doi: 10.1255/tosf.69

Counteracting soil heterogeneity sampling for environmental studies (pesticide residues, contaminant transformation) – TOS is critical

Z. Kardanpour[a,b], O.S. Jacobsen[b] and K.H. Esbensen[a,b,c]
[a]**ACABS Research Group, Aalborg University, Denmark. E-mail: zk@bio.aau.dk**
[b]**Geological Survey of Denmark and Greenland (GEUS). Copenhagen. Denmark**
[c]**ACRG Research Group, Telemark University College, Norway**

This Ph.D. project aims at development of an improved methodology for soil heterogeneity characterization for 'next generation' sampling/monitoring and spatial modeling practices a.o. allowing more realistic pesticide variability in environmental contaminant assessment studies. Such studies typically take place in the laboratory. The key question therefore is: Are current sampling techniques able to counteract the inherent soil heterogeneity met with in the field? Analysis of traditional soil sampling approaches from a Theory of Sampling perspective, the answer is a resounding negative. This contribution summarises the extensive sampling aspects involved in the overall project context, also involving chemometric data analysis with a special twist.

Soil heterogeneity

Soil heterogeneity characteristics in the natural environment do not follow normal statistical distributions and most certainly not regular spatial distributions. There is a need for scientifically based procedures and principles for parameterisation of the intrinsic variability in many types of agricultural, urban and natural soil systems. Conventional computer simulations are critically dependent on the specific choices of model and statistical distribution characteristics, not necessarily always realistic. To a large extent, knowledge in these fields of research is based on laboratory-scale batch experiments involving either soil 'as is' and often with samples of only a few grams of the soil matrix, Figure 1.

Even when procedures for species transport, kinetic studies and analysis follow established scientific and international standards, it is increasingly recognised that small scale experiments do not fully reflect the effective variability and heterogeneity of the salient soil and geological formations at larger, more relevant and more realistic fields scales.[1,2] This will unavoidably cause problems for the later scaling-up to field scale of the processes and effects studied, especially regarding the possibility for valid *volume generalisation*.

The soil matrices involved are in fact often significantly heterogeneous, and a number of the subsequent soil model and interpretation issues are critically related to the empirical variability at scales larger than the laboratory samples (both in vertical and horizontal dimensions). Also taking into account the time scales involved in dynamic studies only add further to system complexity. Within the environmental sciences there is a strong need for an integrated understanding of chemical contaminant transformations (e.g. pesticide degradation), spatial modeling and multivariate data analysis.[3–5] All critical soil characterizing parameters are in need of *effective counteraction* of the variability related to inherent soil heterogeneity when securing valid soil pots for laboratory experimentation, i.e. how to secure representativity of individual pots? Not only that, but how to guarantee that multiple pots containing soil sampled in nature are as identical as possible for replicate laboratory studies?

The main motivation for this Ph.D. has been to develop generic procedures to map the effective heterogeneity of soils *at all relevant scales*. The present paper describes a comprehensive approach

Figure 1. Traditionally "experimental pots" to be used in the laboratory (e.g. pesticide residue, pollutant characterisation or contaminants transformation studies) are 'sampled' directly in the field (*grab sampling*). Ignoring inherent soil heterogeneity leads to compositional differences between pots of an unknown magnitude due to uncontrolled FSE, GSE, ISE. Pots are all too often simply *assumed* to be identical. Observe the drastic mass-reduction from field, sample, pot, often of the order of magnitude 1:1000.

for this purpose, here applied on typical clayey soils with a focus on intrinsic parameters (minerogenic variables, 'soil framework variables'). Clayey soils serve as an exemplar medium, chosen because of the typical non-trivial practical sampling problems and limitations encountered for this type of soil. The main focus is on characterising and comparing grab vs. composite sampling in a full-scale experimental study based on both a short range 2-D design (cm-dm) and a large scale linear profile (dm - ~100m scale). The methodological principles developed are completely general for all soil types however. Results from parallel sandy soil studies are also be presented—together these two soil types cover a significant range of temperate region soil types.

Field experiments—field sampling

This study evaluates a series of experiments testing improved designs of field and laboratory sampling, Figures 2-3, at all stages from the primary field sampling to the final analytical sample preparation, Figures 3-5. The effect of soil heterogeneity at different scales critically affects the validity of the sampling/monitoring procedures involved.

Field samples were collected from the topsoil (A-horizon; 0-25 cm) of a typical clayey soil. For a short range experiment a 50 × 50 cm square of top soil was exposed by carefully removing the uppermost grass layer (approx. 5 cm), Figure 3A. Within this square a total of 68 'standard' soil samples, each of 30-40 g, were collected with the shortest practically possible in-between distance (less than 2 cm) in the pattern shown in Figure 3A. Primary soil samples were extracted using a conventional cut plastic syringe (diameter 1.5 cm, length 10 cm, Figure 3B) and immediately sealed in airtight containers; these samples make up the S-Set samples to be used for '2-D heterogeneity visualisation' (see below). The remaining soil of the square "box" (Figure 3 A) down to 10 cm depth was extracted as a "primary bulk sample", also sealed in a moisture tight plastic bag and transported to the laboratory for further sub-sampling experiments. For a *large scale* variographic characterization, soil samples were collected in identical fashion in the same field with an equidistance of 1 m along a 85 m long profile. A parallel study on sandy soil included a *short scale* replication experiment,[6] is presented in Figure 2.

Contemporary sampling approaches in environmental/soil sciences makes little or no allowance for soil heterogeneity, resulting in significant between-pot heterogeneity which impacts on the discriminating power in laboratory experiments. This is not unavoidable however. Variographic analysis (below) shows the advantage of using increment locations for composite sampling with a distance below the range for both organic and inorganic compounds based on empirical soil sample variograms.

Analytical methods

Field samples were 300-400 gram, while laboratory samples were 20-30 gram moist soil after careful TOS-compliant sub-sampling.[7,8] A focused mass reduction experimental design employed a suite of 16 natural soil parameters including: moisture, organic matter (loss on ignition), pH, soil cations and anions (clayey soil). In addition a large suit of 38 inorganic parameters plus a set of 9 natural and anthropogenic compounds including moisture, organic matter, bacteria counts (CFU), carbon-14 measurement of MCPA sorption and mineralization and glucose respiration were analysed in the sandy soil. Analytical parameters were selected with an aim

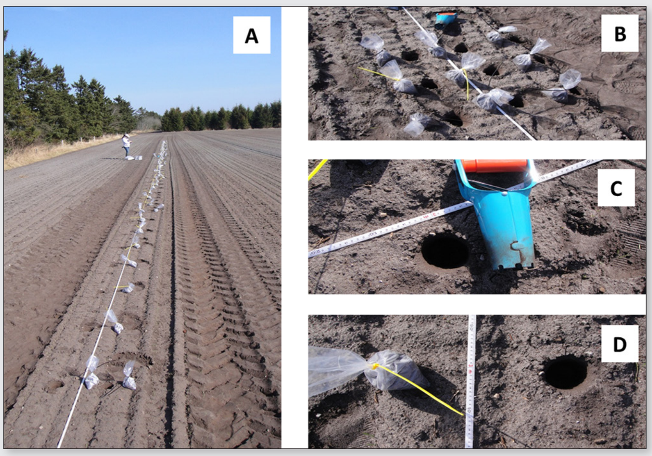

Figure 2. Sandy soil sampling. a) Long range profile for variographic characterisation (~100 m) parallel to the recent ploughing direction; b) Local "grid replication design" (9 samples covering 1 × 1 m); c) Conventional soil sampling hand tool; d) Sample excavation and airtight sample bag.

Figure 3. Laboratory mass reduction for clayey soil samples, A) "Small scale 2-D experiment" (50 × 50 cm), B) Single syringe sample; C) Primary bulk grab sample, D) Single grab sub-sample taken from C); E) Remaining bulk sample laid out for composite sub-sampling, F) Increment size, G) Single composite sample (15 increments); H) Single sample after grinding, entering a bespoke laboratory splitter. Identical procedures were applied to all samples in this study. I) Two sub-samples obtain after splitting.

to study soil heterogeneity with different natural (sandy, clayey), anthropogenic (sandy) and minerogenic (sandy, clayey) parameters with an aim to develop suitable sampling methods for these and similar matrix types.

Laboratory sub-sampling

A general framework is needed for dealing with all operative scale-interdependencies when establishing representative sampling procedures for specific soil types, instead of traditionally having to rely on a universal, standardized sample size and a conventional sampling plan, *supposed* to be able to work well for all soil types as is today's tradition in many fields.

The primary field sample size (200-300 gram) must be reduced to the analytical sample size (1-2 gram), not a trivial mass-handling issue under significant heterogeneity. In order to provide representative sub-samples, TOS principles were applied scrupulously to all mass reduction steps.[7,8] In this project a comparison was directed at grab vs composite sampling, in which two sub-sample sets, a.o. obtained by alternative methods for grinding/splitting, are compared. An embedded 2-D heterogeneity study was finally used for small scale spatial correlation characterisation, supported with a data analytical correlation study (chemometrics). All practical sub-sampling stages are illustrated in Figures 3-5.

Variographic profile characterisation

The cm-dm scale heterogeneity was studied by the '2-D small scale experiments' illustrated above, while the m-100m transect scales were studied by variographic characterisation[6,9] to provide an understanding of how the individual elements are distributed spatially in the field along the 100 m long baseline profile. All data were inspected for possible outliers or trends; outliers have been excluded and in case of a variable trend (possibly to be expected as samples are distributed along the shallow trend incline, Figure 2), de-trended profiles were subjected to variogram characterisation.[9]

Chemometric data analysis

Synoptic overviews of the correlation data structure between 40+ chemical parameters and of the relationships between all variograms were analysed by multivariate data analysis (chemometrics).[11] In here each variogram contributes to a special type of X-matrix in which the objects correspond to the set of variograms, all of which are characterized by a joint set of special variables, 'lag variables' [1, $N_U/2$]. This array is termed the $X_{variogram}$ matrix. The number of 'objects' (N_{obj}) is equal to the number of elements.

Figure 4: Fresh soil mass-reduction steps for clayey soil. A procedure, identical to riffle splitting, was developed for sub-sampling in the laboratory for this type of mildly sticky, non-flowing material.

Figure 5. Manual 'riffle-splitting' simulation for representative sub-sampling in the laboratory. This technique is also known as bed-blending, scaled-down and adapted to both dried, but still cohesive primary sandy soil samples (top) as well as to fine-crushed, dried powder (bottom). This procedure can be described as 'linear bed blending/transverse thin-slice reclaiming'.

Decomposing $X_{variogram}$ results in score plots in which each variogram is depicted in relation to all other variograms, i.e. to which degree variograms are *similar* or *dissimilar* in their characterization of the spatial structures. Whereas standard PCA displays the behavior between correlated variables,[10] the loading plot of the $X_{variogram}$ matrix visualizes the relationships between the variogram lags i.e. which *scales* behave in a coherent fashion, and which display different behaviors. For PCA($X_{variogram}$) the scores and loadings plots render a synoptic characterization of the spatial characteristics for all chemical parameters involved.[11]

According to the nature of the variogram, V(j) values represent squared heterogeneity differences, which means the $X_{variogram}$ data are all expressed in the same 'measurement unit' ('lag distance'). The analogy to ordinary spectral data is clear as conventional spectral values (transmittances, absorbances or otherwise transformed original radiometric data) are also expressed in the same 'measurement units'. This will make interpretation of the $X_{variogram}$ PCA solutions more familiar for those in-the-know regarding multivariate data analysis. Note that these special types of spectra may be in need of auto-scaling, or that may not – which will depend on the empirical variance differences between the variables (or 'lag-variables'). This issue is problem-dependent and cannot be resolved by a general imperative; different data set structures may require specific solutions.

Results and discussion

Aiming for a general approach to exemplify and quantify the effectiveness of heterogeneity characterisation in soil, a set of relevant geochemical parameters was studied at scales from cm to 100 m. In this context both small scale (2-D) and large scale (1-D) variability studies were conducted on different soil types (clayey and sandy soil). This study includes all scales from field sample to analytical aliquot and primary sampling w.r.t. soil type, secondary sub-sampling comparison, further subsampling procedure and evaluation.

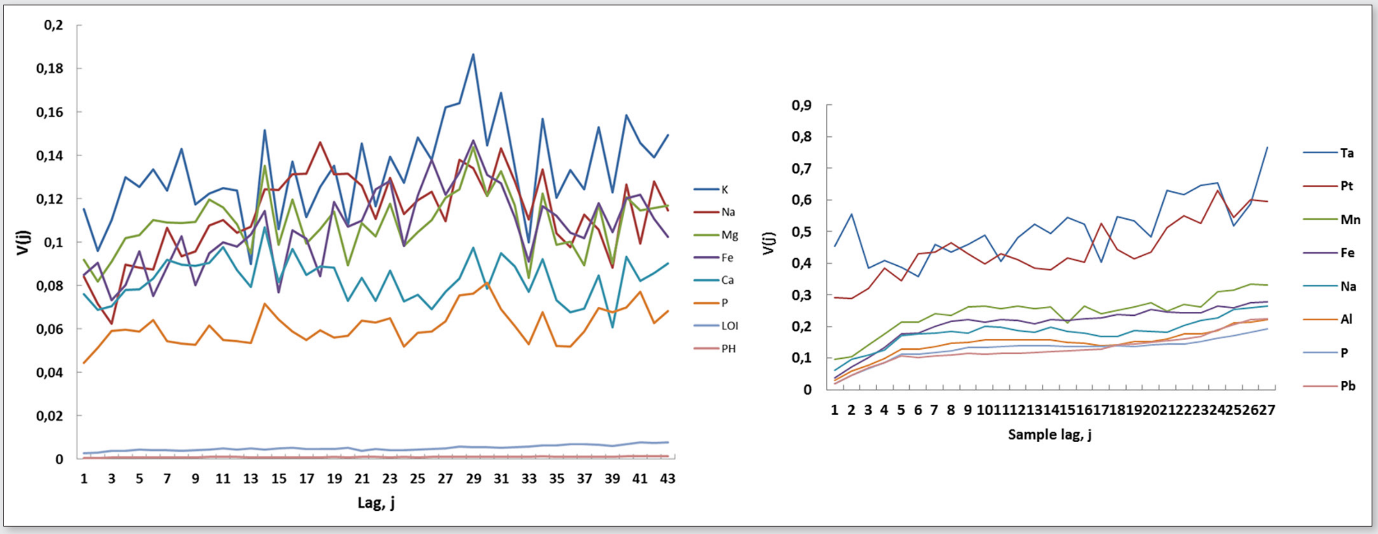

Figure 6. (left) Synoptic variogram plot for eight selected parameters in clayey soil (LOI after de-trending). Two clear groupings can be observed, i.e. two parameters with extremely low and stable variograms and six parameters which show distinctly more irregular variograms at higher sill levels (signifying distinctly larger heterogeneities). All variograms are essentially flat without a clear range for this soil type. (right) Synoptic variogram (sandy soil) for eight selected parameters (Ta, Pt, Mn, Fe, Na, Al, P, Pb) comprising both the highest and the lowest sills encountered (Ta, Pt) vs. (P, Pb) respectively. There would appear to be a general average range of approx. 5 meters. Weak variogram trends do not distrub conventional interpretation.

Evaluation and comparison of subsampling stages were conducted for clayey soil only, the most complex soil type (because of both structural properties and logistics), including a characterisation of different mass reduction (sub-sampling) procedures. Assessing the reproducibility of laboratory grinding/splitting, the TOS-optimized grinding and homogenization step was found to be acceptable for the current purpose. Furthermore, as expected from comparing grab and composite sampling (TOS), for 2/3 of the geochemical soil parameters sub-sampling methods show significant differences when based on grab sampling.

A large scale variability study was directed at two fields with different soil properties with the aim of showing a general comprehensive soil heterogeneity characterization approach wholly based on TOS principles. Figures 6 show variographic characterisation of selected variables for both clayey and sandy soils. These studies are reported in full in the first author's Ph.D. thesis.

It may occasionally be of interest to apply a multivariate approach in order to include all soil parameters simultaneously. A PCA ($X_{variogram}$) approach has been developed that simplify all variogram relationships in conventional scores and loadings plots.[10,11] As one example, Figure 7 shows how it is easy to estimate a general (average) variogram range. This approach is generic and can be applied to any set of parameters in any type of soil. It is only necessary to have enough data (samples) to be able calculate proper variograms.

Combining results of natural organic and anthropogenic parameters with minerogenic parameters from two soil types, the optimal procedure for securing *comparable* field samples (for 'identical' pot samples) for environmental pollutant experiments (samples with minimum inter-sample variability) must be by systematic deployment of *composite sampling* with increment distances less than half the range, Figure 6 - always with a number of increments as high as practical and logistically possible (depending on the total

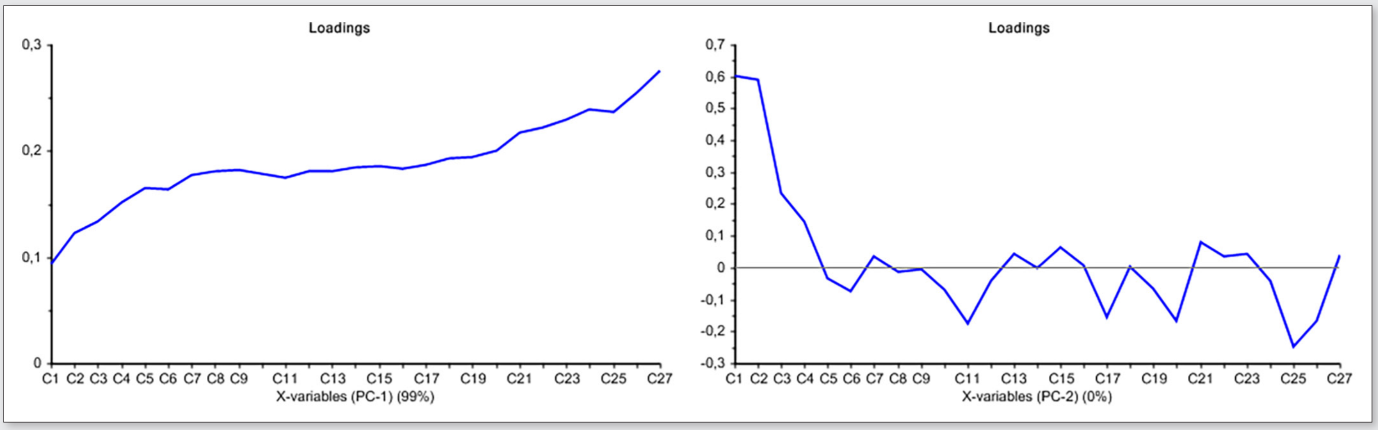

Figure 7. (left) PCA ($X_{variogram}$) loading plot PC-1 (sandy soil, all parameters), and PC-2 (right). The $X_{variogram}$ matrix has not been subjected to pre-treatment (no centering, no scaling), see text for details. The range of the average variogram shape (PC-1 loading spectrum, at left) is ca. 5 meters.

mass required for experimentation). This conclusion strictly speaking only applies to the specific soils investigated here, one of which happens already to be somewhat well mixed (the sandy soil). Other soil types may display much shorter ranges, much higher heterogeneities that is. The general lesson is that increments for composite sampling have to originate only from areas with scale a parameter less than half the range of the salient empirical soil variogram. A heterogeneity-characterising pilot study variogram is sine qua non.

Conclusions

Empirical heterogeneity description is a critical success factor in soil, contamination, pollution and environmental science studies a.o. when natural variability effects are to be reliably managed. The Theory of Sampling (TOS) is a versatile generic framework that is able to deliver the tools for heterogeneity counteraction in the sampling stage(s), which is necessary for designing an unbiased and reproducible sampling procedure.[12]

A pilot experiment focusing on intrinsic heterogeneity characterization will always be advantageous. Different approaches for scale characterization were evaluated: embedded small-scale experimental designs in combination with larger scale 1-D transect sampling can reveal the inherent heterogeneity at scales from sampling volume up to the maximum experimental length scale studied. Thus e.g. for collecting experimental soil samples for laboratory pesticide fate studies based on *realistic* soil samples, this purpose would be served the worst by samples having inter-distance larger than the range. Emphasis should be on securing realistic, representative soil pot samples with the most *similar* characteristics, especially when deploying duplicate or replicate pot samples for such studies. It has also been demonstrated how to use representative mass reduction to get sample sizes down from field to aliquot scales in a fully representative fashion and how to counteract and manage soil heterogeneity in this process.[11,12]

Results from the various replication sampling approaches reveal considerable heterogeneities at scales from 3 cm to 100 meter. The heterogeneity in 1-D profiles can be visualized by a variogram description, the statistics of which (nugget effect, sill, range) offers a full description of all necessary and sufficient spatial characteristics of the heterogeneity. PCA score plots of the special $X_{variogram}$ matrix offer an effective overview of similarity vs. dissimilarity between variograms (especially in the case of many elements), which in the present case mainly reflect different sill levels (in general cases this will also encompass range difference).

We have developed a comprehensive approach to reach all the stated project objectives and evaluated their performances with realistic field and laboratory experiments.[11,12] The methods presented and illustrated in this Ph.D. project have a substantial carrying-over potential to geochemistry and in environmental science, as well as other application areas.

References

1. Boudreault, J.-P., Dubé, J.-S., Sona, M. & Hardy, E. Analysis of procedures for sampling contaminated soil using Gy's Sampling Theory and Practice. *Sci. Total Environ.* **425,** 199–207 (2012).
2. De Zorzi, P. *et al.* Estimation of uncertainty arising from different soil sampling devices: the use of variogram parameters. *Chemosphere* **70,** 745–52 (2008).
3. Barbizzi, S. *et al.* Characterisation of a reference site for quantifying uncertainties related to soil sampling. *Environ. Pollut.* **127,** 131–135 (2004).
4. Dubus, I. G., Brown, C. D. & Beulke, S. Sources of uncertainty in pesticide fate modelling. *Sci. Total Environ.* **317,** 53–72 (2003).
5. Chappell, A. & Viscarra Rossel, R. a. The importance of sampling support for explaining change in soil organic carbon. *Geoderma* **193-194,** 323–325 (2013).
6. DS3077. in **44,** 1–38 (Danish Standard Authority, 2013).
7. Petersen, L., Dahl, C. K. & Esbensen, K. H. Representative mass reduction in sampling—a critical survey of techniques and hardware. *Chemom. Intell. Lab. Syst.* **74,** 95–114 (2004).
8. Wagner, C. & Esbensen, K. H. A critical assessment of the HCGAgrain sampling guide. in *TOS Forum* 16–21 (2014).
9. Esbensen, K. H., Friis-Petersen, H. H., Petersen, L., Holm-Nielsen, J. B. & Mortensen, P. P. Representative process sampling — in practice: Variographic analysis and estimation of total sampling errors (TSE). *Chemom. Intell. Lab. Syst.* **88,** 41–59 (2007).
10. Esbensen, K. H. *Multivariate Data Analysis – in practice.* 597 (CAMO Software, 2010).
11. Z. Kardanpour, O.S. Jacobsen, K. H. E. Soil heterogeneity characterization using PCA (Xvariogram) – Multivariate analysis of spatial signatures for optimal sampling purposes. *Chemom. Intell. Lab. Syst.* (2014).
12. Z. Kardanpour, O. S. Jacobsen, R. K. Juhler, K. H. E. Scale-dependent Soil Heterogeneity Characterization Theory of Sampling(TOS) and Variograms. *Eur. J. Soil Sci.* (2014).