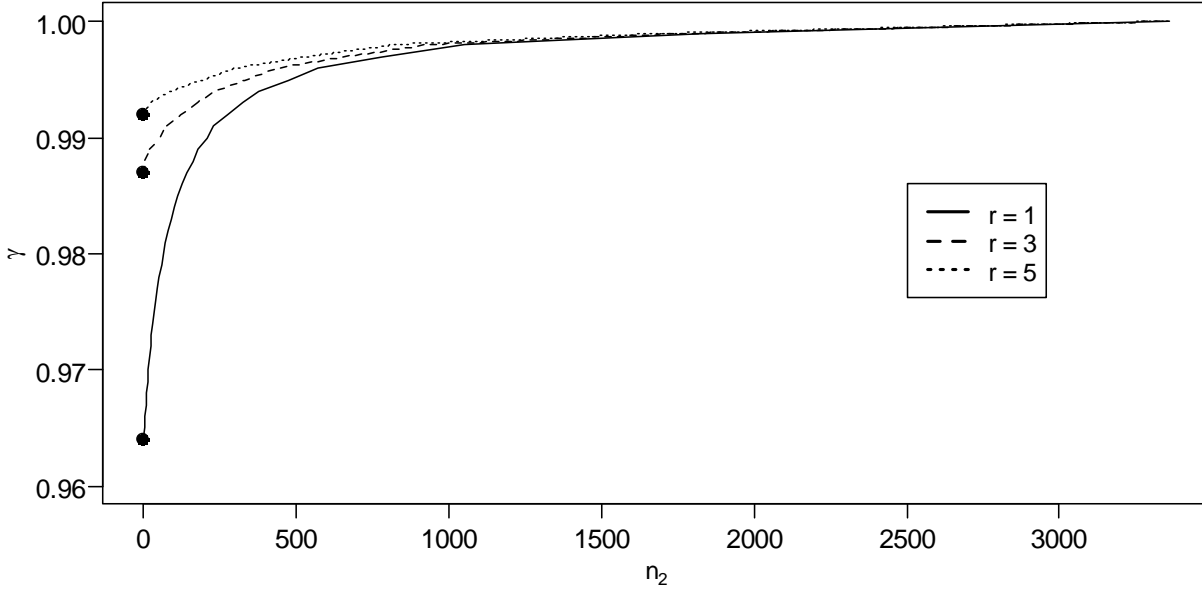
### 4.1 Regarding the fraction of the sampling area that is clean

Certainly it is most desirable to be able to state with high probability that a sampling area has no detectable contamination. Yet, achieving a high probability that no grid cells in the room have detectable contamination typically requires that almost the entire room be sampled---unless there is very strong prior evidence that the room is clean. Table 2 shows the number of random samples required for various levels of  $r$ ,  $C$ , and  $\gamma$ . Note that larger samples sizes are required to achieve higher confidence (i.e. higher  $C$  and/or higher  $\gamma$ ) and that larger values of  $r$  result in fewer random samples. Most striking are the massive amounts of random samples required in order to achieve high probability that none of the grid cells ( $\gamma = 1$ ) contain detectable contamination.

**Table 2:** Values of  $n_2$  required to achieve various levels of confidence for  $N = 3568$ ,  $n_1 = 25$ , and  $P_J = 0.70$ .

$r$	$\gamma = 0.99$		$\gamma = 1.00$	
	$C = 0.95$	$C = 0.99$	$C = 0.95$	$C = 0.99$
1	205	351	3362	3507
3	50	202	3354	3506
5	0	51	3346	3504
10	0	0	3325	3500

Figure 3 shows the number of random samples required to achieve 95% probability that  $100\gamma\%$  of the room does not contain detectable contamination. The plotted circles at  $n_2 = 0$  for each of the three curves shows the level of  $\gamma$  already achieved when the  $n_1 = 25$  judgment samples do not indicate contamination. The figure clearly shows that for  $\gamma > 0.99$ , each additional random sample only “buys” a very small amount of additional confidence that a larger percentage of the room does not contain detectable contamination. Admittedly, stating that there is a high probability that a high fraction of the sampling area (say, 99%) is clean *is* weaker than stating that there is a high probability that the entire sampling area is clean. However, in many instances, the large number of random samples required when  $\gamma = 1$  is neither feasible nor justifiable.

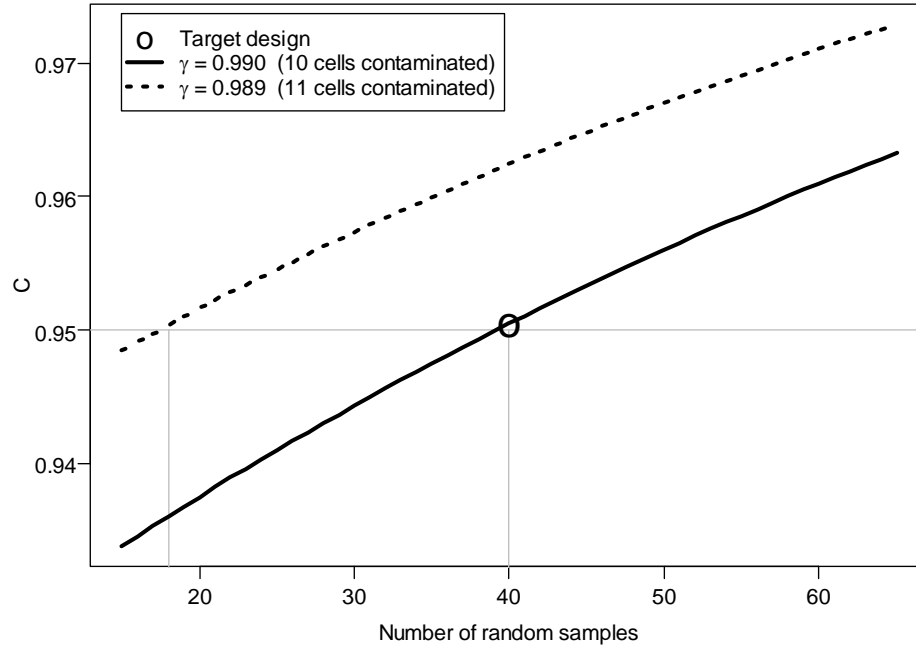


**Figure 3:** Plot of  $\gamma$  versus  $n_2$  for various levels of  $r$ , when  $N = 3568$ ,  $n_1 = 25$ ,  $P_J = 0.70$  and  $C = 0.95$ .

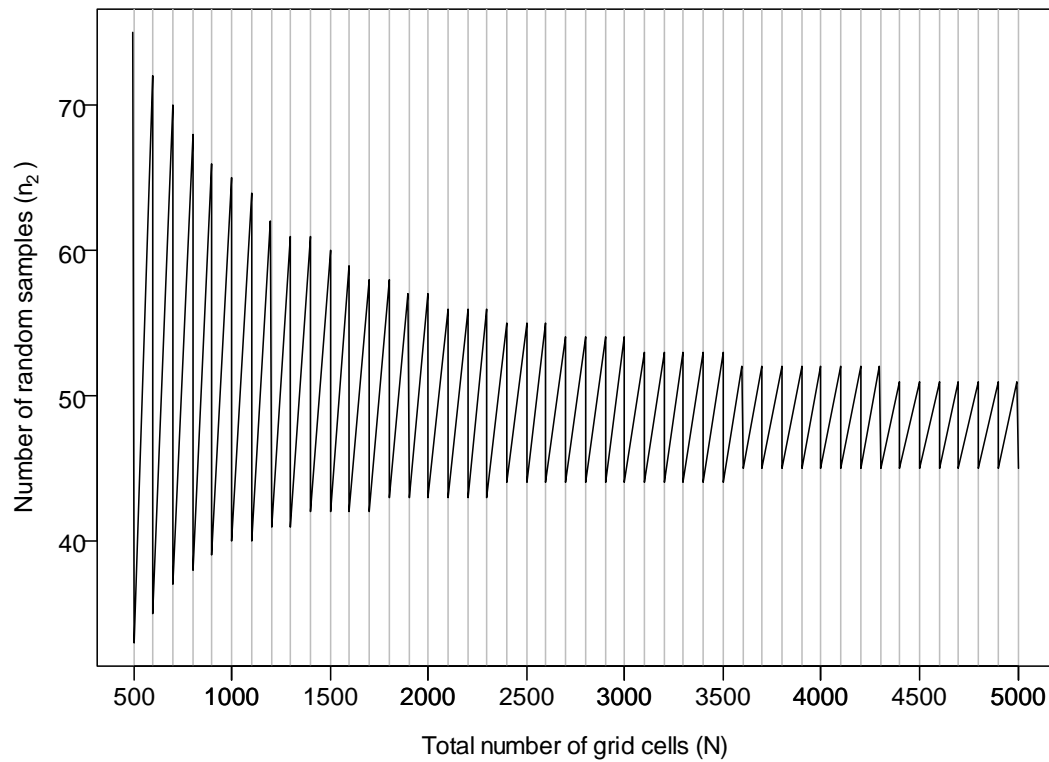
In order to present the confidence statement from an optimistic point of view, we have used the parameter  $\gamma$  to represent the fraction of the room that we wish to demonstrate is free of contamination. However, a close look at equation (2.3) reveals that the model requires that the percentage of the sampling area that is clean be expressed as the largest number of contaminated cells that will be tolerated. For example, in a room with 3568 cells, achieving a high probability that at least 99% of room is clean requires that no more than 35 cells contain contamination. As Figure 3 demonstrates, the lower the number of tolerable "dirty" cells, the higher the number samples required to achieve the desired confidence. Consequently, tolerating one or two additional contaminated cells can appreciably reduce the required number of random samples. Figure 4 demonstrates that the number of random samples can be reduced from 40 to 18 if we are willing to tolerate 11 contaminated cells ( $\gamma = 0.989$ ) instead of 10 contaminated cells ( $\gamma = 0.990$ ) in a relatively small room (or a room with large grid cells) that has only  $N = 1000$  cells. Understanding the relationship between  $\gamma$  and the number of random samples can permit the investigator to evaluate the trade-offs between desired confidence and the total number of samples required to achieve that level of confidence.

Typically, the value of  $N$  is considered fixed, and is determined by the characteristics of the sampling area. However, it is informative to investigate how the number of random samples ( $n_2$ ) changes as the total number of cells ( $N$ ) increases. This relationship is shown in Figure 5. The apparently rapid oscillation in  $n_2$  occurs because as  $N$  increases, the number of tolerable contaminated grid cells,  $\left\lfloor (1-\gamma)N \right\rfloor$ , increases by one when  $N$  reaches a number divisible by

100, as indicated by the gray vertical lines in Figure 5. To illustrate the phenomena, suppose we desire that at least 99% of a room with  $N = 1000$  cells be free from contamination with 95% probability. Specifically, we are willing to tolerate as many as 10 contaminated cells.



**Figure 4:** The effect of slightly decreasing  $\gamma$  on the number of random samples,  $n_2$ . For this figure,  $N = 1000$ ,  $n_1 = 25$ ,  $P_I = 0.70$ , and  $r = 3$ .



**Figure 5:** The number of random samples required ( $n_2$ ) versus the total number of grid cells ( $N$ ) required to achieve  $C = 0.95$  and  $\gamma = 0.99$  when  $n_1 = 25$ ,  $P_I = 0.70$ , and  $r = 3$ .

Suppose we take  $n_1 = 25$  judgment samples, they are taken from cells that are  $r = 3$  more likely contain contamination than the remainder of the room, and we believe there is a 70% chance that none of the 25 judgment samples will indicate contamination. Under these assumptions,  $n_2 = 40$  random samples are required to achieve the desired level of confidence. Suppose instead that the room had 99 more grid cells, i.e.  $N = 1099$ . Note that in order for the room to be at least 99% clean, the number of contaminated cells we are willing to tolerate is still 10 or less. Yet, due to the additional 99 cells, 64 random samples are now required. However, if we add one more cell such that  $N = 1100$ , only 40 random samples are required because 99% of the room being clean now means we are willing to tolerate as many as 11 contaminated cells.

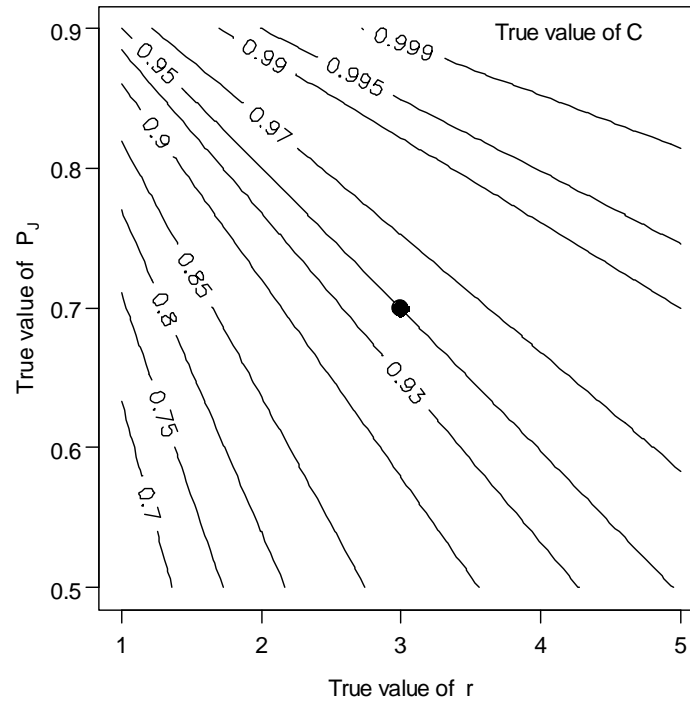
This is not to suggest that the value of  $N$  should be considered variable when creating a sampling design--it should be determined simply by the number of cells that exist within the sampling area. However, it is important to understand that reducing the target value of  $\gamma$  by a small amount so as to tolerate one additional contaminated sample can reduce the sample size appreciably, a trade-off which may be necessary to create a feasible sampling design.

## 4.2 Sensitivity to model inputs and assumptions

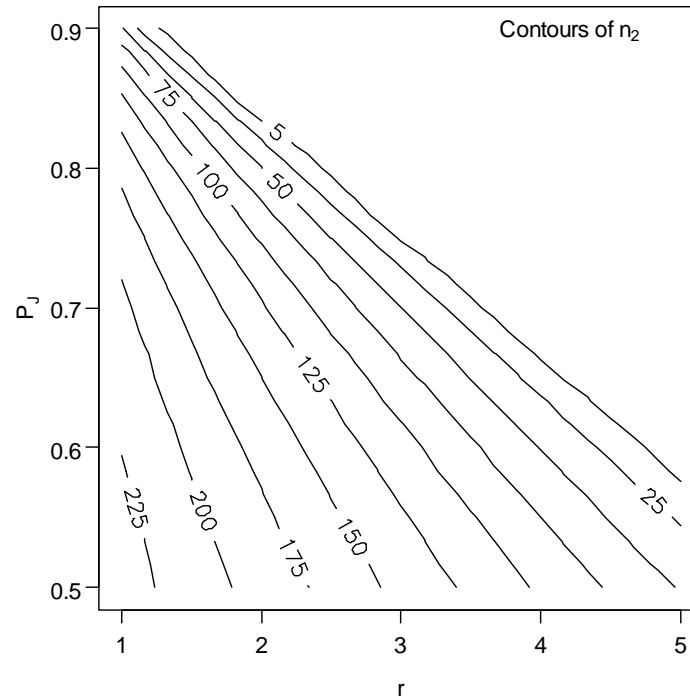
The inputs and assumptions required to use the model are 1) the correct identification of the location and number of high-risk cells from which judgment samples will be taken, 2) the *a priori* probability,  $P_J$ , that all of the judgment samples will not indicate contamination, and 3) the specification of the factor  $r$ , such that a high-risk cell is  $r$  times more likely to be contaminated than a low-risk cell. We will examine the second and third inputs first, and then address the problem of correctly identifying the location and number of high-risk cells.

When using Bayesian methods, one must choose the prior distribution(s) with care. Otherwise, poorly chosen prior parameters can lead to misleading conclusions. In the Bayesian model we consider, the probability  $P_J$  and the factor  $r$  are selected before collecting data. If the true value of  $r$  is lower than the *a priori* estimate of  $r$ , then too few samples will be taken and the confidence level will decrease. Likewise, if the true value of  $P_J$  is lower than the *a priori* estimate, too few samples will be taken and the achieved confidence level will decrease.

Figure 6 shows how the probability ( $C$ ) that 99% of the sampling area does not contain detectable contamination is affected by misspecification of  $r$  and  $P_J$ . In Figure 6, the value of  $n_2 = 50$  was chosen to achieve 95% probability that 99% of the room contains no detectable contamination when  $N = 3568$ ,  $n_1 = 25$ ,  $r = 3$ , and  $P_J = 0.70$ . These “target” parameter values are represented by the solid circle. To investigate the impact of misspecifying  $r$  and  $P_J$ , values of  $C$  were calculated for various combinations of  $r$  and  $P_J$ , using  $N = 3568$ ,  $n_1 = 25$ , and  $n_2 = 50$ . Figure 6 demonstrates that overestimating  $r$  and/or  $P_J$  reduces the achieved value of  $C$ . The inverse is also true: underestimating the values of  $r$  and  $P_J$  results in higher values of  $C$ . Also note that the value of  $C$  can be preserved by overestimating the value of  $r$  and underestimating the value of  $P_J$ , and vice versa.

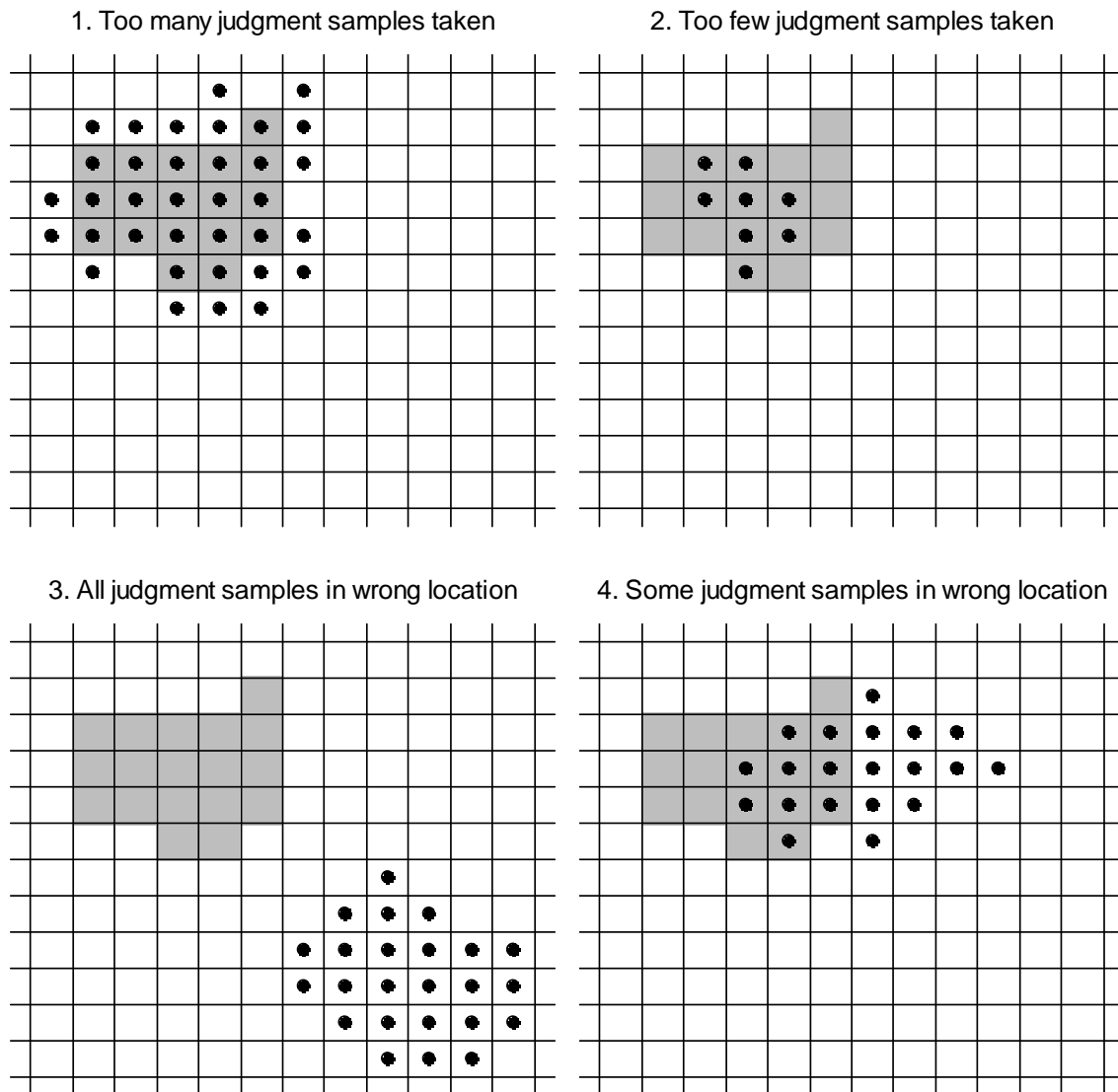


**Figure 6:** Examining the sensitivity of  $C$  for misspecifications of  $r$  and  $P_J$ . The target values of  $r = 3$ ,  $P_J = 0.70$ , and  $C = 0.95$  are shown by the solid circle in the center of the plot. The remaining parameters,  $N = 3568$ ,  $n_1 = 25$ ,  $n_2 = 50$ , and  $\gamma = 0.99$ , were held fixed.



**Figure 7:** Examining the sensitivity of the number of random samples ( $n_2$ ) required to achieve a 95% probability that 99% of the sampling area does not contain detectable contamination, for  $N=3568$  and  $n_1=25$ .

The sensitivity of the number of random samples ( $n_2$ ) to the values of  $r$  and  $P_J$  is shown in Figure 7. The contours show the number of random samples required to achieve 95% probability that 99% of the sampling area does not contain contamination when  $N = 3568$  and  $n_1 = 25$ . The figure demonstrates that lower values of  $r$  and  $P_J$  require a larger number of random samples to achieve the same level of confidence. The blank space in the upper right corner of the figure depicts a region where no random samples are required. Thus, when  $r$  and/or  $P_J$  are sufficiently large, the outcome that none of the 25 judgment samples indicate contamination provides sufficient evidence to achieve the desired level of confidence.



**Figure 8:** Four scenarios depicting possible ways in which the number and location of high-risk cells could be misidentified. The high-risk cells are indicated with solid gray. The judgment samples are indicated by solid black circles.



In addition to the choice of  $r$  and  $P_j$ , the model also relies on the assumption that the number and location of high-risk cells from which the judgment samples are taken have been correctly identified. Figure 8 depicts four possible ways in which the number and location of the high-risk cells could be misidentified. Note that the regions of high-risk cells do not necessarily need to be contiguous, as they are depicted (for convenience) in Figure 8. The consequences of each of these four scenarios could be studied analytically or with simulation. We are currently investigating the consequences of mispecifying judgmental samples and those results will be made available in a separate document.

### 4.3 Other acceptance sampling approaches

We briefly consider two other models that have been proposed for acceptance sampling: 1) a Bayesian approach proposed by Wright (1992) and improved by Grieve (1994) and 2) frequentist acceptance sampling, as discussed by Bowen (1988). Both approaches use the hypergeometric distribution to model the observed number of contaminated cells within the room. In this application, however, we used the binomial distribution because the hypergeometric model was not well suited for linking the information obtained from both the judgment and random samples. The binomial model permits us to assign different probabilities of contamination for judgment and random samples ( $\theta$  and  $\theta/r$ ). Furthermore, the hypergeometric model converges to the binomial model as the total number of grid cells in the sampling area gets large, which is often the case even with relatively small rooms, as demonstrated in Section 3.

The approach proposed by Wright and Grieve (WG) is conceptually very similar to the combined judgment and random sampling methodology (CJR) presented here. In the WG model, a beta-binomial distribution is used as the prior for the number of contaminated cells within the room, the number of observed contaminated cells is modeled by the hypergeometric distribution, and the sample size is determined using the posterior predictive distribution of the number of contaminated cells that were not sampled. Using this framework, Grieve gave a concise approximation to the number of samples required to achieve the desired probability that 100% of the room is clean given that none of the samples indicate contamination. The fundamental difference between the WG methodology and the CJR approach is that we use the binomial distribution to model the number of contaminated cells and some samples (the judgment samples) are assumed to be more likely to be contaminated than the remaining (random) samples.

Table 3 shows the total number of samples required to achieve a 95% probability that 100% of the room is clean for the CJR and WG designs. For the CJR design, the prior distribution for  $\theta$  was beta with shape parameters  $\alpha = 1$  and  $\beta$  as designated in Table 3. For the WG design, the prior distribution was the beta-binomial with shape parameters  $\alpha = 1$  and  $\beta$  and  $N$  as designated in Table 3. (Note that by using  $r = 1$ , the distinction between the judgment and random samples is removed). For the cases out of  $N$  and  $\beta$  considered here, Table 3 demonstrates the virtual equivalence between the CJR design

and the WG design. As part of future research, it would be interesting to compare these two methodologies for values of  $\gamma < 1$ . We suspect that they would be very similar for  $N \geq 100$ .

**Table 3:** Comparison of the total number of samples required to achieve  $C = 0.95$  and  $\gamma = 1.00$  for the combined judgment and random sampling design (CJR) and the Wright/Grieve Bayesian design (WG) for acceptance sampling. For the CJR method,  $r = 1$ .

	$N=100$		$N=500$		$N=10000$	
$\beta$	CJR	WG	CJR	WG	CJR	WG
1	95	95	475	475	9500	9500
9	95	95	475	475	9500	9500
99	91	91	471	471	9495	9496
999	46	46	426	426	9450	9451

The frequentist acceptance sampling (FAS) approach makes no prior assumption regarding the number of contaminated cells in the room, nor does it distinguish between judgment and random samples. Using the hypergeometric distribution to model the number of contaminated cells in the room, the FAS design identifies the smallest sample size such that the probability that no contaminated cells are observed in the samples given that the true number of contaminated cells is  $\left\lfloor (1-\gamma)N \right\rfloor$  or less is at least as large as a given confidence threshold, say, 95%. Although not equivalent, this confidence threshold is the conceptual analog of the Bayesian probability  $C$ . Table 4 compares the CJR and FAS methods by showing the number of samples required to achieve a 95% probability (for the CJR method) or 95% confidence (for the FAS method) that a high percentage of the sampling area does not contain detectable contamination. For the CJR model, a uniform prior distribution was used and the likelihood contribution of the judgment and random samples was made equivalent by setting  $r = 1$ . Conceptually, the uniform prior suggests little prior knowledge about the parameter values. Table 4 shows that the number of samples required by the two methodologies is quite similar, although fewer samples were required for the CJR model in each of the cases that were considered.

**Table 4:** Comparison of the total number of samples required for the combined judgment and random sampling design (CJR) and a frequentist acceptance sampling design (FAS). For the CJR design,  $\alpha = \beta = 1$  (corresponding to the uniform prior),  $r = 1$ , and  $C = 0.95$ . For the FAS design, 95% confidence threshold was used.

	$\gamma = 0.95$		$\gamma = 0.99$	
$N$	CJR	FAS	CJR	FAS
500	53	56	195	225
2000	57	58	264	277
10000	58	59	290	294

## 4.4 Extensions

The combined judgment and random sampling model presented herein could be readily extended to account for the uncertainty in the choice of  $r$  and the choice of  $P_J$ , or even the parameters  $\alpha$  and  $\beta$ . Likely choices for the prior distribution of  $r$  and  $P_J$  would be the gamma and beta distributions, respectively. One way in which the hyperparameters for these distributions could be chosen would be to estimate the values of the first and second central moments. For example, investigators could indicate their belief that the mean value of  $r$  was 2 with a standard deviation of 1.5. Then the theoretical mean and variance of the gamma distribution would be equated to 2 and 2.25, respectively. Solving these two equations would give rise to the shape and scale parameters for the gamma distribution of  $r$ . A similar approach could be used to find the two shape parameters of the beta distribution of  $P_J$ .

The model could also be extended or modified to account for different sampling methodologies. Within any given room, there are a number of different types of surfaces which may require sampling. Different surface types necessitate different sampling methodologies which in turn may require different surface area sizes for the collection of a single sample. Recall from Section 2 that sterile swabs, sterile surface wipes, and vacuum sampling use different sized surface areas for a sample. The model could also be expanded to account for false negatives (not detecting contamination that is present and detectable) and for false positives (detecting contamination that is not present). We would expect that the various sampling methodologies have different false negative and false positive rates.

## 5 Conclusion

In this report we present a Bayesian acceptance sampling model which combines the information obtained from both judgment and randomly placed samples. The locations of the judgment samples divide the sampling areas into two regions of risk: high and low. The high-risk regions are identified by the location of the judgment samples while the remaining low risk cells are available for random sampling. While this simple dichotomous model is admittedly a rough approximation of reality, it provides a convenient framework for combining the information obtained from both judgment and random samples into a single sampling design. After making specific assumptions, the model indicates the number of random samples required to achieve a desired probability that a large fraction of the sampling area does not contain detectable contamination, given that *none* of the samples indicate the presence of contamination. The methodology is intended to be used primarily for clearance sampling after decontamination, or to demonstrate the cleanliness of an area that is presumed unlikely to be contaminated.

The inputs and assumptions required to use the model are 1) the correct identification of the location and number of high-risk cells from which judgment samples will be taken, 2) the *a priori* probability,  $P_J$ , that all of the judgment samples will not indicate contamination, and 3) the specification of the factor  $r$ , such that a high-risk cell is  $r$  times more likely to be contaminated than a low-risk cell. With regard to the first assumption, it appears that the model still performs well even if the number of high-risk cells is overestimated. However, in terms of both sample

size and confidence, the model does appear to be more sensitive to the choices of  $r$  and  $P_j$ . For this reason we recommend that values have  $P_j$  and  $r$  be chosen conservatively, that is, that investigators err on the side of underestimating these two quantities. Unless there were evidence to the contrary, we would not recommend using values of  $P_j$  larger than 0.90 or values of  $r$  larger than 5. However, overly conservative estimates result in a larger number of random samples required to obtain the desired confidence. While trying to construct a feasible sampling plan, investigators should examine the impact of slightly reducing the percentage of the sampling area that does not contain detectable contamination. If multiple sampling methodologies are required, we suggest the conservative approach of letting the square grid cells of the entire sampling area be equal to the surface area of the sampling methodology with the smallest surface area. Then, sampling methodologies that cover larger surface areas would be counted only as a single sample.

This model, like any other, relies on assumptions which at best are still approximations of reality. However, it constitutes a viable attempt to combine the outcomes in judgment and random sampling into a single design and quantify the level of confidence that the sampling area is largely free of detectable contamination.

## Acknowledgements

This research was funded by the National Institute for Occupational Safety and Health. We wish to thank Tom Ferryman and Richard Gilbert of Pacific Northwest National Laboratory (PNNL) for their insights and D.J. Damstedt of PNNL who provided assistance with the graphics.

## Appendix

A review of the Bayesian approach shown below is given by Gelman, et al. (2004). The prior distribution for  $\theta$  is  $Beta(\alpha, \beta)$ , that is,  $p(\theta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}$ . Since  $X|\theta \sim Bin(n_1, \theta)$

and  $Y|\theta \sim Bin(n_2, \theta/r)$ , we have by Bayes' rule and the independence of  $X$  and  $Y$ :

$$p(\theta | X = 0, Y = 0) \propto P(X = 0 | \theta) P(Y = 0 | \theta) p(\theta) \propto (1-\theta)^{n_1} \left(1 - \frac{\theta}{r}\right)^{n_2} \theta^{\alpha-1} (1-\theta)^{\beta-1},$$

and normalizing gives

$$p(\theta | X = 0, Y = 0) = \frac{\theta^{\alpha-1} (1-\theta)^{\beta+n_1-1} \left(1-\frac{\theta}{r}\right)^{n_2}}{\int_0^1 \theta^{\alpha-1} (1-\theta)^{\beta+n_1-1} \left(1-\frac{\theta}{r}\right)^{n_2} d\theta}.$$

The number of contaminated cells in the unsampled area is denoted by  $Z | \theta$  which is distributed as  $\text{Bin}\left(N - n_1 - n_2, \frac{\theta}{r}\right)$ . The conditional posterior predictive distribution for  $Z$  is

$$\begin{aligned} P(Z = z | X = 0, Y = 0) &= \int_0^1 P(Z = z | \theta) p(\theta | X = 0, Y = 0) d\theta \\ &= \int_0^1 \binom{N - n_1 - n_2}{z} \left(\frac{\theta}{r}\right)^z \left(1 - \frac{\theta}{r}\right)^{N - n_1 - n_2 - z} p(\theta | X = 0, Y = 0) d\theta \end{aligned}$$

To find the number of random samples required to achieve the desired level of confidence,  $C$ , we search for the smallest integer value of  $n_2$  that satisfies the following:

$$\begin{aligned} &P(Z \leq (1-\gamma)N | X = 0, Y = 0) \\ &= \sum_{z=0}^{\lfloor (1-\gamma)N \rfloor} \left[ \int_0^1 \binom{N - n_1 - n_2}{z} \left(\frac{\theta}{r}\right)^z \left(1 - \frac{\theta}{r}\right)^{N - n_1 - n_2 - z} p(\theta | X = 0, Y = 0) d\theta \right] \\ &= \frac{\int_0^1 \theta^{\alpha-1} (1-\theta)^{\beta+n_1-1} \sum_{z=0}^{\lfloor (1-\gamma)N \rfloor} \left[ \binom{N - n_1 - n_2}{z} \left(\frac{\theta}{r}\right)^z \left(1 - \frac{\theta}{r}\right)^{N - n_1 - n_2 - z} \right] d\theta}{\int_0^1 \theta^{\alpha-1} (1-\theta)^{\beta+n_1-1} \left(1 - \frac{\theta}{r}\right)^{n_2} d\theta} \geq C \end{aligned}$$

where  $\lfloor (1-\gamma)N \rfloor$  denotes the floor of  $(1-\gamma)N$ , i.e. the largest integer that is less than or equal to  $(1-\gamma)N$ .

Because the integrands in the numerator and denominator of in the previous equation have most of their mass near zero, numerical integration routines applied to the entire unit interval can produce incorrect results. We avoided this problem by taking the sum of the integrals evaluated separately over the four subdomains  $[0, 10^{-5})$ ,  $[10^{-5}, 10^{-4})$ ,  $[10^{-4}, 10^{-3})$ , and  $[10^{-3}, 1]$ . We verified the accuracy of the results by demonstrating agreement for thousands of cases among three different integration routines: 1) QUADPACK routines 'dqags' and 'dqagi', (Piessens 1983) implemented in the statistical language R (R Development Core Team 2007), 2) the QDAGS function from the IMSL Fortran Numerical Library version 5.0 (Visual Numerics 2007), and 3) Simpson's method (Ellis 1991).

## References

- Bowen MW, Bennett CA. 1988. *Statistical Methods for Nuclear Material Management*, NUREG/CR-4604, U.S. Nuclear Regulatory Commission, Washington, DC. 881--887.
- Gelman A, Carlin JB, Stern HS, Rubin DB. 2004. *Bayesian Data Analysis, Second Edition*. Chapman & Hall/CRC, 8 and 40.
- Grieve AP. 1994. A further note on sampling to locate rare defectives with strong prior evidence. *Biometrika*, 88:787--789.
- Ellis R, Gulick D. 1991. *Calculus: One and Several Variables*. Harcourt Brace Jovanovich, Publishers. New York, NY. 479-482.
- NIOSH. 2002. Comprehensive Procedures for Collecting Environmental Samples for Culturing *Bacillus anthracis* (Revised April 2002) <http://www.cdc.gov/niosh/unp-envsamp.html>.
- Piessens R, deDoncker-Kapenga E, Uberhuber C, Kahaner D. 1983. *Quadpack: a Subroutine Package for Automatic Integration*. Springer Verlag.
- R Development Core Team. 2007. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Visual Numerics. 2007. <http://www.vni.com/products/ims1/fortran/overview.html>.
- Wright T. 1992. A note on sampling to locate rare defectives with strong prior evidence. *Biometrika*, 79:685--691.