doi: 10.1255/tosf.38

TOS vs geostatistics—again?

Dominique Francois-Bongarcon and Kim H. Esbensen

Introduction

In a paper presented by the first author at the WCSB1,[1] the intimate relationships that exist between the theory of sampling of broken material (TOS) and the discipline of geostatistics were examined for the first time. In a nutshell, if TOS cannot be mentioned without reference to Pierre Gy's lifetime fundamental contributions, it can neither be fully understood if outside the geostatistical reference frame. TOS calls for geostatistical concepts at the small scale (through Gy's formula and the liberation factor), mixes with it at medium scale (sampling regime of one-dimensional flows) and is also very much needed by larger scale geostatistics (data quality in view of estimation, conditional simulations).

Indeed, in the 1950s, motivated by a growing need for better grade estimation methods in the mining industry worldwide, the late Professor G. Matheron developed the Theory of Regionalized Variables, a.k.a Geostatistics.[2] Pierre Gy was at the same time busy developing his Theory of Sampling, but when his first work was officially published and generally applauded, a certain skepticism also rose about his central statistical demonstrations, it was to Matheron that he turned to validate his numerical development. In a seminal paper,[3] translated to English (and which will be presented in this language for the first time at the WCSB7 in Bordeaux in June 2015), Matheron indeed approached the problem of the calculation of the sampling variance with a fresh eye using a tedious but more rigorous demonstration based on probability calculus, which fully validated the results of Gy's developments.

Geostatistics was at this time "fresh out of the oven", so to speak, and its practical applications were just beginning to be figured out. So when Gy introduced the use of the *variogram* to tackle and analyse the behaviour and sampling of one-dimensional material streams (process), it was with a tool that was not yet fully understood.[4] Decades later, when Francois-Bongarcon established the missing models for the predictive use of the liberation factor,[5,6] it was thanks to a modern understanding of geostatistical concepts.

It is felt important that the relationship between TOS and Geostatistics needs to be examined again, and as always with a critical eye if we are to influence the way TOS is to develop in a proper manner.

The need for a complete solution

Bad vocabulary

"Sampling"! A small word but a full world! While geologists "sample" rocks and mineral deposits using hammers and core drilling, laboratory operators around the world are busy "sub-sampling" crushed material during sample preparation for analysis and metallurgical processes, or in industry a lot of "sampling" is carried out using automatic devices. Meanwhile, surveys are done on human population "samples", for example, for which statisticians have defined proper collection rules. So, is "sampling" a universal human activity?

The truth if the matter is these varieties of sampling are not all equivalent, but the vocabulary has been used freely without much precision, with one single verb (*sampling*) being used to describe very distinct and very different endeavours.

- There are at least three different concepts we can identify under this one verb, which must never be confused: statistical sampling (statistics of independent variables)
- physical sampling of broken ore (TOS) (sampling of heterogeneous matter in general)
- *in situ* sampling (i.e. measurement and interpolation theory, aka geostatistics)

Leaving statistical sampling aside, many parties in technology and industry usually refer to the two latter categories as "sampling" synonymously, as if they were indeed based on, and using the same theories and tools.

In the particular case of the mining and minerals processing industry, indeed there is a need for a complete theory covering both broken ore sampling and *in situ* sampling. That wonderful theory, surprisingly, has not been written yet, and for the time being, TOS is only one important half of it. For the other half, thanks to Matheron, all the tools exist in geostatistics (and many of us are using them in *ad hoc* manners), but no one has ever bothered putting them into a practical theory for the perusal of exploration geologists. This unifying work still needs to be undertaken, and it is definitively not of the resort of TOS alone.

In the meantime, the confusion between the two survives, and misapplications keep appearing, tragically, and apparently on a regular basis.

Erroneous uses of TOS

In this context, there are two classical mistakes, which are repeatedly committed:

- Using TOS's famous variance prediction formula, "Gy's formula", to calculate the precision of sampling *a process* with increments collected *along time*.
- Preparing and assaying the complementary split of a drill hole core interval, to calculate the precision attached to "sampling the half core" (thinking the variance value derived from these pairs of "duplicate samples" can contribute to characterise the precision of the final assay result.

As pointed out in Reference 4, these errors all stem from the implicit use of probabilistic models for which the attached probability space (or, often, even the mere meaning of that concept) is not understood, making any variance calculation a futile, purely mechanical exercise with no usable meaning.

In order to be able to make sensible decisions about the future direction of TOS, it appears advantageous to establish a clear foundation for possible discussions at WCSB7 for example (or here in *TOS forum*). In this context it is essential to weigh the pros and cons of both approaches (TOS / geostatistics). This means that future applications TOS should be carefully examined to make sure that such confusions as above are eliminated, missing theoretical links are developed and implemented in practise. In other words, we need to decide if we want to grow TOS into a complete theory of both *in-situ* and broken ore sampling, or if we prefer to clean up the existing practice from its potentially misleading features and leave it for the geostatistical community to properly figure this out separately, if it ever will.

TOS' future: two options

Status quo–separation

In the "separation" option, the status quo of TOS is maintained, albeit with *some* cleaning-up of concepts and practises, especially making sure the difference between *in-situ* sampling and broken ore sampling can be, and will be better, distinguished by the practitioners, not shying away from relying on proper methods mainly used outside TOS to tackle category 1 above, i.e. not being afraid to call on professional geostatistic collaboration wherever, whenever required. This amounts to "Giving back to Caesar what belongs to Caesar", and hopefully geostatisticians would soon jump on board to study process streams (essentially 1-D geostatistical problems) using state-of-the-art geostatistical methods as a substitute for Gy's sometimes slightly invalid variographic analysis procedures. These include using experimental variogram values instead of a smooth model and the erroneous splitting of variance components along oblique lines in a variogram plot. It is recognised that many colleagues are content and satisfied with the first item, using experimental variograms for heterogeneity characterisation, and process interpretation alone. This is where a healthy debate may find one of its foci (even DFB and KHE do not agree entirely on this matter).

The "pro" of this option would be to collaborate better regarding what can be achieved in terms of understanding, characterising and diagnosing such streams. One may imagine that powerful, more modern geostatistical tools and concepts could provide much more elaborate results than was possible in the 1960s. The "con" option would be, sadly, that a swathe of contemporary activity and experience would escape to another profession.

Integration

Should we decide to opt for integration efforts, then we only need to revamp and modernise our methods whenever they need geostatistical involvement. This will include:

How to handle duplicate half-cores in QA/QC.

- The most important issue is that we should vehemently refrain from considering a half-core as a bona fide sample of the full core (sic).When used by geostatisticians, the data gives the same results and procedures, as it is transparently handled through the nugget effect of the variogram. In fact, often, the practitioner may think he/she is using full core, of which he/she has taken a primary sample, but in reality, he/she is only ever using data defined on a half-cylinder support. The full core is not present anywhere and is not supporting any part of the modelling. There is already a big clue here!
- But when used in QA/QC, it is a different story. Indeed, there is a complete duality between geostatistical auto-correlations and broken ore segregation. Should one be interested in the half-core selection process, even though there is often no good reason for it (see above), then, to be correct, the auto-correlation that exists between a split half core and the other half should be used when quantifying the (useless) variance that one can calculate from such pairs of alledged "sample duplicates" and which includes both sampling and assaying errors, plus the geostatistical nugget effect (auto-correlations). The calculation should treat this as a segregation, but within TOS we do not have the tools to do it—yet.

How to properly calculate/predict the precision attached to the estimation of an average over a period of time from (pseudo-)periodic increments collected from **a 1-D stream**. The variogram should be calculated experimentally, and *then* be *modelled* using a smooth "allowable" function (there are strong geostatistical reasons why doing so is necessary for meaningful results),[2] then a non-illusory estimation variance can be calculated using a kriging program.

How to analyse and diagnose a 1-D stream. The variogram is a naturalistic tool that gives important information on the behaviour of streams. Sills and nugget effects provide key information for those who are in-the-know about the rules of interpreting variograms. Its periodicities can be assessed and important conclusions derived.

But it is possible, with correctly handled geostatistics, to go much further. For instance, the stream can be simulated geostatistically, with formidable benefits in terms of testing certain strategies, whether for the increment collection or within the frame of other types of variability studies.

- **The handling of segregation in large stockpiles** can advantageously be the objective object of geostatistical handling, although this would be a matter of research for the time being. Much work remains here; we have only started this particular strand.

Conclusions

While TOS is unquestionably being applied more and more successfully, witness the series WCSB1–WCSB7 and in many other situations in all of science, technology and industry, it is proposed that a cleaning exercise, as suggested above, would condition the integration effort with great advantage. If this takes flight, it can be started and executed rather quickly and better tools will be developed as a result.

It is hoped that the present opinion piece can serve as a start of a healthy debate about these central issues.

References

1. D. Francois-Bongarcon, "Theory of sampling and geostatistics: an intimate link", in *Proceedings of WCSB1*, July 2001, Esbjerg, Denmark (2001).
2. G. Matheron, *The Theory of Regionalized Variables*. Fascicule 5, Centre de Géostatistique de l'Ecole des Mines de Paris, Fontainebleau, France (1971).
3. G. Matheron, "Comparaison entre les echantillonnages à poids et à effectifs constants", *Revue de l'Industrie Minérale* **August,** 609–621 (1966).
4. D. Francois-Bongarcon, "Some theoretical and practical aspects of 1-D stream sampling", in *Proceedings of WCSB4*. The Southern African Institute of Mining and Metallurgy (2009).
5. D. Francois-Bongarcon, *Geostatistical Tools for the Determination of Fundamental Sampling Variances and minimum Sample Masses, Geostatistics Tróia '92*. Kluwer Academic, Doordrecht, The Netherlands, Vol .2, pp. 989–1000 (1992).
6. D. Francois-Bongarcon and P. Gy, "The most common error in applying 'Gy's Formula' in the theory of mineral sampling, and the history of the liberation factor", in *Towards 2000*. AusImm (2000).