SAMPLING COLUMN

Heterogeneity—the root of all evil (part 2)

Kim H. Esbensen[a] and Claas Wagner[b]
[a]Geological Survey of Denmark and Greenland (GEUS) and Aalborg University, Denmark. E-mail: ke@geus.dk
[b]Sampling Consultant. E-mail: cw@wagnerconsultants.com

All sampling has to deal with materials that are heterogeneous at one scale or another (or at all scales). It is vital to understand the characteristics of heterogeneous materials. Whatever form of primary sampling is to be used, its primary objective must be to *counteract* the effects of the material heterogeneity encountered. Up to five sampling errors are potentially always in play because all sampling processes *interact* with heterogeneous lots. Two sampling errors arise because of the lot heterogeneity (the so-called "correct sampling errors"), and three additional errors are produced by the sampling process itself if not properly understood and eliminated—which is the role of the Theory of Sampling (TOS). This column continues the introduction to the concepts involved in describing, estimating and managing the adverse effects of heterogeneity in sampling.

Constitutional heterogeneity

Compositional heterogeneity (CH_L) is a reflection of the intrinsic compositional differences *between* the ensemble of individual units which makes up all lots (grains, particles, fragments). A material will display a non-zero constitutional heterogeneity *if* it is made up of *different* constituent units. Mixing will have no effect on this type of heterogeneity. It will be the exact same ensemble of units regardless of to what degree they are mixed up—they remain equally different.

The Theory of Samping (TOS) has coined the concept *heterogeneity contribution* for the contribution made to the full lot heterogeneity by each individual fragment (CH realm), or an increment (DH realm, see below). It is advantageous to focus on heterogeneity contributions because this allows the individual fragment masses to be factored in. Large fragments (defined as masses larger than the average fragment mass) may also carry a large concentration deviation from a_L, with the consequence that the heterogeneity contribution from this fragment will be very large. However, should a fragment, identically large in size (mass), happen to have a concentration very close to, or perhaps accidentally equal to a_L, its contribution to the full lot heterogeneity will be insignificant, regardless of its mass; it is simply a large frag-

ment with (almost) precisely the average lot composition. (Had such a fragment been *grabbed* from the lot, by accident its analytical result would have been both accurate and precise, indeed representative—alas such miraculous knowledge is, of course, never known at the time of sampling.) From the perspective of a significant compositional deviation, large fragments will always contribute overwhelmingly to CH_L, while individual small particles (grains of dust for example), will not matter much to the total material heterogeneity because of their infinitesimal weight. Collectively, however, the fine fraction of a lot *may* constitute appreciably to CH_L if the composition of the units in this fraction deviates significantly from a_L.

It follows that the *appearance* of a lot made up of an array of discernible fragments may very well give a false, and only superficial impression of the state of heterogeneity, because large fragments will dominate the visual grain size distribution impression. But the human eye will not in general be able to *see* the analyte concentrations involved. Similarly, a material made up of almost identical grains, e.g. cement, ground coffee, soy beans, wheat grains, sugar, "fines" (these latter because of their similar scattering effects) etc., may nevertheless still import a highly significant heterogeneity contribution, e.g. regarding trace concen-

tration analytes (e.g. toxicants, mycotoxins or genetically modified organisms). Observe that a material may well simultaneously appear very close to homogenous (e.g. sugar or, say, a 99.9% pure chemical), while in reality displaying a very large heterogeneity with respect to an impurity analyte, which necessarily must be distributed very irregularly at such low concentrations (Figure 1).

There is a forceful lesson to be learned from these simple relationships: the visual impression of heterogeneity can be grossly misleading—and because one will never know when this is the case, or not, the visual impression must consequently **never** be used as a basis for heterogeneity assessments.

Heterogeneity vs sampling

If one was obligated to produce a complete heterogeneity characterisation of a specific sampling material it would be necessary to analyse, and weigh, **all** constituent fragments. This is obviously not possible, nor desirable in sampling practice. Because of this impossibility, *sampling* comes to the fore: only a part of the lot will be physically sampled and eventually analysed. Enter the increment!

What constitutes an *ideal sample*? An ideal sample would have to be composed of a subset of individual fragments selected *individually* from the lot, completely *at random*, i.e. based on total

SAMPLING COLUMN

Figure 1. "Homogeneous powder" with a normal grain-size distribution, but the larger-than-average particle sizes have here been dyed blue, allowing detailed insight into grain-size differentiation and segregation behaviour. The original powder visually makes a totally homogeneous impression ("white powder"). Two different DH_L manifestations are shown to the left and centre. The latter was produced by a single 90° rotation around the vertical axis of the container, illustrating that DH_L manifestations are often *transient* phenomena, which are a sensitive function of a number of agitation factors active in production, handling, transportation and by manipulation while being sampled. The constant presence of the gravitational force field and/or the transient presence of centrifugal forces (as in the example above) will impact the specific DH manifestation of very many types of aggregate materials. The right-hand illustration shows the important effect of *pouring segregation* and the resulting problems in trying to acquire a representative single-sample aliquot using a laboratory spatula. Such discrete sampling operations (grab sampling) can never be representative. Illustration courtesy of Peter Paasch-Mortensen (reproduced with permission).

free access to the full geometrical target volume. This latter demand is codified in TOS' Fundamental Sampling Principle (FSP): there must be free access to absolutely every grain in the sampling process, not just in principle but very much also **in practice**. But it is clear that nobody in their right mind would wish to sample an ideal sample in practice; one must have to accept to sample but a set of neighbouring fragments instead (a coherent collection of neighbouring fragments). In TOS, such a set of neighbouring fragments is termed a *group-of-fragments*, or a *group* for short. While any size and disposition of a group can be envisaged, TOS is in practice only interested in the special group-of-fragments that will end up **in** the sampling tool after a unitary sampling operation. Thus for practical reasons, TOS is only interested in those *groups* that make up extracted *increments*. Sampling, therefore, in practice always takes place by extraction of increments of a size that needs to be optimised (much more of **how to** optimise this number in later columns).

This understanding points to two alternative views of any sampling target. All lots can either be envisaged to be made up by the totality of N_F fragments (critically important for TOS' theoretical and conceptual deliberations), or it may be envisaged to comprise N_G groups-of-fragments, increments (obviously of much more interest for the practical sampling situation). This division establishes two scales-of-interest when coming to grips with understanding the complexities of heterogeneity, the fragment scale and the increment scale. There is actually only one more scale of interest, the full lot size (mass), M_{Lot}. Everything needed for a complete understanding of heterogeneity stems from these three observation scales. There is great conceptual and practical power in these simple scale relationships, which were originally outlined by Pierre Gy, the founder of TOS (Figure 2).

It turns out that the jump **from** fragment scale **to** the group scale level is all we need to be able to derive the second key feature of heterogeneity, the spatial, or the distributional, heterogeneity. Enter the *distributional heterogeneity* of the lot, DH_L which is defined in a completely similar fashion as CH_L. Note that distributional here means distribution in space, distribution *within* the geometrical volume occupied by the sampling material.

Distribution heterogeneity

Before sampling, only *virtual groups* exist in the sampling target. The distributional heterogeneity (DH_L) is defined as
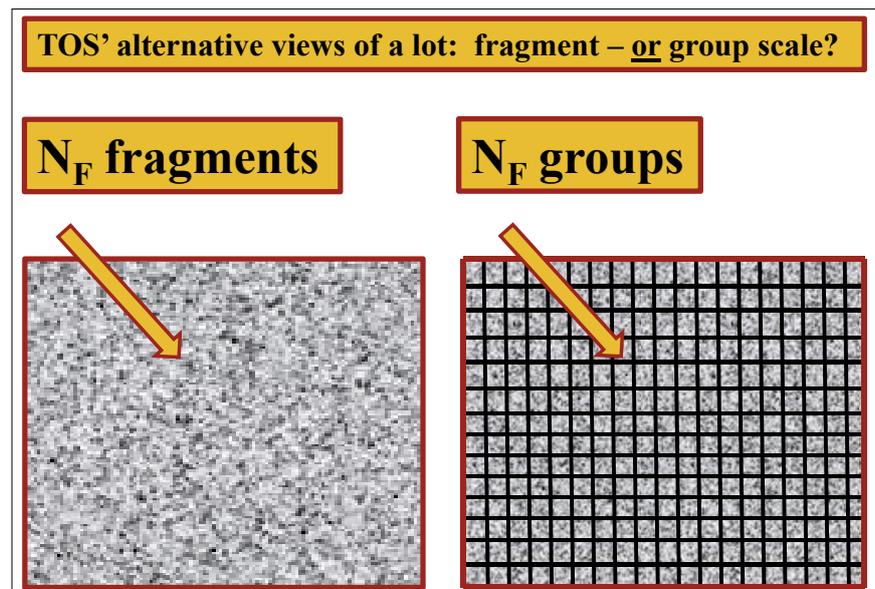

Figure 2. Pierre Gy's inspired *conceptual scale-jump* from fragment-scale to group-scale, which allows definition of distributional heterogeneity (see text). The different observation scales correspond to the different sampling units indicated, fragment vs group. All sampling targets (lots) can be viewed from these alternative vantage points, i.e. either as a collection of N_F fragments, or a collection of N_G groups (increments). The third scale level in TOS corresponds to the sampling target (lot) itself.

SAMPLING COLUMN

the compositional differences *between groups*. Even though now to be viewed only from the point of view of a larger group scale, the entire lot ensemble of different fragments is still available for manipulation, so mixing will now have an effect on the compositional differences between *virtual* groups. Consider an operation like forceful mixing: shaking a pitcher with different types of solid particles; mixing of a segregated slurry tank; whisking egg white and yolk; shaking a cocktail; tumbling mixing in blenders. It is easy to picture in the mind's eye how fragments become increasingly spatially mixed from such an operation. Mixing has a definitive influence on between-group differences, as virtual groups come to be more and more *similar* as mixing progresses.

Note, however, that there is a limit to mixing—"infinite" mixing will not lead to a homogeneous material state, but only to a *minimum residual heterogeneity* state, after which more mixing only results in a steady-state mixing/de-mixing with random deviations confined to centre on this state—thus it may even increase DH locally. Close to this limiting state, there is thus a limit to how much good mixing can do—more mixing far from always leads to a less heterogeneous result, but deviations are of course now confined to be at their smallest levels-as determined by the specific material characteristics... the compositional heterogeneity.

Such between-group variability should logically be termed the *group* Constitutional Heterogeneity, since it is based on compositional differences in a completely identical fashion as for the definition of CH_L. But at this group scale TOS stipulates that the variance of the heterogeneity carried by all groups in the lot is now specifically defined as *Distribution Heterogeneity*. Why DH_L? **Why** isn't CH_L defined as the *compositional heterogeneity* at the group scale level? **How** is it possible to *transform* this into DH_L—isn't this simply just a sleight-of-hand trick? **What** is the relationship between compositional and distributional heterogeneity thus defined?

This is where the inspired insight of the founder of TOS, Pierre Gy, comes into play: what is expressed as identi-

cal compositional differences in the basic definitions has a **different physical meaning** at the two scales involved. Consider a material made up of N_G groups (Figure 2). It follows that the lot volume, the next logical scale level jump, is completely defined as the sum-total collection of all N_G groups. That which statistically can be defined as the variance of all group heterogeneity contributions (*without* spatial considerations) is in the physical 3-D reality precisely the set of groups that physically fills up the spatial lot volume. Thus when statistically summing over all group heterogeneity contributions, one is at the same time summing over the lot volume. Thus this compositional variance physically expresses the spatial differences between all groups, which is the reason why TOS declares that this group-based variance *represents the total spatial heterogeneity of the lot*. While theoretically and formally calculated based on identical compositional difference definitions for CH_L, DH_L actually quantifies the heterogeneity imparted to the lot originating from the different locations of the different groups *within* the lot. The spatial (distributional) heterogeneity of any lot plays out its role at the group-scale level (and at higher scales).

These TOS definitions are usually made invoking a few statistical equations, which have been left out in this didactic treatment. Interested readers can find these definitions in their standard mathematical format, for example in DS 3077[1] or in the basic references listed herein.

It is no coincidence that TOS chooses to define DH_L in this fashion: physical sampling exclusively takes place via extraction of *increments*. An increment *may* occasionally form the whole sample as a singular discrete sampling operation, *grab sampling*, but most emphatically only, and exclusively if, when properly validated as acceptable. In all other instances, increments are sampled with the express purpose of being aggregated to form a composite sample.

While CH_L resides in the scale interregnum between fragment and group, DH_L quantifies the heterogeneity that can

be attributed to the realm between the increment scale and the full lot size. Both these heterogeneity aspects are needed to fully characterise the total heterogeneity of a material, but they cannot be physically separated from one-another. CH_L and DH_L are conceptual, theoretical components that in practice always exist intricately interwoven for any physical material. Not surprisingly, from their closely related definitions, there is also a close mathematical relationship between CH_L and DH_L, see the dedicated TOS literature.

TOS' two interrelated heterogeneity concepts forms the basic element for all of practical sampling, helping practitioners to understand that the primary objective must always first address the negative effects of spatial heterogeneity and segregation. First, after having attended to this imperative, can the other effects related to CH_L be addressed. These aspects of the proper order of dealing with the various aspects of heterogeneity will become clear as these columns starts to address practical sampling in more detail. No worries, we will get there soon enough… ;-)

Structured heterogeneity

Many materials display heterogeneities with a special twist, *structured heterogeneity*, e.g. layered, stratified or otherwise hierarchically organised heterogeneities (Figure 3).

Figure 4 is an example from the food and feed sectors, in which the food elements (minced meat, spices and fat) are present in a very irregular texture, which is indeed the reason behind this particular type of the well-known brand "Chorizo". Sampling of such a material, for example for food compliance or safety purposes, is not a straight-forward issue, perhaps contrary to common thought: "How difficult can it be to sample a salami in the analytical laboratory?" Well, in many sectors, sampling of material with this and *similar* structured heterogeneities takes place with a tubular corer (sampling thief, sampling spear of appropriate dimensions), but from the above it should be clear that a core section of this material runs a severe risk of being non-representative.
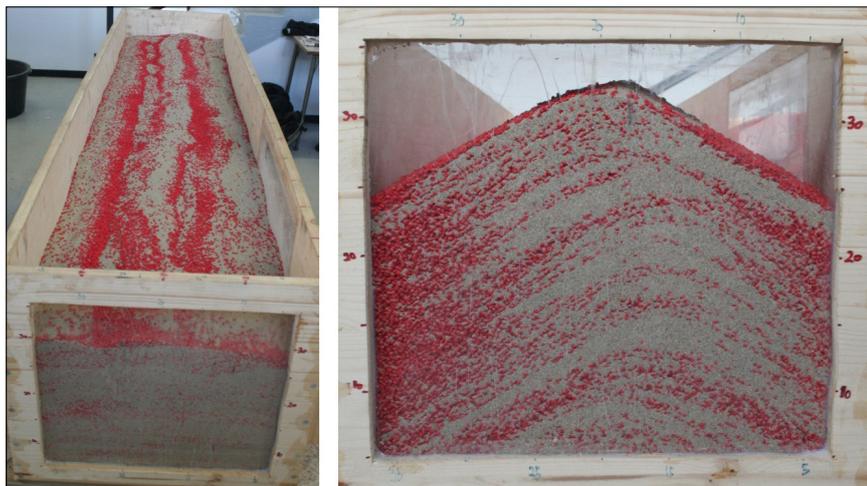
SAMPLING COLUMN

Figure 3. Two examples of strongly *structured heterogeneity*, brought about by different stacking processes, illustrating the type of spatial heterogeneity, DH_L, often present in transportation or storage depots, vessels, trucks, train loads, containers, ship cargo holds but also sometimes in initial reaction and production vessels load-ups. Both lot examples have identical CH_L, illustrating a 10% "analyte" (red plastic pellets) in a 90% sand matrix. Structured heterogeneity is an often occurring characteristic in technology and industry, due to extensive stacking, reclaiming and transportation processes involved in production, processing and manufacturing. It is obvious why grab sampling always comes up short against materials with structured heterogeneity—as indeed against **all** materials with a significant DH_L.

An alternative approach is in fact inspired from the principles of *riffle-splitting*. By slicing a sufficiently high number of slices each covering the full width-thickness of the salami, a division of the material in all aspects identical to riffle-splitting is obtainable.

By selecting a set of slices of equal thickness, a correct TOS mass-reduction can be installed even for materials with very high CH_L and DH_L. Despite the apparent heterogeneity difficulties, a 100% TOS-compliant mass reduction can easily be achieved in principle (albeit sometimes necessitating *some* practical work). The effectiveness, the representativeness of sub-sampling is actually only a matter of practical implementation, i.e. how many thin-slicing cuts one is willing to deploy. More about riffle splitting and laboratory mass reduction (sub-sampling) in future columns.

Two initiating columns have focused of the root of all evil, heterogeneity. It is wrong to address the issues of how to sample heterogeneous materials from the point of view that a fixed increment size (one-size-fits-all lot requirements) will always be able to deal with whatever heterogeneity features are encountered. It is even more wrong to believe that the correct number of increments scales with the total lot mass (the number of necessary increments is a linear function of the lot mass). This is unfortunately a mistake encountered on a massive scale in very many contemporary standards, guidelines and norm-giving documents. Any-and-all sampling procedure that is not scaled with respect to the objectively existing lot heterogeneity will not be able to furnish a representative sample. A thorough analysis of these critical issues was published recently in Esbensen and Wagner.[3]

Acknowledgement

References

1. K.H. Esbensen (chairman taskforce F-205 2008–2013), *DS 3077. Representative sampling—Horizontal Standard*. Danish Standards (2013). http://www.ds.dk
2. L. Petersen, C.K. Dahl and K.H. Esbensen, "Representative mass reduction in sampling—a critical survey of techniques and hardware", *Chemometr. Intell. Lab. Sys.* **74(1),** 95–114 (2004). doi: http://dx.doi.org/10.1016/j.chemolab.2004.03.020
3. K.H. Esbensen and C. Wagner, "Theory of Sampling (TOS) vs Measurement Uncertainty (MU)—a call for integration", *Trends Anal Chem (TrAC)* **57,** 93–106 (2014). doi: http://dx.doi.org/10.1016/j.trac.2014.02.007
4. K.H. Esbensen, C. Paoletti and N. Thiex, "Representative sampling for food and feed materials: a critical need for food/feed safety", *J. AOAC Int.* **98(2),** 249–251 (2015). doi: http://dx.doi.org/10.5740/jaoacint.SGE_Esbensen_intro

Figure 4. Salami ("Chorizo") is an example of highly irregular but still structured material heterogeneity. Tube-coring (using a sampling thief) will not necessarily guarantee a representative sample. Observe, however, how the principle of *riffle-splitting* easily can be applied to this material allowing it to be handled in a fashion identical to that of free-flowing aggregate material. Riffle-splitting can be conducted so as to guarantee representative sub-sampling, see Petersen *et al.*[2]