doi: 10.1255/tosf.74

Distributional assumptions in food and feed commodities: how to develop fit-for-purpose sampling protocols?

C. Paoletti[a] and K.H. Esbensen[b]
[a]European Food Safety Authority (EFSA), Parma, Italy. E-mail: claudia.paoletti@efsa.europa.eu
[b]National Geological Surveys of Denmark and Greenland, Copenhagen and Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University campus Esbjerg (AAUE), Denmark. E-mail: ke@geus.dk

Bulk food and feed sampling is a multi-step procedure in which typically a composite sample is first produced by pooling primary increments, thoroughly mixed and then mass-reduced (possibly in several steps) to obtain an ultimate laboratory sample of suitable size for analysis: the test portion, or the analytical aliquot. Among all sampling steps involved in this pathway, application of composite sampling is the most critical. If the primary sample cannot be proven to be representative, all ensuing steps of mass-reduction, sample preparation and analysis are in vain, for reasons recently explained in full in the horizontal standard DS 3077[1], where the specific requirements for ensuring representativeness, are addressed in full.

Although it is well known that material heterogeneity influences the effectiveness of sampling procedures, most guidelines defining sampling strategies specific for, or routinely applied to, food and feed products are based on stringent distributional assumptions, seldom justified or discussed in sufficient detail, if at all. Indeed most are based on classical statistical distribution requirements – foremost the normal, binomial and Poisson distributions - and almost universally rely on the assumption of randomness[2, 3, 4]. This is an unrealistic and suboptimal state of affairs at best however. Does the supposed randomness relate to constitutional heterogeneity or to distributional heterogeneity for example? How are the unavoidable irregular spatial distributions accounted for? The scientific and industrial communities actually recognizes a strong preponderance of non-random distribution within commodity lots[5, 6, 7, 8], which therefore should be the more realistic pre-requisite for definition of effective sampling protocols. Heterogeneity issues are too often overlooked, instead allowing non-scientific considerations to determine sampling protocols, focusing financial, time, equipment and personnel constraints instead of mandating acquisition of documented representative samples under realistic heterogeneity conditions. We show how the principles promulgated in the Theory of Sampling (TOS), e.g. as practically tested in an EU study on soybean materials[9], actually apply universally in the food and feed realm and should be considered as an *exemplar* for development of valid sampling protocols free from distributional constraints. TOS provides a framework within which identification and development of unbiased sampling plans is driven by empirical observations made on a case-by-case basis and calibrated upon the *specific* heterogeneity characteristics of the material under assessment. Under the guidance of TOS' Fundamental Sampling Principle, systematic application of stratified random sampling will suffice to always 'cover' the entire lot. The appropriate number of increments is not scalable with the size of the lot, contrary to many standard myths perpetuated *ad infinitum*, but only with the degree of heterogeneity in the lot and the a priori chosen degree of confidence, i.e. the acceptable level of risk.

Food and/or feed products constitute no special case in this context: if sampling is not carried out correctly (if biased), subsequent analytical efforts in the laboratory are completely futile[1,5,6,7,10,11,12]. Much work still needs to be done in order to prevent continued use of non-representative sampling protocols that are prevalent in international standards and guidelines, sometimes limited by unsubstantiated distributional assumptions. If providing correct sampling recommendations is a priority for both the scientific community and regulators responsible for consumers' protection, it is necessary to contribute towards a unanimous acceptance of the position that evaluation of the total sampling error (TSE – including laboratory handling errors) is equally important as the evaluation of the analytical error[13]. So far, far too much attention has been devoted only to estimates of the total analytical error (TAE) and many, very specific, and therefore only ad hoc experimental designs, with or without a sufficient number of increments and replicates, has been evaluated only as a function of the specific properties of the analytical method involved. Applying TOS principles allows development fit-for-purpose TSE criteria based on of empirical lot heterogeneity characterisation[5, 6, 7] with which to enter into e.g. risk analysis or compliance testing, on a fully realistic basis.

References

1. DS 3077- Horizontal Standard: Representative Sampling, Esbensen, K.H., "DS 3077 Horizontal—a new standard for representative sampling. Design, history and acknowledgements", *TOS forum* **Issue 1,** 19–22 (2013). doi: 10.1255/tosf.7
2. K. Remund, D. Dixon, D. Wright, & L. Holden, Statistical considerations in seed purity testing for transgenic traits. *Seed Science Res.* 11, 101-120 (2001).
3. M. R. Binns, J. P. Nyrop, & W. van der Werf, *Sampling and monitoring crop protection. The theoretical basis for developing practical decision guides*. CABI Publishing, Oxon, UK (2000).
4. G. W. Cochran, *Sampling Techniques*. 3rd Edition. J. Wiley & Sons Inc., NY, USA (1977).
5. K.H. Esbensen, C. Paoletti and P. Minkkinen, Representative sampling of large kernel lots I. Theory of sampling and variographic analysis. *TrAC* 32, 154-164 (2012).
6. P. Minkkinen, K.H. Esbensen and C. Paoletti, Representative sampling of large kernel lots II. Application to soybean sampling for GMO control. *TrAC* 32, 165-177 (2012).
7. K.H. Esbensen, C. Paoletti and P. Minkkinen, Representative sampling of large kernel lots III. General considerations on sampling heterogeneous foods. *TrAC* 32, 178-184 (2012).
8. C. Wagner, Critical Practicalities in Sampling for Mycotoxins in Feed. *J. AOAC Int*. 98 (2015).
9. C. Paoletti, A. Heissemberger, M. Mazzara, S. Larcher, E. Grazioli, P. Corbisier, N. Hess, G. Berben, P.S. Lübeck, M. De Loose, G. Moran,

C. Henry, C. Brera, I. Folch, J. Ovesna, G. Van den Eede, Kernel lot distribution assessment (KeLDA): a study on the distribution of GMO in large soybean shipments. *Eur. Food Res. Technol.* 224, 129-139 (2006).

10. P. Lischer, Sampling procedures to determine the proportion of genetically modified organisms in raw materials Part I: correct sampling, good sampling practice. *Mitt Lebensm Hyg*. 92, 290-304 (2001).
11. P. Gy, Sampling for Analytical Purposes. Wiley. ISBN 0-471-97956-2 (1998).
12. K. H. Esbensen, C. Wagner, Theory of Sampling (TOS) versus measurement uncertainty (MU) – A call for integration. *TrAC* 57, 93-106 (2014).
13. C. Paoletti and K.H. Esbensen, Distributional assumptions in food and feed commodities – development of fit-for-purpose sampling protocols. *J. AOAC Int*. 98 (2015).