doi: 10.1255/tosf.88

Theory of Sampling and Geostatistics: a technical tribute to two geniuses

Dominique Francois-Bongarcon
Agoratek International Consultants Inc., North Vancouver, Canada. E-mail: dfbgn2@gmail.com

Introduction

The 1950s was a decisive period for the world of earth sciences: two geniuses of mathematical modelling in mining were at work simultaneously, developing theories that are now the basis for practically all we do in this and related areas. While Pierre Gy was developing modelling of microscopic scale variability and its effects on extraction of macroscopic samples of randomly selected broken material (ore, cement, food etc.), now known as the Theory of Sampling (TOS), Georges Matheron was developing models for large scale variability throughout space domains of auto-correlated variables, aka Geostatistics. Both these extraordinary scientists are no longer with us. I here want to pay tribute to both men by focusing on technical issues which bind their works together, although they actually did not regularly meet in real life (although they did at the onset of Reference 1).

As stated in a previous paper,[2] if the theory of the sampling of broken material (TOS) cannot be mentioned without reference to Gy's lifetime fundamental contributions, it can neither be fully understood outside of a geostatistical frame of reference.

TOS calls for some geostatistically flavoured concepts at small scale (Gy's formula), mixes with it at medium scale (sampling regime of one-dimensional flows) and is very much needed by larger scale geostatistics (data quality in view of estimation, variogram nugget effect, conditional simulations).

But these are not the only links between the two, and I will briefly mention a fundamental and more theoretical feature which is also at work, and which takes front stage when consistency is required between those two sets of tools. One that in effect joins the theory of the very small to that of the very large.

Theory of Sampling, TOS (Gy)

Gy's formula for the relative sampling variance is the basis of numerical sampling calculations:

$$Rel.Var. = c f g \ell d^3 (1/M_S - 1/M_L) \tag{1}$$

In this expression, as we know, *c*, *g* and *f* are material constants which can be known or derived from empirical characterisation of the material in question (sometimes only by non-trivial efforts though), *d* is the comminution nominal P95 size.

As to ℓ, the liberation factor, shown recently in Reference 3 to be the ratio of any sample variance to the variance of the liberated sample with the same average number of fragments, it was earlier proposed[4] to be modelled as:

$$\ell = (d_\ell / d)^b \tag{2}$$

where d_ℓ, which is the mineral liberation size, relates directly to the size of the coarsest grains of mineral or metal and *b* is an exponent between 0 and 3 that must normally be calibrated experimentally. Thus the overall exponent alpha of *d* in the formula is alpha = 3 − *b*. This model transforms an unusable formula (1) into a workable one.

After proper experimental calibration of the set of parameters involved, using these formulations, the practical variance of any sample can be predicted with reasonable validity (as always very much dependent upon respecting all TOS' requirement for representative sampling).

Furthermore, letting *V* be the average fragment volume *V* in the lot and rho the average density of the rock in the lot, it can be shown[5] that:

- From the point of view of the practical sampling variance, a lot behaves as if actually constituted by a series of fragments all of volume *V*. This can at times clarify the understanding of sampling enormously but unfortunately has not received a lot of attention to date.
- With $Var_V = c\ell/\text{rho}$ being the variance of the average-volume fragment, i.e. the dispersion variance, within the corresponding type of mineralisation, of a support of a known, calculable size *V* and with enough information on the mineralisation specifics, the effective variance Var_V is also calculable using TOS.

Geostatistics (Matheron)

The basis of geostatistics is Matheron's variogram curve and its modelling. Sampling errors will affect its discontinuity jump at the origin, the famous "nugget effect". It is one of the major achievements of TOS, that this contribution can be evaluated in all its components (the nugget effect consists of all incorrect—as well as all correct sampling errors, to which is added the total analytical error). This rather practical, indirect link between TOS and Geostatistics is well known and forms the basis of a powerful first understanding of the total measurement system uncertainties in practice.

Once a *valid* variogram model is available in a homogeneous domain (a *sine qua non* condition of good application of geostatistical modelling), then the dispersion variance of *any* support of known shape and dimensions throughout the domain can be predicted. The variogram of a different support than that of the data on which the original variogram curves were calculated, can also be derived theoretically.[6] Finally, the estimation variance (providing a precision) of any linear estimator can also be readily evaluated, *ibid*.

Thus it can be stated that geostatistics is "the science of variances", but this notably only reaches its full potential when used in synergy with TOS.

TOS vs Geostatistics

Gy and Matheron's theories need each other, indeed they complement one-another perfectly. Taken together, they offer the sampling practitioner the complete palette of variance study tools, from very small to very large scale. They are inseparable. Each one is incomplete without the other. Together they have empowered the earth science practitioner with an incredibly efficient modelling capability. We should rightfully marvel at the fact that these two theories were developed precisely at the same time by two different minds, independently form each other, in the 1950s.[1]

Now if we take a lot of broken ore from a homogeneous domain in a deposit (e.g. a copper mineralisation in a zone with only Chalcopyrite as copper mineral), we have seen that TOS allows us to calculate the dispersion variance Var_V of the average-volume fragment from appropriate empirically derived mineralogical parameters. The total sill of the variogram calculated based on this support, which depends directly from the sill of the original variogram, normally represents the dispersion variance of that support, i.e. the variance of data sample taken at random over the domain. It must therefore be equal to the variance of the average-volume fragment as predicted by TOS.

This is not a minor statement: provided domaining makes sense mineralogically, while TOS tells us what proportion of the variogram nugget effect is represented by the various sampling errors, it also tells us what the total sill of that variogram *should be*. This is the *ultimate link between TOS and Geostatistics*, and it can be paramount in terms of consistency, especially if comparisons are to be made between sampling variances and variances calculated using variograms, or if one wants an accurate derivation of the proportion of the nugget effect related to sampling errors.

If the mineralogy is well known, including d_ℓ, but parameter "*b*" in Equation (2) is not and the variogram domain is homogeneous enough, the relationship can even be used to adopt for that parameter "*b*" the value that brings full consistency to the sill of the variogram, thus providing a principally new calibration method for models of liberation factor.

It is clear the two theories are in fact but two faces of the full, complete framework of modelling of uncertainty in the earth sciences, and both their authors should equally be entitled to the utmost gratitude from the scientific fields and industries who have benefitted so much from their work, as well as to the highest respect from their practitioners.

Final, personal note

I have had the immense personal privilege of working with both Gy and Matheron and have had the singular opportunity of studying and researching their teachings at some significant depth. Recognised geniuses in their respective areas of technical research and expertise, they were both equally great human beings. They shared the same kindness and patience, impeccable professional ethics and a common and outstanding social intellect. They will never be forgotten as individuals, no more than will their fundamental technical contributions.

Directly or indirectly, these two giants of mathematical modelling have taught me important lessons:

- Models are only models and if they can easily be invalidated, they cannot be proven, but only validated, in the long term, by the "sanction of practice" (Matheron[7]).
- Keep a critical mind about the underlying theories (including theirs!).
- Internal consistency of models is paramount and should never be sacrificed lest the models can grossly mislead us.

As their heirs, we all have a duty to use, disseminate, further clarify, promote and harmonise their teachings. As to the industries that have taken advantage of the two theories for some 50–70 years, it is suggested they should be more visibly grateful, and each practitioner in the industry should relentlessly fight for their further overall recognition. Indeed, how many companies in the industry would be here today was it not for their use of TOS and Geostatistics? And without the magnificent influence of these two giant mentors, how many of us would be as successful and enjoying our works, as is the case in our respective technical domains?

They have been, in turns, friends and mentors. Occasionally, they opened up on their inner feelings. I will always remember these more tender moments, when one was lamenting the cruel betrayals of professional life, discounting geostatistics as only a "social thing", with a snarl, while insisting the only important thing he had ever written were his two mathematical books, or when the other, still shedding tears, would evoke the war and the terrible fate that had been that of persons close to him. They were compassionate but morally strong beings, and the feelings they shared were as inspiring as their technical insights. Both of them could recount situations of life where they had chosen the hard way of ethics and moral duty over the easier path. As former friends, we shall also cherish the more intimate lessons of life they had shared with us.

Pierre Gy was particularly happy when we finally launched the WCSB conferences, in his honour, as it was for him a guarantee his works would survive him, a fear he had had for very long. Figure 1 shows both of us during the happy times of WCSB1 in Esbjerg, Denmark, in 2003.

Figure 1. The author and Pierre Gy at WCSB1, Esbjerg, 2003.

References

1. D. Francois-Bongarcon and F. Pitard, 2015, "Comparison between samples with constant mass and samples with constant fragment population size (and calculations of their sampling variances) by G. Matheron", translated from French to English, clarified and further commented, *TOS forum* **5,** 231–238 (2015). doi: http://dx.doi.org/10.1255/tosf.80
2. D. Francois-Bongarcon, "Theory of Sampling and Geostatistics: an intimate link", *Chemometr. Intell. Lab. Syst.* **74,** 95–114 (2004). doi: http://dx.doi.org/10.1016/j.chemolab.2004.03.015
3. D. Francois-Bongarcon, "Introduction and first ever rigorous derivation of the liberation factor", *TOS forum* **5,** 165–168 (2015). doi: http://dx.doi.org/10.1255/tosf.79
4. D. Francois-Bongarcon, "The modeling of the liberation factor and its calibration", in *Proceedings of the 2nd World Conference on Sampling and Blending*, Twin Waters, QD, Australia, May 2005.
5. D. Francois-Bongarcon, "Theory of Sampling and Geostatistics: the ultimate link", proposed for publication at the 7th World Conference on Sampling and Blending (2017).
6. G. Matheron, "The theory of regionalized variables and its applications", *Les Cahiers du CMM de Fontainebleau*, Fascicule 5. Ecole des Mines de Paris (1971).
7. G. Matheron, *Estimating and Choosing: An Essay on Probability in Practice*. Springer-Verlag (1989) and in *Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau*, Ecole des Mines de Paris.