Sampling with discrete contamination

Geoffrey Lyman, Florent Bourgeois

Geoffrey J. Lyman
Materials Sampling & Consulting, Australia

Florent S. Bourgeois
Materials Sampling & Consulting Europe, France

Sampling with discrete contamination

ABSTRACT

The sampling variance for a process stream which carries fluctuating levels of the sought-after analyte and is subject to mass flow variation can be estimated from the covariance function of the analyte fluctuation and the covariance function of the mass flow when these covariance functions are well-defined and can be considered to be a stationary property of the process stream.

However, in the case of sampling a flow of material (a one-dimensional lot) or from material removed from the hold of a ship (a three-dimensional lot) which does not possess a covariance function for the analyte of interest, a different approach must be taken. An important example of such a case is a shipment of grain that is contaminated by some component such as genetically modified organisms (GMOs) or by mycotoxins. Depending on the manner of contamination, the regions of the lot that carry contamination can be considered as randomly located distributions of concentration. The distributions themselves may be stochastic in that their mean concentrations and extents may be statistically defined rather than fixed.

This paper develops the sampling variance for 'slugs' of contamination with a uniform concentration distribution and regular spacing of the sample increments, based on the assumption that the origins of the slugs are uniformly and randomly located (a Poisson point process).

INTRODUCTION

When analyte and mass flow covariance functions are defined for a process stream, sampling theory provides the means to determine the sampling variance. If these functions are not well-defined, one may assume that the analyte variations are far more random in nature and it is natural to use a model where the elevated analyte concentrations occur randomly in time or mass.

The most common method of sampling grain involves taking increments from the lot on a regular basis. In a one-dimensional flow of material, sample increments are taken at regular intervals of time or accumulated mass flow. In a three-dimensional case, the increments may also be taken at regularly spaced locations throughout the lot as it is being unloaded. Random stratified sampling in the 1D or 3D case can also be undertaken and, apart from some minor corrections for end effects in the stratum, the following analysis is applicable as well.

The structure of the sampling problem considered here is similar to the problem addressed by Paoletti et al. (2003) who sought to evaluate the effect of lot heterogeneity with respect to GMO content on the sampling variance as a function of the number of increments taken from the lot using the KeSTE tool. The formulation proposed hereafter is in some respects more general.

The development shows that it is possible to arrive at an analytic expression for the sampling variance and this expression is verified by simulation, which also captures the distribution function of the sample concentration. The results are applied to a practical sampling design problem for a shipload of grain.

DEVELOPMENT

The sampling period is divided into strata of extent l and the total sampling period is normalised to unity. A sample increment is taken instantaneously over an interval *v. n* such increments are taken, placed centrally within each stratum. The slugs of contaminated material, in which the concentration of contaminant is taken to be c_0, are of a constant duration u and m such slugs occur randomly in the total sampling interval; the slugs may overlap. The situation is illustrated in Figure 1.

In reality, the slugs of contaminant are not expected to be of uniform concentration; such a variation can be superimposed on the results presented herein. This analysis seeks to develop the simplest case to provide a first means of exploring sampling variance. It is substantially more difficult to vary the size of the slugs; such a case is best dealt with by simulation.

Given these rules for the location of samples and slugs, it is necessary to find the expected value of the sample concentration and the sampling variance.

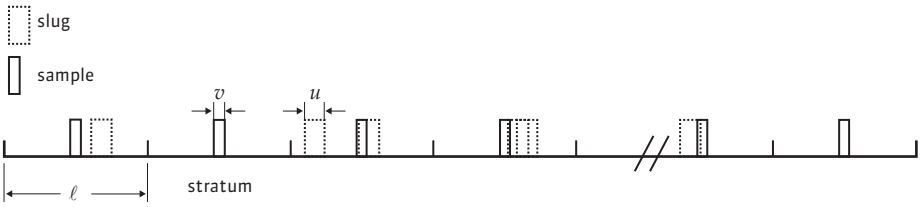

Figure 1 Illustration of disposition of samples and slugs of contaminated material

The statistical analysis of this problem is most effectively viewed as a Bernoulli process in which slugs are placed into the structure of the sampling increments. The process is one of 'throwing' slugs of extent u onto a line that has n equally spaced sample increments of extent v. A Bernoulli process trial has only two possible outcomes, i.e., the slug hits the sample increment with probability q, which we define as the success event, or the slug does not hit the sample increment with probability (1-q), the fail event. Each slug has a probability:

$$q = \frac{u+v}{l} \tag{1}$$

of hitting an increment, regardless of where the increment is placed in the stratum. With m slugs, the number of hits on the increments will be binomial with:

$$\Pr\{j \mid m,q\} = \binom{m}{j} q^j (1-q)^{m-j} = \frac{m!}{(m-j)!\,j!} q^j (1-q)^{m-j} \tag{2}$$

This is the classic result that gives the probability of having j successes from m independent trials, given the probability q of success in a single trial. The expected value and variance of the random variable j are:

$$\mathrm{E}\{j \mid m,q\} = mq \tag{3}$$

$$\mathrm{var}\{j \mid m,q\} = mq(1-q) \tag{4}$$

The concentration probability distribution for the random and uniform intersection of the sample increment of extent v with the slug of extent u is needed.

For $u \le v$, the concentration probability density function is:

$$\begin{aligned} p(c) &= \frac{2v}{c_0(v+u)} && 0 \le c < c_0 \frac{u}{v} \\ &= \frac{v-u}{v+u}\delta\left(c_0 \frac{u}{v}\right) && c = c_0 \frac{u}{v} \end{aligned} \tag{5}$$

The expected value of the concentration is:

$$\mathrm{E}\{c\} = c_0\left(\frac{u}{v+u}\right) \tag{6}$$

and the variance is:

$$\mathrm{var}\{c\} = {c_0}^2 \left(\frac{u}{v+u}\right)^2 \left[\frac{2uv-u^2}{3v^2}\right] \tag{7}$$

For $u \geq v$, the density, expected value and variance are:

$$\begin{aligned} p(c) &= \frac{2v}{c_0(u+v)} && 0 \leq c < c_0 \\ &= \frac{u-v}{u+v}\delta(c_0) && c = c_0 \end{aligned} \tag{8}$$

$$\mathrm{E}\{c\} = c_0\left(\frac{u}{u+v}\right) \tag{9}$$

$$\mathrm{var}\{c\} = {c_0}^2 \left(\frac{u}{u+v}\right)^2 \left[\frac{2uv-v^2}{3u^2}\right] \tag{10}$$

When the slug and sample increment have the same length, $u = v$, both density functions yield the same values, that is:

$$\begin{aligned} p(c) &= \frac{1}{c_0} && 0 \leq c < c_0 \\ &= 0 && c = c_0 \end{aligned} \tag{11}$$

$$\mathrm{E}\{c\} = \frac{c_0}{2} \tag{12}$$

$$\mathrm{var}\{c\} = \frac{{c_0}^2}{12} \tag{13}$$

The sample that is collected consists of n increments and if in the particular instance of the sampling there were j hits on slugs of material, the sample concentration will be:

$$c_S = \frac{1}{n}\sum_{k=1}^{j} c_k \tag{14}$$

where the c_k are independent identically distributed random variables following density $p(c)$. To find the expected value of c_S, expectations over the concentration distribution and the number of hits by slugs must be determined. For fixed u and v, the concentration distribution is fixed, so the expectations can be taken sequentially. One then has:

$$\begin{aligned}\mathrm{E}\{c_S\} &= \frac{1}{n}\Big[\Pr\{j=1\}\mathrm{E}\{c\} + \Pr\{j=2\}2\mathrm{E}\{c\} + \mathsf{L}\ \Big] \\ &= \frac{1}{n}\mathrm{E}\{c\}\mathrm{E}\{j\} \\ &= c_0\frac{u}{v+u}\frac{u+v}{\mathsf{l}}\frac{m}{n} \\ &= c_0\frac{mu}{n\mathsf{l}}\end{aligned} \tag{15}$$

and this is the true concentration (mass per unit length or time) of contaminant in the lot. The conclusion is that the sampling is unbiased.

The determination of the variance of the sample concentration is a more difficult task. The derivation of the result is provided in Appendix A. The result is, for $u<v$:

$$\begin{aligned}\mathrm{var}\{c_S\} &= \frac{1}{n^2}\left[mqc_0{}^2\left(\frac{u}{v+u}\right)^2\left[\frac{2uv-u^2}{3v^2}\right] + c_0{}^2\left(\frac{u}{v+u}\right)^2 mq(1-q)\right] \\ &= c_0{}^2\frac{mq}{n^2}\left(\frac{u}{v+u}\right)^2\left[\frac{2uv-u^2}{3v^2}+1-q\right]\end{aligned} \tag{16}$$

and for $u\geq v$:

$$\begin{aligned}\mathrm{var}\{c_S\} &= \frac{1}{n^2}\left[mqc_0{}^2\left(\frac{u}{u+v}\right)^2\left[\frac{2uv-v^2}{3u^2}\right] + c_0{}^2\left(\frac{u}{v+u}\right)^2 mq(1-q)\right] \\ &= c_0{}^2\frac{mq}{n^2}\left(\frac{u}{u+v}\right)^2\left[\frac{2uv-v^2}{3u^2}+1-q\right]\end{aligned} \tag{17}$$

This result is valid only when $\mathsf{l}>u+v$.

The standard deviation relative to the expected sample concentration is then, for $u<v$,

$$\frac{\sigma_{c_S}}{\mathrm{E}\{c_S\}} = \sqrt{\frac{1}{nm}\left[\left(\frac{L}{u+v}\right)\left(\frac{2uv-u^2}{3v^2}+1\right)-n\right]} \tag{18}$$

and for $u\geq v$:

$$\frac{\sigma_{c_S}}{\mathrm{E}\{c_S\}} = \sqrt{\frac{1}{nm}\left[\left(\frac{L}{u+v}\right)\left(\frac{2uv-v^2}{3u^2}+1\right)-n\right]} \tag{19}$$

where $L=n\mathsf{l}$ is the total length of the sampling period.

These results demonstrate that the sampling of a fixed lot of material becomes relatively more precise as the slugs of material occur in higher numbers, the slugs become larger in extent, and the number of increments taken increases. This is an entirely common sense outcome of the analysis.

As an example, take a total lot length L=1, m=20 slugs, n=100 sample increments, slug length u=0.0001 and sample increment length v=0.0005. The contamination level is then $\frac{mu}{L} = 0.2\%$ and the sampling collects $\frac{nv}{L} = 5\%$ of the flow. Then the relative uncertainty is:

$$\begin{aligned}\frac{\sigma_{c_S}}{\mathrm{E}\{c_S\}} &= \sqrt{\frac{1}{nm}\left[\left(\frac{L}{u+v}\right)\left(\frac{2uv-u^2}{3v^2}+1\right)-n\right]} \\ &= \sqrt{\frac{1}{100\times 20}\left[\left(\frac{1}{0.0006}\right)\left(\frac{2\times 0.0001\times 0.0005}{3\times 0.0005^2}+1\right)-100\right]} \\ &= 0.946\end{aligned} \qquad (20)$$

STUDY BY SIMULATION

The sampling process described above can of course be simulated and the simulation results can be used to verify the mathematical development. It is always prudent in such matters to check a simulation against an analytical analysis and vice versa. The simulation has the advantage of revealing the distribution of the sample concentrations which the analytical approach cannot define at the present level of analysis. The simulations presented here are aimed at mycotoxin sampling issues rather than GMO contamination; the mycotoxin issue presents an even more difficult sampling problem than the GMO problem.

Consider first a case in which there are 20 slugs of contamination in the process stream to be sampled. The average level of contamination is taken to be 2 ppb, the average concentration in the slug is varied as is the number of increments taken from the lot. The size of the slug and increment sizes are varied as well. This represents a highly heterogeneous lot of grain.

It should be noted that in carrying out simulations such as these, it is very important to employ a highly reliable random number generator. The random number generator in an application such as Excel is notoriously poor.

Table 1 Simulation results for sampling with discrete slugs of contaminant (10 000 simulation)

Case	1	2	3	4	5	6	7	8
Slugs (m)	20	20	20	20	20	20	20	20
Increments (n)	100	200	100	200	100	100	100	100
Slug size (u)	0.0001	0.0001	0.0001	0.0001	0.0005	0.0050	0.0030	0.001
Increment size (v)	0.0005	0.0005	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
Sample fraction [%]	5.0	10.0	1.0	2.0	1.0	1.0	1.0	1.0
Slug conc (ppb)	1000	1000	1000	1000	200	20	33	100
Average conc (ppb)	2	2	2	2	2	2	2	2
Av conc (simulation)	2.01	1.98	2.07	1.99	1.96	1.91	1.94	1.97
RSD theory	0.940	0.645	1.812	1.271	0.940	0.222	0.339	0.658
RSD simulation	0.939	0.657	1.770	1.283	0.936	0.223	0.335	0.662

Table 1 shows the simulation results for the conditions chosen and Figures 2 to 9 show the histograms of concentrations and the cumulative distributions of concentrations. In each case the average contaminant concentration is held at 2 ppb and with the exception of cases 1 and 2, the increment size is 0.0001 of the mass of the lot. Taking 100 such increments recovers 0.01 or 1% of the mass of the lot as primary increments. These primary increments can then be divided down in an appropriate manner to arrive at an analysis sample.

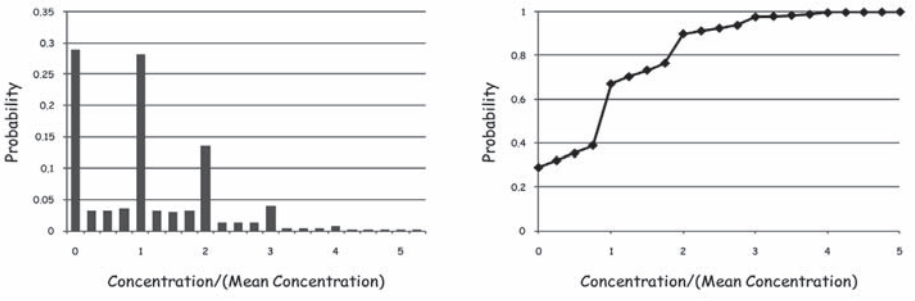

Figure 2 Histogram of sample concentrations and cumulative distribution from simulation, Case 1

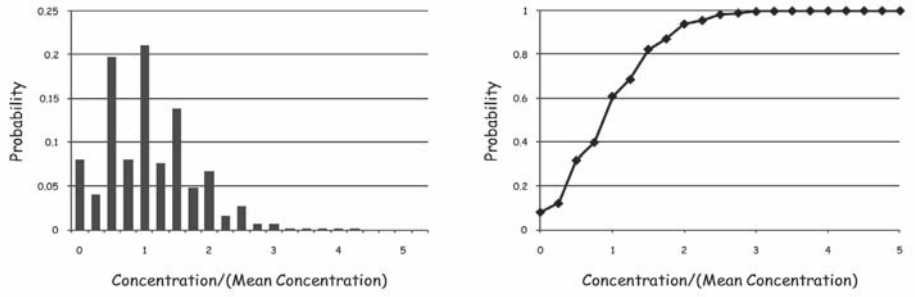

Figure 3 Histogram of sample concentrations and cumulative distribution from simulation, Case 2

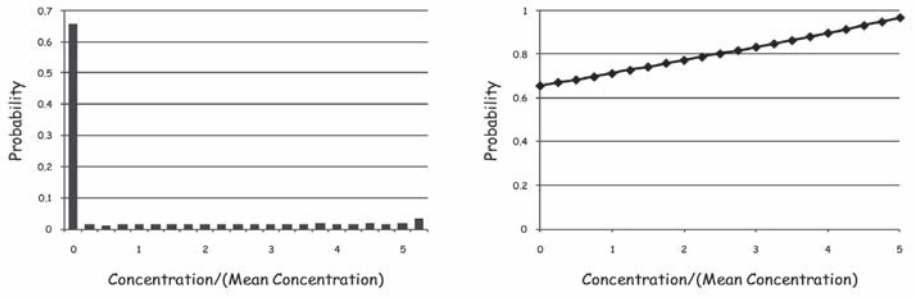

Figure 4 Histogram of sample concentrations and cumulative distribution from simulation, Case 3

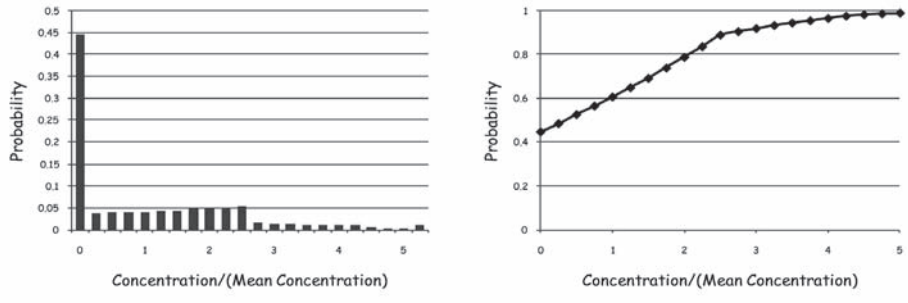

Figure 5 Histogram of sample concentrations and cumulative distribution from simulation, Case 4

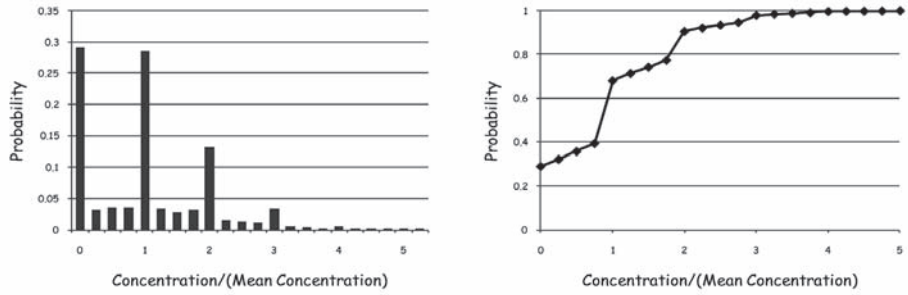

Figure 6 Histogram of sample concentrations and cumulative distribution from simulation, Case 5

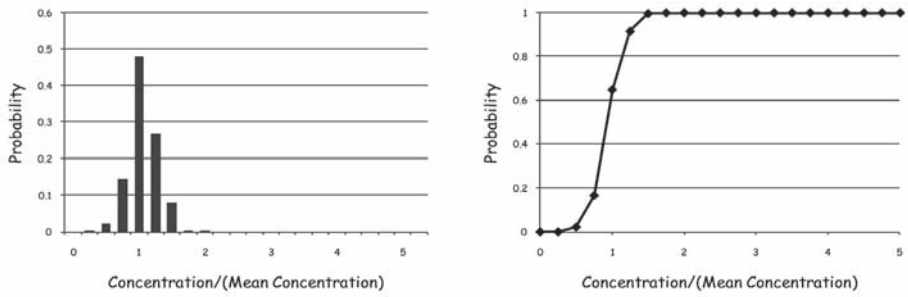

Figure 7 Histogram of sample concentrations and cumulative distribution from simulation, Case 6

When there are 100 increments taken, the extent of each sampling stratum is 0.01 of the lot mass. When the slug size is 0.005, the extent of a slug is half the extent of the stratum and the lot is moving towards homogeneity although the slugs are randomly placed in the lot. As the slug size decreases the concentration within the slug is increased to maintain the average concentration of the contaminant at 2 ppb.

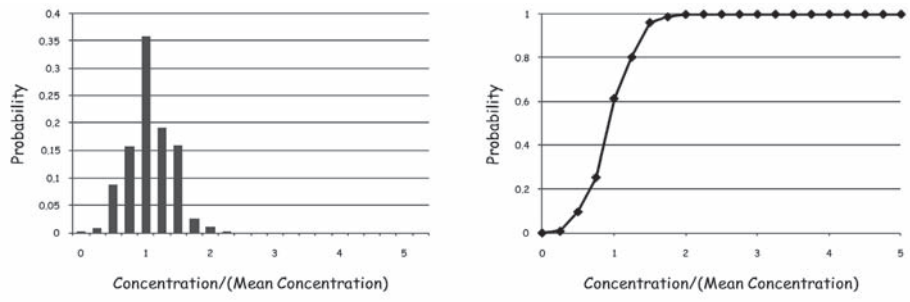

Figure 8 Histogram of sample concentrations and cumulative distribution from simulation, Case 7

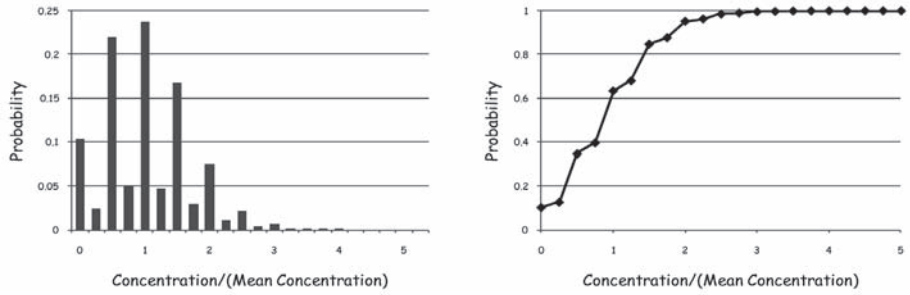

Figure 9 Histogram of sample concentrations and cumulative distribution from simulation, Case 8

In Table 1, the slug size is the value of u and the increment size is the value of v, expressed as a fraction of the lot. The sample fraction is $n \times v$ expressed as a percentage of the lot mass. Av conc (simulation) is the average concentration from simulation and is found by averaging the simulation outcomes over the 10 000 simulations. RSD theory is the value from Equation 16 or 17 as appropriate. RSD simulation is the RSD found by simulation and can be compared to RSD theory. There is clearly a good agreement between the simulations and the theoretical expressions, cross-validating the theory and simulations.

The results show that $n(u+v)$ must be large enough, relative to L for a given value of m, to bring the variance of sampling down to a reasonable value. If u is small, then one must take either larger increments or more of them.

DESIGN EXAMPLE

Consider the loading of a shipment at a rate of 2000 tph. The grain is sampled by a cross-stream cutter having an aperture of 19 mm and moving at 0.5 m/s. The increment mass is:

$$m_I = \frac{Tw}{3.6s} \tag{21}$$

where m_I is the increment mass in kg, T is the mass flow in tph, w is the aperture in metres and s is the cutter speed in m/s. The result is:

$$\begin{aligned} m_I &= \frac{2000 \times 0.019}{3.6 \times 0.5} \\ &= 21.1 \end{aligned} \tag{22}$$

Assume that an increment is taken every 40 seconds and that a sampling unit, L, corresponds to 2000 tonnes of grain. There are then 90 increments taken during the sampling. The fraction of the sampling unit taken as primary increments is:

$$\begin{aligned} f &= \frac{nm_I}{L} \\ &= \frac{90 \times 21.1}{2.0 \times 10^6} \\ &= 0.00095 \end{aligned} \tag{23}$$

where L, the extent of the sampling unit is taken as 2000 tonnes. This sampling protocol is the one defined by the Canadian Grain Commission for wheat sampling (Canadian Grain Commission, 2009).

Next, assume the average level of contamination is $\bar{c} = 10$ ppb and the average contamination level in a slug is $c_0 = 1600$ ppb. Let the slug occupy a volume of ten litres and weigh 8 kg. Then the number of slugs must be:

$$m = \frac{\bar{c}L}{c_0 u} \tag{24}$$

so:

$$\begin{aligned} m &= \frac{10 \times 2 \times 10^6}{1600 \times 8} \\ &= 1563 \end{aligned} \tag{25}$$

The increment extent, v and the slug extent, u are then determined as $21.1/2 \times 10^6$ and $8/2 \times 10^6$ kg respectively.

The RSD from theory is:

$$RSD = \sqrt{\frac{1}{nm}\left[\left(\frac{L}{u+v}\right)\left(\frac{2uv-u^2}{3v^2}+1\right)-n\right]} \tag{26}$$

which evaluates to 0.767.

The results of simulation of this scenario are provided in Figure 10. The RSD from simulation is 0.768. The 95% confidence interval, determined from the simulation results, for the result of sampling 2000 tonnes of the shipment is 0.0 to 27.1 ppm, while the true result is 10 ppm. There are about 13% of false negative samples.

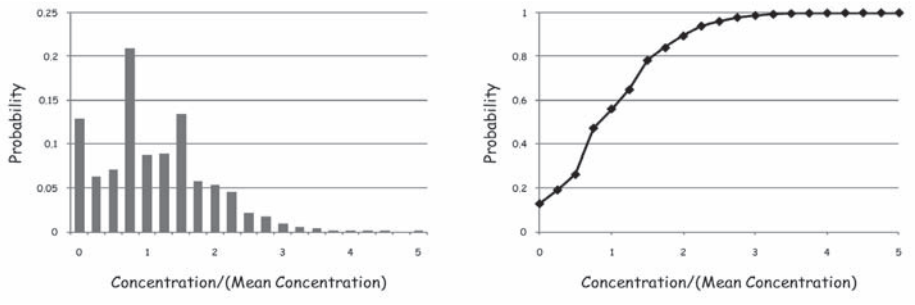

Figure 10 Histogram of sample concentrations and cumulative distribution from simulation, design study

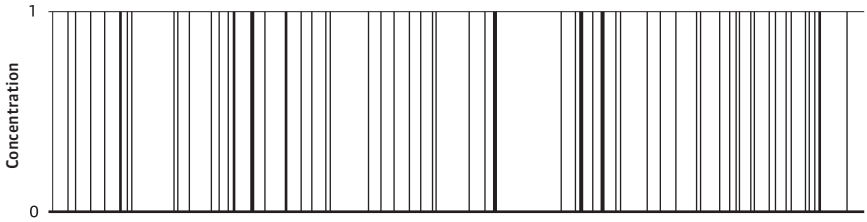

Figure 11 Random slugs for one realisation of the design example

In this example, the slugs of contaminated grain are relatively small and form a volume fraction of 0.625% of the lot. The probability that two slugs overlap is therefore very small. Figure 11 shows a set of 1563 random slugs of contamination in a 2000 tonne shipment. The contaminant must be mixed into much larger volumes of grain before there is a reasonable probability of overlap. Figure 12 shows a portion of a realisation of the concentration of contaminant as a function of time for slugs where the mixing has lead to 22 tonnes being contaminated. In this case there has been an overlap of lots of contamination according to the model. However, the grain is still heterogeneous.

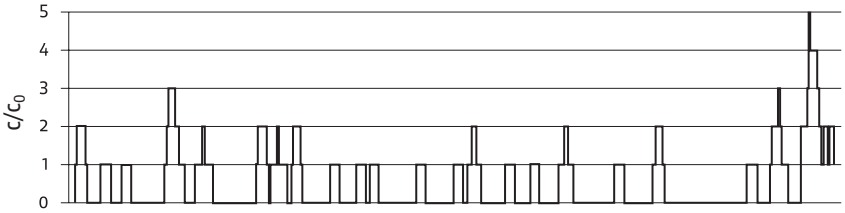

Figure 12 Concentration as a function of time for 22 tonne slugs (portion of data set)

Depending on the size of the shipment, more than one sample may be collected and analysed. For a 10 000 tonne shipment, five subsamples would be collected and analysed separately. The RSD, excluding any analytical variance component would then reduce by a factor of $1/\sqrt{5} = 0.447$. In the above case, this would bring the RSD to 0.343 (using the theoretical result). The 95% confidence interval must be found by appropriate simulation, but will be significantly reduced.

VARIOGRAMS OF CONCENTRATION

In general, there is no analytical expression for the variogram of concentration. However, when the slug size is small and the number of slugs is not too large ($um<<1$), there will be almost no overlap of the slugs and the slugs form a Boolean random set. In such a case, it is known that the variogram rises from zero to a sill of:

$$\gamma(\infty) = f(1-f) \tag{27}$$

where:

$$f = 1 - e^{-um} \tag{28}$$

over the distance u, the extent of the slug. f is the fraction of the grain that is contaminated.

Therefore, unless the mixing of the contaminant into the grain is nearly complete or the sampling and analysis of the grain is extremely frequent, one does not expect to observe a variogram of the contaminant concentration.

Even when the contamination is spread out over a significant fraction of the lot under consideration, if the contamination is adventitious, the variogram of concentration is not expected to be well-defined as a consequence of the random nature of the placement of the contaminated material.

CONCLUSIONS

The mathematical analysis of the sampling problem is exact for uniform distributions of contaminants in the 'slug' of contaminated material and for non-overlapping slugs. More complex sampling problems can be simulated within reasonable time. The design problem simulation with 10 000 replications of the sampling required about 20 minutes of computer time on a late model laptop.

The practical circumstances of sampling as in the design problem indicate that the sampling problem is difficult and that the RSDs due to lack of mixing (distributional heterogeneity) for current practical protocols are relatively large. In many cases the primary increments will not be taken as frequently as in the design example, giving the sampling result a very high variance.

The sampling variance that will be encountered in cases in which the number of slugs is low and the average local concentration is higher than say 2000 ppb will generally lead to unsatisfactory uncertainties in sampling. In such a case, the sampling of the grain flow must be very frequent, leading to a substantial fraction of the primary flow being recovered as a flow of primary increments. Consequently, multistage sampling systems must be used to bring the sample volume down to a manageable mass.

Blending within the handling system by means of which the slugs of contaminated material are mixed throughout uncontaminated material has a very beneficial effect on the sampling requirements. However, there is little if any information on the extent of mixing in grain shipments contaminated by mycotoxins. This problem requires some urgent attention to establish the characteristics of mycotoxin contamination in vessels on loading or unloading.

The problem remains as one for which the bulk food commodity regulators seek a good solution. The results from this analysis extend the ability to assess sampling and contamination scenarios in a robust manner.

The results herein may also assist in dealing with environmental problems of site contamination.

REFERENCES

Paoletti, C., Donatelli, M., Kay, S., & Van den Eede, G. (2003) 'Simulating kernel lot sampling: the effect of heterogeneity on the detection of GMO contaminations', *Seed Science and Technology*, vol. 31, pp. 629–638.

Canadian Grain Commission (2009) *Sampling Systems Handbook and Approval Guide,* Canadian Grain Commission Industry Services, <www.grainscanada.gc.ca>.

APPENDIX A – THE VARIANCE OF THE SAMPLE CONCENTRATION

Equation 14 of the text provides the expression for the concentration of the analyte in a sample composed of n increments from the flow of material. The objective is to evaluate the variance of c_S. One has:

$$\mathrm{var}\{c_S\} = \frac{1}{n^2}\mathrm{var}\left\{\sum_{k=1}^{j} c_k\right\} \tag{A1-1}$$

where the c_k are independent identically distributed random variables with the expected value given by Equation 6 and variance given by Equation 7 or 10 of the text as the case may be. By definition one may write:

$$\mathrm{var}\{c_S\} = \mathrm{E}\left\{\left[c_S - \mathrm{E}\{c_S\}\right]^2\right\} \tag{A1-2}$$

and introduce the conditional variable $c_S \mid j$. This is the sample concentration given j slugs hits. Then:

$$\begin{aligned}\mathrm{var}\{c_S\} &= \mathrm{E}\left\{\left[c_S - \mathrm{E}(c_S \mid j) + \mathrm{E}(c_S \mid j) - \mathrm{E}\{c_S\}\right]^2\right\} \\ &= \mathrm{E}\left\{\left[c_S - \mathrm{E}(c_S \mid j)\right]^2 + 2\left[c_S - \mathrm{E}(c_S \mid j)\right]\left[\mathrm{E}(c_S \mid j) - \mathrm{E}\{c_S\}\right] + \left[\mathrm{E}(c_S \mid j) - \mathrm{E}\{c_S\}\right]^2\right\}\end{aligned} \tag{A1-3}$$

The expectation of the middle term may be shown to be zero by explicitly multiplying out the terms and taking expectations. The first term $\mathrm{E}\left\{\left[c_S - \mathrm{E}(c_S \mid j)\right]^2\right\}$ is an expectation taken over both the concentration distribution and the distribution of the number of hits. These expectations can be taken sequentially by writing:

$$\mathrm{E}\left\{\left[c_S - \mathrm{E}(c_S \mid j)\right]^2\right\} = \mathrm{E}_j\left\{\mathrm{E}_c\left(\left[c_S \mid j - \mathrm{E}(c_S \mid j)\right]^2\right)\right\} \tag{A1-4}$$

The inner expectation is:

$$\mathrm{E}_c\left(\left[c_S \mid j - \mathrm{E}(c_S \mid j)\right]^2\right) = \frac{1}{n^2} j\,\mathrm{var}\{c\} \tag{A1-5}$$

as the conditional variables all involve a sum of j identical variance terms. Taking the expectation over the number of hits, one has:

$$\begin{aligned}\mathrm{E}\left\{\left[c_S - \mathrm{E}(c_S \mid j)\right]^2\right\} &= \mathrm{E}_j\left\{\frac{1}{n^2} j\,\mathrm{var}\{c\}\right\} \\ &= \frac{1}{n^2}\mathrm{E}\{j\}\,\mathrm{var}\{c\}\end{aligned} \tag{A1-6}$$

The last term $\mathrm{E}\left\{\left[\mathrm{E}\left(c_S \mid j\right)-\mathrm{E}\left\{c_S\right\}\right]^2\right\}$ involves expectations only over the number of hits and since:

$$\mathrm{E}\left\{c_S\right\}=\frac{1}{n}\mathrm{E}\left\{c\right\}\mathrm{E}\left[j\right] \tag{A1-7}$$

and:

$$\mathrm{E}\left\{c_S \mid j\right\}=\frac{j}{n}\mathrm{E}\left\{c\right\} \tag{A1-8}$$

then:

$$\begin{aligned}\mathrm{E}_j\left\{\left[\mathrm{E}\left(c_S \mid j\right)-\mathrm{E}\left\{c_S\right\}\right]^2\right\} &= \mathrm{E}_j\left\{\frac{\mathrm{E}^2\left\{c\right\}}{n^2}\left[j-\mathrm{E}\left\{j\right\}\right]^2\right\} \\ &= \frac{\mathrm{E}^2\left\{c\right\}}{n^2}\mathrm{var}\left\{j\right\}\end{aligned} \tag{A1-9}$$

The final result is:

$$\mathrm{var}\left\{c_S\right\}=\frac{1}{n^2}\mathrm{E}\left\{j\right\}\mathrm{var}\left\{c\right\}+\frac{\mathrm{E}^2\left\{c\right\}}{n^2}\mathrm{var}\left\{j\right\} \tag{A1-10}$$

For $u<v$:

$$\begin{aligned}\mathrm{var}\left\{c_S\right\} &= \frac{1}{n^2}\left[mqc_0^2\left(\frac{u}{v+u}\right)^2\left[\frac{2uv-u^2}{3v^2}\right]+c_0^2\left(\frac{u}{v+u}\right)^2 mq\left(1-q\right)\right] \\ &= c_0^2\frac{mq}{n^2}\left(\frac{u}{v+u}\right)^2\left[\frac{2uv-u^2}{3v^2}+1-q\right]\end{aligned} \tag{A1-11}$$

For $u\geq v$:

$$\begin{aligned}\mathrm{var}\left\{c_s\right\} &= \frac{1}{n^2}\left[mqc_0^2\left(\frac{u}{u+v}\right)^2\left[\frac{2uv-v^2}{3u^2}\right]+c_0^2\left(\frac{u}{v+u}\right)^2 mq\left(1-q\right)\right] \\ &= c_0^2\frac{mq}{n^2}\left(\frac{u}{u+v}\right)^2\left[\frac{2uv-v^2}{3u^2}+1-q\right]\end{aligned} \tag{A1-12}$$