Geoffrey J. Lyman
Materials Sampling & Consulting, Australia

Florent S. Bourgeois
Materials Sampling & Consulting Europe, France

Sheryl Tittlemier
Grain Research Laboratory, Canadian Grain Commission, Canada

Use of OC curves in quality control with an example of sampling for mycotoxins

ABSTRACT

An 'operating characteristics' (OC) curve is a simple tool that has been in use in quality control for many years but does not seem to be widely applied in the particulate sampling field. The OC curve provides the probability that a lot of material will be deemed to meet a specification (will be found to have an assay that falls above (or below) a specified level, given the true assay of the lot). In the application considered herein, it provides the probability that a grain shipment will be accepted, given the true value of the assay for the lot. It directly measures the probability of a type II error.

To construct the OC curve for a given sampling protocol, it is necessary to know all the relevant components of variance and their distribution, as a function of the level of contamination in the shipment. This may be quite a challenge in many circumstances as the assumption of normality of distributions may be poor when dealing with substances such as mycotoxins.

The paper introduces the method of OC curve construction and reviews the method developed by Whitaker for the construction of OC curves for mycotoxins in a wide range of commodities. It is shown that his method excludes a potentially critical component of uncertainty. Further, the discussion concludes that the estimation of the distribution of the missing component of uncertainty is potentially prohibitively expensive and logistically very difficult.

The final conclusion is that more intensive sampling methods should be employed for mycotoxins.

INTRODUCTION

The monitoring of the quality of bulk commodities is a critical element of international and domestic trade and both seller and buyer are concerned with the risks involved. The seller needs to be able to quantify the risk of shipping a lot that will fail to meet contractual specifications when it is sampled upon unloading at its destination and the buyer needs to quantify the risk of accepting a lot that does not actually meet specification.

When the lot of material being shipped or received is relatively homogeneous and the sampling, sample preparation and analysis variances are relatively small, the assessment of the risks is not a difficult problem, especially when the uncertainties are normally distributed. The shipper of a cargo of coal where the cargo is made up from blended stockpiles, the inputs to which were sampled as the piles were being built, can aim for an ash content that is just below the contract specification and assess his risk, knowing that the uncertainty in the cargo ash content will be normally distributed with a relatively small variance. He can improve his revenue by selling his client as much ash and moisture as he is allowed to.

In contrast, the monitoring of bulk shipments of foodstuffs for mycotoxins, pesticide residues, heavy metals or low levels of genetically modified varieties of the commodity usually presents a more difficult situation due to the fact that the lot being shipped may be heterogeneous and the subsequent sampling, sample preparation and analysis procedure is subject to relatively large uncertainties. Distributional heterogeneity (DH) of the commodity demands intensive sampling of the primary flow and the intrinsic heterogeneity (IH) of the commodity with respect to the analyte must be well understood to ensure that the sample preparation protocol limits consequent uncertainties as far as possible. In analysing for trace components, the analytical uncertainty can be expected to be skewed.

When uncertainties do not follow a Gaussian distribution, it is not sufficient to estimate a sampling variance; the full distribution of the uncertainty must be defined and such definition demands a much fuller understanding of the heterogeneity of the commodity being shipped. A tool which brings all the statistical information together to permit the explicit determination of both sellers and buyers risk is needed. The statistical tool which meets this need is called and operating characteristics (OC) curve.

The OC curve directly quantifies the seller's and buyer's risk as a function of the true analyte content of the shipment.

The first part of the paper contains some remarks on the occurrence of mycotoxins which are intended to be helpful in assessing the IH and DH of commodities with respect to a mycotoxin. The second part reviews the construction of the OC curve for a commodity using a simple example and a more complex example. The last part of the paper examines the issue of sampling wheat for mycotoxins with the objective of defining all the elements of the problem that must be brought together to arrive at a robust solution.

SOME MYCOTOXINS AND THEIR OCCURRENCE

Mycotoxins are secondary metabolites of fungi. There are a large number of moulds (microfungi) that produce mycotoxins in grains, but due to their occurrence and potential health hazards, some of these toxins are more important than others. The presence of mycotoxins in grains or other foods is regulated in many countries. Some of the more economically - and food safety - relevant mycotoxins and their sources are listed in Table 1.

DON is produced by *Fusarium* species and is associated with the fungal cereal disease *Fusarium* head blight (FHB). FHB develops when cereals at the flowering stage are infected with *Fusarium* fungal species. Infection is favoured by moist, warm conditions during flowering. FHB results in a reduction in yield and quality of grain, as well as the production of DON. DON has been measured in a number of cereals and cereal-based foods. Reported concentrations are generally less than 1000 μg/kg, but do range up into the 1000s mg/kg (Murphy et al., 2006).

With respect to the distribution of DON, FHB tends to infect not just a small portion of a crop, and thus affected kernels can be present in most seed heads. The more seriously affected kernels generally contain the most DON, but kernels bearing no symptoms of *Fusarium* damage may also contain the mycotoxin. The range of DON concentrations observed in kernels varied by a factor of approximately only 270. Sinha & Savard (1997) measured DON in apparently uninfected kernels of wheat at concentrations of 1 to 1.2 mg/kg. Pink shrivelled kernels that appeared most damaged contained DON at concentrations up to 274 mg/kg.

Table 1 Selection of mycotoxins, range of international regulatory limits in food (van Egmond, Schothorst & Jonker, 2007), sources in grains, and potential health effects caused by exposure

Mycotoxin	International regulatory limits	Associated fungal species	Potential effects
Deoxynivalenol (DON)	300–2000 μg/kg	*Fusarium graminearum* and *Fusarium culmorum*	Also called 'vomitoxin' as it causes vomiting and food refusal in animals; immunotoxic in animal studies.
Zearalenone	50–1000 μg/kg	*Fusarium* species	Estrogenic; high exposure reported to cause infertility in livestock.
Ochratoxin A (OTA)	3–50 μg/kg	*Aspergillus* species and *Penicillium verrucosum*	Nephrotoxic; possible human carcinogen.
Aflatoxins B1, B2, G1, and G2	0–35 μg/kg	*Aspergillus* species	Hepatotoxic; class 1 human carcinogen; immunotoxic in animal studies.

As with DON, zearalenone is also produced by *Fusarium* species. It is most often observed in maize, but it also has been measured in other cereals, such as wheat. A multi-year survey of maize from northern Italy measured concentrations of zearalenone from < 50 to 2500 μg/kg. This study also showed that if the preharvest season of maize was characterised by humid and rainy weather, zearalenone production was favoured (Pietri et al., 2004).

Unlike *Fusarium* mycotoxins, OTA and aflatoxins are produced during storage of grains when temperature and moisture conditions are favourable. In some instances, aflatoxins may also be produced on crops in the field. In addition to temperature and moisture, the strain of fungal species is also a factor in the production of these mycotoxins. OTA and aflatoxins are generally not associated with any visible symptoms on grain, as opposed to DON and its relationship with *Fusarium* damaged kernels.

Aflatoxins have been deemed the most significant mycotoxins in foods and feed (Sweeney & Dobson, 1998), but they occur in more mild and tropical climates and are not as relevant in temperate climates. Aflatoxin B_1 is often the most abundant, detected in foods such as maize, peanuts, and tree nuts; it is also the most potent liver carcinogen of the aflatoxins. Concentrations of aflatoxins are generally in the low μg/kg to 100s μg/kg range.

OTA is present in commodities from a variety of climates. It is produced in grains by *Penicillium verrucosum* (Figure 1) in temperate climates, and in coffee, grapes, cocoa beans, and spices by *Aspergillus ochraceus* in more tropical climates. As with the aflatoxins, OTA is observed in the low μg/kg range.

With respect to sampling, the most critical difference between DON and OTA and aflatoxins in grain is the wide range of concentrations observed on individual kernels for OTA and aflatoxins. In a study on aflatoxins in corn, aflatoxin B_1 in individual corn kernels ranged from 3 to 207 000 μg/kg (Shotwell, Goulden, & Hesseltine, 1974). Similarly, OTA in individual wheat kernels has been observed to range from below the limit of quantitation of 20 μg/kg up to the 100 000s μg/kg (Canadian Grain Commission, unpublished data). The IH of aflatoxins and OTA in grain appears to be much larger than for DON in grain.

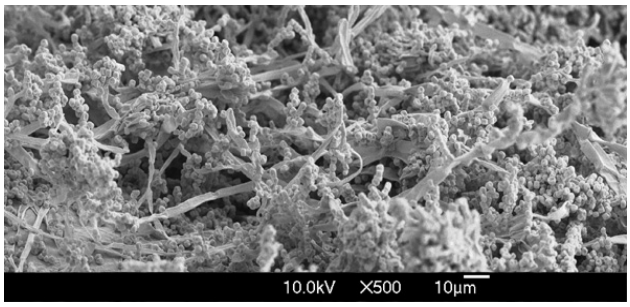

Figure 1 Scanning electron micrograph of Penicillium verrucosum on wheat

The greater heterogeneity of OTA in grains as compared to DON is also demonstrated in a study performed by Biselli, Persin & Syben (2008). In this work, multiple samples were taken from a 26 t truckload of wheat and analysed for DON and OTA. The mycotoxin produced in the field – DON – was found to be more homogenously distributed throughout the truckload than OTA, which is produced during storage. From the point of view of sampling, the mode of occurrence of the fungal infection may be an important factor in the heterogeneity of mycotoxin in a bulk lot of grain.

The difference in heterogeneity requires that sampling of bulk grains for OTA and aflatoxins employ more intense strategies than for DON. OTA and aflatoxins, when sufficiently established in a lot of wheat or corn, give rise to an extremely heterogeneous distribution of mycotoxin in a bulk lot of grain. For example, OTA developing in storage will produce local 'hot spots' of fungal colonies and mycotoxin contamination. Unless handling mixes the fungal colonies and OTA within the lot, or throughout a much larger volume of grain, the hot spots will persist. The DH of the grain with respect to regulatory limits or other targets will be higher. This situation makes the detection of the hot spots by cross-stream sampling, a common sampling method in handling bulk grain, extremely difficult.

OC CURVE DEVELOPMENT

The OC curve was developed along with statistical quality control in the 1930's and 40's and in the simple form for a one-sided decision making process plots the probability of acceptance of a lot vs the true (not measured) quality of the lot based on a prescribed test of the quality of the lot and on an acceptance rule.

A buyer of coal will sample the coal on delivery according to a fixed protocol and analysis method and his measurement of the lot ash content will carry an uncertainty of 0.1% ash at one standard deviation. He will accept the lot if the measurement is less than 10% ash. The distribution of the uncertainty in the ash content is Gaussian (normal).

The OC curve is simple to construct. For example, if the true ash content of the lot is 10.2%, a buyer may be interested in determining the probability that the measured ash content comes out at 10% or less. This is simply the probability that the uncertainty is greater than two standard deviations. Tables indicate that the probability is 0.0228. So the buyer has a probability of 0.0228 of accepting a lot of 10.2% ash. He accepts a lot of 10.1% ash with a probability of 0.1587 and one of 10% with probability 0.5000. Plotting the full OC curve, Figure 2 results.

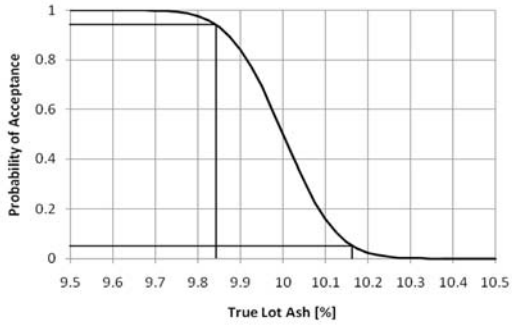

Figure 2 OC curve for coal sampling example

Since the buyer can only accept or reject the lot, the probability of rejecting the lot is $1-p$ when the probability of acceptance is p.

If the seller uses the same sampling and analysis method as the buyer, this curve will also apply to his risks. His risk of having his cargo incorrectly rejected if he supplies coal at $10-0.1\times1.65=9.835\%$ ash is 0.05 and he can expect to have coal at 10.165% ash rejected 95% of the time.

The important fact about the OC curve is that the quality value in the plot refers to the true (but unknown) quality of the lot.

The buyer also has the option of changing his acceptance rule. If what he really wants to have is coal that will only contain more than 10% ash with a probability of 0.05, he must move the acceptance level down to 9.835% ash, shifting the whole curve to the left by 0.165% ash. Another option for control of the buyer's risk is to improve the sampling and analysis to reduce the standard deviation (SD) of the measurements. Reducing the SD will increase the steepness of the curve.

The OC curve provides a visual yet quantitative representation of the buyer's and seller's risks under various measurement scenarios (measurement uncertainty distribution) and decision making policies. An optimal situation for the buyer results when the cost of having to deal with higher ash coal is balanced by the cost of the cleaner coal that must be demanded from the seller. Construction of OC curves for various levels of measurement uncertainty can also lead to an optimal sampling protocol where the higher cost of the more accurate sampling is balanced by a lower probability of having to deal with higher ash coal.

It can be noted in passing that the OC curve must be known in situations where there is motivation to use control charts or to design cumulative sum procedures for detecting the drift of a quality parameter away from a target. The OC curve captures the full distribution of uncertainty in the final measurement and is consequently fundamental to the issue of quality control.

This first example regarding the ash content of coal is based on an assumption that the measurement uncertainty is constant. Such an assumption is reasonable when dealing with ash contents over a rather narrow range. However when dealing with foodstuffs, the level of contaminant may vary relatively more and the uncertainties due to material heterogeneity (IH and DH) may be much larger. As an example for this scenario, green coffee bean sampling for OTA will be considered. A study by Vargas, et al. (2004; 2005; 2006) provides good documentation of how one might quantify sampling, sample preparation and sample analysis variances and the distribution of the sampling result as a function of the OTA concentration to enable the construction of the OC curve for the commodity. Whitaker has used the basic experimental method to build OC curves for a variety of mycotoxins (OTA, aflatoxins and DON) in a range of commodities (corn, peanuts and almonds).

The basic procedure for the coffee investigation involved the following steps:

- Identify multiple lots of coffee that have a range of OTA levels.
- Sample each lot according to a specific protocol in a mechanically correct manner, collecting a composite sample from each lot that is sufficiently large.
- Homogenise the composite sample.
- Create 16 nominally identical subsamples from the composite.
- Prepare each subsample by grinding to < 0.84 mm.
- Subsample the ground subsample to obtain two nominally identical 25 g solid aliquots for analysis for eight out of 16 subsamples. Take only one subsample for the other eight.
- Extract each 25 g solid aliquot with 200 ml solvent and filter.
- From each filtrate prepare 100 ml of solution for analysis; duplicate this step for one out of three filtrates.
- Analyse the liquid aliquots for OTA.

For each of the 16 subsamples, there is an average of two final analyses; 25 lots of coffee were selected resulting in a total of 800 analyses for the programme. A diagram of the testing procedure together with the means of estimating the variance components is provided in Appendix A.

The discussion of Appendix A shows that of the three variance components estimated for each lot, the first component is due to the IH of the unground coffee, the third component captures the analytical variance at the liquid aliquot level and the second component includes the variance due to splitting ground coffee from 1 kg to 25 g and making the primary extraction from the ground coffee. The components of variance are described as (1) the sampling variance, (2) the preparation variance and (3) the analytical variance.

Each variance component was submitted to regression to estimate the variance as a function of concentration. The objective of this was to provide a breakdown of the variance component so that the dependence of the variance component on sample or subsample mass could be introduced. For example, the regression relationship between the sampling variance and the concentration was:

$$\hat{s}_S^2 = 1.35c^{1.090} \quad (1)$$

It was then assumed that the variance for a sample of M_S kg of unground coffee would be given by:

$$\hat{s}_S^2 = \frac{1}{M_S} 1.35c^{1.090} \quad (2)$$

Since this variance component is associated with the IH of the unground coffee, this is a reasonable assumption. For the sample preparation variance, the regression result provided:

$$\hat{s}_p^2 = 0.272c^{1.457} \quad (3)$$

and it was assumed that using m_{ss} g as a subsample instead of 25 g would provide a variance of:

$$\hat{s}_p^2 = \frac{25}{m_{ss}} \times 0.272c^{1.457} \quad (4)$$

This variance is associated with both the splitting variance from 1 kg to 25 g and the extraction variance. The splitting variance is approximately correct but it must then be assumed that the extraction variance is also inversely proportional to the mass extracted. For the analytical variance, the regression gave:

$$\hat{s}_a^2 = 0.00825c^{1.609} \quad (5)$$

and it was then assumed that for n aliquots from the 200 ml extract, the variance would be:

$$\hat{s}_a^2 = \frac{1}{n} 0.00825c^{1.609} \quad (6)$$

and this is correct as the determinations are independent.

The final model for the total variance was:

$$\hat{s}_t^2 = \frac{1}{n} 0.00825c^{1.609} + \frac{25}{m_{ss}} \times 0.272c^{1.457} + \frac{1}{M_S} 1.35c^{1.090} \quad (7)$$

This relationship does add up within a few units of variance to the total sampling variance regression on concentration.

Having modelled the variances as a function of concentration, the last step is to find a distribution function that adequately describes the observed data. For each of the 25 lots analysed, the 16 analyses can be used to calculate the empirical distribution function (EDF) for the data set. Then a distribution can be trialled as a representation of the EDF. Vargas and Whitaker used a log-normal distribution to represent the distribution function, using the mean and variance of the data set to determine the two parameters of the log normal. It would have been better to use the data set mean and the modelled variance from Equation 7 as this is what one will have in constructing the OC curve.

The construction of the OC curve can now be made using the log-normal cumulative distribution function, the model for the total variance, Equation 7 and an accept/reject rule.

Assume for this illustration that the lot will be accepted if the observed analysis of the lot is 5 µg/kg or less and rejected otherwise.

The relationship between the parameters (μ,σ) of the log normal and the mean, $\overline{c}$, and variance, $s_t^2(\overline{c})$, of the distribution is:

$$\sigma^2(\overline{c}) = \ln\left(\frac{s_t^2(\overline{c})}{\overline{c}^2} + 1\right)$$

$$\mu(\overline{c}) = \ln\overline{c} - \frac{\sigma^2}{2} \quad \textbf{(8)}$$

The cumulative distribution function for the normal distribution is:

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2}\, dt \quad \textbf{(9)}$$

Given that the true concentration of OTA in a lot is $\overline{c}$, the probability that a value c_l of 5 ppb or less is observed for a lot is:

$$\Pr\left[c_l < 5 \mid \overline{c}\right] = \Phi\left(\frac{\ln 5 - \mu(\overline{c})}{\sigma(\overline{c})}\right) \quad \textbf{(10)}$$

The last equation defines the OC curve for the OTA in coffee, as investigated by Vargas and Whitaker.

There is however one serious shortcoming of the study they made, as noted in the Appendix. The experimental programme fails to account for the sampling variance due to the DH of the lot sampled. Only one sample was taken from each lot. Had a second sample been taken, it would have a different OTA content. The samples are taken from every 4th 60 kg bag of beans and there were 320 bags or more in the lot. The observed OTA variation between the 1 kg subsamples is due entirely to the IH of the coffee with respect to OTA. A bag probably holds a high percentage of uninfected bean mixed with a small fraction of infected beans. The infected beans will carry OTA concentrations much higher than the average for the bag and the distribution of concentration from one infected bean to the next may follow a log-normal distribution. It has been demonstrated that wheat kernels carrying a log-normal distribution of OTA will produce log-normal distributions of OTA concentration when diluted with clean wheat and sampled (Lyman et al., 2009).

Excluding the existence of another component of variance that derives from the DH of the coffee in the bags, as done in the Vargas and Whitaker methodology, omits a potentially important and possibly dominant component of variance.

FULL OC CURVE DEVELOPMENT FOR MYCOTOXINS

The foregoing section has demonstrated that the Whitaker methodology is not sufficient to capture all of the variability associated with the sampling of the lot of coffee and to construct the OC curve for the sampling for mycotoxins. Omission of the variance component due to the DH present between and within bags and use of an OC curve based on the omission is likely to substantially under-estimate the risks to buyer and seller.

What is required is a methodology that includes the sampling variance due to DH within and between bags as well as that due to the IH of the primary sample. This is the normal situation in sampling a commodity; there is a variance component due to the long term variability or spatial variation of the analyte and another expressing the fundamental variance for the analyte which depends on IH of the commodity in question with respect to the target analyte.

In many cases for bulk commodities, the instantaneous variation of the analyte concentration passing along a conveyor will follow a Gaussian random function. It then follows that the uncertainty due to this DH will be Gaussian as well. The uncertainty due to the IH of the commodity may be Gaussian as well and unless non-Gaussian components of variance are introduced in subsampling and sampling, the entire variance can be dealt with using a normal distribution to construct an OC curve.

However, in dealing with mycotoxins, it appears that distributions of sampling uncertainty associated with the IH of the commodity are not normal. This fact can be deduced from the many studies carried out by Whitaker (2011) for various commodities. The observed pattern of results is that the IH sampling variance component is the dominant source of variance and it is modelled with log-normal, negative binomial or compound gamma distributions. The OC curves determined are distinctly non-normal. These observations are valid for aflatoxins as well as OTA.

While Whitaker has characterised the distributions that account for sample IH, sample preparation and sample analysis, there is virtually no information available for sampling uncertainty distribution attached to primary sampling. The best information that is available is simulation results such as presented in Lyman & Bourgeois (2011). The distributions of primary sampling variance can be extremely non-Gaussian but tend towards Gaussian when the mycotoxins are more uniformly distributed in the lot being sampled.

To combine the sampling uncertainty due to DH of the lot and the uncertainty from primary sample IH and subsequent preparation and analysis, the final probability density is the convolution of the two densities. This convolution can always be calculated numerically to obtain the result.

An experimental programme that would permit the estimation of the distribution of the primary sampling uncertainty due to the DH of the lot being sampled and the distribution of sampling, sample preparation and analysis uncertainty would have the following structure:

- The primary sampling of a lot would be replicated many times (e.g. 25).
- Each independent primary sample would be split down to produce not one but 16 or more nominally identical subsamples of the mass required by the prescribed sampling protocol. 24 of these subsamples would be analysed via the protocol and one would be dealt with in a manner similar to that used by Whitaker.

Adding this extra layer of replication to the sampling process and carrying it out over 25 lots having different average levels of contamination would involve an additional 600 analyses. For bagged lots of a commodity, such a procedure is possible. However, for commodities traded in bulk such as the major cereals, such a proposal is almost untenable due to the fact that firstly, it would be very difficult to target consignments of a commodity having a particular level of contamination and secondly, the replication of the sampling at the primary stage could require the installation of new sample handling systems at the loading or unloading facility. Using a figure of 50 dollars for a single analytical determination, the cost for 1400 analyses is 70 000 dollars. The additional cost of setting up the sampling programme would easily match this figure.

These considerations suggest that determination of the full OC curve for mycotoxin determination in a given bulk commodity will be determined only if serious efforts are made and additional correct sampling equipment is located at loading facilities. The best that can be said at present is that it is likely that the OC curves determined by Whitaker must be considered to be subject to an additional source of variance of an unknown magnitude. Such a situation blows out the risk attached to decision making on a lot to a substantial extent.

CONCLUSIONS

It has been demonstrated that the current methodology for construction of OC curves for foodstuff contaminated by mycotoxins lacks inclusion of an additional source of variance. Estimation of the distribution of the additional source of variance experimentally will demand a serious effort and expense and must be determined for each commodity of interest and each mycotoxin.

Simulations studies of the possible range of distributions capturing the DH of commercial consignments may be the only practical means of estimating full OC curves for various commodities. However, the simulations must be based on a series of assumptions, the validity of which can be questioned.

The real challenge in the superintendence of bulk lots subject to mycotoxin contamination is to improve sampling systems while ensuring correctness and to develop sample preparation protocols that minimise uncertainties within reasonable bounds.

REFERENCES

Biselli, S., Persin, C. & Syben, M. (2008) 'Investigation of the distribution of deoxynivalenol and ochratoxin A contamination within a 26 t truckload of wheat kernels', *Mycotoxin Research*, vol. 24, pp. 98–104.

Lyman, G.J., Bourgeois, F.S., Gawalko, E. & Roscoe, M. (2009) 'Estimation of the sampling constants for grains contaminated by mycotoxins and the impact on sampling precision', *Fourth World Conference on Sampling & Blending*, The Southern African Institute of Mining and Metallurgy.

Lyman, G.J. & Bourgeois, F.S. (2011) 'Sampling with discrete contamination', *WCSB5 5th World Conference on Sampling and Blending, 25-28 October, Santiago, Chile,* Gecamin, Santiago, in press.

Murphy, P.A., Hendrich, S., Landgren, C. & Bryant, C.M. (2006) 'Food mycotoxins: an update', *Journal of Food Science*, vol. 71, pp. R51–R65.

Pietri, A., Bertuzzi, T., Pallaroni, L. & Piva, G. (2004) 'Occurrence of mycotoxins and ergosterol in maize harvested over 5 years in Northern Italy', *Food Additives and Contaminants*, vol. 21, pp. 479–487.

Shotwell, O.L., Goulden, M.L. & Hesseltine, C.W. (1974) 'Aflatoxin: distribution in contaminated corn', *Cereal Chemistry*, vol. 51, pp. 492–499.

Sinha, R.C. & Savard, M.E. (1997) 'Concentration of deoxynivalenol in single kernels and various tissues of wheat heads', *Canadian Journal of Plant Pathology*, vol. 19, no. 1, pp. 8–12.

Sweeney, M.J. & Dobson, A.D. (1998) 'Mycotoxin production by Aspergillus, Fusarium and Penicillium species', *International Journal of Food Microbiology*, vol. 43, pp. 141–158.

van Egmond, H., Schothorst, R. & Jonker, M. (2007) 'Regulations relating to mycotoxins in food', *Analytical and Bioanalytical Chemistry*, vol. 389, pp. 147–157.

Vargas, E.A., Whitaker, T.B., Santos, E.A., Slate, A.B., Lima, F.B., & Franca, R.C.A (2004) 'Testing green coffee for ochratoxin A, Part I: Estimation of variance components', *Journal of AOAC International*, vol. 87, pp. 884–891.

Vargas, E.A., Whitaker, T.B., Santos, E.A., Slate, A.B., Lima, F.B., & Franca, R.C.A. (2005) 'Testing green coffee for ochratoxin A, Part II: Observed distribution of ochratoxin A test results', *Journal of AOAC International*, vol. 88, pp. 780–787.

Vargas, E.A., Whitaker, T.B., Santos, E.A., Slate, A.B., Lima, F.B., & Franca, R.C.A. (2006) 'Testing green coffee for ochratoxin A, Part III: Performance of ochratoxin A sampling plan', *Journal of AOAC International*, vol. 89, pp. 1021–1026.

Whitaker, T.B. (2011) *Publications by Thomas B. Whitaker*, list of publications by the Whitaker from 1964 to 2010, <http://www.bae.ncsu.edu/usda/www/researchpaperstbw.htm>.

APPENDIX A–COFFEE TEST PROGRAMME

Whitaker's test programme design is illustrated in Figure A1. The coffee was an Arabica grade 7 (not a particularly good grade).

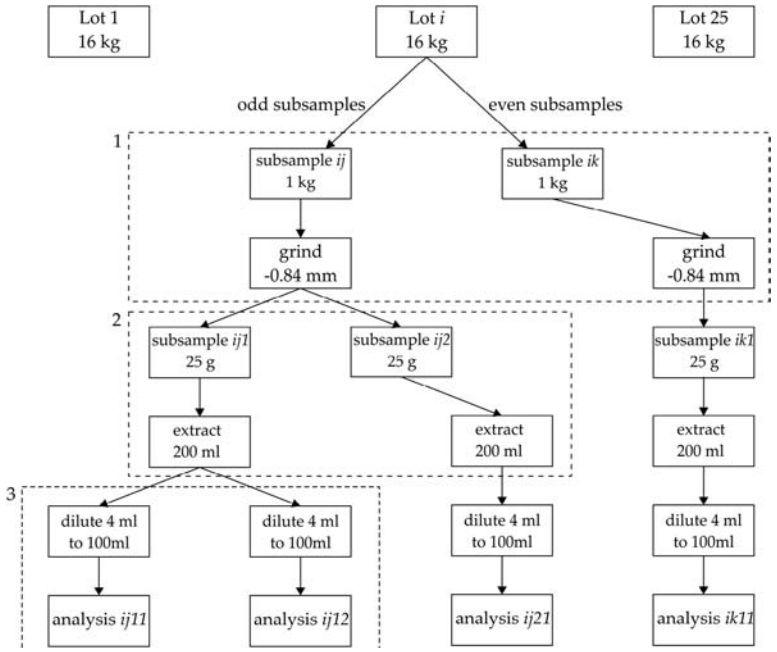

Figure A1 Test programme design for development of OC curve for coffee

From each of the 25 lots selected, a 16 kg sample was collected, homogenised and split down to 1 kg subsamples. Each 1 kg subsample was ground to pass through a 0.84 mm sieve and one or two 'solid aliquots' of 25 g split out. Each aliquot was extracted with 200 ml of solvent and the filtrate collected. 4 ml of filtrate was then diluted to 100 ml and used for analysis.

If the true analysis of the i^{th} lot is denoted as c_i, three components of variance can be associated with an analytical result. An analytical result can be uniquely identified as a_{ijst} where $i = 1, K, 25$; $j = 1, K, 16$; $s = 1, 2$ for j odd and $s = 1$ for j even and $t = 1, 2$ for j odd and $s = 1$ and $t = 1$ otherwise. For each lot, there are eight duplications of the analysis at the liquid aliquot level which permit estimation of the variance of the analysis step (level 3 in the diagram). There are also eight duplications at level 2 and eight duplications at level 1. This permits estimation of variance components for each level.

The analytical results can be written as:

$$\begin{aligned} a_{ij11} &= c_i + \varepsilon_{1ij} + \varepsilon_{2ij1} + \varepsilon_{3ij1} \\ a_{ij12} &= c_i + \varepsilon_{1ij} + \varepsilon_{2ij1} + \varepsilon_{3ij2} \\ a_{ij21} &= c_i + \varepsilon_{1ij} + \varepsilon_{2ij2} + \varepsilon_{3ij3} \\ a_{i(j+1)11} &= c_i + \varepsilon_{1i(j+1)} + \varepsilon_{2i(j+1)1} + \varepsilon_{3i(j+1)1} \end{aligned} \qquad \textbf{(A1)}$$

The variance of ε_{1i} is related to the sampling constant for the coffee in its unground state and the coffee when ground to -0.6 mm. One may write:

$$\mathrm{var}\left\{\varepsilon_{1i\cdot}\right\} = c_i^2 K_{S1i}\left[\frac{1}{1000} - \frac{1}{16000}\right] \tag{A2}$$

and this variance component accounts for the mass reduction from 16 kg to 1 kg. K_{S1i} is the sampling constant associated with the unground coffee.

The variance of ε_{2i}.. is associated with the division of the 25 g solid aliquot from the 1 kg subsample and the extraction of the 200 ml filtrate from the 25 g sample and can be expected to depend on the concentration c_i:

$$\begin{aligned}\mathrm{var}\left[\varepsilon_{2i\cdot\cdot}\right] &= \sigma_{2i}^2 \\ &= c_i^2 K_{S2i}\left[\frac{1}{25} - \frac{1}{1000}\right] + \sigma_{2i}'^2\end{aligned} \tag{A3}$$

where K_{S2i} is the sampling constant for the ground coffee and $\sigma_{2i}'^2$ is the variance associated with the extraction step.

The variance of ε_{3i}.. is associated with the dilution and analysis step for the filtrate.

$$\mathrm{var}\left[\varepsilon_{3i\cdot\cdot}\right] = \sigma_{3i}^2 \tag{A4}$$

The final analytical result for the i^{th} lot and j^{th} subsample should be taken as a weighted sum of the 4 analytical results. These 16 results can then be used to construct and empirical distribution function for the coffee having the average concentration $\hat{c}_i$, where the latter value is the estimate of c_i.

The final step in the analysis is to fit a distribution to the 16 values that is parameterised by the total variance and the mean concentration for the lot. In this case a log-normal density was fitted. The log-normal requires only two parameters, the mean value and the variance. It is this parameterised distribution that is used to construct the OC curve.

There is one critical shortcoming of this methodology; it does not account for the variance due to the DH of a lot and sampling incorrectness issues; a single 16 kg sample represents the lot. To determine the component of variance associated with the sampling of the lot, replicate samples from each shipment of coffee would be required adding another level of testing to the design. Vargas and Whitaker state that the original lot was sampled by taking probe samples of about 200 g from every 4th bag in the lot. The sample then consisted of about 80 increments and the lot mass can be estimated at 320 bags or 19.2 tonnes. This experimental plan cannot resolve the primary sampling variance for the coffee sampling protocol and this variance could be as large or larger than any variance component identified.

Consequently, the OC curve developed is missing what perhaps is the most important component of variance. The first component of variance identified, termed the sampling variance by Vargas and Whitaker is, according to Equation A1, related only to the IH of the coffee and not the DH of the lot.