WILEY
InterScience®
DISCOVER SOMETHING GREAT

Representative process sampling for reliable data analysis—a tutorial

Lars Petersen* and Kim H. Esbensen

Applied Chemometrics, Analytical Chemistry, Applied Biotechnology, Bioenergy & Sampling research group (ACABS), Aalborg University Esbjerg, Niels Bohrs Vej 8, DK-6700 Esbjerg, Denmark

Process sampling of moving streams of particulate matter, fluids and slurries (over time or space) or stationary one-dimensional (1-D) lots is often carried out according to existing tradition or protocol not taking the theory of sampling (TOS) into account. In many situations, sampling errors (sampling variances) can be reduced greatly however, and sampling biases can be eliminated completely, by respecting a simple set of rules and guidelines provided by TOS. A systematic approach for description of process heterogeneity furnishes in-depth knowledge about the specific variability of any 1-D lot. The *variogram* and its derived auxiliary functions together with a set of error generating functions provide critical information on:—process variation over time or space,—the number of extracted increments to composite into a final, optimal sample,—the frequency with which to extract increments—and which sampling scheme will be optimal (random, stratified random or systematic selection). In addition variography will delineate cyclic behaviors as well as long-term trends thereby ensuring that future sampling will not accidentally be performed with a sampling rate coincident with the frequency of any hidden cycle, eliminating the risk of under-estimating process variation. A brief description of selected hardware for extraction of samples from 1-D lots is provided in order to illustrate the key issues to consider when installing new, or optimizing existing sampling devices and procedures. A number of practical examples illustrate the use of TOS and variography to design optimal sampling protocols for a variety of typical process situations. Copyright © 2006 John Wiley & Sons, Ltd.

1. INTRODUCTION

In a previous paper [1] an introduction to sampling of zero-dimensional (0-D) material lots and the theory of sampling (TOS) by Pierre Gy [2–4] was given, which also underlies all of process sampling. 0-D lots are collections of material consisting of separate 'free' units or fragments, without any mutual, ordered correlation (fully independent fragments); 0-D lots can be (or have been) mixed mechanically. The previous paper summarized TOS in a number of practical sampling unit operations (SUO's).

A *one-dimensional (1-D) object* is simply a lot where two dimensions are negligible in size compared to the third, and where there is a distinct spatial or temporal correlation along this defining dimension. This may appear either as an

ordered series of discrete units (time or space) or as a moving/flowing material stream. Any such elongated or moving material body is, strictly speaking, a three-dimensional object, but for all practical purposes it can be regarded as a 1-D object since the singular dominating dimension is overwhelmingly much larger than the other two. For instance, a lot consisting of powder, traveling on a conveyor belt at 1 m/s during an 8 h shift can be regarded as, say, a $0.5 \times 0.2 \times 28\,800$ m lot. From TOS' point of view the strict definition of a 1-D lot is any lot from which a sample, or an increment, forms a complete 'slice' covering two of its three physical dimensions.

This paper deals exclusively with proper sampling of such elongated streams, process streams or ordered series of discrete units while focusing both on the theoretical principles and the practical aspects involved.

The specific sample volume from a 1-D lot, termed 'increments' or 'units', have the possibility of being more-or-less similar to neighboring increments—i.e. being correlated—and this feature may apparently, at first, make

*Correspondence to: L. Petersen, ACABS, Aalborg University Esbjerg, Niels Bohrs Vej 8, DK-6700 Esbjerg, Denmark.
E-mail: Lape@aaue.dk

sampling a bit more tricky, although a full understanding of this feature will soon be seen to make 1-D sampling much easier in practice.

To sample 1-D lots *correctly* [1,5] calls for one of three possible sampling schemes (random, stratified random or systematic), and a number of rules and guidelines regarding the physical extraction of material as given in TOS. These rules and guidelines ensure that sampling will be *unbiased* ('correct') and that the sampling error (variance) is *minimized*. TOS describes the use of a feature called the *variogram* to characterize the auto-correlation (and hence the heterogeneity) of all 1-D lots as a function of inter-increment (unit) distance and this will be shown to greatly help the design of optimal sampling procedures and strategies.

Furthermore, many production and other industrial processes may vary periodically with a more-or-less distinct cycle or according to a somewhat distinct increasing or decreasing trend. This can, for example be due to: 8 h shifts, different shipments of raw material, change of operators, automated process control etc. In such cases, the possibility of committing serious sampling mistakes dramatically increases when not based upon the rich information potential of an initial variographic analysis. The variogram will reveal all, often hidden, cyclic periods or trends, and hence will help avoiding the danger of extracting samples at a frequency *coincident* with the period of a cycle, for example.

This paper provides an introduction to the concept of material heterogeneity in its 1-D form—heterogeneity is the ultimate source of all sampling errors. A brief description of the different errors that contribute to the overall sampling error in the 1-D regimen follows. This tutorial gives a thorough introduction to *variography*, together with a brief description of selected, typical process sampling hardware. Finally, three practical examples of variography for sampling optimization in distinctly different contexts are presented.

2.   BASIC DEFINITIONS AND TERMS

The *lot* is defined as the total volume/mass of material to be sampled. This could typically be in the form of a huge stockpile, a railroad car, a barrel, a small laboratory bag—the rules of correct sampling apply regardless of the physical size of the lot, whereas the difficulty of the actual sampling process depends on the physical size of the lot. In process sampling the lot is an elongated stockpile (extended in space), a material stream (extended in time), but may as well be an *ordered* series of, for instance, railroad cars, production batch units (sacks, barrels etc.) or similar.

Fragments are defined as the smallest inseparable physical parts of the lot (fragments can only be altered by crushing/comminution). This is most often a single particle, for instance a grain of sand (or a fragment of a grain), a granule or a mineral (in a rock) etc.

Increments are defined as a spatially coherent *group of fragments*, characterized as being extracted simultaneously (i.e. in a single sampling operation) by the sampling tool. This may, for instance, be a scoop of extracted material. If several (less than the final sample size) increments are combined, a *composite* sample is in the making.

A *sub-sample* is a sample *reduced* by a representative mass reduction process [6].

The *critical component* (A) is the component of interest; this component can be physical, for example a grain-size bin, or chemical, for example the *analyte*. The critical component is often described by its *grade*, a, either in the whole lot, a_L, or in the individual fragments/units of the lot, a_i. It may also be expressed as a concentration, a percentage etc. The objective of all sampling processes is to be able estimate the *sample grade*, a_S, without bias and with minimum sampling variance—a_S will then be *representative* of a_L.

Autocorrelation is a measure of the degree of spatial (or time) correlation between increments or samples along the defining dimension, expressed as a function of the distance (or time) between the units in question. TOS expresses this feature in a slightly different fashion than, for example regarding time series (and similar) analyses, where the auto-correlation coefficient (or the cross-correlation coefficient between two or more time-series) is the measure. TOS instead focuses on the *variogram*, which includes a measure of auto-correlation as well.

3.   LOT DIMENSIONALITIES

To be able to select an optimal sampling scheme in a given sampling situation it is important to recognize the *lot type*, or the intrinsic sampling dimensionality. Four different lot types exist: 0-, 1-, 2- and 3-D. The major difference is between 0-D and the other three, since 0-D describes lots with no internal correlation between the individual (virtual or real) increments or fragments (see Section 2). A 0-D lot is, in other words, a lot where the fragments or increments are considered random; whereas, in the 1-D, 2-D or 3-D cases, the individual increments (possible sampling volumes) or fragments are *fixed* (and spatially correlated) along the defining dimension(s), either in space or time. TOS deals primarily with 0- and 1-D lots; these are the only kinds which can be sampled correctly in total accord with TOS rules and guidelines. There is, however, a very interesting challenge to apply as much of TOS as possible also in the 2-D and 3-D scenarios and, hence, to achieve the most reliable and representative sampling possible. If the depth-of-interest is fixed and well related to the overall sampling purpose, 2-D sampling can be carried out with complete coverage *in the plane*.

There are principally three different kinds of 1-D lots [2,3]:

- A moving or stationary, stream of *particulate material*. Examples: conveyor belts transporting fragmental materials or powders, slurries in ducts etc.
- A moving or stationary string of *fluids* (i.e. gasses, liquids). Examples: rivers or produced/manufactured fluids in pipelines.
- A moving or stationary stream made of *discrete chronological units*. Examples: railroad cars, truck loads, 'units' (bags, drums, packages . . .) from a production or a manufacturing line.

4. INTRODUCTION TO HETEROGENEITY

Heterogeneity is fundamentally divided in two parts: constitutional heterogeneity (CH) and distributional heterogeneity (DH). The CH describes the heterogeneity dependent on the physical or chemical *differences between* the individual fragments. CH can only be reduced by *altering the material* (for instance by comminution). The DH describes the aspect of heterogeneity dependent upon the *spatial distribution* (and on CH) of the individual fragments or groups-of-fragments in the lot (stratification or segregation and the formation of local groups of fragments with a significant high/low concentration of the critical element). DH can be reduced by using 'correct' sampling methods (more on this later). DH can never be larger than CH. CH is *never* zero. Dependent on the purpose and scale of observation CH *may* be close to negligible, but it is never nil. Homogeneity is defined as the theoretical limit case of zero heterogeneity. Indeed if such a thing as a homogeneous material exists sampling would not be needed (as all sampling errors would be zero). However the concept of homogeneity is only relevant in theory, since true homogeneity never exists in practice for any naturally occurring material. Even in extreme cases one can argue that there are principal differences at molecular or nuclear levels in the form of isomers or isotopes, but it is very rare that TOS is invoked at such ultimate small scales.

4.1. Constitutional heterogeneity, CH

Normally interest is only on the heterogeneity *between* the individual fragments (the lowest practical level of observation) and not so much *within* the fragments. This is where one of the major theoretical achievements of TOS comes to the fore: defining the contribution to the total heterogeneity by an individual fragment. In doing this, TOS characterizes the individual fragments according to the component of interest (the analyte), A, described by the proportion (or grade), a_i, and the fragment mass, M_i. If a lot consists of N_F individual fragments, F_i, with individual masses, M_i, with an average fragment mass, $M_{\bar{i}}$ and the lot has a grade, a_L, of component A, and a mass, M_L, then the heterogeneity contribution from each individual fragment, h_i, can be calculated according to Reference [3]:

$$h_i = \frac{(a_i - a_L)}{a_L} \cdot \frac{M_i}{M_{\bar{i}}} = N_F \frac{(a_i - a_L)}{a_L} \cdot \frac{M_i}{M_L}$$

This definition of heterogeneity contribution is dimensionless and hence any intensive unit can be used in characterizing the material, for example concentration, size. It furthermore compensates for variation in the fragment masses; larger fragments result in a larger influence on the total heterogeneity than smaller ones. This viewpoint constitutes a major distinction from 'classical statistics' where all units contribute equally.

CH is simply defined as the variance of the heterogeneity contributions resulting from the distribution of all individual fragments [3]:

$$CH_L = s^2(h_i) = \frac{1}{N_F} \sum_i h_i^2 = N_F \sum_i \frac{(a_i - a_L)^2}{a_L^2} \cdot \frac{M_i^2}{M_L^2}$$

The above formula is applicable only if dealing with the whole of a finite and isolated population of units. If only a subset of a much larger population is utilized, $(N_F - 1)$ is used instead of N_F.

Calculation of CH_L involves knowing the total number of fragments in the lot, N_F; but, this is of course never a known quantity in practical situations. Instead, at the cost of some approximations, the *constant factor of the constitution heterogeneity*, IH_L, which is independent of the size of the lot, can be used. IH_L is defined [3]:

$$IH_L = CH_L \cdot M_{\bar{i}} = \frac{CH_L \cdot M_L}{N_F} = \sum_i \frac{(a_i - a_L)^2}{a_L^2} \cdot \frac{M_i^2}{M_L}$$

IH_L has the unit dimension of mass and can be calculated via an approximate material's parameter approach, according to the equation below—the so-called 'Gy's formula' or estimated experimentally. The individual descriptions and guidelines for the estimation of the parameters involved can be found in References [1–3,7].

$$IH_L = c\,f\,g\,l\,d^3$$

'Gy's formula' (c, f, g and l are material characteristics)

Gy's formula is *always* able to provide an *estimate* of IH_L (N.B. only when dealing with particulate materials). The reliability of the estimate depends heavily on the quality of the estimates for the material parameters, of which the composition factor (also called the mineralogical factor), c, and the top particle size (d) is the most influential since c can vary between one and infinity (but is *constant* for any material in a specific state) and the particle size (d) is to the third exponent. The remaining factors may often be used at default values, or estimated more precisely for higher overall estimate quality. It is noteworthy that this famous formula generates an estimate to an *order-of-magnitude* only, which is most often all that is needed in practice [2,3].

4.2. Distributional heterogeneity

When considering more than one fragment, a move into the realm of DH of the lot, DH_L is made. If the lot is divided into a number of smaller volumes (groups-of-fragments), N_G, coinciding with the volume of the sampling tool and consider the differences in concentration of the critical components between these volumes (index n), a_n, DH_L can be calculated. In a strict analog to the above definition of heterogeneity carried by a single fragment, a group-of-fragments (index n), G_n, similarly carries an amount of heterogeneity, h_n, which can be calculated from the mass of the individual groups, M_n, the average group mass, $M_{\bar{n}}$, and the average grade of the groups, $a_{\bar{n}}$:

$$h_n = \frac{(a_n - a_L)}{a_L} \cdot \frac{M_n}{M_{\bar{n}}} = N_G \frac{(a_n - a_L)}{a_L} \cdot \frac{M_n}{M_L}$$

The distribution heterogeneity can now be calculated as the variance of these group heterogeneity contributions:

$$DH_L = s^2(h_n) = \frac{1}{N_G} \sum_n h_n^2 = N_G \sum_n \frac{(a_n - a_L)^2}{a_L^2} \cdot \frac{M_n^2}{M_L^2}$$

Unlike for CH_L, which only is a function of the material properties, DH_L can actively be *altered*, for example by

choosing a smaller volume of observation (a smaller sampling tool volume) and in this way reduce it, or the lot can be mixed. More on this can be found in Reference [1].

It is critically important to notice that DH_L can in fact easily be estimated in practice by extracting and analyzing a number of groups 'covering the lot' and calculating the resulting empirical variance. Often a relatively small number of groups will suffice.

In order establish a first link between CH and DH, the following relationship also involves the average constitution heterogeneity of the groups, CH_n^-:

$$CH_L = CH_n^- + DH_L$$

The average constitution heterogeneity of the groups, CH_n^-, is equivalent to the compositional heterogeneity of the material which was extracted as a coherent increment by the sampling tool; that is the material *in* the sampling tool volume, for which CH of course can never be strictly zero. From this follows that CH_L is always larger than or equal to DH_L, and that both of these are always larger than or equal to zero (equal to zero only in extreme and theoretical cases). Thus for natural materials, DH_L is *never* strictly zero, due to the heterogeneous nature of all materials:

$$(DH_L)_{nat} > 0$$

No matter how much work and effort is put into homogenizing a material, DH_L can never be reduced below a certain minimum:

$$(DH_L)_{min} > 0$$

A universal ranking of these relationships results in the following:

$$CH_L \geq (DH_L)_{nat} > (DH_L)_{min} > 0$$

In Reference [3] was also shown the following relationship between CH_L and $(DH_L)_{min}$:

$$(DH_L)_{min} = \frac{N_G - 1}{N_F - 1} CH_L$$

from which it can be appreciated that the 'reduction factor' of CH_L is related to the ratio between the number of groups, N_G (inversely proportional to the number of fragments in a group) and the lot size, represented by the total number of fragments, N_F.

In order to understand heterogeneity in full conceptual detail, Gy further defines two conceptual factors describing the two principal phenomenological aspects of heterogeneity: the *grouping factor*, γ, and the *segregation factor*, ξ. The grouping factor is a measure of the ratio of the number of fragments to the number of groups, that is a measure of group size:

$$\gamma = \frac{N_F - N_G}{N_G - 1} \quad \text{which re-arranges into:} \quad \frac{1}{1 + \gamma} = \frac{N_G - 1}{N_F - 1}$$

Hence, the above expression of minimum distribution heterogeneity becomes:

$$(DH_L)_{min} = \frac{1}{1 + \gamma} CH_L$$

If N_F is equal to N_G (representing an ideal sampling procedure of taking only one fragment at the time—ideal but rarely realizable in practice of course), then $\gamma = 0$ and, hence, $(DH_L)_{min} = CH_L$. If, on the other hand, groups are large, ultimately taking the whole lot ($N_G = 1$), the grouping factor goes towards infinity, and the minimum DH becomes zero.

In the case where interest is no longer only on the minimum DH (which is brought about by 'correct' sampling practices), but instead on the *natural* DH, it will be necessary to include the segregation factor in the analysis. This factor represents the added heterogeneity brought about by the spatial distribution of fragments, groups or increments due to *segregation* (and/or other forms of stratification) to the overall DH:

$$(DH_L)_{nat} = \frac{1 + \gamma \xi}{1 + \gamma} CH_L \quad \text{where } 0 \leq \xi \leq 1$$

If the lot material is completely mixed (the *ideal* case of a homogenous state), ξ *would be zero*, while, if the lot is completely segregated (stratified or otherwise), ξ would be one (see Figure 1).

Ultimately the above relationships can be combined to form [3]:

$$DH_L = (1 + \xi\gamma) \cdot (N_G - 1) \frac{N_F}{N_F - 1} \sum_i \frac{(a_i - a_L)^2}{a_L^2} \cdot \frac{M_i^2}{M_L^2}$$
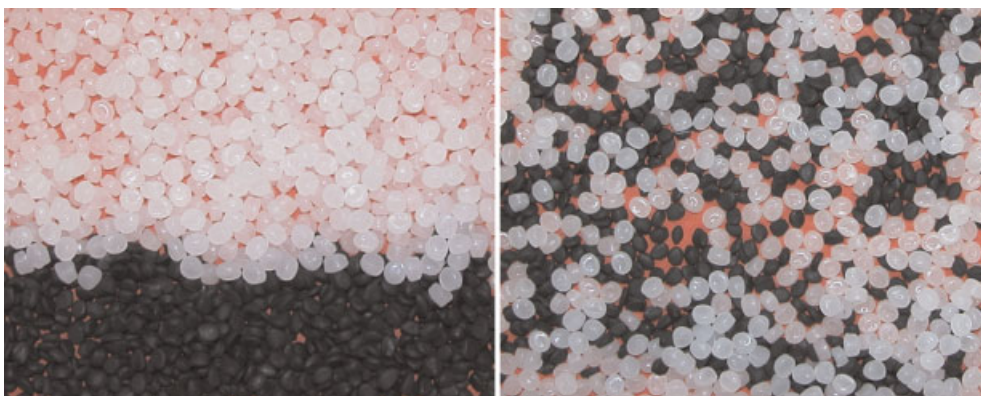
Figure 1. Compositionally identical material in two different segregation states: to the left is shown a state of very high segregation (ξ close to 1.0); to the right is shown the opposite situation characterized by a low value of ξ. This figure is available in colour online at www.interscience. wiley.com/journal/cem

In the case of a very large number of fragments, this reduces to:

$$DH_L = (1 + \xi\gamma) \cdot (N_G - 1) \sum_i \frac{(a_i - a_L)^2}{a_L^2} \cdot \frac{M_i^2}{M_L^2} \quad \text{or}$$

$$DH_L = \frac{(1 + \xi\gamma) \cdot (N_G - 1)}{N_F} CH_L$$

4.3. Interpretation of TOS

These formulas, and the analysis above, signify that basically only three factors are responsible for the magnitude of the DH:

- CH_L (constant for a given material in a given grain size distribution state)
- γ (depends on the size of the extracted increments/observation module)
- η (depends on the spatial distribution of fragments in the lot)

In order to extract samples with the lowest variation (sampling error) it is necessary to minimize DH_L. For a *given material state*, first considering the easiest case of reducing the two phenomenological factors, this can principally be achieved in only two ways:

- Decreasing the size of the extracted increments (reduces γ); that is increasing the number of increments that are combined to form a given sample mass, M_S.
- Mixing/homogenizing the lot (reduces ξ).

If these measures are insufficient for a given sampling error specification, it is necessary to reduce the CH itself, which necessitates active, physical reduction of the fragment sizes, *comminution* (grinding or crushing) or increasing the total sample mass, M_S.

All the above relate to understanding the origins of heterogeneity *within* one increment, selected as a complete cross section of the 1-D lot. Extraction of any one ('full-slice') increment is covered in full by the principles of correct, representative 0-D sampling, all aspects of which were reviewed above and treated in full detail in References [1–4].

4.4. Heterogeneity of 1-dimensional lots

In the special case of 1-D sampling, consideration of still further aspects to characterize heterogeneity is necessary. This especially involves understanding the nature of the non-random heterogeneity fluctuations *along* elongated or moving lots. Interest is now no longer in the heterogeneity *within* the units of observation (as treated above) but specifically in the heterogeneity related to the differences *between* them.

Focus will now be on a new set of N_U discrete units, U_m, making up the 1-D lot. The units are discrete (complete cross-stream) collections/materializations of material or similar; for instance: railroad cars, drums, bags or scoopfuls of material, *chronologically* (*or linearly*) *ordered* (with $m = 1, 2, 3, \ldots, N_U$).

The heterogeneity contribution, h_m, of unit U_m, is defined (strictly analogous to the definitions for fragment and group

contributions above):

$$h_m = \frac{(a_m - a_L)}{a_L} \cdot \frac{M_m}{M_{\overline{m}}} = N_U \frac{(a_m - a_L)}{a_L} \cdot \frac{M_m}{M_L}$$

where M_m is the unit mass, $M_{\overline{m}}$ the average unit mass, a_m is the grade of the unit and N_U the number of units in the lot.

The heterogeneity contribution, h_m, from a unit is composed of three (four if including the total analytical error (TAE); see below) parts when dealing with 1-D processes:

- A random, discontinuous, *short range fluctuation* term, h_{1m}. This term describes the natural randomness of the CH
- A non-random, continuous, *long range fluctuation* term, h_{2m}, that describes trends in the process/lot (between units) over time/distance
- A non-random, continuous, *cyclic term*, h_{3m}, describing cyclic or periodic behavior of the process/lot

Thus:

$$h_m = h_{1m} + h_{2m} + h_{3m}$$

It may be argued—at least in practical situations—that one needs to include the errors associated with the physical handling and assaying (analysis) of samples. This is done by a fourth part:

- A *random fluctuation* term, h_{4m}, taking into account all measurement errors stemming from weighing, sample processing and analysis. This is also named the TAE.

Thus:

$$h_m = h_{1m} + h_{2m} + h_{3m} + h_{4m}$$

Characterization of the heterogeneity of a 1-D lot must include information on the chronological order of the units. If this information is disregarded, one is restricted to employing only techniques from the previous 0-D section, calculating the DH, DH_L, as of any *static* population of units, U_m, or the constant factor of the constitution heterogeneity, IH_L can be calculated. If it is specifically sought to include the correlation information, the so-called *variogram* (technically known as a semi-variogram) is utilized. The variogram is explained further in the following sections.

5. INTRODUCTION TO VARIOGRAPHY

In order to characterize the autocorrelation between units of the process/lot or the 1-D heterogeneity of the population, the *semi-variogram* (for the remainder of this work referred to simply as the 'variogram') is derived. This describes the variation observed between units as a function of the *distance between them* (in time or space). In addition, the variogram also yields information in the forms of the 'nugget effect', the 'sill' and the 'range', which will all be explained in more detail below.

5.1. Calculation of the variogram

To calculate a variogram a sufficient number of units extracted *equidistantly*, spanning the process interval of interest is needed. An example could be a production process over a 24-h period, sampled every 20 min to characterize the variation, including three 8 h shifts. Often also much shorter

time-spans are investigated, for instance, during the filling of a number of bags from a batch (blending) process, or something much longer, like daily or season of variation, for periods up to an entire year or even more.

In order to calculate a variogram, a dimensionless and relative *lag parameter*, j, describing the distances between two extracted units is defined:

$$j = \frac{\theta}{\theta_{\min}}$$

where θ is an inter-sample interval (measured in units of: minutes, hours, meters, kilometers, product number, dependent on the given situation) and $\theta_{\min}$ is the smallest interval sampled. Figure 2 explains the sizes and calculations of all intervals of interest.

If N_U units are extracted and analyzed, N_U-1 *unit pairs* with space θ, N_U-2 unit pairs with space 2θ are available. In practice, it is not necessary to go higher than half the measurement series, $N_U/2$ from the different unit pairs are illustrated in Figure 2.

The interesting measure in the variogram, is the average of the squared *difference in heterogeneity*, $\bar{D}$, between the N_U-pairs of units for *all* lags (j):

$$\bar{D} = \frac{1}{N_U - j} \sum_m (h_{m+j} - h_m)^2$$

$\bar{D}$ is thus a function of j, $\bar{D}(j)$. $\bar{D}$ represents the heterogeneity variability *along* the extended dimension of principal interest, which is exactly the new, additional heterogeneity contribution sought. This measure must be comparable to the standard statistical variance (where all values are compared to the mean and not to another value in the population), so it is necessary to insert 2 in the denominator of the above formula and hereby get the defining variogram master equation:

$$V(j) = \frac{1}{2(N_U - j)} \sum_m (h_{m+j} - h_m)^2$$

The variogram is only calculated for values of $j = N_U/2$ (rounded down), since calculation for higher values result in some of the central experimental values not being included in the calculations—see Figure 3. An empirical rule-of-thumb states that preferentially the variogram should not be calculated for any j with less than 20–25 pairs, indicating that a variographic experiment should always consist of extracting a minimum of some 40–50 increments or units [2,3].

The variogram can equally well be calculated based on units-of-measurement such as grade, mass, concentration, total content of critical component etc. In these cases it is often needed to make the variogram dimensionless and relative, in order to make comparison easier and more meaningful (independent of measurement units). A *relative variogram* can easily be calculated by inserting a squared 'lot value' into the master equation, for instance the grade
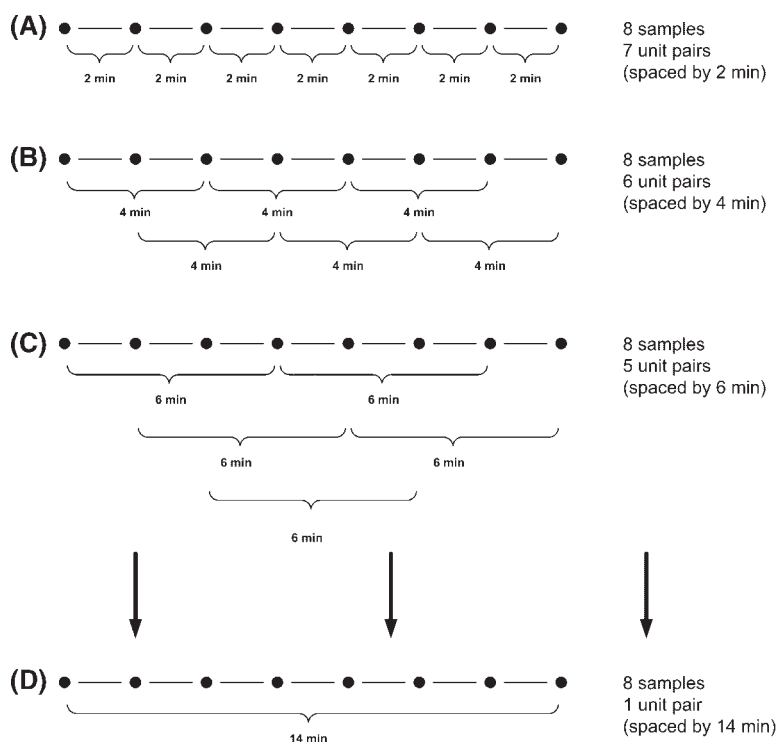
Figure 2. Example of *unit pairs* in a variographic experiment. In this small example, only eight units are extracted at 2 min intervals. Then seven unit pairs exist that are spaced by 2 min (A), six unit pairs spaced by 4 min (B), five unit pairs spaced by 6 min (C) and this continues until finally 1 unit pair spaced by 14 min (D) exist. The units shown in the figure are the *same* physically extracted samples in A, B, C and D, only the *calculation of intervals* is different.
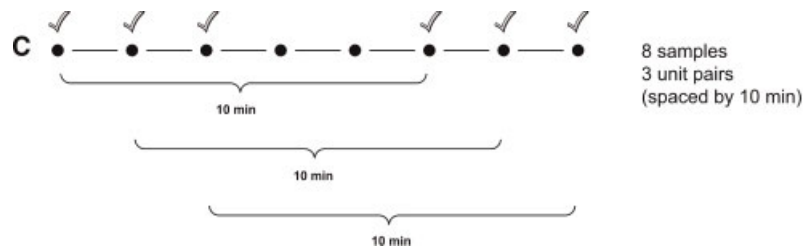
Figure 3. Lags of j higher than $N_U/2$ is not used, as this will result in some of the experimental values not being included in the calculations—In the illustrated example the two central values are not included in the calculations if j is 5 (out of a total of eight units).

of the lot, a_L:

$$V(j) = \frac{1}{2(N_U - j)a_L^2} \sum_m [a_{m+j} - a_m]^2$$

The variogram can alternatively be based on heterogeneity contributions. Calculation of the relative variogram in not necessary in this case as the heterogeneity contributions are already relative and dimensionless (the mean of these contributions $m(h_m)$ is always zero).

Whether to use an absolute or a relative variogram is a matter of the specific subsequent use hereof.

5.2. Interpretation of variograms

The shape of the variogram provides a wealth of information on the variation of the process (or elongated stationary lot) at

hand. In practice, often one of four primary types of variograms is encountered:

1. The increasing variogram
2. The flat variogram
3. The cyclic variogram
4. The decreasing variogram

The four basic types of variograms are outlined in Figure 4—but deviations from, or combinations between, these forms are often observed in practice.

The *increasing variogram* is a reflection of units which are autocorrelated when the inter-sample distance(s) is *within* a specific range, meaning that $\bar{D}$ becomes larger as the lag distance j increases, until beyond a certain distance (the 'range') where the differences level out. The *flat variogram* is
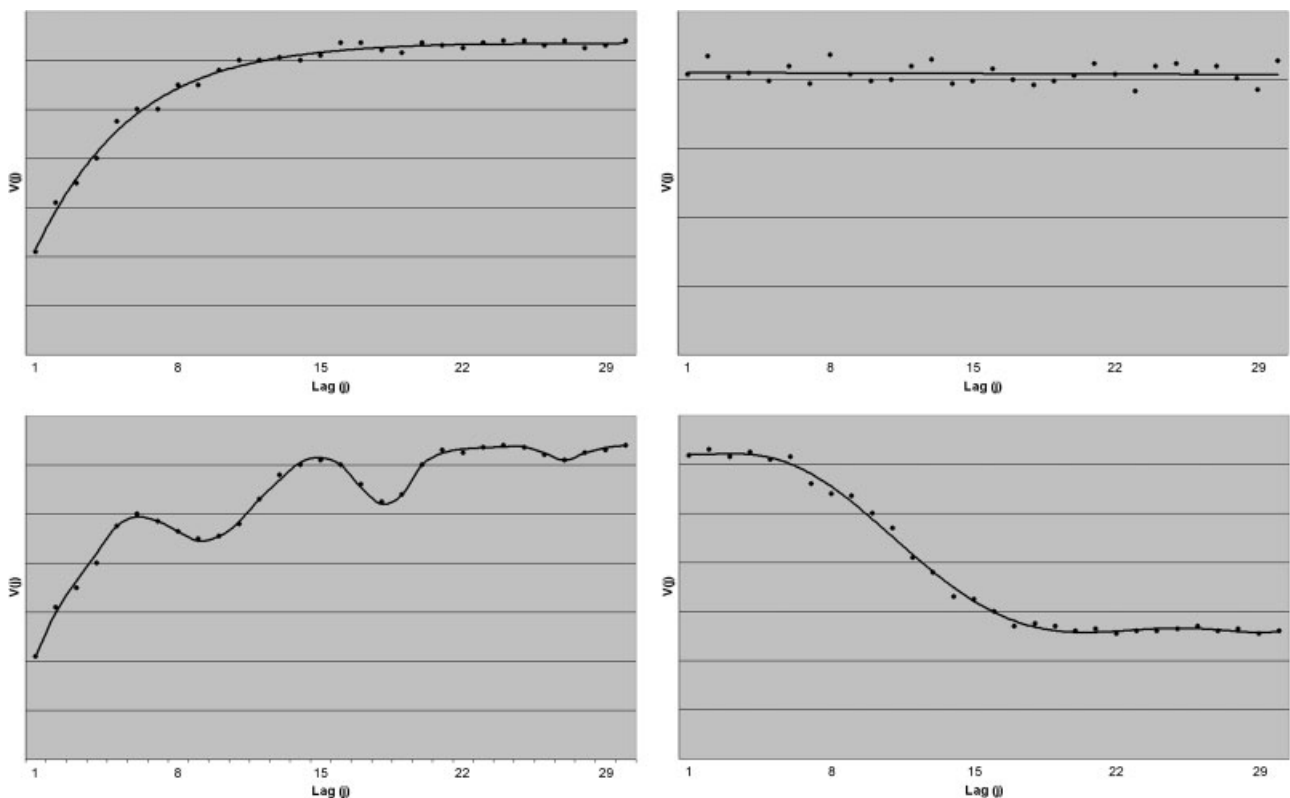
Figure 4. The four basic shapes of the variogram. Top-left: The increasing variogram. Top-right: The flat variogram. Bottom-left: The cyclic variogram and bottom-right: A decreasing variogram. All examples depict $V(j)$ as a function of the lag parameter j.

observed when no autocorrelation exists between the units even at the lowest j's (or if units are extracted at either too large or too small intervals compared to the existing autocorrelation). The *cyclic variogram* is observed when the variability of the process is influenced by a factor causing a cyclic behavior. This could be an automatic controller regulating a process between some upper and lower control limits, or a systematic change of raw-material, or similar. In the above example (Figure 4), a minimum at approximately $j = 9$, 18 and 27 is observed, and thus, deduce a cycle with a period of nine times θ. The cyclic variogram in the figure additionally shows an increasing trend. The *decreasing variogram* is extremely rare and could for example be caused by the variographic experiment being too short to capture the periodic variation of a process.

After determining the principal type of variogram, information pertaining to the process or the lot can be extracted. To illustrate this, the increasing variogram is used as an example, since this is most often met with for production processes and for very many natural processes (*geostatistics*). When the first five points are extrapolated backwards to intercept the ordinate axis, the result is a measure for the so-called 'nugget effect' (a term stemming from geostatistics), which is an indication of the minimum practical error (MPE). This in turn is an indication of the minimum sampling variance expected in practice. MPE *includes* the fundamental sampling error (FSE), the grouping and segregation error (GSE), the TAE and the incorrect sampling errors (ISE)—if present. These errors are all explained in detail in the literature [1–4,7,8].

When the variogram becomes flat (for increasing lags), the 'sill' of the variogram (a term originally from geology meaning a 'flat-lying slab') has been reached; the sill can also be calculated as the variance of all of the individual

heterogeneity contributions in the set, N_U, which characterizes the global heterogeneity of the lot *disregarding* autocorrelation. Lastly the 'range' of the variogram which is the lag distance beyond which there is no further autocorrelation observable, can be found. This can be converted directly into an interval (minutes, meters etc.) as the case may be. The above terms are shown in Figure 5 below.

From these derived variogram features the following critical new information can be incorporated into a *sampling scheme* with great benefits:

- Avoid extraction of increments/units with the same frequency (or a multiple hereof) as the period of the variogram cycle (if present), since this will provide a serious underestimation of the real process variation. It is recommended always to extract samples with a higher frequency than two per period (this is also known as the Nyqvist frequency).
- The overall sampling variation can be reduced significantly by extracting increments/units at a suitably high frequency, that is by sampling *below* the range.
- Extracting increments/units at a frequency *above* the range will only provide statistically independent samples (not a desirable feature in process sampling).
- A reliable estimate of the MPE (the total variation of samples) comes from the 'nugget effect', which is a very rich information source.

5.2.1. A note on estimating the 'nugget effect'

Oftentimes estimating the nugget effect with a high enough degree of certainty is difficult. This can be due to units extracted at a too low frequency, and it is always recommended to perform a small additional 'short-range' experimental series to obtain data for a more reliable
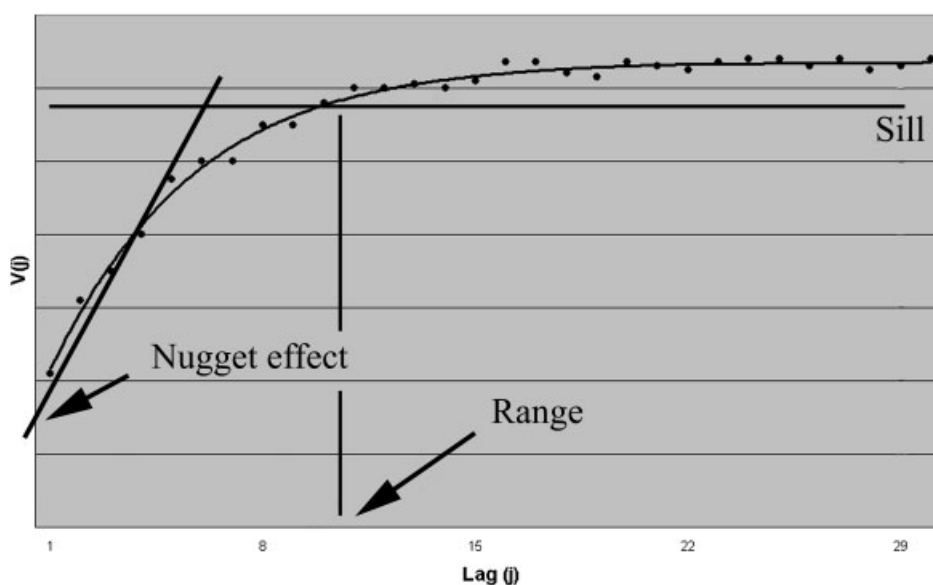
Figure 5. A generic variogram with illustrations of the 'nugget effect', or the MPE and the "range" which indicates the lag (j) above which autocorrelation is no longer discernable. The 'sill', which indicates the maximum variance difference between the units, would normally—for long measurement series—be practically coincident with the flat variogram part after the range has been reached. If calculated only for the points shown in the graph it would be positioned slightly below this flat section.

estimation of the nugget effect. This experiment is performed by extracting, for example 30–50 samples with the *highest possible* (equidistant) frequency, followed by calculating a new (short-range) variogram, used specifically only to estimate the nugget effect at a much higher degree of certainty.

5.3. Auxiliary functions and their use in variography

The variogram only provides a set of limited integer values of the lag parameter j. To remedy this, and in order to estimate the continuous selection error (CE), see further below, a set of so-called *auxiliary functions* and *error generating functions* is needed [2]. The generation of these functions can be performed in two ways: a simple point-by-point calculation based directly on the individual points of the variogram (recommended by Gy and Pitard [2,3]) or by an algebraic modeling, which demands higher computational power and programming skills; this latter approach is chosen by geostatistics.

The variogram can be broken down into four component parts—corresponding to the description of the four heterogeneity components above:

$$V(j) = V_1(j) + V_2(j) + V_3(j) + V_4(j)$$

where:

$V_1(j)$ is the *discontinuous random* part of $V(j)$ or the variogram of the random, discontinuous short range fluctuation term, h_{1m}. $V_1(j)$ describes the random fluctuations of the constitutional and distributional heterogeneities *within* each unit of the total set of units. Since these fluctuations are random, the mean would normally be approximately zero and their variance $s^2(h_{1m})$ is a constant.

$V_2(j)$ is the *non-random continuous* part of $V(j)$ or the variogram of the continuous long range fluctuation term, h_{2m}. This part of the variogram describes *trends* in the process/lot.

$V_3(j)$ is the *non-random, continuous, cyclic* part of $V(j)$ or the variogram of the cyclic term h_{3m}.

$V_4(j)$ is the *residual* part of $V(j)$ not described by any of the above. Most often this term is equal to zero or is very small and the variance of the fluctuations, $s^2(h_{4m})$, is a constant. It is the part of the total physical-chemical error associated with *estimation*—for example sample weighing, analysis etc., but excluding the specific sampling errors already delineated.

The 'nugget effect', equal to the intercept of the variogram with the ordinate axis, termed $V(0)$, actually involves two of the above four parts: $V(0) = V_1(0) + V_4(0)$.

After calculation and interpretation of the variogram a set of so-called 'auxiliary functions', that are helpful in expressing the sampling variance, can be calculated or derived. Four of these auxiliary functions exist:

- The integral $S(j)$ of the variogram $V(j)$
- The *average* integral $w(j)$ of $S(j)$
- The double integral $S'(j)$ of the variogram $V(j)$
- The *average* double integral $w'(j)$ of $S'(j)$
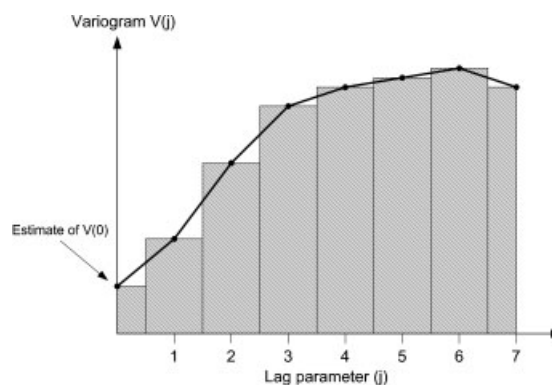
Figure 6. Point-by-point integration of a variogram (or any other function). The hatched area is approximately equal to the area between the curve and the abscissa axis, which is approximately equal to the area under an algebraically fitted, integrated mathematical model through the points. The value of $V(0)$ is estimated by backward extrapolation, since it is not experimentally known—see text.

In common for the above is their ability to smoothen the variogram, dampening the influence of random, periodic and residual terms.

5.4. Point-by-point calculation of the auxiliary functions

Before going into details, it is necessary to elaborate on the preferred method of variogram integration: the point-by-point method. This is performed as indicated in Figure 6, and completely substitutes—as recommended by Gy and Pitard [2,3]—the mathematical modeling favored in geostatistics. This is based on the fact that oftentimes the (simple) mathematical models are unable to capture and satisfactory describe, especially periodic, phenomena; they also provide greater demands to the computer and programming. It may also be argued that such smooth, continuous mathematical functions need not necessarily be corresponding to anything physical in nature, or with regard to a specific production process.

The point-by-point calculations can be done according to the following procedure [2,3]:

- Estimation of $V(0)$, either by backward extrapolation or an *added experiment* of very closely spaced increment extractions (distributed *within* the j-interval [0,1])
- Calculation of the integral of the variogram according to:
$$S(j) = S(j-1) + \tfrac{1}{2}V(j-1) + \tfrac{1}{2}V(j) \quad \text{for } j \geq 1$$
$$S(0) = 0 \qquad\qquad\qquad\qquad\qquad\quad \text{for } j = 0$$

- Calculation of the average integral according to:
$$w(j) = \tfrac{S(j)}{j} \qquad \text{for } j \geq 1$$
$$w(0) = V(0) \quad \text{for } j = 0$$

- Calculation of the double integral according to:
$$S'(j) = S'(j-1) + \tfrac{1}{2}V(j-1) + \tfrac{1}{2}V(j) \quad \text{for } j > 1$$
$$S'(0) = 0 \qquad\qquad\qquad\qquad\qquad\quad \text{for } j = 0$$

- Calculation of the average double integral according to:
$$w'(j) = \tfrac{2S'(j)}{j^2} \quad \text{for } j > 0$$
$$w'(0) = 0 \qquad \text{for } j = 0$$

The auxiliary functions are used to estimate the so-called continuous selection error (CE). CE is comprised of three error parts, CE_1 stemming from the short-range random (stochastic) variation of the process, CE_2 stemming from the long-range trend development of the process and CE_3 stemming from cyclic variations of the process. More on the estimation of CE in the section on error generating functions is described below.

5.5. The error generating functions and calculation of the continuous selection error (CE)

In a practical situation typically a sample, S, made up by a number, N_U, of increments, describing a lot, L, is needed. The objective of the sample is to characterize the lot with regard to the average of the critical component (the analyte). This is not the same scenario if the goal is, for example to map the 1-D variance or map the internal heterogeneity variance of the lot—in which case the increments are analyzed *individually*.

For typical process sampling purposes increments can be extracted according to basically three different *sampling schemes*:

- Systematic sampling, denoted 'sy', where increments are extracted *equidistant* over the runtime of the process (lot)—perhaps with a random starting point
- Stratified random sampling, denoted 'st', where the runtime of the process is divided into a number of equally sized intervals, and an increment is extracted at random *within* each of these intervals
- Totally random selection, denoted 'ra', of the increments over the runtime of the process.

The variogram and its auxiliary functions provide a lot of information on the sampling error as a function of the distance between selected increments. This information can be used at a great advantage, and this can be seen directly from the so-called error generating functions. The sampling variance—or the sampling error—is a function of the number of increments making up the sample and the sampling scheme chosen. The error generating functions, denoted W, are listed below:

$$W(j)_{sy} = 2w(j/2) - w'(j) \quad \text{and} \quad s^2(CE)_{sy} = \frac{W(j)_{sy}}{N_U}$$

$$W(j)_{st} = w'(j) \quad \text{and} \quad s^2(CE)_{st} = \frac{W(j)_{st}}{N_U}$$

$$W(j)_{ra} = s^2(h_m) = CH_L = \text{const.} \quad \text{and} \quad s^2(CE)_{ra} = \frac{W(j)_{ra}}{N_U}$$

In order to calculate the variances of the continuous integration error, a few more calculations is needed in addition to the above-described auxiliary functions (see previous section):

- The value $2w(j/2)$ is estimated according to:

  j_0 is an integer

  if j is even, then : $j = 2j_0$ and $2w(j/2) = 2w(j_0)$

  if j is odd, then : $j = 2j_0 + 1$ and $2w(j/2) = \frac{2S(j_0+1/2)}{(j_0+1/2)}$

where

$$S(j_0 + 1/2) = S(j_0) + 1/4 V(j_0) + 1/4 V(j_0 + 1/2)$$

$$V(j_0 + 1/2) \text{ is estimated by linear interpolation}$$

This completes the theoretical introduction to the specific process sampling features in TOS. While the above may perhaps at first appear slightly complicated, it is in practice very easy to use the variogram and its attendant error generating functions to estimate the sampling error incurred with a particular sampling scheme contemplated. Examples of *interpretation* of the information present in a variogram is presented below.

5.6. Recommended procedure for performing a variographic experiment

In all practical situations the following procedure is recommended in order to achieve the most reliable information on the process/lot at hand [2]:

- Extract 60–100 units (increments) at a constant interval covering the *expected* autocorrelation or cyclic behaviors of the process or lot. The rules and guidelines of correct sampling described in TOS should be followed throughout to avoid biased results. The interval chosen should be based on all available information as to the process' variability, and should include the currently favored sampling interval (even if not based on TOS). Deviating results will always lead to identification of an improved sampling rate.
- Prepare and analyze all of these units according to the existing sample preparation and analysis method. Especially when performing secondary sub-sampling (mass reduction), it is always recommend to pay close attention to the choice of method or device, as this greatly can minimize the variance of the results. Please see Reference [6] for a thorough analysis of existing techniques and hardware for representative mass reduction.
- Calculate the individual heterogeneity contributions, h_m, for all units.
- Calculate a variogram $V(j)$ of lag (j) up to 30 or 50 of the heterogeneity contributions (dependent on the number of original units).
- Extract another 30–50 units at constant intervals as quickly after one-another as possible (the short-range variogram basis), prepare/analyze and calculate another variogram of lag up to 15 or 25 (dependent on the number of units). This will result in the best possible estimate of the 'nugget effect' and hence of the MPE.
- Calculate the auxiliary functions according to the point-by-point method.
- Finally calculate the error generating functions.

A theoretical and several practical examples of calculation of variograms, auxiliary functions and error generating functions follow below.

5.7. An exemplar variographic experiment

First consider a showcase example with simulated, but realistic data:

A factory is producing a three-component powder mixture product in a continuously fed mixing process. Normally one unit (of approximately 1 kg) is extracted at random within every hour by a correctly designed cross-stream sampling device. The device does the sampling automatically and it is installed immediately before the packing machine. The
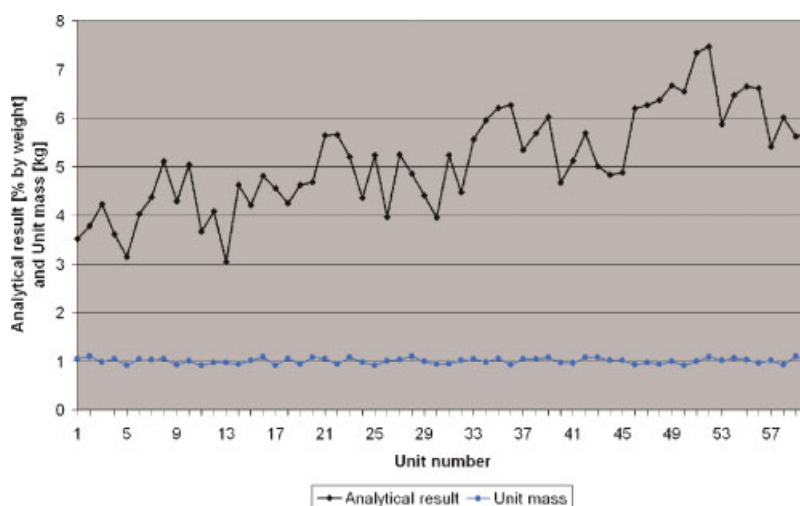
Figure 7. Analytical results, a_m, and mass, M_m, of 60 units extracted at 1 min intervals over 1 h. The mass of the individual increments (lower graph) is reasonably constant. The average content of 'A' is 5.14%; also note that an increasing trend is present. This figure is available in colour online at www.interscience.wiley.com/journal/cem

engineer in charge has for some time been wondering if the composition of the final product is constant and she wants to know if the sampling scheme used is optimal for this quality assurance (QA) purposes, so she orders a variographic experiment performed. In this experiment 60 units are extracted over 1 h (exactly every minute), and the units are prepared and analyzed individually according to the normal procedure (everything to be in complete accord with TOS principles), to arrive at an estimate for the content of component 'A', considered the most important of the three in the mixture. According to the specifications and the settings of the production machinery, component 'A' should be present in 5.14% by weight.

Figure 7 shows the analytical results of these 60 units (average 5.14%) together with the weight of the individual units:

The heterogeneity contributions of the individual units are now calculated according to the formula above and are displayed in Figure 8:

From the heterogeneity contributions the variogram is calculated and $V(0)$ is estimated to be 0.007 by backward linear extrapolating the first five points to intercept the ordinate axis. The variogram $V(j)$ and the auxiliary functions $w(j)$ and $w'(j)$ are shown in Figure 9.

From Figure 9 a local minimum at $j = 13$ can be observed and a tentative repetition at $j = 25$–27 (this can only be observed for $V(j)$, since the auxiliary functions already are smoothed out and thus did not displaying this cycle). This indicates the existence of some periodic phenomenon with a period of approximately 13 min. If the raw data (Figure 7) is inspected the cycle is very hard to spot, if at all, and this underlines one of the major features of a variographic experiment.

Now the corresponding error generating functions can be calculated to get information on which sampling strategy that will provide the best results (lowest sampling variance) in a prospective sampling situation. This can be seen in Figures 10 and 11; some of the calculated values can be seen in Table I.
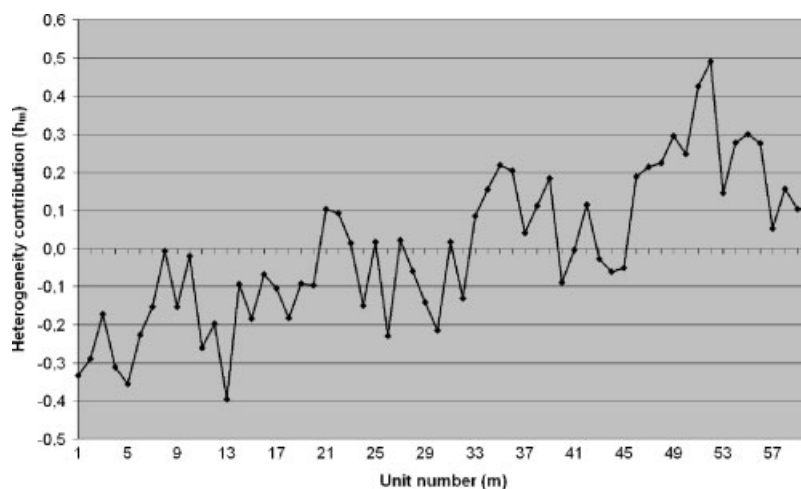
Figure 8. Individual heterogeneity contributions, h_m, of the 60 units.
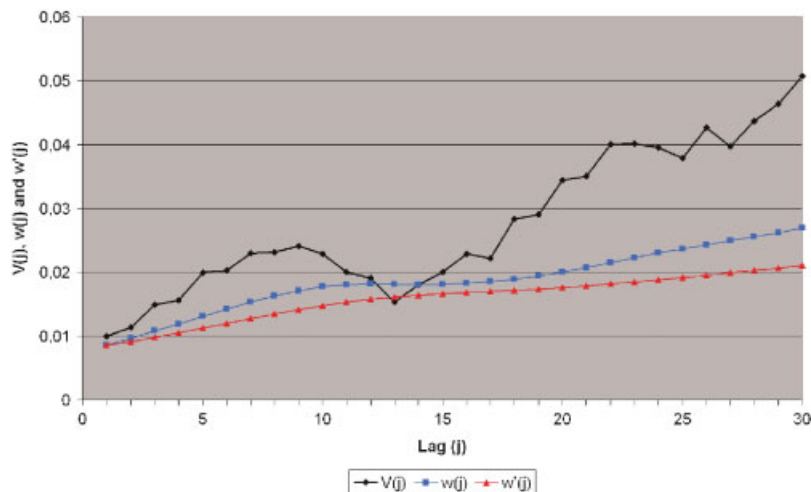
Figure 9. Plot of the variogram, $V(j)$, the average first order integral, $w(j)$ and the second order integral, $w'(j)$. The value of $V(0) = w(0) = w'(0)$ is estimated as the interception of the ordinate axis by an extrapolated line from the first five values of $V(j)$ to be 0.007. Observe how it is always more reliable to base this extrapolation on w or w'. Only values up to $j = 30$ (half the total number of units) are shown. Note that the variogram both displays a trend-like feature as well as a periodicity. This figure is available in colour online at www.interscience.wiley.com/journal/cem
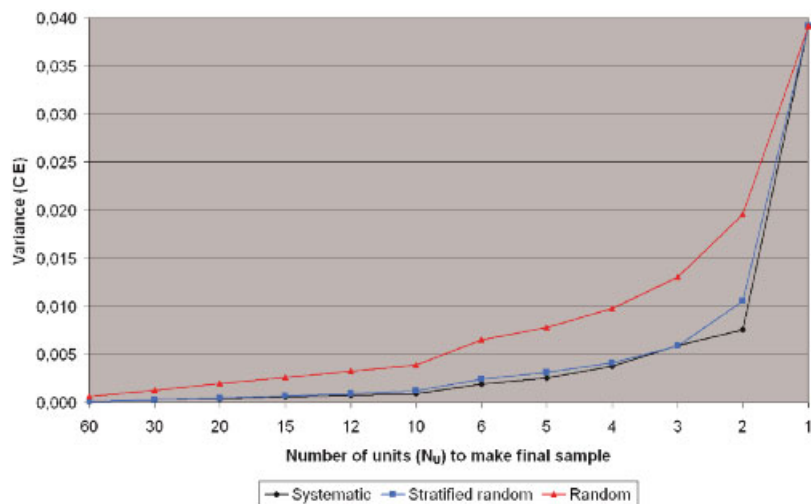
Figure 10. Plot of the error generating functions of the three sampling modes: systematic sampling (sy), stratified random sampling (st) and completely random sampling (ra), for the data series in Figure 7. The abscissa axis is reversed, here showing the number of units going into the final sample (*composite sampling*), N_U, instead of the lag parameter, j. This figure is available in colour online at www.interscience.wiley.com/journal/cem

From Figures 10 and 11 it can be observed that the higher the number of units (increments), N_U, that are combined to form the final sample (i.e. *composite sampling*) the lower the sampling variance (CE). It is also observed that systematic sampling seems to provide the lowest variation in this case and that completely random sampling provides the highest—by far. Further it is noted that increasing from one unit to using just two (Figure 10) provides a reduction in variance of approximately 50% using random selection, an approximately 70% reduction using stratified random selection, and an approximately 80% reduction using systematic sampling. If 10 units are included in the final sample instead, a 90%

reduction for random selection, a 97% reduction for stratified selection and a 98% reduction using systematic selection is the result. If need be, the sampling variation can be reduced by as much as 99.7% by using all 60 units and systematic or stratified sampling compared to the existing procedure of extracting randomly one sample per hour.

The standard deviation can be calculated instead of the variance (as the square root). The standard deviation is relative, and in order to make for easy comparison, it can be made absolute by multiplying by the average grade of 'A' (5.14% in this case). In Figure 12, all standard deviations (of the three selection modes) are calculated and multiplied by
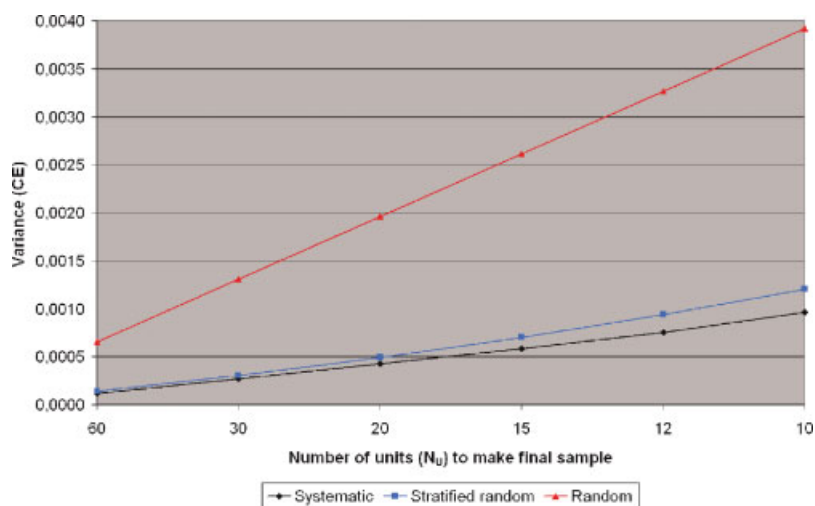
Figure 11. Close-up of the section of 60 to 10 units in the final sample seen in Figure 10 above. This figure is available in colour online at www. interscience.wiley.com/journal/cem

Table I. Calculated values by the point-by-point method

j	$V(j)$	$S(j)$	$w(j)$	$S'(j)$	$w'(j)$	$2w(j/2)$	W_{sy}	W_{st}	W_{ra}	N_U	S^2_{sy}	s^2_{st}	s^2_{ra}
0	0.007	0.000	0.007	0.000	0.009	0.016	0.007	0.009	0.039	—	0.00012	0.00014	0.00065
1	0.010	0.009	0.009	0.004	0.009	0.017	0.008	0.009	0.039	60	0.00027	0.00030	0.00131
2	0.011	0.019	0.010	0.018	0.010	0.018	0.009	0.010	0.039	30	0.00043	0.00049	0.00196
3	0.015	0.032	0.011	0.044	0.011	0.019	0.009	0.011	0.039	20	0.00058	0.00070	0.00261
4	0.016	0.048	0.012	0.084	0.011	0.020	0.009	0.011	0.039	15	0.00075	0.00094	0.00327
5	0.020	0.065	0.013	0.141	0.012	0.022	0.010	0.012	0.039	12	0.00096	0.00120	0.00392
6	0.020	0.086	0.014	0.216	0.009	0.016	0.007	0.009	0.039	10	0.00012	0.00014	0.00065
	—	—	—	—	—	—	—	—	—	—	—	—	—
30	0.051	0.809	0.027	9.477	0.021	0.036	0.015	0.021	0.039	2	0.00756	0.01053	0.01960
—	—	—	—	—	—	—	—	—	—	—	—	—	—

Results are calculated up to $j = 60$, but not all are shown here.

three and the average content to get an approximate measure of the expected sampling variation in absolute units:

$$\text{Expected variation} = 3 \times \sqrt{s^2} \times \bar{a_m}$$

From the above variogram analysis it may be concluded that it is possible to reduce the absolute sampling variation from over 3% to 0.5% simply by combining 10 units into a composite sample, if the units are extracted systematically or stratified randomly. It is of course imperative to apply all of the correct 0-D TOS principles when reducing the mass of this 10-increment composite sample to obtain the final sample [1–3] (and Section 6 below).

What to suggest to a process engineer about to design a new sampling scheme based on the above variographic analysis? Well, it is known that every cut by the automated (correct) sampling device provides approximately 1 kg of material. It is also known that handling *very* large amounts of material is troublesome and laborious, since this increases the amount of work necessary to handle the material and perform the correct representative mass reduction [6], so, it is needed to stay within reasonable masses of our final (composite) sample. The plot of the error generating functions suggests using either the systematic or the stratified random sampling approach (systematic is slightly

better). It is further observed from the variogram that the process is cyclic with a period of approximately 13 min, and one must therefore be aware of the danger of sampling systematically in the presence of cycles. Considering this, it is recommended to use stratified random sampling, as this always overcomes the danger of picking increments/units coincident with the cycle. So, the final sampling scheme recommended:

• Divide the process into 6 min intervals (10 increments per hour) and make a random cut within each of these intervals. Every hour the automated sampling device is emptied and the mass of the final sample is reduced to whatever size needed by the analytical laboratory using a correctly designed device (for instance a correctly designed riffle splitter with a large number of chutes). *If a smaller variation is critical, then use a higher number of units (perhaps 30), even though this complicates material handling.*

5.7.1. Note on didactic data set
The above data was constructed using a linearly increasing trend (from 1 to 4) over the 60 samples, adding a cyclic value of period 13 and amplitude 1.5. Finally, a random noise
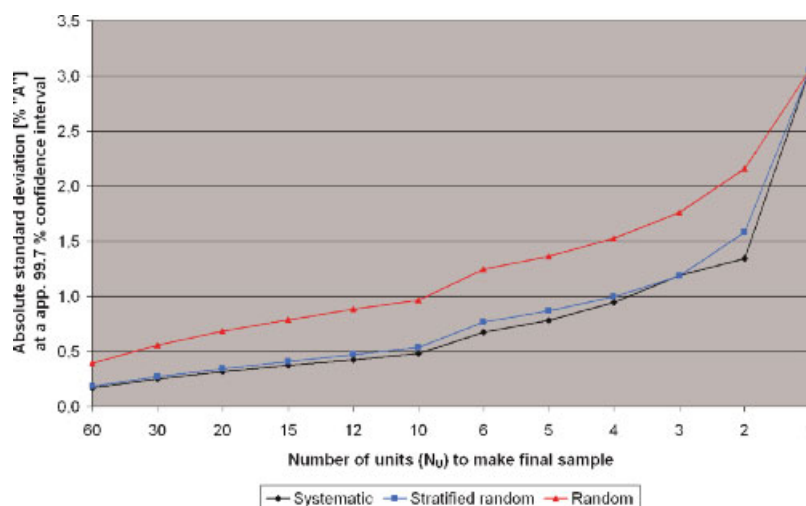
Figure 12. Absolute standard deviation of content of 'A' in samples extracted by the different selection modes as function of the number of units making up the final sample. This figure is available in colour online at www.interscience. wiley.com/journal/cem

component with value between 0 and 2 was added to make the data more realistic.

6.   SUMMARY OF ALL SAMPLING ERRORS

Basically, TOS defines seven different errors connected to sampling. These will not be explained in detail here, only summarized. The interested reader is referred to the literature on TOS for further information [1–4].

The 'zero-dimensional errors' covers errors that have to do with sampling of 0-D lots (see Section 3):

- The Fundamental Sampling Error (FSE): This is the error due to the CH of the material and is constant for a given material in a given physical state (comminution/crushing reduces FSE)
- The Grouping and Segregation Error (GSE): This error exists because of two factors, as the name implies: segregation (mixing reduces this) and grouping (composite sampling reduces this)
- The Increment Delimitation Error (IDE): When the delimitation of an increment does not cover *entirely, with parallel sides, a full cross-section* of the lot (ensuring that all parts of the cross section of the lot are equally represented in the sample), IDE results. It can (and should) be eliminated completely
- The Increment Extraction Error (IEE): If the extracted increment does not comply with that delineated, IEE crops out. A simple rule states that a fragment (physical piece of material) is to be extracted if it has its center of gravity inside the delineated increment
- The Increment Preparation Error (IPE): This error does not follow any statistical distributions—it covers all forms of alteration of the increment after extraction. Examples are: loss of fines, water evaporation, degradation of biological material etc. Only *careful* field, plant and laboratory practice can guard against IPE

The additional '1-D errors' covers two special errors that only arise when sampling 1-D lots.

- The CE comprises of three parts, of which only two are new in addition to the 0-D case [9]:

  ○ CE1: This error describes random fluctuations and is nothing more than a compounding of all the zero-dimensional errors pertaining to sampling each increment
  ○ CE2: Also known as the 'Time Fluctuation Error' (TFE). This error describes contributions due to the existence of *trends* in the lot/process
  ○ CE3: Also known as the 'Cyclic Fluctuation Error' (CFE) describes errors stemming from the presence of *periodic variation* of the lot/process.

All the above errors combine to form the total sampling error (TSE).

$$TSE = FSE + GSE + IDE + IEE + IPE + TFE + CFE$$

Note: The global estimation error (GSE) also includes the TAE, which is strictly connected to the analytical measurement and has nothing to do with the sampling process itself:

$$GSE = TSE + TAE$$

7.   EQUIPMENT OF 1-DIMENSIONAL SAMPLING

This section serves to inform on only a *very small selection* of the vast amount of (correct and incorrectly designed) process sampling equipment and devices, and to give an idea of which aspects to consider when installing a new device or altering an existing one. The reader is referred to the comprehensive, authoritative treatment in Pitard [3].

7.1.   Cross-stream sampling device

This is the automated counterpart to the so-called 'stopped belt sampling', where a perfectly delineated increment with parallel sides is extracted from a lot of material in a conveyor belt. If designed correctly, a cross-stream sampling device also extracts a full cross-section with parallel sides from the
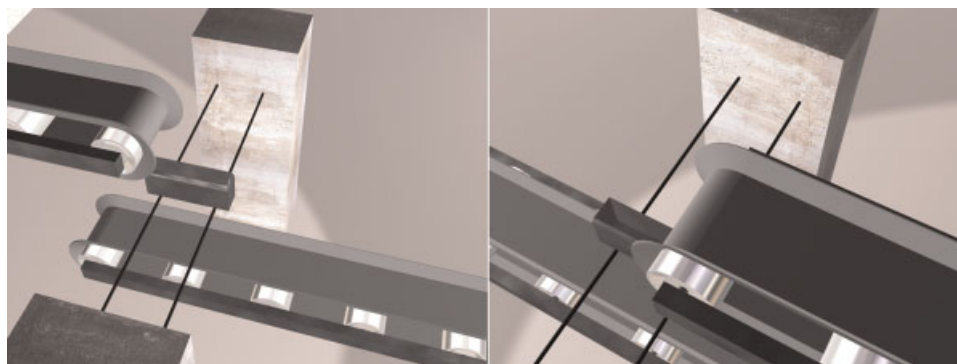
Figure 13. Cross-stream sampling device extract a full cross-section of a falling stream of material by traversing the sample box completely through the stream from side to side at a *constant speed*. In this example, the device is placed in a transfer end-point between two conveyor belts.

lot [10,11]. It is simply a rectangular box that traverses at constant speed through the entire stream of free falling material. It is normally installed at a discharge or end-point of a conveyor belt. Dependent on the material to be sampled (material cohesion, stickiness etc.), some models have a scraper-blade mounted to ensure that all material is removed from the belt. Figure 13 illustrates the principle of a cross-stream sampling device:

7.2. Flap sampling device

Another widely used device is the so-called 'flap sampling device'. Unfortunately, this device will very often yield *biased samples* according to TOS principles. A flap sampling device functions by diversion of the entire stream from the main flow into a 'sample flow' by a vertically (or much worse, horizontally) installed flap. Figure 14 illustrates the function of both an incorrectly and a correctly designed vertical flap sampling device.

The design on the left side in Figure 14 is so-called *structurally incorrect*, since only a part of the stream is
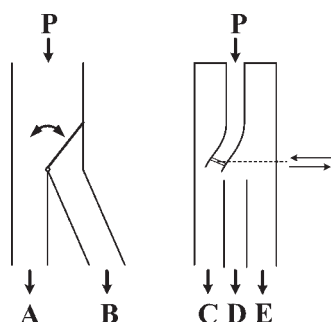
Figure 14. Schematics of two different flap-sampling device designs. The model on the left is *incorrectly* designed, since the right-hand side of the material flow (P) will be overrepresented in the sample (B) when the flap is moved from right to left diverting the material into the sample container, and back. The model on the right is *correctly* designed, as all parts of the process stream (P) are represented equally in the final sample (D) as a flexible sample hose is traversed across the sample outlet. The left-over material (A, C and E respectively) is for both models fed back to the main process stream.

diverted as the flap is beginning to move. This means that the right side of the material flow is unavoidably always overrepresented. The degree of overrepresentation is dependent on the speed of flap movement and on the time of the device in the open position. The motion of the incorrectly designed flap causes an increment delimitation error as the delimited increment does *not* have parallel sides. If any form of flow segregation or stratification exists in the flowing material (flow, rolling or gravitation segregation etc.), the flap sampling device will overestimate some size or compositional fraction of the flow.

This can however be easily remedied by altering the device to work according to the design in the right side of Figure 14 instead. Here, the entire stream is led through a flexible hose that normally rests at either side of the pipeline (C or E). When a sample is extracted, the flexible hose is moved completely across the sample outlet (D), which now cuts a correctly delimited increment with parallel sides. The sample outlet is naturally covering the whole dimension perpendicular to the paper plane in Figure 14, and it is most often designed as a rectangular box.

7.3. Valves

Valves are often used as sampling devices in pipelines. This usually results in a sampling bias unknown to most users. Imagine a liquid with suspended particles moving in a horizontal pipeline. If the flow is 'homogeneous' (i.e. turbulent and totally mixed), resulting in a perfectly identical cross-section of the flow in both vertical and horizontal directions, using valves as sampling outlets would be permitted. However, if the stream is not completely mixed (for instance by having laminar flow along the walls of the pipe, or displaying local flow eddies and vortices), a serious sampling error will always result, as heavier and lighter particles are bound to be disproportionally represented in an increment materialized through a valve (which in this context *functions* as a flap valve). The problems associated with segregation can be minimized, but never absolutely eliminated, by inserting the valve in a vertical pipeline instead (Figure 15).

Many different designs and engineering solutions exist in the industrial sector, but the authors have not yet been able to
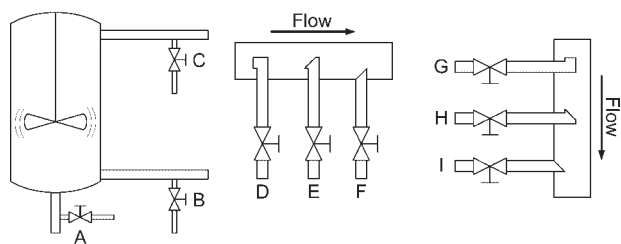
Figure 15. Different valve designs. The position of a valve (A, B, C) is dependent on how the material in the tank is behaving and how well it is mixed. One can argue that the position of the valve is of insignificant importance *if* the material is indeed well-mixed (and sampling is performed isokinetically to avoid classification of the different particle sizes), but process industry *knows* that mixing in reactor vessels is never *this* perfect. The design of the valve on the other hand does have an impact on the sampling quality. None of the designs (D, E, F) will ensure a correct sample, unless the flow is *totally* turbulent and thus completely mixed (and, again, sampling is performed isokinetically). Often, it proves best to install the valve in a vertical pipeline (G, H, I) as this minimizes significantly (but does *not* necessarily eliminate completely), the problems associated with *horizontal* segregation of particles in the flow.

find a truly correct valve sampling device for pipeline flows (one that samples a perfect 'slice' of the flow like the cross stream sampling device described above); the best (representing existing process technology) would be installed in upward-flowing pipelines.

8. PRACTICAL EXAMPLES

In this section examples of sampling 1-D lots, and the use of variography to characterize the lots and find an optimal sampling strategy, are presented.

8.1. On-line measuring of pressure, temperature and oxygen concentration in a power plant

This first data set is used to illustrate all of the information features of a variographic analysis. In a major Danish coal-fuelled power plant, several parameters in the feed water are routinely monitored on-line. A data set for a variographic analysis is extracted, in order to see if any hidden trends or periodic fluctuations are present in the process. The data are collected over a 24-h period, from 23.00 until 23.00. The extracted parameters are:

- Steam pressure in feed water tank [bar]—Logged on average every 215 s (402 measurements total)
- Oxygen content in feed water [µg/kg]—Logged on average every 32 s (2670 measurements total)
- Feed water flow [kg/s]—Logged on average every 47 s (1842 measurements total)

The data set is way too large for normal variographic analysis, so it is compressed by averaging every measurement inside all 20 min windows, yielding a suitable total of 72 units. Since all measurements are performed on-line, without loss of generality the sample weight contributions are here neglected by setting them all to unity. Normally, on-line data points like these should be coupled with a flow measurement, to ensure a single readout per defined volume that passes the sampling/measurement probe. Hence, the sample mass in the calculation of the variogram should be substituted by a flow measurement. These data were, however, not present in this data set.

Figure 16 shows a plot of the raw data.

From the plot of the raw data, a relatively smooth time-trend is observed. This is in a dominant way due to averaging over 20 min intervals. The grand means are:

- Steam pressure: 13.4 bar
- O_2 content: 74.3 µg/kg
- Flow rate: 183.7 kg/s
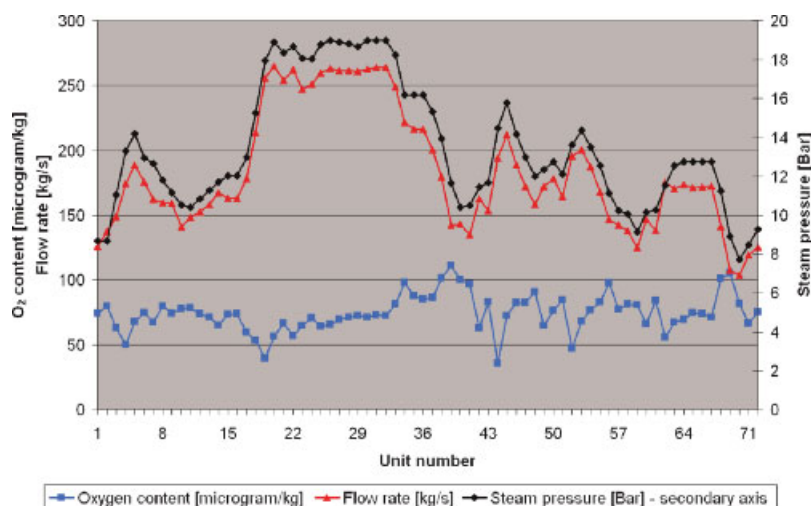
Figure 16. Data from on-line measurements of oxygen content, flow rate and steam pressure in a power plant feed water system over 24 h; all three series has been averaged to one value every 20 min. There seems to be a close relationship between steam pressure and flow rate. Steam pressure is depicted on the secondary ordinate axis (right). This figure is available in colour online at www.interscience.wiley.com/journal/cem
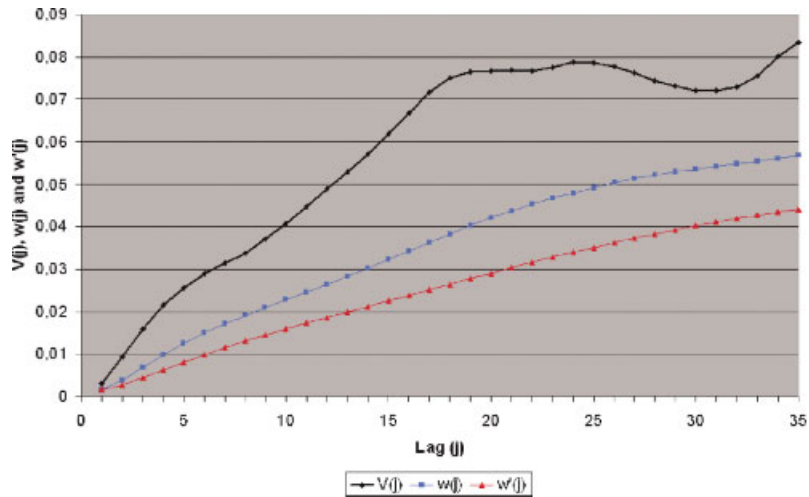
Figure 17. Variogram and auxiliary functions for steam pressure. A minimum at $j = 30$ is observed. This indicates the existence of a cyclic fluctuation with a rather long period ($j = 30 = 600$ min $= 10$ h). $V(0)$ can be estimated to be *very close* to zero in this case. This figure is available in colour online at www.interscience.wiley.com/journal/cem

Figure 17 shows the resulting variogram and auxiliary functions. A minimum at $j = 30$ is observed, which indicates that the process is following a period of 10 h (30 × 20 min = 600 min). Either this period is already known to the plant operators or not; in the latter case, the variogram would be very valuable indeed for establishing this new process insight alone.

Now the error generating functions are calculated in order to see if an optimal sampling procedure for this process can be found, *illustrating* a scenario where it would be preferable to find the *average* steam pressure and not have to continuously monitor this process on-line perhaps. The pertinent error generating functions are shown in Figure 18.

Increasing the number of units going into the final sample drastically increases the reliability of the result. Systematic sampling apparently provides a slightly lower variation as does stratified random sampling. However, both these selection methods are far superior to the complete random selection, and since the process shows periodic behavior, it is suggested to use stratified random selection of perhaps eight units making up a composite sample. This would yield an expected sampling variation of steam pressure (for eight units in a composite sample) of:
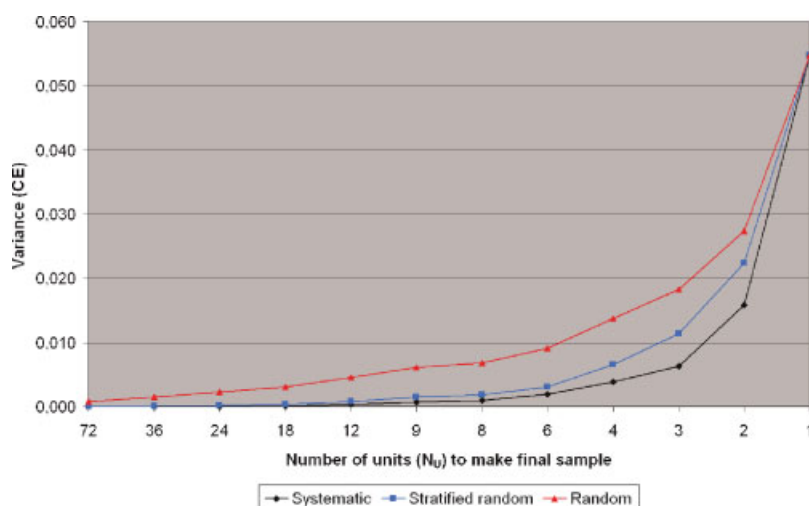
Systematic :

$$\text{Expected variation} = \sqrt{0.00099} \cdot 13.4\,\text{bar} \cdot 3 = 1.26\,\text{bar}$$

Stratified :

$$\text{Expected variation} = \sqrt{0.00182} \cdot 13.4\,\text{bar} \cdot 3 = 1.71\,\text{bar}$$

Random :

$$\text{Expected variation} = \sqrt{0.00685} \cdot 13.4\,\text{bar} \cdot 3 = 3.33\,\text{bar}$$

Figure 18. Error generating functions (variance of CE) for steam pressure plotted as a function of the number N_U of units combined to form the final sample. From this plot it is suggested to use, for instance six units picked by systematic random selection (due to the periodic process), as this reduces the error sampling variation by approximately 95%. This figure is available in colour online at www.interscience.wiley.com/journal/cem
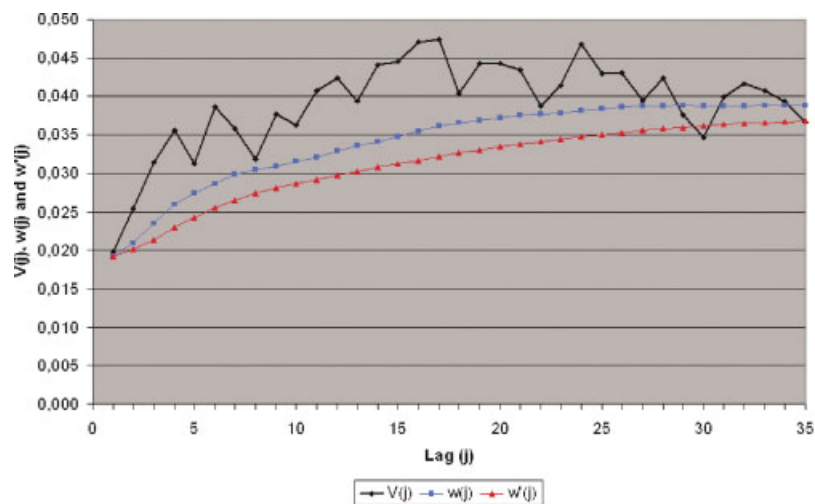
Figure 19. Variogram and auxiliary functions for the power plant feed water oxygen contents over 24 h (20 min averaged samples). The variogram does *not* suggest the existence of periodic behavior of the process. On the other hand it can be seen that the variogram increases until at app. $j = 10$, where the process seems to have passed the range, meaning that samples further apart than 200 min are not correlated to one another. $V(0)$ can best be estimated as 0.018, from back-extrapolation of the five last points of either of the auxiliary functions. This figure is available in colour online at www.interscience. wiley.com/journal/cem

Analyzing flow rate behavior yields almost exactly the same results (not shown here).

Regarding the oxygen content a very different result is obtained. Figure 19 shows the variogram and auxiliary functions for the oxygen content measurements.

The error generating functions (Figure 20) suggest basically the same as with steam pressure: select perhaps eight units to reduce sampling variation heavily. In this case it is safe to select the systematic selection scheme, as there does not appear to be any cycles in the development of the process for this on-line analyte.

This would yield an expected sampling variation of oxygen content (for eight units in a composite sample) of:

Systematic :

$$\text{Expected variation} = \sqrt{0.00322} \cdot 74.3\,\mu g/kg \cdot 3c$$

$$= 12.65\,\mu g/kg$$

Straified :

$$\text{Expected variation} = \sqrt{0.00351} \cdot 74.3\,\mu g/kg \cdot 3$$

$$= 13.21\,\mu g/kg$$

Random :

$$\text{Expected variation} = \sqrt{0.00464} \cdot 74.3\,\mu g/kg \cdot 3$$

$$= 15.18\,\mu g/kg$$

For comparison, the variogram and auxiliary functions of the full data set of 2670 measurements of oxygen content has been calculated. This can be seen in Figure 21.

A high degree of consistency between the results from the full and the reduced data set is found, as they both have

ranges of approximately 100 min. None of these variograms show any periodic behavior of the feed water oxygen content, signifying completely different time-functionalities for this variable.

8.2. On-line measurement of particle sizes by acoustic chemometrics

In a major Danish-international company, a particulate material is produced by a granulation process. In this process the particle size of the product is measured on-line by *acoustic chemometrics*: at a specific process point, particulate aggregates collide with the exterior reactor vessel walls producing compound acoustic signals picked up by suitable vibration sensors. The resulting raw signals are further processed, domain transformed and ultimately used as input data in a PLS prediction with regard to grain size, the principles behind which can be found in References [12–15]. A particular data set has been gathered over approximately 2000 min, logged every 5 min. Predicted particle sizes in this example have been *transformed* (encrypted) to a relative particle size of [−1.6 to 1.6] for confidentiality reasons. There were a few (26) missing values, which have all been replaced by the average particle size. The raw data are plotted in Figure 22:

The objective of the acoustic chemometric monitoring is straightforward: the particle size must stay within (very) narrowly defined specifications (customer demands). On-line prediction of this critical parameter is of obvious interest.

The pertinent variogram and auxiliary functions can be seen in Figure 23.

This variogram does not show any clear signs of periodicity. A *tentative* indication of a minimum around $j = 175$ *might* perhaps be argued, but this is not doubled
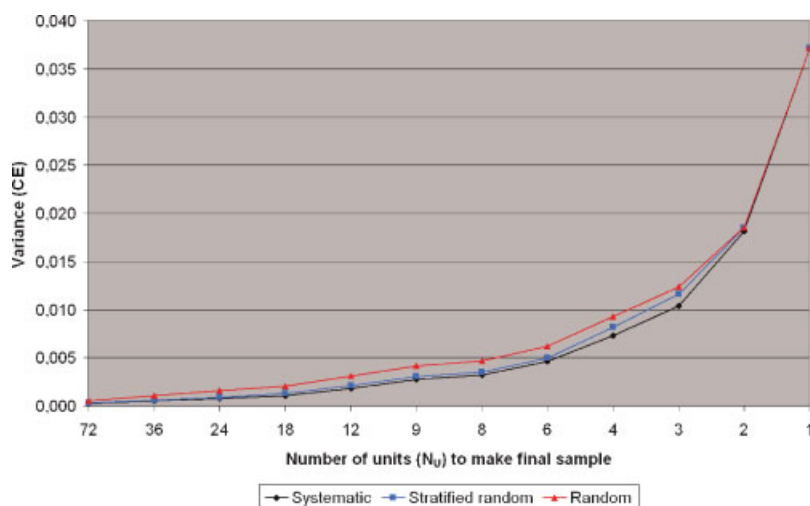
Figure 20. Error generating functions (variance of CE) for measurement of oxygen content as a function the number of units making up the final composite sample. As with steam pressure, it is suggested to combine six or eight units, this time selected systematically, however, to form the final sample. This figure is available in colour online at www.interscience.wiley.com/journal/cem
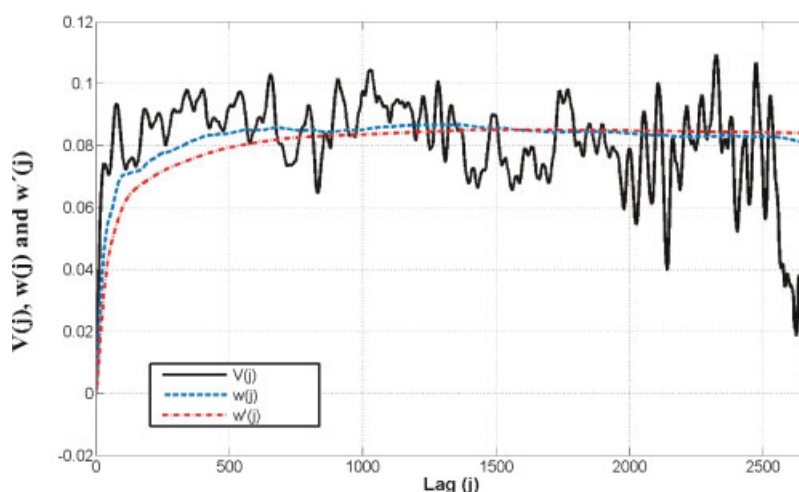
Figure 21. Variogram and auxiliary functions for the full dataset (2670 measurements) for oxygen content. The variogram reaches its sill at $j =$ app. 200, and therefore it is concluded that no autocorrelation exist between units spaced by more than approximately 107 min (200×32 s). This is consistent with the results for the reduced dataset above (where the range = approximately 100 min). This figure is available in colour online at www.interscience.wiley.com/journal/cem

around $j = 350$. Here it must be concluded that at least no obvious periodicity is present.

An optimal sampling strategy can be specified, in order to minimize the sampling variance. In this case the variation is again reduced by increasing the number of increments.

The goal of invoking on-line acoustic chemometrics is to substitute current manual sampling at regular intervals of samples for laboratory size-distribution analysis (a process which is far from delivering results in real-time). The above results of the variographic experiments and its conclusions have been used to develop an optimal reference sampling strategy for the critical particle size multivariate calibration.

8.3.   Variographic check of grain impurities during unloading

In order to find the true average content of dust and impurities in a grain silo (an unwanted component in grain lots meant for livestock consumption), a base-line variographic experiment was performed. A grain storage silo was emptied at constant flow from a hatch at the bottom where units were collected by correct, representative cross-stream cuts every minute (for a total of 75 min). Analysis for dust and impurities was performed by an automated device (model MLN from Rationel Kornservice A/S, see Reference [16]). This is a 'by-the-book' example
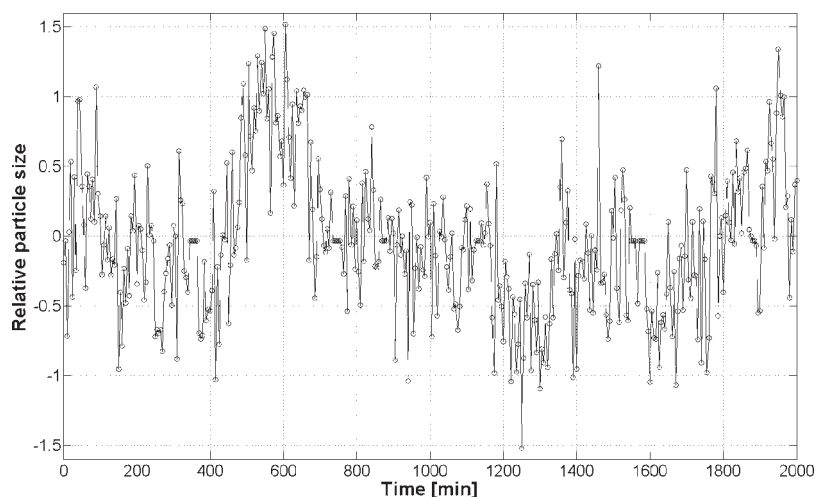
Figure 22. Particle size predictions (estimates) from acoustic chemometrics on-line measurements in a granulated product. Approximately 2000 measurements were performed every 5 min for 34 h. All measurements pertain to the same formulation production campaign.
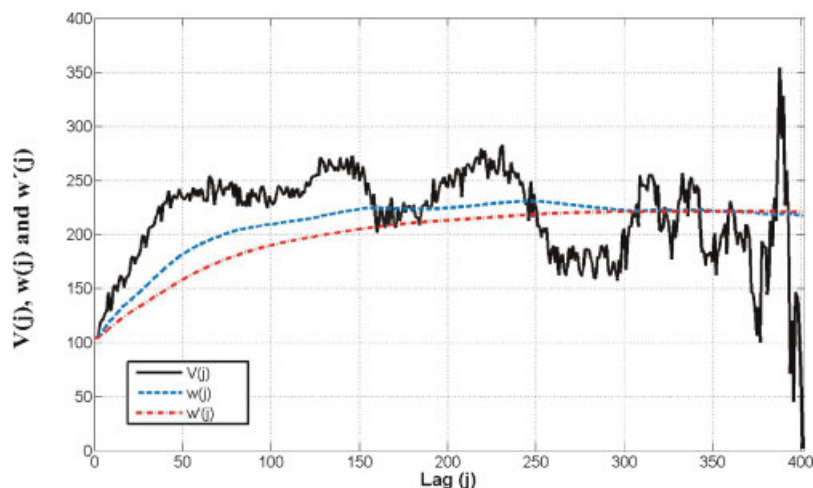
Figure 23. Variogram and auxiliary functions for on-line measurement of particle sizes. The variogram has a range of $j = 50$, signifying that units spaced by more than 250 min are no longer correlated with each other. This figure is available in colour online at www.interscience.wiley.com/journal/cem

of reducing the dimensionality of a difficult 3-D lot to a 1-D lot, opening up for much more reliable sampling [1]. There is a certain similarity of the present experiment and the flow segregation study described by Miserque and Pirard [17].

After leaving the silo, the unsampled material (stock) was stored in a large trailer, and subsequently pumped back into the silo.

An 'easier', but far less reliable and very often used alternative, would have been extracting just one (or maybe a few) samples either from the top of the silo, or from the bottom hatch—employing the proverbial 'grab-sampling'. This is an approach used by many not conversant with TOS' scientific principles, to be emphatically discouraged [1]. Below, sampling errors arising from this approach are given. Figure 24 shows the raw data.

Figure 25 shows the corresponding variogram and auxiliary functions.

At first glance, there does not seem to exist much autocorrelation between these units as the variogram quickly reaches what appears to be a relatively flat sill for $j = 4$–10. Close inspection of the variogram also reveals a subtle *trend* related to the silo discharge process, however. On the whole, this trend reflects that the variance for the highest j-lags is some 40% higher than for j below, say, 4. This means that the average *difference* between the increments of the silo discharge is increasing throughout the emptying of the silo—attesting to a significant heterogeneity in the vertical dimension, perhaps as a consequence of flow segregation (originating either from the current discharging or from previous loading(s) of the silo), or from a direct influence of gravity. Here the increasing variogram type produces insight into the internal distribution of impurities in the silo, which in the present didactic experiment, of course, is also revealed by the display of the individual sample contents (Figure 24), but in a much less obvious manner.
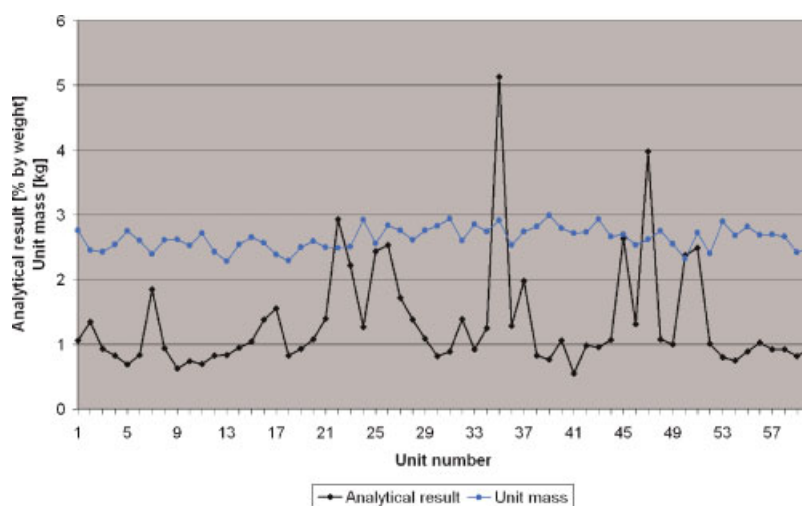
Figure 24. Raw data from the silo experiment. The average dust and impurities content is 1.30%, while deviations run as high as 5.20% (extreme low values of impurities are of less interest in the context of the objective). This figure is available in colour online at www.interscience.wiley.com/journal/cem
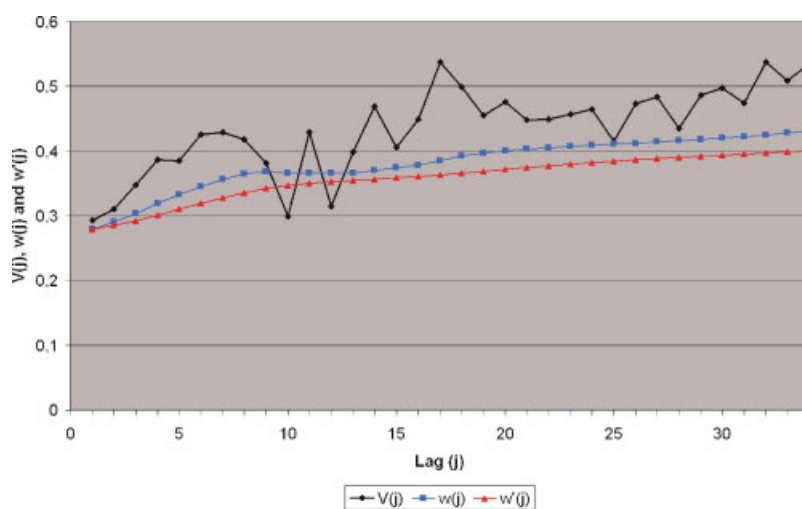
Figure 25. Variogram and auxiliary functions for measurements of dust and impurities sampled during the base-line emptying of a grain silo (3-D to 1-D lot dimensionality reduction). $V(0)$ is estimated to be approximately 0.27 by linear extrapolation of the first five points of either of the auxiliary functions or for that matter, from $V(j)$. This figure is available in colour online at www. interscience.wiley.com/journal/cem

As in the other examples, the error generating functions (Figure 26) are calculated in order to see what kind of sampling variance reduction can be expected from using composite sampling and one of the three selection modes.

Figure 26 indicates, for example eight units in a final composite sample. Further it is observed that in this case there is no significant difference between the three selection modes. This is due to the fact that there is no significant autocorrelation between the units, in which context all selection modes will perform equally well. This is a nice substantiation of the general TOS findings on this issue [2,3].

The absolute standard deviations (99.7% confidence interval) are likewise calculated for a sample made from just one unit (a 'grab sample') and a composite sample made from eight units, in order to see the effect of reducing the lot

dimensionality and of using composite sampling:

1 unit (grab sampling) :

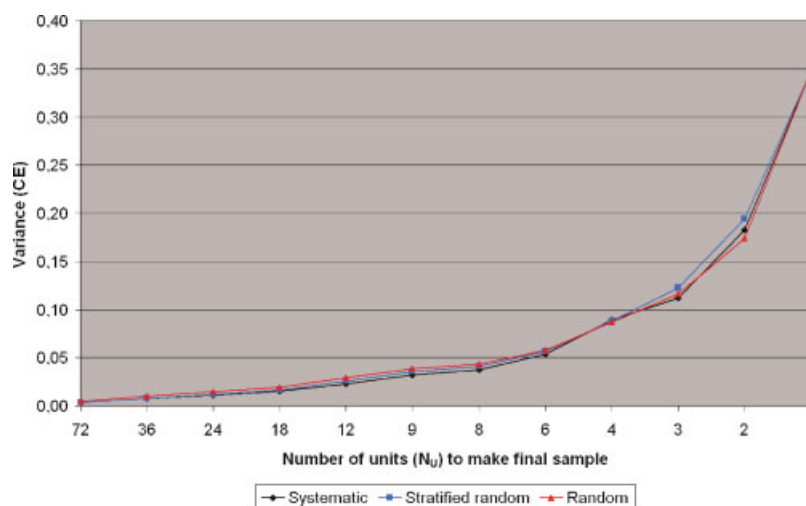Expected variation $= \sqrt{0.34725} \cdot 1.30\% \cdot 3 = 2.30\%$

8 units (random):

Expected variation $= \sqrt{0.04341} \cdot 1.30\% \cdot 3 = 0.81\%$

8 units (systematic)

Expected variation $= \sqrt{0.03739} \cdot 1.30\% \cdot 3 = 0.75\%$

Had a grab sample (one sample only) been utilized, a sampling error of 2.30% would have been incurred, whereas a threefold reduction can be achieved (0.75%–0.81%) by basing the analysis for impurities/dust on a composite

Figure 26. Error generating functions as a function of the number of units making up a final composite sample by three different selection modes. This figure is available in colour online at www.interscience.wiley.com/journal/cem

sample made up of eight increments. Agricultural accounting demands highly accurate and precise determination of impurities/dust, and will gladly accept the extra work included in the above composite sampling approach in order to reduce TSE significantly (to below 1%).

This example, like the others above, underlines that large reductions of sampling errors from using composite sampling, combined with one of the three sampling modes can always be quantified, and clearly illustrates the significant benefits from performing a variographic experiment. It is always valuable to know to which degree the units of the lot are correlated with their 'neighbors' for optimal planning of a sampling scheme.

9. DISCUSSION AND CONCLUSIONS

This tutorial has outlined the most important aspects regarding sampling of 1-D lots or processes. By several didactic examples the benefit of a thorough *variographic analysis* of any lot, prior to implementing a new sampling routine, has been shown. From variographic experiments specific information on the following issues can be obtained:

- Lot/process variation over time or space
- *Autocorrelation* in the lot/process (MPE, range, sill)
- Delineation of *trends* in the lot/process
- Delineation of short- or long-term *periodic behavior* of the lot/process
- Sampling variance as a function of the sample selection mode
- Sampling variance as a function of the number of units in composite sampling

The following can be concluded:

- Generally, for heterogeneous materials, systematic sample selection for composite sampling provides the lowest sample variances. Contrastingly:
- Completely random sample selection provides the highest errors and may for this reason as well be abandoned forthwith—Use one of the other two instead

- If the lot/process shows any form of periodic behavior, it is recommended to use stratified random sample selection—to make sure samples are not extracted with the same frequency as the cycle (or a multiple thereof)

The specific examples presented here are only that—*examples*: the general 1-D sampling principles presented are completely general however. All 1-D sampling procedures must be based on a correct variographic experiment if to have credibility and reliability.

This tutorial has shown that TOS provides the necessary and sufficient theoretical as well as practical, framework for fully representative process sampling. The procedures outlined are easy to understand and while not in any way being prohibitive with respect to the physical and the analytical efforts needed for total Q.A. control over the often dominating sampling errors.

9.1. TOS process sampling software
It is relatively easy to construct a spreadsheet (MS Excel, etc.) that will produce all the above variographic plots and calculations following the outlined point-by-point calculation method, at least when the number of units is reasonably low (say less than 100 or so). When dealing with much larger data sets, an alternative may also be needed. Codes/spreadsheets have been developed for both EXCEL (for small data sets) and MATLAB (for large data sets). In addition a (beta-version) of a stand-alone program (VARIO) that handles (almost) any data set size has been developed by ACABS. This software is available at the homepages: www.acabs.dk (freeware).

equipment, and grain for the experimental series on grain impurities. *Hans Henrik Friis-Pedersen* is gratefully acknowledged for his critical participation in the silo experiment and especially for developing the VARIO program. *Peter Paasch-Mortensen*, NovoZymes, is greatly acknowledged for putting the encrypted acoustic chemometric data set at our disposition for the purpose of this tutorial. Lastly, but not leastly, the authors thank *Rationel Kornservice A/S, Esbjerg, Denmark*, for providing the necessary sample cleaner equipment and professionally educating the authors in grain sample preparation and cleaning.

REFERENCES

1. Petersen L, Minkkinen P, Esbensen KH. Representative sampling for reliable data analysis: theory of sampling. *Chemom. Intell. Lab. Syst*. 2005; **77**: 261–277.
2. Gy P. *Sampling for Analytical Purposes*. Wiley: Chichester, UK, 1998.
3. Pitard FF. *Pierre Gy's Sampling Theory and Sampling Practice* (2nd edn). CRC Press LLC: Boca Raton, USA, 1993.
4. Smith PL. *A Primer for Sampling Solids, Liquids and Gases—Based on the Seven Sampling Errors of Pierre Gy*. ASA SIAM: USA, 2001.
5. Gy PM. The analytical and economic importance of correctness in sampling. *Anal. Chim. Acta* 1986; **190**: 13–23.
6. Petersen L, Dahl CK, Esbensen KH. Representative mass reduction: a critical survey of techniques and hardware. *Chemom. Intell. Lab. Syst*. 2004; **74**: 95–114.
7. Minkkinen P. Evaluation of the fundamental sampling error in the sampling of particulate solids. *Anal. Chim. Acta* 1987; **196**: 237–245.
8. Heikka R, Minkkinen P. Comparison of some methods to estimate the limiting value of the variogram, $v_h(j)$, for the sampling interval $j = 0$ in sampling error estimation. *Anal. Chim. Acta* 1997; **346**: 277–283.
9. Gy P. Sampling of discrete materials-III: quantitative approach—Sampling of one dimensional objects. *Chemom. Intell. Lab. Syst* 2004; **74**: 39–47.
10. Gy P, Marin L. Unbiased sampling from a falling stream of particulate material. *Int. J. Miner. Process* 1978; **5**: 297–315.
11. Gy P. Does your mechanical sampler do what it's supposed to? *Coal mining process* 1981; 71–74.
12. Esbensen KH. *Multivariate Data Analysis—In Practice* (5th edn). CAMO ASA: Norway, 2002.
13. Halstensen M. Experimental multivariate sensor technology and developments of system prototypes for industrial multi-phase characterisation: selected forays, Doctoral thesis. Telemark University College (HIT): Porsgrunn, Norway, 2001. ISBN: 82-471-5381-5.
14. Halstensen M, Esbensen KH. New developments in acoustic chemometric prediction of particle size distribution—''the problem is the solution''. *J. Chemometrics* 2000; **14**: 463–481.
15. Esbensen KH, Halstensen M, Lied TT, Saudland A, Svalestuen SDS, Hope B. Acoustic chemometrics—from noise to information. *Chemom. Intell. Lab. Syst*. 1998; **44**: 61–76.
16. Rationel Kornservice A/S website. http://www.rako.dk/ [28 March 2006].
17. Miserque O, Pirard E. Segregation of the bulk blend fertilizers. *Chemom. Intell. Lab. Syst*. 2004; **74**: 215–224.