

Gy's Discrete Heterogeneity and the Grouping and Segregation Variance

Geoffrey J Lyman

Materials Sampling & Consulting. E-mail: geoff.lyman@gmail.com

This paper returns to Gy's work to make a recapitulation of his derivation of the constitutional and distributional heterogeneity of a particulate material with a careful delineation of the assumptions that he employed to arrive at expressions for the fundamental sampling variance and the variance due to distributional heterogeneity or the grouping and segregation variance. Gy derives a link between the constitutional and distributional heterogeneity based on the assumption that increments are 'similar'. This in fact requires that potential increments contain the same number of particles which is very restrictive. Gy's expressions are explored with numerical examples which demonstrate that these may be valid only under limited circumstances. The second part of the paper provides a derivation that presents an alternative approach to the variance of sampling of a highly segregated particulate material. It is shown that when all particles have the same mass as can be expected under Gy's assumptions, the new expression coincides with Gy's. The new derivation is essentially free of assumptions regarding particle numbers and masses in increments and appeals to common sense regarding the extraction of samples by mechanically correct samplers. The new approach provides an expression for the variance due to grouping and segregation which involves the properties of the particulate material and the same variable used by Gy to express the extent of segregation of the lot. The grouping factor γ used by Gy is eliminated from consideration.

Editor's ingress: Regrettably, this paper could not be presented either orally or on-line at WCSB10. It has, therefore, not been subjected to the open scientific discussion which is customary in academe, but only to a closed proxy through reviews and rebuttals via the Editor's desk. This paper deserves more, however, not least because of the author's career-long contributions to the science and technology of sampling. The Editor has, therefore, decided to call for a public discussion as part of the next regular issue of TOS Forum. Readers of this ingress may consider themselves invited; more information and directions will be forthcoming on the homepage of the International Pierre Gy Sampling Association (<https://intsamp.org>), or e-mail the Editor with your comments (khe.consult@gmail.com).

Introduction

The understanding of sampling uncertainty due to particle grouping and segregation is a long-standing issue in particulate sampling theory. The derivation by Gy is complex indeed and not made easier due to the volume of notation that must be used to arrive at a result. The issue starts with Gy's analysis of distributional heterogeneity and then follows on to his discussion of the selection variances; no easy study to digest. Few seem to have done so.

To understand Gy's definitions of distributional heterogeneity, it is necessary to go back to his development in Chapter 19 of his book of 1982 or equivalent expositions. It can be shown that there are assumptions made which limit the magnitude of the distributional heterogeneity and so subsequent variances such as the grouping and segregation variance. These assumptions are also necessary to arrive at Gy's link between constitutional and distributional heterogeneity.

After a comprehensive review of Gy's derivation of the fundamental and grouping and segregation variances, an alternative approach to deriving a variance due to particle segregation is presented herein based on the concept of correct extraction of constant volume increments. The new result is the same as Gy's result when assumptions regarding increment masses and particle numbers as suggested by Gy are invoked. This new approach provides considerable insight into the challenge of describing sampling variance due to segregation.

Gy's Notation and Assumptions

While this is repetitive of Gy's section 19.1 [1], it is necessary to define and appreciate the basis of Gy's arguments and derivations. Italics have been used to identify terms that have a specific meaning. Summation indices have been changed in line with usual mathematical usage and additional subscripts added to ensure that a variable is uniquely defined. There is little alternative to such a complex notation.

The *lot* is defined as a set of *units* $U_i; i = 1, \dots, N_U$ such as

- the set L_F of *fragments* $F_i; i = 1, \dots, N_F$ where N_F is the number of *fragments* in the *lot*
- the set L_G of *groups* $G_j; j = 1, \dots, N_G$ of *fragments* in the *lot* which are complementary (no overlap or intersection)

The *groups* may be thought of as the *potential increments* that can be taken from the *lot*.

It is **assumed** that all groups have about the same *magnitude*, the definition of the *magnitude* is not made clear until later.

L	A lot of particulate material having a mass M_L , containing a mass of <i>critical analyte</i> (CA) A_L with a_L the <i>critical content</i> (concentration, CC)
F_i	A <i>fragment</i> belonging to L having a mass M_{F_i} , a mass of CA A_{F_i} with a CC a_{F_i} ; $i = 1, \dots, N_F$
U_i	A non-specified <i>unit</i> belonging to L containing a mass M_{U_i} , a mass of CA A_{U_i} with CC a_{U_i} ; $i = 1, \dots, N_U$
G_j	A <i>group of fragments</i> belonging to L each containing a mass M_{G_j} , a mass of CA A_{G_j} with a CC a_{G_j} ; $j = 1, \dots, N_G$
$F_{G_j,i}$	A <i>fragment</i> belonging to G_j with mass $M_{F_{G_j,i}}$ with a mass of CA $A_{F_{G_j,i}}$ and with a CC $a_{F_{G_j,i}}$; $i = 1, \dots, N_{G_j}$

Material balances dictate that the following relationships hold

$$M_L = \sum_{i=1}^{N_G} M_{G_i} = \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} M_{F_{G_j,i}} = \sum_{i=1}^{N_F} M_{F_i} \quad (1)$$

$$A_L = \sum_{j=1}^{N_G} A_{G_j} = \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} A_{F_{G_j,i}} = \sum_{i=1}^{N_F} A_{F_i} \quad (2)$$

$$a_L = \frac{A_L}{M_L} = \frac{\sum_{j=1}^{N_G} A_{G_j}}{\sum_{j=1}^{N_G} M_{G_j}} = \frac{\sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} A_{F_{G_j,i}}}{\sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} M_{F_{G_j,i}}} = \frac{\sum_{i=1}^{N_F} A_{F_i}}{\sum_{i=1}^{N_F} M_{F_i}} \quad (3)$$

The assumption that the groups are of 'similar magnitude' is a strong one. Gy¹ makes reference to the material of Chapter 16 which is concerned with increment delimitation and is clearly oriented to sampling as in 'stopped belt' sampling or sampling from a one-dimensional moving stream. In such a case, unless the mass per unit length of the lot varies substantially, the increments are of similar *volume*. When a tool is used to collect an increment in a correct manner, the increments will tend to be of similar volume as well. Sampling with an auger or taking a drill core creates increments of similar volume.

If magnitude relates to similar number of particles in each group, then all groups must have similar size distributions. Similar masses for the groups suggests a similar mean density. Similar composition is not a restriction to make as this defeats the analysis where differences in composition are being sought.

Gy further defines average properties of the groups.

$\bar{F}$	The average fragment belonging to L having a mass $M_{\bar{F}}$, a mass of CA $A_{\bar{F}}$ with a CC $a_{\bar{F}}$. Note that the average is a number average. The average mass of a fragment must be $M_{\bar{F}} = \frac{M_L}{N_F}$ The average CA must be $A_{\bar{F}} = \frac{A_L}{N_F}$ The average CC must be $a_{\bar{F}} = \frac{A_{\bar{F}}}{M_{\bar{F}}} = \frac{A_L}{M_L}$
$\bar{G}$	The average group belonging to L each containing a mass $M_{\bar{G}}$, a mass of CA $A_{\bar{G}}$ with a CC $a_{\bar{G}}$ The average mass of a group is $M_{\bar{G}} = \frac{M_L}{N_G}$ The average CA must be $A_{\bar{G}} = \frac{A_L}{N_G}$ The average CC must be $a_{\bar{G}} = \frac{A_{\bar{G}}}{M_{\bar{G}}} = \frac{A_L}{M_L} = a_L$
$\bar{F}_{G_j}$	The average fragment belonging to G_j with mass $M_{\bar{F}_{G_j}}$ with a mass of CA $A_{\bar{F}_{G_j}}$ with a CC $a_{\bar{F}_{G_j}}$; $j = 1, \dots, N_G$ The average mass of a fragment in the j group is $M_{\bar{F}_{G_j}} = \frac{M_{G_j}}{N_{G_j}}$ The average CA of a fragment in the j^{th} group is $A_{\bar{F}_{G_j}} = \frac{A_{G_j}}{N_{G_j}}$ The average CC of a fragment in the j^{th} group is $a_{\bar{F}_{G_j}} = \frac{A_{\bar{F}_{G_j}}}{M_{\bar{F}_{G_j}}} = \frac{A_{G_j}}{M_{G_j}} = a_{G_j}$

So we have a situation where the groups can vary in composition, but their average composition must be that of the lot. We must think of the groups as having a statistical distribution of composition about the average composition of the lot.

Next, the definition of heterogeneity needs to be considered. Gy defines the heterogeneity carried by a particle as the difference between the analyte content of the particle and the lot weighted by the particle mass and then divided by the analyte mass in the average particle of the lot, so

$$h_i = \frac{(a_{F_i} - a_L)M_{F_i}}{A_F} = \frac{(a_{F_i} - a_L)M_{F_i}}{a_F M_F} = \frac{(a_{F_i} - a_L)M_{F_i}}{a_L \frac{M_L}{N_F}} = \frac{N_F (a_{F_i} - a_L)M_{F_i}}{a_L M_L} \quad (4)$$

He defines the heterogeneity carried by the i^{th} fragment in the j^{th} group in an analogous manner

$$h_{FG_{ji}} = \frac{(a_{FG_{ji}} - a_L)M_{FG_{ji}}}{a_F M_F} = \frac{N_F (a_{FG_{ji}} - a_L)M_{FG_{ji}}}{a_L M_L} \quad (5)$$

The next question is the heterogeneity carried by the average fragment within the j^{th} group. This is defined as

$$h_{\bar{F}G_j} = \frac{(a_{\bar{F}G_j} - a_L)M_{\bar{F}G_j}}{a_F M_F} = \frac{N_F}{a_L M_L} (a_{\bar{F}G_j} - a_L)M_{\bar{F}G_j} = \frac{N_F}{a_L M_L} (a_{G_j} - a_L) \frac{M_{G_j}}{N_{G_j}} \quad (6)$$

Gy now considers the average of the heterogeneity carried by the i^{th} fragment in the j^{th} group and this is

$$\frac{1}{N_{G_j}} \sum_{i=1}^{N_{G_j}} h_{FG_{ji}} = \frac{N_F}{a_L M_L} \frac{1}{N_{G_j}} \sum_{i=1}^{N_{G_j}} (a_{FG_{ji}} - a_L)M_{FG_{ji}} = \frac{N_F}{a_L M_L} \frac{1}{N_{G_j}} \left[\sum_{i=1}^{N_{G_j}} a_{FG_{ji}} M_{FG_{ji}} - a_L \sum_{i=1}^{N_{G_j}} M_{FG_{ji}} \right] \quad (7)$$

But the first sum on the right is the mass of analyte in the j^{th} group which is $a_{G_j} M_{G_j}$ and the second sum is the mass of the j^{th} group M_{G_j} so

$$\frac{1}{N_{G_j}} \sum_{i=1}^{N_{G_j}} h_{FG_{ji}} = \frac{1}{N_{G_j}} \frac{N_F}{a_L M_L} (a_{G_j} - a_L) M_{G_j} = h_{\bar{F}G_j} \quad (8)$$

which means that heterogeneity carried by the average particle of the j^{th} group is the numerical average of the heterogeneities carried by the fragments in the j^{th} group. So the heterogeneities can be averaged numerically as they relate to masses of analyte and masses of particles unlike particle compositions.

The heterogeneity carried by a group of particles is defined as

$$h_{G_j} = \frac{(a_{G_j} - a_L)M_{G_j}}{a_G M_G} = \frac{N_G}{a_L M_L} (a_{G_j} - a_L) M_{G_j} \quad (9)$$

Gy then makes a comparison between the heterogeneity carried by the j^{th} group, h_{G_j} and that carried by the average particle of the group, $h_{\bar{F}G_j}$. From (6) and (9) we have

$$h_{\bar{F}G_j} = \frac{N_F}{a_L M_L} (a_{G_j} - a_L) \frac{M_{G_j}}{N_{G_j}} \quad (10)$$

and

$$h_{G_j} = \frac{N_G}{a_L M_L} (a_{G_j} - a_L) M_{G_j} \quad (11)$$

so if

$$\frac{N_F}{N_{G_j}} = N_G \quad (12)$$

these heterogeneities are equal. This requires that all groups have the same number of particles

$$N_{G_j} = \frac{N_F}{N_G} \quad (13)$$

The constitutional heterogeneity of the lot, CH_L is now defined as the variance of the heterogeneity carried by the fragments of the lot. Since the expected value of the heterogeneity is zero,

$$\begin{aligned} CH_L &= \sigma^2(h_i) = \frac{1}{N_F} \sum_{i=1}^{N_F} h_i^2 \\ &= \frac{1}{N_F} \sum_{i=1}^{N_F} \left[\frac{N_F (a_{F_i} - a_L) M_{F_i}}{a_L M_L} \right]^2 \\ &= \frac{N_F}{a_L^2 M_L^2} \sum_{i=1}^{N_F} (a_{F_i} - a_L)^2 M_{F_i}^2 \end{aligned} \quad (14)$$

$$\begin{aligned}
CH_L &= \sigma^2(h_i) = \frac{1}{N_F} \sum_{i=1}^{N_F} h_i^2 \\
&= \frac{1}{N_F} \sum_{i=1}^{N_F} \left[\frac{N_F (a_{F_i} - a_L) M_{F_i}}{a_L M_L} \right]^2 \\
&= \frac{N_F}{a_L^2 M_L^2} \sum_{i=1}^{N_F} (a_{F_i} - a_L)^2 M_{F_i}^2
\end{aligned}$$

This is a particle mass squared weighted second moment of the particle compositions about the mean composition of the lot. Noting that $(a_{F_i} - a_L) M_{F_i}$ is the excess/deficit mass of analyte from the mean composition, the summation is total of this excess/deficit squared and the sum vanishes only if all particles have the same analyte content. It is easy to visualise a probability density function for $(a_{F_i} - a_L) M_{F_i}$ which straddles the origin. The sum above divided by the number of fragments is the variance of this density function which is discrete for a finite number of particles. Note that CH_L is a very large number in general as it is proportional to the number of particles in the lot; it is an extensive (depends on lot mass) rather than intensive (independent of lot mass) property of the lot.

We note here that Pitard² defines a practical measure of the heterogeneity of the lot by multiplying CH_L by the average mass of a fragment. The quantity was also recognised by Gy as the 'invariant of heterogeneity', hence the notation 'IH'. Pitard writes

$$IH_L = \frac{M_L}{N_F} CH_L = \frac{1}{M_L} \sum_{i=1}^{N_F} \left(\frac{a_{F_i} - a_L}{a_L} \right)^2 M_{F_i}^2 \quad (15)$$

which has units of mass but is an intrinsic property of the lot. The i^{th} fragment mass can be characterised by a volume v_i and density ρ_i , $M_{F_i} = v_i \rho_i$ and M_{F_i}/M_L is the mass fraction x_i of the particle in the lot, so

$$IH_L = \sum_{i=1}^{N_F} x_i v_i \rho_i \left(\frac{a_{F_i} - a_L}{a_L} \right)^2 \quad (16)$$

The distributional heterogeneity DH_L is defined as the variance of the heterogeneities h_{G_j} carried by the groups G_j in the lot.

$$\begin{aligned}
DH_L &= \sigma^2(h_{G_j}) = \frac{1}{N_G} \sum_{j=1}^{N_G} h_{G_j}^2 = \frac{1}{N_G} \sum_{j=1}^{N_G} \left[\frac{N_G}{a_L M_L} (a_{G_j} - a_L) M_{G_j} \right]^2 \\
&= \frac{N_G}{a_L^2 M_L^2} \sum_{j=1}^{N_G} (a_{G_j} - a_L)^2 M_{G_j}^2
\end{aligned} \quad (17)$$

Gy makes a comparison of DH_L and CH_L under the assumption that all groups are 'similar' and contain essentially the same number of fragments. However, if the groups are not so constrained, then we could choose the groups in any reasonable way. One such choice of groups is to put every particle into its own group. Then

$$DH_L = \sigma^2(h_{G_j}) = \frac{N_F}{a_L^2 M_L^2} \sum_{j=1}^{N_F} (a_{F_j} - a_L)^2 M_{F_j}^2 \quad (18)$$

and this is just CH_L . But the only similarity between the groups is that they contain one particle and with a size distribution, it is not really true that the groups are similar, certainly not in regard to their mass. Note also that the sampling has been carried out by number. The value of DH_L like CH_L is a large number due to the proportionality to the number of fragments in the lot.

To conclude the analysis of DH according to Gy, we follow his analysis of section 19.3.4. It is possible to write that

$$h_i = \frac{N_F (a_{F_i} - a_L) M_{F_i}}{a_L M_L} \equiv h_{FG,i} = \frac{N_F (a_{FG,i} - a_L) M_{FG,i}}{a_L M_L} \quad (19)$$

as each term corresponds to a analyte content of a fragment and a mass of a fragment and both are defined over the full set of fragments in the lot.

One may then introduce an expression or identity

$$h_i \equiv h_{FG,i} = (h_{FG,i} - h_{G_j}) + h_{G_j} \quad (20)$$

If we then square both sides and sum, we have

$$\begin{aligned}
\sum_{i=1}^{N_F} h_i^2 &= \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} \left[(h_{FG,i} - h_{G_j}) + h_{G_j} \right]^2 \\
&= \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} \left\{ (h_{FG,i} - h_{G_j})^2 + 2h_{G_j} (h_{FG,i} - h_{G_j}) + h_{G_j}^2 \right\}
\end{aligned} \quad (21)$$

Let us examine the middle term. We write

$$2 \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} h_{G_j} \left(\frac{N_F (a_{FG_{ji}} - a_L) M_{FG_{ji}}}{a_L M_L} - \frac{N_G}{a_L M_L} (a_{G_j} - a_L) M_{G_j} \right) =$$

$$2 \sum_{j=1}^{N_G} h_{G_j} \left(\sum_{i=1}^{N_{G_j}} \frac{N_F (a_{FG_{ji}} - a_L) M_{FG_{ji}}}{a_L M_L} - \sum_{i=1}^{N_{G_j}} \frac{N_G}{a_L M_L} (a_{G_j} - a_L) M_{G_j} \right) \quad (22)$$

Then the term in brackets is

$$\sum_{i=1}^{N_{G_j}} \frac{N_F (a_{FG_{ji}} - a_L) M_{FG_{ji}}}{a_L M_L} - \sum_{i=1}^{N_{G_j}} \frac{N_G}{a_L M_L} (a_{G_j} - a_L) M_{G_j}$$

$$= \frac{N_G}{a_L M_L} \left[\frac{N_F}{N_G} \sum_{i=1}^{N_{G_j}} (A_{FG_{ji}} - a_L M_{FG_{ji}}) - \sum_{i=1}^{N_{G_j}} (A_{G_j} - a_L M_{G_j}) \right] \quad (23)$$

$$= \frac{N_G}{a_L M_L} \left[(A_{G_j} - a_L M_{G_j}) \left(\frac{N_F}{N_G} - N_{G_j} \right) \right]$$

and if the number of particles in all groups is $N_{G_j} = \frac{N_F}{N_G}$ for all j , the term vanishes but not otherwise. If it does not vanish

we have

$$\sum_{i=1}^{N_F} h_i^2 = \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} \left\{ (h_{FG_{ji}} - h_{G_j})^2 + 2h_{G_j} (h_{FG_{ji}} - h_{G_j}) + h_{G_j}^2 \right\}$$

$$= \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} (h_{FG_{ji}} - h_{G_j})^2 + 2 \frac{N_G}{a_L M_L} \sum_{j=1}^{N_G} h_{G_j} \left[(A_{G_j} - a_L M_{G_j}) \left(\frac{N_F}{N_G} - N_{G_j} \right) \right] + \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} h_{G_j}^2 \quad (24)$$

$$= \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} (h_{FG_{ji}} - h_{G_j})^2 + 2 \left(\frac{N_G}{a_L M_L} \right)^2 \sum_{j=1}^{N_G} (A_{G_j} - a_L M_{G_j})^2 \left(\frac{N_F}{N_G} - N_{G_j} \right) + \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} h_{G_j}^2$$

The middle term can have negative and positive terms depending on the number of particles in the j^{th} group being higher or lower than a mean value of N_F/N_G . If the groups all contain the same number of particles, we have

$$\sum_{i=1}^{N_F} h_i^2 = \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} \left\{ (h_{FG_{ji}} - h_{G_j})^2 + h_{G_j}^2 \right\} = \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} (h_{FG_{ji}} - h_{G_j})^2 + N_{G_j} \sum_{j=1}^{N_G} h_{G_j}^2 \quad (25)$$

and dividing by N_F

$$CH_L = \frac{1}{N_F} \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} (h_{FG_{ji}} - h_{G_j})^2 + \frac{N_{G_j}}{N_F} \sum_{j=1}^{N_G} h_{G_j}^2 \quad (26)$$

and again, if all groups are the same size or (12) holds,

$$CH_L = \frac{1}{N_F} \sum_{j=1}^{N_G} \sum_{i=1}^{N_{G_j}} (h_{FG_{ji}} - h_{G_j})^2 + DH_L \quad (27)$$

and regardless of the first term on the right, which is always positive, shows that $DH_L \leq CH_L$. Gy states that the first term is the average constitutional heterogeneity of the groups.

So the magnitude of the DH is capped by the constitutional heterogeneity when all groups have the same number of particles which, as stated above, precludes the groups having significantly different size distributions and being of similar mass or volume as they would be if the groups correspond to correctly extracted increments from a moving stream or a static zero- or one-dimensional lot. The same cap may also exist depending on the behaviour of the terms that don't vanish when groups don't have the same number of particles.

Segregation by size is perhaps the first and most obvious type of particle segregation that is encountered.

Following Gy further, he defines the grouping factor, γ as

$$\gamma = \frac{N_F - N_G}{N_G - 1} \approx \frac{N_F}{N_G} - 1 ; N_G \gg 1 \quad (28)$$

and this is of the order of the number of particles in a group which can be envisioned as the number of particles in a potential increment, the groups corresponding to potential increments within the lot. This number is potentially HUGE. Gy suggests that the minimal DH is given by

$$DH_{L \min} = \frac{CH_L}{1 + \gamma} \quad (29)$$

which means that it is potentially zero. The maximal DH is of course the CH . He states that this maximum is achieved when the composition of all groups is *homogeneous*. This can be achieved, as previously deduced, when each group contains only one particle or when all particles in a group carry the same critical analyte concentration. He suggests that the maximal distributional heterogeneity can be achieved when, simultaneously

- the material under scrutiny is perfectly liberated (all particles consist of a single mineral phase)
- the various minerals are perfectly segregated
- the particles are grouped according to their mineral composition

and the groups carry the same number of particles, so that the relationship developed between DH and CH is valid.

Let us go back to the original definition of the distributional heterogeneity

$$DH_L = \sigma^2(h_{G_i}) = \frac{1}{N_G} \sum_{j=1}^{N_G} h_{G_j}^2 = \frac{N_G}{a_L^2 M_L^2} \sum_{j=1}^{N_G} (a_{G_j} - a_L)^2 M_{G_j}^2 \quad (30)$$

and that for the constitutional heterogeneity

$$CH_L = \sigma^2(h_i) = \frac{1}{N_F} \sum_{i=1}^{N_F} h_i^2 = \frac{N_F}{a_L^2 M_L^2} \sum_{i=1}^{N_F} (a_{F_i} - a_L)^2 M_{F_i}^2 \quad (31)$$

and do some arithmetic with some simple cases.

A good case is a binary mixture of two minerals with different densities and different particle sizes. The two minerals are segregated and grouped. The DH is then calculated as

$$DH = 2 \times \left[\left(\left(\frac{a_{G_1} - a_L}{a_L} \right) \times \frac{M_{G_1}}{M_L} \right)^2 + \left(\left(\frac{a_{G_2} - a_L}{a_L} \right) \times \frac{M_{G_2}}{M_L} \right)^2 \right] \quad (32)$$

$$CH = (N_{G_1} + N_{G_2}) \left[N_{G_1} \left(\left(\frac{a_{F_1} - a_L}{a_L} \right) \times \left(\frac{M_{F_1}}{M_L} \right) \right)^2 + N_{G_2} \left(\left(\frac{a_{F_2} - a_L}{a_L} \right) \times \left(\frac{M_{F_2}}{M_L} \right) \right)^2 \right] \quad (33)$$

Let's take the case in Table 1 and carry out a calculation. Further, we can vary the number of particles of the Phase 1 and see what happens to the values of the CH and DH . For the conditions in the table the calculation follows.

Table 1. Heterogeneity Calculation Example (particles are all taken to be cubes)

Phase 1	Value	Unit
density	5	g/cm ³
size	1	cm
analyte conc	0.3	
number of particles	50	
particle mass	5	g
group mass	250	g
Phase 2		
density	2.6	g/cm ³
size	2	cm
analyte conc	0.075	
number of particles	20	
particle mass	20.8	g
group mass	416	g
Lot		
total particles	70	
analyte conc	0.159	
total mass	666	g
DH		
from Phase 1	0.1095	
from Phase 2	0.1095	
total	0.4378	
CH		
from Phase 1	0.002189	
from Phase 2	0.005473	
total	0.5363	
ratio DH/CH	0.816	
gamma	68	

$$DH = 2 \left[\left(\left(\frac{0.3 - 0.159}{0.159} \right) \times \frac{250}{666} \right)^2 + \left(\left(\frac{0.075 - 0.159}{0.159} \right) \times \frac{416}{666} \right)^2 \right] \quad (34)$$

$$= 2[0.1095 + 0.1095] = 0.4378$$

$$CH = (50 + 20) \times \left[50 \left(\left(\frac{0.3 - 0.159}{0.159} \right) \times \left(\frac{5}{666} \right) \right)^2 + 20 \left(\left(\frac{0.075 - 0.159}{0.159} \right) \times \left(\frac{20.8}{666} \right) \right)^2 \right] \quad (35)$$

$$= 70 \times [0.002189 + 0.005473] = 0.5363$$

The calculations carried out for a range of particle numbers in Phase 1 are plotted in Figure 1. When there are 20 particles

in each phase, the DH is indeed equal to the CH and because we have segregated according to phase and the groups are homogeneous, we have maximal DH in that the DH equals the CH .

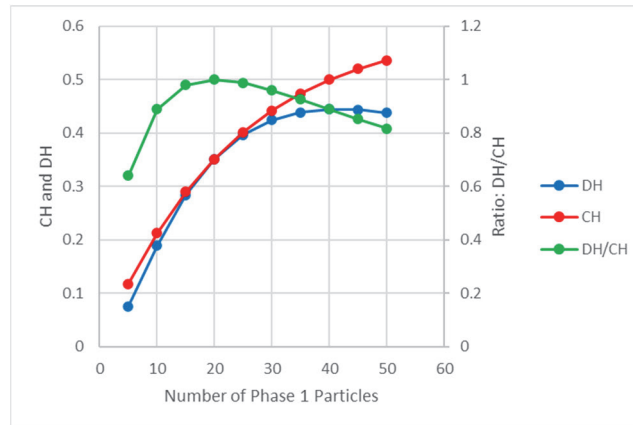


Figure 1. Variation of DH_L and CH_L for the mixture considered in Table 1, with additional variation of the number of particles in Phase 1.

A case of some interest to gold miners and processors is one in which we have relatively fine gold mixed with barren solids of a larger size. Consider 50 micron gold mixed with 150 micron barren material taken to have a density 2.6 g cm^{-3} . The particles will be taken to have a shape factor of 0.5. For 250 gold particles in a nominal 30 g sample, the grade is 9.67 g/t and there are 7×10^6 barren particles. The DH and CH are calculated as follows.

$$DH = 2 \left[\left(\left(\frac{1 - 9.67 \times 10^{-6}}{9.67 \times 10^{-6}} \right) \times \frac{0.000297}{30.71} \right)^2 + \left(\left(\frac{0 - 9.67 \times 10^{-6}}{9.67 \times 10^{-6}} \right) \times \frac{30.71}{30.71} \right)^2 \right] \quad (36)$$

$$= 2[0.99998 + 0.99998] = 3.9999$$

$$CH = (7.0 \times 10^6 + 250) \times \left[\frac{250 \left(\left(\frac{1 - 9.67 \times 10^{-6}}{9.67 \times 10^{-6}} \right) \times \left(\frac{1.188 \times 10^{-6}}{30.71} \right) \right)^2}{7.0 \times 10^6 \left(\left(\frac{0.0 - 9.67 \times 10^{-6}}{9.67 \times 10^{-6}} \right) \times \left(\frac{4.39 \times 10^{-6}}{30.71} \right) \right)^2} + \right] \quad (37)$$

$$= 7.0 \times 10^6 \times [0.0039999 + 1.43 \times 10^{-7}] = 28001.5$$

We see in this case that the DH is very small compared to the CH and that the magnitude of CH is governed largely by the gold phase. Note that here also we have two groups of liberated material but the DH is very small compared to the CH . There is clearly more to this issue of constitutional and distributional heterogeneity than is apparent at first glance.

To complete Gy's linking of the DH and CH , we consider his definition of a segregation factor, ξ , as

$$DH_L = \frac{1 + \gamma \xi}{1 + \gamma} CH_L = (1 + \gamma \xi) \frac{N_G - 1}{N_F - 1} CH_L \quad (38)$$

He states that $0 \leq \xi \leq 1$ with the maximal value applying when the material is perfectly segregated. In the case of the gold example, using γ as given by (28), we have $\gamma = N_F - 2$ as we have two groups

$$DH_L = (1 + \xi(N_F - 2)) \frac{1}{N_F - 1} CH_L \quad (39)$$

$$= \left(\frac{1}{N_F - 1} + \xi \left(\frac{N_F - 2}{N_F - 1} \right) \right) CH_L$$

and since N_F is very large,

$$DH_L \approx \xi CH_L \quad (40)$$

and the two groups are liberated and separate, so we would expect $\xi = 1$, but DH_L is very small.

It must be emphasised at this point that Gy's analysis corresponds to what is referred to in the statistical literature as a *non-parametric* analysis. He has made no assumptions in regard to the statistical distribution of particle grades or sizes, except insofar as he has considered the choice of groups that are homogeneous. It may be that to progress further, it will be necessary to bring in some statistical distribution concepts.

Chapter 20 of Gy¹ is concerned with the development of the discrete selection model. This chapter is rather heavy-going but what can be taken away from the development is that the final results are based on particles and units being selected with a uniform probability. Gy states at 20.5.2 that for 'correct' sampling, the expected mass of a sample is P times the mass of the lot, where P is the selection probability. This uniformity of mass makes the samples very similar.

The development leads to the conclusion that the selection of particles leads to the sampling variance being related to the CH of the material and selection of groups leads to the sampling variance being related to the DH . But it must be recalled that the CH to DH link is predicated by groups containing the same numbers of particles. This is a substantial assumption

but seems to be valid under the definition of correct sampling if the material is not substantially segregated. The final result of the development is the definition of the variance due to DH being related to the variance due to CH by the relationship

$$\sigma_{GE}^2 = \gamma \xi \sigma_{FE}^2 \quad (41)$$

where we recall that the value of the grouping factor γ is of order of the number of particles in a group, allowing the variance due to grouping to be much larger than the variance due to the fundamental error. However we note from Pitard³ that the experience of Gy was that the product $\gamma \xi$ tended to be about unity, so that the fundamental variance plus the grouping and segregation variance is double the fundamental variance.

An Alternative Approach to Quantifying Distributional Heterogeneity

Gy makes considerable appeal to the mechanically correct sampling of particulate materials from a flowing stream or stationary lot, be it zero- or one-dimensional. In this he considers the use of a sampling tool that can be used to define potential increments from the lot. In the case of a flowing stream, we have the cross-stream sampler which effectively cuts a swath of material from the conveyor belt as the solids fall off the end of the belt. One can also consider the fundamental concept of stopped belt sampling off a conveyor or 'ribbon' sampling from a lot that has been spread out in a linear manner on a surface with a more or less constant mass per unit length.

While we usually consider the mass flow along a belt, it is also relevant to consider the volumetric flow along a conveyor belt. If the flow is coming from a feeder that has a device shaping the flow on the belt such as a gate at the outlet of a feeder from a bin, the flow will tend to be a constant volume flow and the sampling tool will cut a specific volume from the flow as an increment. This is certainly true in stopped belt sampling.

Gy's sampling theory has some significant constraints under it, as illustrated by the assumptions needed to connect CH and DH . It is possible to remove the constraints if we will admit to a parametric development of sampling theory. The author^{4,5} has made such a development that is in agreement with Gy's result for the fundamental sampling variance (error). This alternative development is based on the concept that when a well-mixed heterogeneous material is sampled in a zero- or one-dimensional manner, the number of particles of any one type that arrive in an increment follows a Poisson distribution with an expected value derived from the properties of the lot. With the concept of the material being laid out on a conveyor belt and flowing at a constant mass flow rate, this assumption corresponds to the particles of any one type being placed on the belt in a totally random manner without spatial constraints. Their time of arrival at the end of the belt is a Poisson process. This situation clearly corresponds to Gy's selection of particles by number. We find that the fundamental sampling variance based on this conceptualisation agrees exactly with Gy's results. The new approach allows sampling theory to be taken further because we can use the properties of the Poisson distribution to calculate the full sampling distribution for a material if the size and composition distributions of the particles are known or can be reasonably assigned.

So perhaps it is legitimate to explore a variation of this conceptualisation of correct sampling that led to agreement with Gy's work. In the same way as Gy's did, we will consider the potential increments taken from the flow as 'units' having particular properties.

So what happens to our sampling scenario when the increments are defined volumetrically? First of all, their expected mass will no longer have to be constant as at more or less constant void fraction, the increment mass will reflect the mean density of the particles. It will also be possible to have increments that have different particle size distributions.

To explore this issue further, let us imagine sampling a set of groups of particles which have been made up of single mineral species. We will take the mass fraction of the i^{th} particle type in the lot to be z_i with a density ρ_i and some size distribution $g_i(d)$. Consider that the groups are laid out in a row which has a constant width and each particle class occupies a length fraction q_i in the direction perpendicular to the width of the layout. One might imagine that these are arranged on a conveyor belt so that the depth of each class is the same.

Next, we have to consider what the bulk density of the sections on the belt might be. If the particle size distributions were all the same, the void fraction on each class would tend to be the same, but the bulk density would be proportional to the density of the particles in the class. With a void fraction ε_i the bulk density for the i^{th} class would be

$$\bar{\rho}_i = (1 - \varepsilon_i) \rho_i \quad (42)$$

For unit width and depth of the classes, the fraction of the length occupied by the i^{th} class will be, for a total number of classes M_T

$$q_i = \frac{\frac{z_i}{\bar{\rho}_i}}{\sum_{j=1}^{M_T} \frac{z_j}{\bar{\rho}_j}} \quad (43)$$

We can imagine that multiple strips of the classes are laid out in a random order so that there are many strips of the classes along the belt.

Now, if this belt were moving and we were to carry out a 'sampling experiment' taking a finite number, N , of *small* increments at random points in time, the number of times a given class was collected would follow a multinomial distribution. We will simplify the situation a bit by taking the probability of interception of increments consisting of adjacent particle classes to be zero. This could be taken into account, but this simplifying assumption makes the results easier to see. With small increments, the probability of intercepting more than one particle type is reduced.

The sampling tool will be taken to collect equal volumes of particles into each increment; this is analogous to 'stopped belt' sampling of a flow on a conveyor with a constant bed depth. The increment volume can be taken to be a volume δ . The mass of this increment is then $\delta \bar{\rho}_i$ so it carries a mass of target analyte of $a_i \bar{\rho}_i \delta$ where a_i is the mass fraction of the target analyte in the i^{th} class.

The number of times that the i^{th} class is 'hit', R_i , is a multinomial random variable. Denoting the number of hits on the i^{th} class in one realisation of the sampling experiment as r_i , the sample composition for the experiment is,

$$a_s = \frac{\sum_{i=1}^{M_r} a_i \bar{\rho}_i r_i \delta}{\sum_{j=1}^{M_r} \bar{\rho}_j r_j \delta} \quad (44)$$

The expected value of R_i is Nq_i when the number of increments taken is N . The expected value of the sample assay is then

$$\begin{aligned} E\{A_s\} &= \frac{\sum_{i=1}^{M_r} a_i \bar{\rho}_i q_i \delta}{\sum_{i=1}^{M_r} \bar{\rho}_i q_i \delta} \\ &= \frac{\sum_{i=1}^{M_r} a_i z_i}{\sum_{i=1}^{M_r} z_i} \\ &= a_L \end{aligned} \quad (45)$$

so the sampling is unbiased to a first approximation. Note that when reference is made to a random variable, an upper case letter is used to denote the random variable and the corresponding lower case letter is used denote a realisation of that random variable.

To find the variance of the sample analyte content, the usual rule of propagation of variance is used so that

$$\text{var}\{A_s\} = \sum_{i=1}^{M_r} \left(\frac{\partial A_s}{\partial R_i} \bigg|_{E\{R\}} \right)^2 \text{var}\{R_i\} + \sum_{i=1}^{M_r} \sum_{j=1, j \neq i}^{M_r} \frac{\partial A_s}{\partial R_i} \bigg|_{E\{R\}} \frac{\partial A_s}{\partial R_j} \bigg|_{E\{R\}} \text{cov}\{R_i, R_j\} \quad (46)$$

From the properties of the multinomial distribution, the variance of R_i is

$$\text{var}\{R_i\} = Nq_i(1 - q_i) \quad (47)$$

and the covariance between the number of hits is

$$\text{cov}\{R_i, R_j\} = -Nq_i q_j \quad (48)$$

The partial derivatives evaluated at the expected values of the number of hits can be shown to be

$$\frac{\partial A_s}{\partial R_i} \bigg|_{E\{R\}} = \frac{\bar{V} \bar{\rho}_i}{N} (a_i - a_L) \quad (49)$$

and

$$\bar{V} = \sum_{i=1}^{M_r} \frac{z_i}{\bar{\rho}_i} \quad (50)$$

is the bulk specific volume of the solids (bulk volume of unit mass of solids mixture). The variance is then

$$\begin{aligned} \text{var}\{A_s\} &= \sum_{i=1}^{M_r} \frac{\bar{V}^2 \bar{\rho}_i^2}{N^2} (a_i - a_L)^2 Nq_i(1 - q_i) \\ &\quad - \sum_{i=1}^{M_r} \sum_{j=1, j \neq i}^{M_r} \frac{\bar{V}^2 \bar{\rho}_i \bar{\rho}_j}{N^2} (a_i - a_L)(a_j - a_L) Nq_i q_j \end{aligned} \quad (51)$$

We write the last term as

$$\sum_{i=1}^{M_r} \sum_{j=1, j \neq i}^{M_r} \frac{\bar{V}^2 \bar{\rho}_i \bar{\rho}_j}{N^2} (a_i - a_L)(a_j - a_L) Nq_i q_j = \sum_{i=1}^{M_r} \sum_{j=1}^{M_r} \frac{\bar{V}^2 \bar{\rho}_i \bar{\rho}_j}{N^2} (a_i - a_L)(a_j - a_L) Nq_i q_j - \sum_{i=1}^{M_r} \frac{\bar{V}^2 \bar{\rho}_i^2}{N^2} (a_i - a_L)^2 Nq_i^2 \quad (52)$$

The first term on the right can be split as

$$\begin{aligned} \sum_{i=1}^{M_r} \sum_{j=1}^{M_r} \frac{\bar{V}^2 \bar{\rho}_i \bar{\rho}_j}{N^2} (a_i - a_L)(a_j - a_L) Nq_i q_j &= \frac{\bar{V}^2}{N} \sum_{i=1}^{M_r} \bar{\rho}_i (a_i - a_L) \frac{z_i}{\bar{\rho}_i \bar{V}} \sum_{j=1}^{M_r} \bar{\rho}_j (a_j - a_L) \frac{z_j}{\bar{\rho}_j \bar{V}} \\ &= \sum_{j=1}^{M_r} \bar{\rho}_j (a_j - a_L) \frac{z_j}{\bar{\rho}_j \bar{V}} \frac{\bar{V}}{N} \sum_{i=1}^{M_r} z_i (a_i - a_L) \\ &= 0 \end{aligned} \quad (53)$$

Each of the sums equate to zero.

Then

$$\begin{aligned}\text{var}\{A_s\} &= \sum_{i=1}^{M_T} \frac{\bar{V}^2 \bar{\rho}_i^2}{N^2} (a_i - a_L)^2 N q_i (1 - q_i) + \sum_{i=1}^{M_T} \frac{\bar{V}^2 \bar{\rho}_i^2}{N^2} (a_i - a_L)^2 N q_i^2 \\ &= \sum_{i=1}^{M_T} \frac{\bar{V}^2 \bar{\rho}_i^2}{N^2} (a_i - a_L)^2 N q_i \\ &= \frac{\bar{V}}{N} \sum_{i=1}^{M_T} z_i \bar{\rho}_i (a_i - a_L)^2\end{aligned}\quad (54)$$

This result shows that the variance due to small scale particulate heterogeneity does not depend on particle masses (or volumes) in the same way as the fundamental sampling variance of the sampled material and that the variance is inversely proportional to the number of increments collected which corresponds to Gy's statement that taking lots of small increments is *the* way to defeat the variance due to grouping and segregation. Note also that if all the class bulk densities are similar then $\bar{V}\bar{\rho} = 1$ and

$$\text{var}\{A_s\} \approx \frac{1}{N} \sum_{i=1}^{M_T} z_i (a_i - a_L)^2 \quad (55)$$

So now, what about the particle number dependence of the grouping factor γ ? Another way to write the fundamental sampling variance defined by Gy is

$$\frac{\sigma_{FE}^2(A_s)}{a_L^2} = \frac{1}{M_S} \sum_{i=1}^{M_T} z_i \rho_i v_i \left(\frac{a_i - a_L}{a_L} \right)^2 \quad (56)$$

where v_i is the typical fragment volume in the i^{th} particle class and M_S is the sample mass. This is Gy's result for the case of the lot mass much larger than the sample mass. So from (41),

$$\begin{aligned}\sigma_{GE}^2 &= N_{PI} \zeta \sigma_{FE}^2 \\ &= \frac{N_{PI} \zeta}{M_S} \sum_{i=1}^{M_T} z_i \rho_i v_i (a_i - a_L)^2\end{aligned}\quad (57)$$

where N_{PI} is the number of particles in a group (potential increment). Now, as the particle volumes go down, the number of particles in our 'groups' goes up, but this is appropriately balanced by the volume factor in the above expression for the fundamental sampling variance; volume down, numbers up. So as σ_{FE}^2 goes down with decreasing particle volume, the number of particles goes up in inverse proportion to the volumes and we have a balance which might just come out close to the result in (54) or (55). Some simple exploration is appropriate.

Consider a very simple case in which all particles have the same mass, $\bar{m}$, but differing analyte contents. In such a case, the grouping and segregation variance can be written as

$$\begin{aligned}\sigma_{GE}^2 &= N_{PI} \zeta \sigma_{FE}^2 = \frac{N_{PI} \zeta}{M_S} \sum_{i=1}^{M_T} z_i \bar{m} (a_i - a_L)^2 \\ &= \frac{\zeta N_{PI} \bar{m}}{M_S} \sum_{i=1}^{M_T} z_i (a_i - a_L)^2\end{aligned}\quad (58)$$

But, with N_{PI} the number of particles in an increment, $\bar{m} N_{PI} = \bar{m}_I$, the mean mass of an increment, so

$$\frac{N_{PI} \bar{m}}{M_S} = \frac{\bar{m}_I}{M_S} = \frac{1}{N} \quad (59)$$

making

$$\begin{aligned}\sigma_{GE}^2 &= N_{PI} \zeta \sigma_{FE}^2 = \frac{N_{PI} \zeta}{M_S} \sum_{i=1}^{M_T} z_i \bar{m} (a_i - a_L)^2 \\ &= \frac{\zeta}{N} \sum_{i=1}^{M_T} z_i (a_i - a_L)^2\end{aligned}\quad (60)$$

To the extent of the simplification, the new relationship is in agreement with Gy's. Certainly, the order of magnitude is correct, and we recall that Gy assumed his increments were 'similar' and having similar numbers of particles in each 'group' or potential increment.

From a practical point of view, there is a need to be able to gauge the possible magnitude of grouping and segregation variance arising from *DH* compared to the fundamental sampling variance that is governed by *CH*. On the one hand, we have the relative variance due to constitutional or intrinsic heterogeneity (*IH*) given by

$$\frac{\sigma_{FE}^2(A_s)}{a_L^2} = \frac{1}{M_S} \sum_{i=1}^{M_T} z_i \rho_i v_i \left(\frac{a_i - a_L}{a_L} \right)^2 \quad (61)$$

which can be written in terms of the sampling constant K_S

$$K_S = \sum_{i=1}^{M_T} z_i \rho_i v_i \left(\frac{a_i - a_L}{a_L} \right)^2 \quad (62)$$

as

$$\frac{\sigma_{FE}^2(A_S)}{a_L^2} = \frac{K_S}{M_S} \quad (63)$$

On the other hand

$$\frac{\text{var}_{GE}\{A_S\}}{a_L^2} = \frac{\bar{V}}{N} \sum_{i=1}^{N_I} z_i \bar{\rho}_i \left(\frac{a_i - a_L}{a_L} \right)^2 \quad (64)$$

for which we can define

$$D_S = \sum_{i=1}^{N_I} z_i \bar{V} \bar{\rho}_i \left(\frac{a_i - a_L}{a_L} \right)^2 \quad (65)$$

If we form the ratio of the variances and apply the segregation factor, we have

$$\begin{aligned} \frac{\text{var}_{GE}}{\sigma_{FE}^2} &= \xi \frac{M_S}{N} \frac{D_S}{K_S} \\ &= \xi m_{inc} \frac{D_S}{K_S} \end{aligned} \quad (66)$$

where m_{inc} is the increment mass that is used to collect a total sample mass, M_S . We can also make the point that the sample mass collected will in practical terms be large enough to make the relative standard deviation due to IH somewhere around 2% or 0.02, so $K_S/M_S = 4 \times 10^{-4}$ and then

$$\frac{\text{var}_{GE}}{\sigma_{FE}^2} = \xi \times 2.5 \times 10^3 \frac{D_S}{N} \quad (67)$$

Now, for the first example above with the values in Table 1 and zero voidage, we find $D_S = 0.603$. If we were to take say 50 increments in forming the sample, the ratio becomes

$$\begin{aligned} \frac{\text{var}_{GE}}{\sigma_{FE}^2} &= \xi \times 2.5 \times 10^3 \frac{0.603}{50} \\ &= 30.2 \xi \end{aligned} \quad (68)$$

This suggests that for the grouping and segregation variance to be comparable to the variance due to IH, ξ must be quite small. If it is not, the variance due to grouping and segregation will be larger than that due to IH. It is possible to calculate the value of the sampling constant for this material which comes out to 5.1 g. The sample mass to provide a 2% relative error is then 12.8 kg. The particles are 10 and 20 mm cubes and both phases carry significant analyte concentrations.

Moving to the second example with the fine gold, if we do the same calculations, the value of D_S is very large due to the low grade; $D_S = 755986$. The sampling constant is $K_S = 0.123$ g. The sample mass that would be used to confine the relative standard deviation to 2% is 308 g. The variance ratio is now, having taken 50 increments

$$\begin{aligned} \frac{\text{var}_{GE}}{\sigma_{FE}^2} &= \xi \times 2.5 \times 10^3 \frac{7.56 \times 10^5}{50} \\ &= \xi \times 3.78 \times 10^7 \end{aligned} \quad (69)$$

This result suggests that the propensity for the gold ore to segregate is enormously greater than the simple material of example 1. We might conclude that materials containing a small quantity of a liberated high-grade phase is much easier to segregate than some other type of material and consequently also impossible to mix to a condition of particulate homogeneity.

Discussion

Could this be a step towards better understanding and quantifying the grouping and segregation variance? We have the segregation factor, ξ , left with which we can adjust to move from particulate homogeneity to total segregation. But we have defined a maximal value of the grouping and segregation variance for total segregation that derives from consideration of sampling with increments of equal volume which is almost exactly true in mechanically correct sampling.

Size distribution effects can be taken into account with the introduction of a voidage factor for each particle class. The problem of dealing with the seemingly unmanageable number of particles in the increment has been more or less eliminated but finding a means of determining the segregation factor remains. This issue is akin to the issue that arose with the liberation factor before it was concluded that it was a concept that was hard to deal with and could in any event be removed when the size distribution of the target phase(s), liberated or not, could be estimated^{3,4,5}.

Maybe this is what Gy had in mind; we shall not know as sadly he has passed away without further clarification of these concepts which are quite critical to his integration of sampling theory. His arguments concerning the grouping and segregation factors are 'hard work'. This argument does seem to encapsulate his remarks on segregation in terms of explaining the need to take many increments and his statement that the factor γ related to particle numbers. However, it still suggests that the variance can be rather large, but it does provide a mitigation of the dependence of γ on the number of particles.

To the author's knowledge there has been no quantitative analysis made previously in the sampling literature of Gy's development for segregation variance so there is no other literature that has a bearing on the analysis. We have only

Pitard's publications which follow Gy¹ closely. The author hopes that the step-wise analysis of Gy's derivation and the alternative approach to the derivation of a segregation variance based on the sampling of a mineral mixture by volume, which is more or less what happens with correct sampling equipment, will clarify thinking about the segregation variance and make the concept clearer to sampling practitioners. The new derivation does not solve the problem of the practical estimation of the sampling variance but illustrates what a state of total segregation is and hopefully permits the sampling community to better conceptualise/visualise the spectrum of states of segregation that can exist in practice.

ORCID iDs

Geoffrey Lyman: <https://orcid.org/0000-0001-7570-9223>

References

1. P. Gy., Sampling of Particulate Materials, Elsevier, 1982
2. F. F. Pitard, Pierre Gy's Sampling Theory and Sampling Practice, Volume 1, CRC Press, 1989
3. F. F. Pitard, Theory of Sampling and Sampling Practice, 3rd edition, CRC Press, 2019
4. G. J. Lyman, Theory and Practice of Particulate Sampling; an Engineering Approach, Materials Sampling & Consulting, Brisbane, 2019, Chapter 4.
5. G.J. Lyman, Mathematical developments in particulate sampling theory, P M Gy Oration, Ninth World Conference on Sampling and Blending, Beijing, China May 2019